AWS ParallelCluster User Guide (v3)

AWS ParallelCluster

AWS ParallelCluster: AWS ParallelCluster User Guide (v3)

Table of Contents

What is AWS ParallelCluster

AWS ParallelCluster is an AWS supported, open source cluster management tool that helps you to deploy and manage High-Performance Computing (HPC) clusters in the AWS Cloud. It automatically sets up the required compute resources, scheduler, and shared filesystem. You can use AWS ParallelCluster with AWS Batch and Slurm schedulers.

With AWS ParallelCluster, you can quickly build and deploy proof of concept and production HPC compute environments. You can also build and deploy a high level workflow on top of AWS ParallelCluster, such as a genomics portal that automates an entire DNA sequencing workflow.

You can access AWS ParallelCluster using these methods:

- AWS ParallelCluster command line interface (CLI)
- AWS ParallelCluster API
- PCUI (added with release 3.5.0)
- AWS ParallelCluster Python library API (added with release 3.5.0)
- As an AWS CloudFormation custom resource (added with release 3.6.0)

Pricing

When using the AWS ParallelCluster command line interface (CLI) or API, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

How AWS ParallelCluster works

AWS ParallelCluster was built not only as a way to manage clusters, but as a reference on how to use AWS services to build your HPC environment. The following topics describe AWS ParallelCluster processes, the AWS services that AWS ParallelCluster uses and how, and the internal directories.

Topics

AWS ParallelCluster processes

This section applies to clusters that are deployed with Slurm. When used with this scheduler, AWS ParallelCluster interacts with the underlying job scheduler to manage compute node provisioning and removal.

For HPC clusters that are based on AWS Batch, AWS ParallelCluster relies on the capabilities provided by AWS Batch to manage compute nodes.

`clustermgtd`

The cluster management daemon (`clustermgtd`) performs these tasks:

- Clean up inactive partitions
- Manage Slurm reservations and nodes associated with Capacity Blocks (see the following section)
- Manage static capacity to make sure it is always up and healthy
- Sync scheduler with Amazon EC2.
- Clean up orphaned instances
- Restore scheduler node status upon an Amazon EC2 termination that happens outside of the suspend workflow
- Manage unhealthy Amazon EC2 instances (those that fail Amazon EC2 health checks)
- Manage scheduled maintenance events

- Manage unhealthy scheduler nodes (those that fail scheduler health checks)

Management of Slurm reservations and nodes associated with Capacity Blocks

ParallelCluster supports On-Demand Capacity Reservations (ODCR) and Capacity Blocks for Machine Learning (CB). Unlike ODCR, CB can have a future start time and is time-bound.

`clustermgtd` searches for unhealthy nodes in a loop, terminates any Amazon EC2 instances that are down, and replaces them with new instances if they are static nodes.

AWS ParallelCluster manages static nodes associated with Capacity Blocks differently– it creates a cluster even if the CB is not yet active, and automatically launches instances once the CB is active.

The Slurm nodes that correspond to compute resources associated with CBs that are not yet active are kept in the maintenance state until the CB start time is reached. These Slurm nodes remain in a reservation/maintenance state associated with the Slurm admin user, which means they can accept jobs, but the jobs remain pending until the Slurm reservation is removed.

`clustermgtd` automatically creates or deletes Slurm reservations– it puts the related CB nodes in a maintenance state based on the CB state. When the CB becomes active, the Slurm reservation is removed, the nodes start and become available for the pending jobs or for new job submissions.

When the CB end time is reached, the nodes are moved back to a reservation/maintenance state. It's up to users to resubmit/requeue the jobs to a new queue/compute resource when the CB is no longer active and instances are terminated.

`clusterstatusmgtd`

The cluster status management daemon (`clusterstatusmgtd`) manages the compute fleet status update. Every minute it fetches the fleet status stored in a DynamoDB table and manages any STOP/START request.

`computemgtd`

The compute management daemon (`computemgtd`) processes run on each of the cluster compute nodes. Every five (5) minutes, the compute management daemon confirms that the head node can be reached and is healthy. If five (5) minutes pass during which the head node cannot be reached or is not healthy, the compute node is shut down.

AWS services used by AWS ParallelCluster

The following Amazon Web Services (AWS) services are used by AWS ParallelCluster.

Topics

Amazon API Gateway

Amazon API Gateway is an AWS service that makes it possible to create, publish, maintain, monitor, and secure REST, HTTP, and WebSocket APIs at any scale

AWS ParallelCluster uses API Gateway to host the AWS ParallelCluster API.

For more information about Amazon API Gateway, see https://aws.amazon.com/api-gateway/ and https://docs.aws.amazon.com/apigateway/.

AWS Batch

AWS Batch is an AWS managed job scheduler service. It dynamically provisions the optimal quantity and type of compute resources (for example, CPU or memory-optimized instances) in AWS Batch clusters. These resources are provisioned based on the specific requirements of your batch jobs, including volume requirements. With AWS Batch, you don't need to install or manage additional batch computing software or server clusters to run your jobs effectively.

AWS Batch is used only with AWS Batch clusters.

For more information about AWS Batch, see https://aws.amazon.com/batch/ and https://docs.aws.amazon.com/batch/.

AWS CloudFormation

AWS CloudFormation is an infrastructure-as-code service that provides a common language to model and provision AWS and third-party application resources in your cloud environment. It is the main service used by AWS ParallelCluster. Each cluster in AWS ParallelCluster is represented as a stack, and all resources required by each cluster are defined within the AWS ParallelCluster CloudFormation template. In most cases, AWS ParallelCluster CLI commands directly correspond to AWS CloudFormation stack commands, such as create, update, and delete. Instances that are launched within a cluster make HTTPS calls to the AWS CloudFormation endpoint in the AWS Region where the cluster is launched.

For more information about AWS CloudFormation, see https://aws.amazon.com/cloudformation/ and https://docs.aws.amazon.com/cloudformation/.

Amazon CloudWatch

Amazon CloudWatch (CloudWatch) is a monitoring and observability service that provides you with data and actionable insights. These insights can be used to monitor your applications,

respond to performance changes and service exceptions, and optimize resource utilization. In AWS ParallelCluster, CloudWatch is used for a dashboard, to monitor and log Docker image build steps and the output of the AWS Batch jobs.

Before AWS ParallelCluster version 2.10.0, CloudWatch was used only with AWS Batch clusters.

For more information about CloudWatch, see https://aws.amazon.com/cloudwatch/ and https://docs.aws.amazon.com/cloudwatch/.

Amazon CloudWatch Events

Amazon CloudWatch Events (CloudWatch Events) delivers a near real-time stream of system events that describe changes in Amazon Web Services (AWS) resources. Using simple rules that you can quickly set up, you can match events and route them to one or more target functions or streams. In AWS ParallelCluster, CloudWatch Events is used for AWS Batch jobs.

For more information about CloudWatch Events, see https://docs.aws.amazon.com//eventbridge/latest/userguide/eb-cwe-now-eb.

Amazon CloudWatch Logs

Amazon CloudWatch Logs (CloudWatch Logs) is one of the core features of Amazon CloudWatch. You can use it to monitor, store, view, and search the log files for many of the components used by AWS ParallelCluster.

Before AWS ParallelCluster version 2.6.0, CloudWatch Logs was only used with AWS Batch clusters.

For more information, see Integration with Amazon CloudWatch Logs.

AWS CodeBuild

AWS CodeBuild (CodeBuild) is an AWS managed continuous integration service that compiles source code, runs tests, and produces software packages that are ready to deploy. In AWS ParallelCluster, CodeBuild is used to automatically and transparently build Docker images when clusters are created.

CodeBuild is used only with AWS Batch clusters.

For more information about CodeBuild, see https://aws.amazon.com/codebuild/ and https://docs.aws.amazon.com/codebuild/.

Amazon DynamoDB

Amazon DynamoDB (DynamoDB) is a fast and flexible NoSQL database service. It is used to store the minimal state information of the cluster. The head node tracks provisioned instances in a DynamoDB table.

DynamoDB is not used with AWS Batch clusters.

For more information about DynamoDB, see https://aws.amazon.com/dynamodb/ and https://docs.aws.amazon.com/dynamodb/.

Amazon Elastic Block Store

Amazon Elastic Block Store (Amazon EBS) is a high-performance block storage service that provides persistent storage for shared volumes. All Amazon EBS settings can be passed through the configuration. Amazon EBS volumes can either be initialized empty or from an existing Amazon EBS snapshot.

For more information about Amazon EBS, see https://aws.amazon.com/ebs/ and https://docs.aws.amazon.com/ebs/.

Amazon Elastic Compute Cloud

Amazon Elastic Compute Cloud (Amazon EC2) provides the computing capacity for AWS ParallelCluster. The head and compute nodes are Amazon EC2 instances. Any instance type that supports hardware virtual machine (HVM) can be selected. The head and compute nodes can be different instance types. Moreover, if multiple queues are used, some or all of compute nodes can also be launched as a Spot Instance. Instance store volumes found on the instances are mounted as striped Logical Volume Manager (LVM) volumes.

For more information about Amazon EC2 , see https://aws.amazon.com/ec2/ and https://docs.aws.amazon.com/ec2/.

Amazon Elastic Container Registry

Amazon Elastic Container Registry (Amazon ECR) is a fully managed Docker container registry that makes it easy to store, manage, and deploy Docker container images. In AWS ParallelCluster, Amazon ECR stores the Docker images that are built when clusters are created. The Docker images are then used by AWS Batch to run the containers for the submitted jobs.

Amazon ECR is used only with AWS Batch clusters.

For more information, see https://aws.amazon.com/ecr/ and https://docs.aws.amazon.com/ecr/.

Amazon EFS

Amazon Elastic File System (Amazon EFS) provides a simple, scalable, and fully managed elastic NFS file system for use with AWS Cloud services and on-premises resources. Amazon EFS is used when the `EfsSettings` are specified. Support for Amazon EFS was added in AWS ParallelCluster version 2.1.0.

For more information about Amazon EFS, see https://aws.amazon.com/efs/ and https://docs.aws.amazon.com/efs/.

Amazon FSx for Lustre

FSx for Lustre provides a high-performance file system that uses the open-source Lustre file system. FSx for Lustre is used when the `FsxLustreSettings` properties are specified. Support for FSx for Lustre was added in AWS ParallelCluster version 2.2.1.

For more information about FSx for Lustre, see https://aws.amazon.com/fsx/lustre/ and https://docs.aws.amazon.com/fsx/.

Amazon FSx for NetApp ONTAP

FSx for ONTAP provides a fully managed shared storage system built on NetApp's popular ONTAP file system. FSx for ONTAP is used when `FsxOntapSettings` properties are specified. Support for FSx for ONTAP was added in AWS ParallelCluster version 3.2.0.

For more information about FSx for ONTAP, see https://aws.amazon.com/fsx/netapp-ontap/ and https://docs.aws.amazon.com/fsx/.

Amazon FSx for OpenZFS

FSx for OpenZFS provides a fully managed shared storage system built on the popular OpenZFS file system. FSx for OpenZFS is used when the `FsxOpenZfsSettings` properties are specified. Support for FSx for OpenZFS was added in AWS ParallelCluster version 3.2.0.

For more information about FSx for OpenZFS, see https://aws.amazon.com/fsx/openzfs/ and https://docs.aws.amazon.com/fsx/.

AWS Identity and Access Management

AWS Identity and Access Management (IAM) is used within AWS ParallelCluster to provide a least privileged IAM role for Amazon EC2 for the instance that is specific to each individual cluster. AWS ParallelCluster instances are given access only to the specific API calls that are required to deploy and manage the cluster.

With AWS Batch clusters, IAM roles are also created for the components that are involved with the Docker image building process when clusters are created. These components include the Lambda functions that are allowed to add and delete Docker images to and from the Amazon ECR repository. They also include the functions allowed to delete the Amazon S3 bucket that is created for the cluster and CodeBuild project. There are also roles for AWS Batch resources, instances, and jobs.

For more information about IAM, see https://aws.amazon.com/iam/ and https://docs.aws.amazon.com/iam/.

AWS Lambda

AWS Lambda (Lambda) runs the functions that orchestrate the creation of Docker images. Lambda also manages the cleanup of custom cluster resources, such as Docker images stored in the Amazon ECR repository and on Amazon S3.

For more information about Lambda, see https://aws.amazon.com/lambda/ and https://docs.aws.amazon.com/lambda/.

Amazon RDS

Amazon Relational Database Service (Amazon RDS) is a web service that makes it easier to set up, operate, and scale a relational database in the AWS Cloud.

AWS ParallelCluster uses Amazon RDS for AWS Batch and Slurm.

For more information about Amazon RDS, see https://aws.amazon.com/rds/ and https://docs.aws.amazon.com/rds/.

Amazon Route 53

Amazon Route 53 (Route 53) is used to create hosted zones with hostnames and fully qualified domain names for each of the compute nodes.

For more information about Route 53, see https://aws.amazon.com/route53/ and https://docs.aws.amazon.com/route53/.

Amazon Simple Notification Service

(Amazon SNS) is a managed service that provides message delivery from publishers to subscribers (also known as producers and consumers).

AWS ParallelCluster uses Amazon SNS for API hosting.

For more information about Amazon SNS, see https://aws.amazon.com/sns/ and https://docs.aws.amazon.com/sns/.

Amazon Simple Storage Service

Amazon Simple Storage Service (Amazon S3) stores AWS ParallelCluster templates located in each AWS Region. AWS ParallelCluster can be configured to allow CLI/SDK tools to use Amazon S3.

AWS ParallelCluster also creates an Amazon S3 bucket in your AWS account to store resources that are used by your clusters, such as the cluster configuration file. AWS ParallelCluster maintains one Amazon S3 bucket in each AWS Region that you create clusters in.

When you use AWS Batch cluster, an Amazon S3 bucket in your account is used for storing related data. For example, the bucket stores artifacts created when a Docker image and scripts are created from submitted jobs.

For more information, see https://aws.amazon.com/s3/ and https://docs.aws.amazon.com/s3/.

Amazon VPC

An Amazon Virtual Private Cloud (VPC) defines a network used by the nodes in your cluster.

For more information about Amazon VPC, see https://aws.amazon.com/vpc/ and https://docs.aws.amazon.com/vpc/.

Elastic Fabric Adapter

Elastic Fabric Adapter (EFA) is a network interface for instances that you can use to run applications requiring high levels of inter-node communications at scale on AWS.

For more information about Elastic Fabric Adapter, see https://aws.amazon.com/hpc/efa/.

EC2 Image Builder

EC2 Image Builder is a fully managed AWS service that helps you to automate the creation, management, and deployment of customized, secure, and up-to-date server images.

AWS ParallelCluster uses Image Builder to create and manage AWS ParallelCluster images.

For more information about EC2 Image Builder, see https://aws.amazon.com/image-builder/ and https://docs.aws.amazon.com/imagebuilder/.

Amazon DCV

Amazon DCV is a high-performance remote display protocol that provides a secure way to deliver remote desktops and application streaming to any device over varying network conditions. Amazon DCV is used when the HeadNode section / Dcv settings are specified. Support for Amazon DCV was added in AWS ParallelCluster version 2.5.0.

For more information about Amazon DCV, see https://aws.amazon.com/hpc/dcv/ and https://docs.aws.amazon.com/dcv/.

AWS ParallelCluster internal directories

There are several internal directories that AWS ParallelCluster uses to share data within the cluster. The following directories are shared between the head node, compute nodes, and login nodes:

- `/opt/slurm`
- `/opt/intel`
- `/opt/parallelcluster/shared (only with compute nodes)`
- `/opt/parallelcluster/shared_login_nodes (only with login nodes)`
- `/home (unless specified in SharedStorage)`

> **Note**
>
> By default these directories are created on the head nodes EBS volume and shared as NFS exports to the compute and login nodes. Starting from AWS ParallelCluster 3.8 you can enable AWS ParallelCluster to create and manage an Amazon EFS filesystem to host and share these directories by setting the SharedStorageType parameter to efs.

When the cluster scales out, NFS exports via the EBS volume may pose performance bottlenecks. Using EFS, you can avoid NFS exports as your cluster scales out and avoid the performance bottlenecks associated with them.

Setting up AWS ParallelCluster

The following topics describe how to set up AWS ParallelCluster. You will learn how to install the necessary tools and how to use them, how to implement and manage multiple user access to clusters, and the best practices.

Topics

Prerequisites

Before you can start setting up and using AWS ParallelCluster, make sure that you've completed the following prerequisites.

Setting up an AWS account

Set up an AWS account to use AWS ParallelCluster.

Sign up for an AWS account

If you do not have an AWS account, complete the following steps to create one.

To sign up for an AWS account

1. Open https://portal.aws.amazon.com/billing/signup.
2. Follow the online instructions.

   Part of the sign-up procedure involves receiving a phone call or text message and entering a verification code on the phone keypad.

When you sign up for an AWS account, an *AWS account root user* is created. The root user has access to all AWS services and resources in the account. As a security best practice, assign administrative access to a user, and use only the root user to perform tasks that require root user access.

AWS sends you a confirmation email after the sign-up process is complete. At any time, you can view your current account activity and manage your account by going to https://aws.amazon.com/ and choosing **My Account**.

Create a user with administrative access

After you sign up for an AWS account, secure your AWS account root user, enable AWS IAM Identity Center, and create an administrative user so that you don't use the root user for everyday tasks.

Secure your AWS account root user

1. Sign in to the AWS Management Console as the account owner by choosing **Root user** and entering your AWS account email address. On the next page, enter your password.

   For help signing in by using root user, see Signing in as the root user in the *AWS Sign-In User Guide*.

2. Turn on multi-factor authentication (MFA) for your root user.

   For instructions, see Enable a virtual MFA device for your AWS account root user (console) in the *IAM User Guide*.

Create a user with administrative access

1. Enable IAM Identity Center.

   For instructions, see Enabling AWS IAM Identity Center in the *AWS IAM Identity Center User Guide*.

2. In IAM Identity Center, grant administrative access to a user.

   For a tutorial about using the IAM Identity Center directory as your identity source, see Configure user access with the default IAM Identity Center directory in the *AWS IAM Identity Center User Guide*.

Sign in as the user with administrative access

- To sign in with your IAM Identity Center user, use the sign-in URL that was sent to your email address when you created the IAM Identity Center user.

  For help signing in using an IAM Identity Center user, see Signing in to the AWS access portal in the *AWS Sign-In User Guide*.

Assign access to additional users

1. In IAM Identity Center, create a permission set that follows the best practice of applying least-privilege permissions.

   For instructions, see Create a permission set in the *AWS IAM Identity Center User Guide*.
2. Assign users to a group, and then assign single sign-on access to the group.

   For instructions, see Add groups in the *AWS IAM Identity Center User Guide*.

Create a key pair

To deploy clusters, AWS ParallelCluster launches Amazon EC2 instances to create the cluster head node and compute nodes. To perform cluster tasks, such as running and monitoring jobs, or managing users, you must be able to access the cluster head node. To verify you can access the head node instance using SSH, you must use an Amazon EC2 key pair. To learn how to create a key pair, see Create a key pair in the *Amazon Elastic Compute Cloud* User Guide for Linux Instances.

Installing the AWS ParallelCluster command line interface (CLI)

AWS ParallelCluster is distributed as a Python package and is installed using the Python `pip` package manager. For instructions on how to install Python packages, see Installing packages in the *Python Packaging User Guide*.

Ways to install AWS ParallelCluster:

- Install AWS ParallelCluster in a virtual environment (recommended)
- Installing AWS ParallelCluster in a non-virtual environment using pip
- Install AWS ParallelCluster as a standalone application

You can find the version number of the most recent CLI on the releases page on GitHub. In this guide, the command examples assume that you have installed a version of Python that is later than version 3.6. The `pip` command examples use the `pip3` version.

Manage both AWS ParallelCluster 2 and AWS ParallelCluster 3

If you use both AWS ParallelCluster 2 and AWS ParallelCluster 3 and want to manage the CLIs for both packages, we recommend that you install AWS ParallelCluster 2 and AWS ParallelCluster 3 in different virtual environments. This ensures that you can continue using each version of AWS ParallelCluster and any associated cluster resources.

Install AWS ParallelCluster in a virtual environment (recommended)

We recommend that you install AWS ParallelCluster in a virtual environment to avoid requirement version conflicts with other `pip` packages.

Prerequisites

- AWS ParallelCluster requires Python 3.7 or later. If you don't already have it installed, download a compatible version for your platform at python.org.

To install AWS ParallelCluster in a virtual environment

1. If `virtualenv` isn't installed, install `virtualenv` using `pip3`. If `python3 -m virtualenv help` displays help information, go to step 2.

   ```
   $ python3 -m pip install --upgrade pip
   $ python3 -m pip install --user --upgrade virtualenv
   ```

   Run `exit` to leave the current terminal window and open a new terminal window to pick up changes to the environment.

2. Create a virtual environment and name it.

   ```
   $ python3 -m virtualenv ~/apc-ve
   ```

   Alternatively, you can use the `-p` option to specify a specific version of Python.

   ```
   $ python3  -m virtualenv -p $(which python3) ~/apc-ve
   ```

3. Activate your new virtual environment.

```
$ source ~/apc-ve/bin/activate
```

4. Install AWS ParallelCluster into your virtual environment.

```
(apc-ve)~$ python3 -m pip install --upgrade "aws-parallelcluster"
```

5. Install Node Version Manager and the latest Long-Term Support (LTS) Node.js version. AWS Cloud Development Kit (AWS CDK) requires Node.js for CloudFormation for template generation.

> **Note**
>
> If your Node.js installation isn't working on your platform, you can install an LTS version prior to the latest LTS version. For more information, see the Node.js release schedule and the AWS CDK prerequisites.
> Example Node.js installation command:
>
> ```
> $  nvm install --lts=Hydrogen
> ```

```
$ curl -o- https://raw.githubusercontent.com/nvm-sh/nvm/v0.38.0/install.sh | bash
$ chmod ug+x ~/.nvm/nvm.sh
$ source ~/.nvm/nvm.sh
$ nvm install --lts
$ node --version
```

6. Verify that AWS ParallelCluster is installed correctly.

```
$ pcluster version
{
  "version": "3.13.2"
}
```

You can use the `deactivate` command to exit the virtual environment. Each time you start a session, you must reactivate the environment.

To upgrade to the latest version of AWS ParallelCluster, run the installation command again.

```
(apc-ve)~$ python3 -m pip install --upgrade "aws-parallelcluster"
```

Installing AWS ParallelCluster in a non-virtual environment using pip

You can also install AWS ParallelCluster in a non-virtual environment using pip, a package manager for Python packages.

Prerequisites

- AWS ParallelCluster requires Python 3.7 or later. If you don't already have it installed, download a compatible version for your platform at python.org.

Install AWS ParallelCluster

1. Use `pip` to install AWS ParallelCluster.

   ```
   $ python3 -m pip install "aws-parallelcluster" --upgrade --user
   ```

   When you use the `--user` switch, `pip` installs AWS ParallelCluster to `~/.local/bin`.

2. Install Node Version Manager and the latest Long-Term Support (LTS) Node.js version. AWS Cloud Development Kit (AWS CDK) requires Node.js for CloudFormation for template generation.

   Note

   If your Node.js installation isn't working on your platform, you can install an LTS version prior to the latest LTS version. For more information, see the Node.js release schedule and the AWS CDK prerequisites.

   ```
   $  nvm install --lts=Hydrogen
   ```

   ```
   $ curl -o- https://raw.githubusercontent.com/nvm-sh/nvm/v0.38.0/install.sh | bash
   $ chmod ug+x ~/.nvm/nvm.sh
   $ source ~/.nvm/nvm.sh
   $ nvm install --lts
   $ node --version
   ```

3. Verify that AWS ParallelCluster installed correctly.

```
$ pcluster version
{
  "version": "3.13.2"
}
```

4. To upgrade to the latest version, run the installation command again.

```
$ python3 -m pip install "aws-parallelcluster" --upgrade --user
```

Install AWS ParallelCluster as a standalone application

Install AWS ParallelCluster as a standalone application on your environment. Follow the instructions for installing AWS ParallelCluster on an available OS in the following section.

Prerequisites

- An environment with an operating system compatible with an available version of the installer.

> **Note**
>
> AWS ParallelCluster requires NodeJS. AWS ParallelCluster Installer includes a bundled version of NodeJS (v18), which is installed if it does not already exist. If your system is not compatible with NodeJS v18, you should install NodeJS before installing AWS ParallelCluster.

Linux

Linux x86 (64-bit)

Install AWS ParallelCluster on your environment.

1. Download the latest pcluster installer (checksum).
2. Unzip the installer bundle and install AWS ParallelCluster by using the following commands:

```
$ unzip pcluster-installer-bundle-<VERSION>-Linux_x86_64-signed.zip -d pcluster-
installer-bundle
$ cd pcluster-installer-bundle
```

```
$ chmod +x install_pcluster.sh
```

3. Run the following install script.

```
$ bash install_pcluster.sh
```

4. Verify that AWS ParallelCluster is installed correctly.

```
$ pcluster version
{
    "version": "3.13.2"
}
```

Troubleshooting `pcluster` installation errors

- If the AWS ParallelCluster version isn't returned in step 4, restart the terminal or `source` the `bash_profile` to update the `PATH` variable to include the new binary directory as shown in the following example:

```
$ source ~/.bash_profile
```

- If you use your `pcluster` installation to create clusters with `CustomActions` specified as HTTPS resources, rather than S3 URIs, you might see a `WARNING` message indicating that these resources might not be verified ([`SSL: CERTIFICATE_VERIFY_FAILED`]). This is caused by a known issue and you can ignore this warning if you trust the authenticity of the specified resources.

Previous installer bundle versions

- None

Steps to take after installation

This section assumes that you've installed AWS ParallelCluster. You will learn how to verify that AWS ParallelCluster installed correctly, how to update to the latest version of AWS ParallelCluster, and how to uninstall.

You can verify that AWS ParallelCluster was installed correctly by running `pcluster version`.

```
$ pcluster version
{
"version": "3.13.2"
}
```

AWS ParallelCluster is updated regularly. To update to the latest version of AWS ParallelCluster, run the installation command again. For more information about the latest version of AWS ParallelCluster, see the AWS ParallelCluster release notes.

```
$ pip3 install aws-parallelcluster --upgrade --user
```

To uninstall AWS ParallelCluster, use `pip3 uninstall`.

```
$ pip3 uninstall aws-parallelcluster
```

If you don't have Python and `pip3`, use the procedure for your environment.

Installing the PCUI

The AWS ParallelCluster UI (PCUI) is a web-based user interface that mirrors the AWS ParallelCluster `pcluster` CLI, while providing a console-like experience. You install and access the PCUI in your AWS account. When you run it, the PCUI accesses an instance of the AWS ParallelCluster API hosted on Amazon API Gateway in your AWS account. For more information about the PCUI, see AWS ParallelCluster UI.

Prerequisites:

- You must have an AWS account
- You must have access to the AWS Management Console

For more information, see Setting up an AWS account.

Topics

- Identify the AWS ParallelCluster and PCUI version
- PCUI costs

Install the PCUI

To install an instance of the AWS ParallelCluster UI (PCUI), choose an AWS CloudFormation quick-create link for the AWS Region in which you create clusters. The quick-create URL takes you to a **Create Stack Wizard** where you provide quick-create stack template inputs and deploy the stack. For more information about CloudFormation quick-create stacks, see Creating quick-create links for stacks in the *AWS CloudFormation User Guide*.

> **Note**
>
> You can only create and edit clusters or build images with the same AWS ParallelCluster version that you use to install the PCUI.

Use an AWS CloudFormation quick-create link to deploy a PCUI stack with nested Amazon Cognito, API Gateway, and Amazon EC2 Systems Manager stacks.

1. Sign in to the AWS Management Console.
2. Deploy the PCUI by choosing an AWS Region quick-create link from the list here. This takes you to the CloudFormation **Create Stack Wizard** in the console.

   - us-east-1
   - us-east-2
   - us-west-1
   - us-west-2
   - eu-west-1
   - eu-west-2
   - eu-west-3
   - eu-central-1
   - eu-north-1
   - me-south-1
   - sa-east-1

- ca-central-1
- ap-northeast-1
- ap-northeast-2
- ap-south-1
- ap-southeast-1
- ap-southeast-2
- us-gov-west-1

3. Enter a valid email address for **Admin's Email** and a **ParallelCluster version**.

   After deployment completes successfully, the PCUI will send a temporary password to this email address that you can use to access the PCUI. If you delete the email before you save or use the temporary password, you must delete the stack and reinstall the PCUI.

4. Keep the rest of the form blank or enter values for (optional) parameters to customize the PCUI build.
5. Note the stack name for use in later steps.
6. Navigate to **Capabilities**. Agree to the CloudFormation capabilities.
7. Choose **Create**. It takes about 15 minutes to complete the AWS ParallelCluster API and PCUI deployment.
8. View the stack details as the stack is created.
9. After the deployment completes, open the admin email that was sent to the address you entered and that contains the temporary password. Use that to access the PCUI. (Remember, if you permanently delete the email before you log in to the PCUI, you must delete the PCUI stack you created and reinstall the PCUI.
10. In the AWS CloudFormation console list of stacks, choose the link to the stack name that you noted in a previous step.
11. In **Stack details**, choose **Outputs** and select the link for the key named ***`Stackname`*****URL** to open the PCUI (where ***`Stackname`*** is the name that you noted in a previous step).
12. Enter the temporary password. Follow the steps to create your own password and log in again.
13. You are now on the home page of the PCUI in the AWS Region that you selected.
14. To get started using the PCUI, see Configure and create a cluster with the PCUI.

> **Note**
>
> PCUI sessions have a default duration of 5 minutes, which is the minimum value provided by Cognito as of PCUI 2023.12.0. So, it is expected that a user removed from Cognito User Pools will still able to access the system until the session expires.

Stack parameters

AdditionalPoliciesPCAPI:

Description: (Optional) ARN of the additional IAM policy to be attached to the default execution role for the ParallelCluster Lambda function. Only one policy can be specified.

Type: String

Default: ''

AllowedPattern: "^(arn:.*:iam::.*:policy\\/([a-zA-Z0-9_-]+))|()$"

AdminUserEmail:

Description: Email address of administrative user to setup by default (only with new Cognito instances).

Type: String

Default: ''

CognitoCustomDomain:

Description: (Optional) Custom domain name for Cognito. If omitted, the default Cognito domain name will be used.

Type: String

Default: ''

CognitoCustomDomainCertificateArn:

Description: (Optional) ARN of the ACM Certificate issued for the Cognito custom domain. This is required only if `CognitoCustomDomain` is specified.

Type: String

Default: ''

CustomDomain:

Description: (Optional) Custom domain name. If omitted, the default domain name will be used.

Type: String

Default: ''

CustomDomainCertificateArn:

Description: (Optional) ARN of the ACM Certificate issued for the custom domain. This is required only if `CustomDomain` is specified.

Type: String

Default: ''

IAMRoleAndPolicyPrefix:

Description: Prefix applied to the name of every IAM role and policy (max length: 10). [ParallelCluster >= 3.8.0]

Type: String

Default: ''

MaxLength: 10

ImageBuilderSubnetId:

Description: (Optional) Select the subnet to use for building the container images. The subnet must be public and auto-assign public IPs. If not selected, the default Subnet will be used.

Type: String

Default: ''

ImageBuilderVpcId:

Description: (Optional) Select the VPC to use for building the container images. If not selected, the default VPC will be used.

Type: String

Default: ''

InfrastructureBucket:

Description: (Optional) S3 bucket where CloudFormation files are stored. Change this parameter only when testing changes made to the infrastructure itself.

Type: String

Default: ''

LambdaSecurityGroupIds:

Description: Comma separated list of security groups to be associated with the PCUI Lambda function.

Type: CommaDelimitedList

Default: ''

LambdaSubnetIds:

Description: Comma separated list of subnet IDs to be associated with the PCUI Lambda function. These subnets should be private and associated with your VPC endpoint.

Type: CommaDelimitedList

Default: ''

PermissionsBoundaryPolicy:

Description: ARN of the IAM policy to use as permissions boundary for every IAM role created by ParallelCluster UI infrastructure.'

Type: String

Default: ''

AllowedPattern: "^(arn:.*:iam::.*:policy\\/([a-zA-Z0-9_-]+))|()$"

PermissionsBoundaryPolicyPCAPI:

Description: ARN of the IAM policy to use as permissions boundary for every IAM role created by ParallelCluster API infrastructure. [ParallelCluster >= 3.8.0]

Type: String

Default: ''

AllowedPattern: "^(arn:.*:iam::.*:policy\\/([a-zA-Z0-9_-]+))|()$"

PublicEcrImageUri:

Description: When specified, the URI of the Docker image for the Lambda of the ParallelCluster UI container.

Type: String

Default: public.ecr.aws/pcm/parallelcluster-ui:2024.11.0

SNSRole:

Description: SNSRole ARN of a previously deployed PCUI Cognito Stack. Leave blank to create a new one.

Type: String

Default: ''

UserPoolAuthDomain:

Description: UserPoolAuthDomain of a previously deployed PCUI Cognito User Pool. Leave blank to create a new one.

Type: String

Default: ''

UserPoolId:

Description: UserPoolId of a previously deployed PCUI Cognito User Pool. Leave blank to create a new one.

Type: String

Default: ''

Version:

Description: Version of AWS ParallelCluster to deploy.

Type: String

Default: 3.11.1

VpcEndpointId:

Description: Enter a VPC endpoint with type interface for the execute-api service to enable private PCUI implementation. When enabled, the API will only accept requests from within the given VPC.

Type: String

Default: ''

Configure a custom domain

Learn how to configure a custom domain for the AWS ParallelCluster UI (PCUI). The UI is hosted on Amazon API Gateway in your AWS account. You can configure a custom domain in the API Gateway console.

Prerequisites:

- You have an AWS account.
- You own a domain.
- The domain root is resolvable. For example, if you want to use `xyz.example.com` as your custom domain, then the root domain `example.com` must be resolvable. This means that you must have at least an A record for `example.com` pointing to an IP. If you don't have such record, add a dummy record pointing to 0.0.0.0.
- You have created or imported the custom domain certificate for Cognito into AWS Certificate Manager (ACM) in us-east-1.
- You have the necessary permissions to change the Domain Name System (DNS) settings.

Topics

Deploy PCUI

Complete the following steps to create, or update, your AWS ParallelCluster UI (PCUI) deployment so that it will use your custom domain.

In this example we assume that you want to deploy PCUI with a custom domain `xyz.example.com` and the Amazon Cognito interface with a custom domain `auth-xyz.example.com`.

> **Note**
>
> Please note that customizing the Amazon Cognito domain is not required and can be left with the default value. We include this customization for completeness.

> **Note**
>
> Custom domains for Amazon Cognito are not supported in the AWS GovCloud (US) Regions.

To accomplish this, deploy the PCUI stack with the following parameters:

- **CustomDomain:** *`xyz.example.com`*.
- **CustomDomainCertificateArn:** the ACM certificate Amazon Resource Name (ARN) for *`xyz.example.com`*.
- **CognitoCustomDomain:** *`auth-xyz.example.com`*.
- **CognitoCustomDomainCertificateArn** to the ACM certificate Amazon Resource Name (ARN) for *`xyz.example.com`*.

Configure DNS

Complete the following steps to configure your domain name service (DNS) so that AWS ParallelCluster UI (PCUI) and Amazon Cognito respond on the desired custom domains.

This example assumes that you want to deploy PCUI with custom domain `xyz.example.com` and the Amazon Cognito interface with custom domain `auth-xyz.example.com`.

1. Deploy the PCUI stack with the following parameters:

   - **CustomDomainEndpoint:** Create an A record in your DNS for *`xyz.example.com`* that points to the alias specified in the output.

- **CognitoCustomDomainEndpoint:** Create an A record in your DNS for *auth-xyz.example.com* that points to the alias specified in the output.

2. Wait about 10 minutes so that the DNS changes can be distributed.
3. When the DNS changes are distributed, PCUI responds on *xyz.example.com/pcui* and the Amazon Cognito authentication page responds on *auth-xyz.example.com*.

Amazon Cognito user pool options

The following sections refer to CloudFormation quick-create links or quick-create URLs. A quick-create URL takes you to a **Create Stack Wizard** where you provide quick-create stack template inputs and deploy the stack. For more information about CloudFormation quick-create stacks, see Creating quick-create links for stacks in the *AWS CloudFormation User Guide*.

To maintain an Amazon Cognito user pool that you can use with multiple AWS ParallelCluster UI (PCUI) instances, consider the following options:

- Use an existing PCUI instance that links to an Amazon Cognito user pool created from a nested CloudFormation stack. This is what is created when you deploy the PCUI by using the quick-create link and keep all Amazon Cognito parameters blank.
- First, deploy a standalone Amazon Cognito user pool. Then, deploy a new PCUI instance that's linked to the standalone Amazon Cognito user pool that you just deployed. This way, you separate the Amazon Cognito deployment from the PCUI deployment. Note that non-nested PCUI CloudFormation stacks are easier to update.

Use an existing Amazon Cognito user pool with a new PCUI instance

Complete the following steps to use an existing Amazon Cognito user pool with a new PCUI instance

Use an existing Amazon Cognito user pool with a new PCUI instance

1. In the **CloudFormation console**, select the PCUI stack that contains the Amazon Cognito user pool that you want to use with multiple PCUI instances.
2. Navigate to the nested stack that created the Amazon Cognito userpool.
3. Select the **Outputs** tab.
4. Copy the values of the following parameters:

- `UserPoolId`
- `UserPoolAuthDomain`
- `SNSRole`

5. Deploy a new PCUI instance by using the quick-create link, and fill in all `External PCUI Amazon Cognito` parameters with the outputs that you copied. This prevents the new PCUI stack from creating a new pool and links it to the existing Amazon Cognito user pool that was created from a nested stack. You can deploy subsequent new PCUI instances that have the same parameter values, and you can link them to the Amazon Cognito user pool.

Create a standalone Amazon Cognito userpool

To create a standalone Amazon Cognito userpool

1. Launch an Amazon Cognito-only stack by choosing a quick-create link labeled with same AWS Region in which you deploy your PCUI instances from this list:

   - us-east-1
   - us-east-2
   - us-west-1
   - us-west-2
   - eu-west-1
   - eu-west-2
   - eu-west-3
   - eu-central-1
   - eu-north-1
   - me-south-1
   - sa-east-1
   - ca-central-1
   - ap-northeast-1
   - ap-northeast-2
   - ap-south-1
   - ap-southeast-1

- ap-southeast-2
- us-gov-west-1

2. After stack creation completes, select the **Outputs** tab and copy the values of the following parameters:

   - `UserPoolId`
   - `UserPoolAuthDomain`
   - `SNSRole`

3. Deploy a new PCUI instance by choosing an PCUI quick-start link and filling in all `External PCUI Amazon Cognito` parameters with the values that you copied. The new PCUI instance links to the standalone Amazon Cognito user pool and doesn't create a nested stack or a new user pool. You can deploy subsequent new PCUI instances that have the same parameter values, and you can link them to the standalone Amazon Cognito user pool.

Identify the AWS ParallelCluster and PCUI version

To identify the AWS ParallelCluster and PCUI version:

1. In the CloudFormation console, select a PCUI stack.
2. Select the **Parameters** tab.
3. The AWS ParallelCluster version is the value of the parameter **Version**.
4. The PCUI version is at the end of the **PublicEcrImageUri** value. For example, if the value is `public.ecr.aws/pcui/parallelcluster-ui-awslambda:2023.02`, then the version is `2023.02`.

> **Note**
>
> To update the PCUI to the latest AWS ParallelCluster version, launch a new stack by choosing a quick-create link.

PCUI costs

The PCUI is built on a serverless architecture and you can use it within the AWS Free Tier category for most cases. The following table lists the AWS services that the PCUI depends on and their free-tier limits. Typical usage is estimated to cost less than one dollar each month.

Service	AWS Free Tier
Amazon Cognito	50,000 monthly active users
Amazon API Gateway	1 million rest API calls
AWS Lambda	1 million free requests each month and 400,000 GB-seconds of compute time each month
EC2 Image Builder	No cost, except EC2
Amazon Elastic Compute Cloud	15-minute one-time container image build
AWS CloudFormation	5 GB data (ingestion, archive storage, and data scanned by Logs Insights queries)

Getting started with AWS ParallelCluster

To get started with AWS ParallelCluster configure and create a cluster using the AWS ParallelCluster command line interface (CLI) or web-based user interface (UI). The PCUI was added in release 3.5.0.

Topics

- Configure and create a cluster with the AWS ParallelCluster command line interface
- Configure and create a cluster with the PCUI
- Connect to a cluster

Configure and create a cluster with the AWS ParallelCluster command line interface

After you install AWS ParallelCluster, complete the following configuration steps.

1. Verify that your AWS Account has a role that includes the permissions needed to run the `pcluster` CLI. For more information, see AWS ParallelCluster example `pcluster` user policies.
2. Set up your AWS credentials. For more information, see Configuring the AWS CLI in the *AWS CLI user guide*.

   ```
   $ aws configure
   AWS Access Key ID [None]: AKIAIOSFODNN7EXAMPLE
   AWS Secret Access Key [None]: wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY
   Default region name [us-east-1]: us-east-1
   Default output format [None]:
   ```
3. The AWS Region where the cluster is launched must have at least one Amazon EC2 key pair. For more information, see Amazon Elastic Compute Cloud key pairs in the *Amazon Elastic Compute Cloud User Guide for Linux Instances*.

When you use the AWS ParallelCluster command line interface (CLI), you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

Configure and create your first cluster

To create your first cluster, use the `pcluster configure` CLI command to initiate a wizard that prompts you for all of the information that's required to configure and create your cluster. The details of the sequence differ when using AWS Batch as the scheduler compared to using Slurm.

Slurm

```
$  pcluster configure --config cluster-config.yaml
```

From the list of valid AWS Region identifiers, choose the AWS Region where you want your cluster to run.

Note

The list of AWS Regions shown is based on the partition of your account, and only includes AWS Regions that are enabled for your account. For more information about enabling AWS Regions for your account, see Managing AWS Regions in the *AWS General Reference*. The example shown is from the AWS Global partition. If your account is in the AWS GovCloud (US) partition, only AWS Regions in that partition are listed (`gov-us-east-1` and `gov-us-west-1`). Similarly, if your account is in the AWS China partition, only `cn-north-1` and `cn-northwest-1` are shown. For the complete list of AWS Regions supported by AWS ParallelCluster, see Supported AWS Regions for AWS ParallelCluster.

```
Allowed values for AWS Region ID:
1. af-south-1
2. ap-east-1
3. ap-northeast-1
4. ap-northeast-2
5. ap-south-1
6. ap-southeast-1
7. ap-southeast-2
8. ap-southeast-3
9. ap-southeast-5
10. ap-southeast-7
11. ca-central-1
12. eu-central-1
13. eu-north-1
14. eu-south-1
15. eu-west-1
16. eu-west-2
17. eu-west-3
18. il-central-1
19. me-south-1
20. sa-east-1
21. us-east-1
22. us-east-2
23. us-west-1
24. us-west-2
AWS Region ID [ap-northeast-1]:
```

Choose the key pair from those that are registered with Amazon Elastic Compute Cloud in the selected AWS Region:

```
Allowed values for Amazon EC2 Key Pair Name:
1. your-key-1
2. your-key-2
Amazon EC2 Key Pair Name [your-key-1]:
```

Choose the scheduler to use with your cluster.

```
Allowed values for Scheduler:
1. slurm
2. awsbatch
Scheduler [slurm]:
```

Choose the operating system.

```
Allowed values for Operating System:
1. alinux2
2. alinux2023
3. ubuntu2404
4. ubuntu2204
5. ubuntu2004
6. rhel8
7. rhel9
Operating System [alinux2]:
```

Choose the head node instance type:

```
Head node instance type [t2.micro]:
```

Choose the queue configuration. Note: Instance type can't be specified for multiple compute resources in the same queue.

```
Number of queues [1]:
Name of queue 1 [queue1]:
Number of compute resources for queue1 [1]: 2
Compute instance type for compute resource 1 in queue1 [t2.micro]:
Maximum instance count [10]:
```

Enable EFA to run applications that require high levels of inter-instance communication at scale on AWS at no additional charge:

- Choose an instance type that supports Elastic Fabric Adapter (EFA).
- Enable EFA.
- Specify an existing Placement Group name. If you leave it blank, AWS ParallelCluster creates one for you.

```
Compute instance type for compute resource 2 in queue1 [t2.micro]: c5n.18xlarge
Enable EFA on c5n.18xlarge (y/n) [y]: y
Maximum instance count [10]:
Placement Group name []:
```

After the previous steps are completed, decide whether to use an existing VPC or let AWS ParallelCluster create a VPC for you. If you don't have a properly configured VPC, AWS ParallelCluster can create a new one for you. It either places both the head and compute nodes in the same public subnet, or only the head node in a public subnet with all compute nodes in a private subnet. If you let AWS ParallelCluster create a VPC, you must decide if all nodes are to be in a public subnet. For more information, see Network configurations.

If you configure your cluster to use instance types that have multiple network interfaces or a network card, see Network configurations for additional networking requirements.

It's possible to reach your quota for the number of VPCs allowed in a AWS Region. The default quota is five VPCs for a AWS Region. For more information about this quota and how to request an increase, see VPC and subnets in the *Amazon VPC User Guide*.

Important

VPCs created by AWS ParallelCluster do not enable VPC Flow Logs by default. VPC Flow Logs enable you to capture information about the IP traffic going to and from network interfaces in your VPCs. For more information, see VPC Flow Logs in the *Amazon VPC User Guide*.

If you let AWS ParallelCluster create a VPC, make sure that you decide whether all nodes are to be in a public subnet.

> **Note**
>
> If you choose `1. Head node in a public subnet and compute fleet in a private subnet`, AWS ParallelCluster creates a NAT gateway that results in additional cost, even if you specify free tier resources.

```
Automate VPC creation? (y/n) [n]: y
Allowed values for Availability Zone:
1. us-east-1a
2. us-east-1b
3. us-east-1c
4. us-east-1d
5. us-east-1e
6. us-east-1f
Availability Zone [us-east-1a]:
Allowed values for Network Configuration:
1. Head node in a public subnet and compute fleet in a private subnet
2. Head node and compute fleet in the same public subnet
Network Configuration [Head node in a public subnet and compute fleet in a private
 subnet]: 1
Beginning VPC creation. Please do not leave the terminal until the creation is
 finalized
```

If you don't create a new VPC, you must select an existing VPC.

If you choose to have AWS ParallelCluster create the VPC, make a note of the VPC ID so you can use the AWS CLI to delete it later.

```
Automate VPC creation? (y/n) [n]: n
Allowed values for VPC ID:
  #  id                    name                                 number_of_subnets
---  --------------------  -----------------------------------  -------------------
  1  vpc-0b4ad9c4678d3c7ad  ParallelClusterVPC-20200118031893                    2
  2  vpc-0e87c753286f37eef  ParallelClusterVPC-20191118233938                    5
VPC ID [vpc-0b4ad9c4678d3c7ad]: 1
```

After the VPC has been selected, decide whether to use existing subnets or create new ones.

```
Automate Subnet creation? (y/n) [y]: y
```

```
Creating CloudFormation stack...
Do not leave the terminal until the process has finished
```

AWS Batch

```
$ pcluster configure --config cluster-config.yaml
```

From the list of valid AWS Region identifiers, choose the AWS Region where you want your cluster to run.

> **Note**
>
> The list of AWS Regions shown is based on the partition of your account. It only includes AWS Regions that are enabled for your account. For more information about enabling AWS Regions for your account, see Managing AWS Regions in the *AWS General Reference*. The example shown is from the AWS Global partition. If your account is in the AWS GovCloud (US) partition, only AWS Regions in that partition are listed (`gov-us-east-1` and `gov-us-west-1`). Similarly, if your account is in the AWS China partition, only `cn-north-1` and `cn-northwest-1` are shown. For the complete list of AWS Regions supported by AWS ParallelCluster, see Supported AWS Regions for AWS ParallelCluster.

```
Allowed values for AWS Region ID:
1. af-south-1
2. ap-east-1
3. ap-northeast-1
4. ap-northeast-2
5. ap-south-1
6. ap-southeast-1
7. ap-southeast-2
8. ap-southeast-3
9. ap-southeast-5
10. ap-southeast-7
11. ca-central-1
12. eu-central-1
13. eu-north-1
14. eu-south-1
15. eu-west-1
```

```
16. eu-west-2
17. eu-west-3
18. il-central-1
19. me-south-1
20. sa-east-1
21. us-east-1
22. us-east-2
23. us-west-1
24. us-west-2
AWS Region ID [us-east-1]:
```

The key pair is selected from the key pairs registered with Amazon EC2 in the selected AWS Region. Choose the key pair:

```
Allowed values for Amazon EC2 Key Pair Name:
1. your-key-1
2. your-key-2
Amazon EC2 Key Pair Name [your-key-1]:
```

Choose the scheduler to use with your cluster.

```
Allowed values for Scheduler:
1. slurm
2. awsbatch
Scheduler [slurm]: 2
```

When `awsbatch` is selected as the scheduler, `alinux2` is used as the operating system. The head node instance type is entered:

```
Head node instance type [t2.micro]:
```

Choose the queue configuration. The AWS Batch scheduler only contains a single queue. The maximum size of the cluster of compute nodes is entered. This is measured in vCPUs.

```
Number of queues [1]:
Name of queue 1 [queue1]:
Maximum vCPU [10]:
```

Decide whether to use existing VPCs or let AWS ParallelCluster create VPCs for you. If you don't have a properly configured VPC, AWS ParallelCluster can create a new one. It either uses both

the head and compute nodes in the same public subnet, or only the head node in a public subnet with all nodes in a private subnet. It's possible to reach your quota on the number of VPCs allowed in a Region. The default number of VPCs is five. For more information about this quota and how to request an increase, see VPC and subnets in the *Amazon VPC User Guide*.

> **Important**
>
> VPCs created by AWS ParallelCluster do not enable VPC Flow Logs by default. VPC Flow Logs enable you to capture information about the IP traffic going to and from network interfaces in your VPCs. For more information, see VPC Flow Logs in the *Amazon VPC User Guide*.

If you let AWS ParallelCluster create a VPC, make sure that you decide whether all nodes are to be in a public subnet.

> **Note**
>
> If you choose `1. Head node in a public subnet and compute fleet in a private subnet`, AWS ParallelCluster creates a NAT gateway that results in additional cost, even if you specify free tier resources.

```
Automate VPC creation? (y/n) [n]: y
Allowed values for Availability Zone:
1. us-east-1a
2. us-east-1b
3. us-east-1c
4. us-east-1d
5. us-east-1e
6. us-east-1f
Availability Zone [us-east-1a]:
Allowed values for Network Configuration:
1. Head node in a public subnet and compute fleet in a private subnet
2. Head node and compute fleet in the same public subnet
Network Configuration [Head node in a public subnet and compute fleet in a private
 subnet]: *1*
Beginning VPC creation. Please do not leave the terminal until the creation is
 finalized
```

If you don't create a new VPC, you must select an existing VPC.

If you choose to have AWS ParallelCluster create the VPC, make a note of the VPC ID so you can use the AWS CLI or AWS Management Console to delete it later.

```
Automate VPC creation? (y/n) [n]: n
Allowed values for VPC ID:
  #  id                    name                                number_of_subnets
---  --------------------  ----------------------------------  ------------------
  1  vpc-0b4ad9c4678d3c7ad  ParallelClusterVPC-20200118031893                   2
  2  vpc-0e87c753286f37eef  ParallelClusterVPC-20191118233938                   5
VPC ID [vpc-0b4ad9c4678d3c7ad]: 1
```

After the VPC has been selected, make sure that you decide whether to use existing subnets or create new ones.

```
Automate Subnet creation? (y/n) [y]: y
```

```
Creating CloudFormation stack...
Do not leave the terminal until the process has finished
```

When you have completed the preceding steps, a simple cluster launches into a VPC. The VPC uses an existing subnet that supports public IP addresses. The route table for the subnet is `0.0.0.0/0 => igw-xxxxxx`. Note the following conditions:

- The VPC must have `DNS Resolution = yes` and `DNS Hostnames = yes`.
- The VPC must also have DHCP options with the correct `domain-name` for the AWS Region. The default DHCP Option Set already specifies the required AmazonProvidedDNS. If specifying more than one domain name server, see DHCP options sets in the *Amazon VPC User Guide*. When using private subnets, use a NAT gateway or an internal proxy to enable web access for compute nodes. For more information, see Network configurations.

When all settings contain valid values, you can launch the cluster by running the create command.

```
$ pcluster create-cluster --cluster-name test-cluster --cluster-configuration cluster-
config.yaml
{
  "cluster": {
```

```
    "clusterName": "test-cluster",
    "cloudformationStackStatus": "CREATE_IN_PROGRESS",
    "cloudformationStackArn": "arn:aws:cloudformation:eu-west-1:xxx:stack/test-cluster/
abcdef0-f678-890a-5abc-021345abcdef",
    "region": "eu-west-1",
    "version": "3.13.2",
    "clusterStatus": "CREATE_IN_PROGRESS"
  },
  "validationMessages": []
}
```

Follow the cluster progress:

```
$ pcluster describe-cluster --cluster-name test-cluster
```

or

```
$ pcluster list-clusters --query 'clusters[?clusterName==`test-cluster`]'
```

After the cluster reaches the `"clusterStatus": "CREATE_COMPLETE"` status, you can connect to it by using your normal SSH client settings. For more information about connecting to Amazon EC2 instances, see the Amazon EC2 User Guide in the *Amazon EC2 User Guide*. Or you can connect the cluster through

```
$ pcluster ssh --cluster-name test-cluster -i ~/path/to/keyfile.pem
```

To delete the cluster, run the following command.

```
$ pcluster delete-cluster --region us-east-1 --cluster-name test-cluster
```

After the cluster is deleted, you can delete the network resources in the VPC by deleting the CloudFormation networking stack. The stack's name starts with "parallelclusternetworking-" and contains the creation time in "YYYYMMDDHHMMSS" format. You can list the stacks using the `list-stacks` command.

```
$ aws --region us-east-1 cloudformation list-stacks \
   --stack-status-filter "CREATE_COMPLETE" \
   --query "StackSummaries[].StackName" | \
   grep -e "parallelclusternetworking-"
```

```
    "parallelclusternetworking-pubpriv-20191029205804"
```

The stack can be deleted using the `delete-stack` command.

```
$ aws --region us-east-1 cloudformation delete-stack \
    --stack-name parallelclusternetworking-pubpriv-20191029205804
```

The VPC that `pcluster configure` creates for you *isn't* created in the CloudFormation networking stack. You can delete that VPC manually in the console or by using the AWS CLI.

```
$ aws --region us-east-1 Amazon EC2 delete-vpc --vpc-id vpc-0b4ad9c4678d3c7ad
```

Configure and create a cluster with the PCUI

The PCUI is a web-based user interface that mirrors the AWS ParallelCluster `pcluster` CLI, while providing a console-like experience. You install and access the PCUI in your AWS account. When you run it, the PCUI accesses an instance of the AWS ParallelCluster API hosted on Amazon API Gateway in your AWS account.

> **Note**
>
> The PCUI wizard may not have UI options for all the supported features in the latest supported AWS ParallelCluster version. You can manually edit the configuration file as needed or use the AWS ParallelCluster CLI.

This section guides you through how to configure and create a cluster using the PCUI.

Prerequisites:

- Access to a running instance of PCUI. For more information, see Installing the PCUI.

Configure and create a cluster

1. In the PCUI **Clusters** view, choose **Create cluster**, **Step by step**.
2. In **Cluster**, **Name**, enter a name for your cluster.
3. Choose a **VPC** with a public subnet for your cluster, then choose **Next**.
4. In **Head node**, choose **Add SSM session**, then choose **Next**.

5. In **Queues**, **Compute resources**, choose **1** for **Static nodes**.
6. For **Instance type**, remove the selected default instance type, choose **t2.micro**, then choose **Next**.
7. In **Storage**, choose **Next**.
8. In **Cluster configuration**, review the cluster configuration YAML and choose **Dry run** to validate it.
9. Choose **Create** to create your cluster, based on the validated configuration.
10. After a few seconds, the PCUI automatically navigates you back to **Clusters**, where you can monitor the cluster create status and **Stack events**.
11. Choose **Details** to see cluster details, such as the version and status.
12. Choose **Instances** to see the list of Amazon EC2 instances and status.
13. Choose **Stack events** to view cluster stack events, and a AWS Management Console link to the CloudFormation stack that creates the cluster.
14. In **Details**, after cluster creation completes, choose **View YAML** to view or download the cluster configuration YAML file.
15. After cluster creation completes, choose **Shell** to access the cluster head node.

    > **Note**
    >
    > When you choose **Shell**, AWS ParallelCluster opens an Amazon EC2 Systems Manager session and adds an `ssm-user` to `/etc/sudoers`. For more information, see Turn on or turn off `ssm-user` account administrative permissions in the *Amazon EC2 Systems Manager User Guide*.

16. To clean up, in the **Clusters** view, select the cluster, and choose **Actions**, **Delete cluster**.

Connect to a cluster

When you use AWS ParallelCluster, you can connect to the cluster head node to run jobs, view results, manage users, and monitor the cluster and job status. To connect to the cluster head node instance, use one of the following methods:

- You can use `ssh` with a key pair to log in. Specify the private key in `HeadNode` / `KeyName` in the cluster configuration. For more information, see Connect to your Linux instance using SSH in the *Amazon EC2 User Guide for Linux Instances*.

- You can use the `pcluster ssh` command line interface (CLI) command to log in. Specify the private key in the cluster configuration `HeadNode` / `KeyName`. For more information, see `pcluster ssh`.
- You can use an SSM session to connect to the cluster head node. You must add the `AmazonSSMManagedInstanceCore` managed policy to `HeadNode` / `AdditionalIamPolicies` in the cluster configuration to connect by using an SSM session. For more information, see SSM session manager in the *SSM User Guide*.
- You can use Amazon DCV to connect to the cluster head node. For more information, see Connect to the head and login nodes through Amazon DCV.
- When you use the PCUI, you can also use an Amazon EC2 Connect command that the UI provides to connect to the cluster head node.

Multiple user access to clusters

Learn to implement and manage multiple user access to a single cluster.

In this topic, an AWS ParallelCluster user refers to a system user for compute instances. An example is an `ec2-user` for an Amazon EC2 instance.

AWS ParallelCluster multi-user access support is available in all the AWS Regions where AWS ParallelCluster is currently available. It works with other AWS services, including Amazon FSx for Lustre and Amazon Elastic File System.

You can use an AWS Directory Service for Microsoft Active Directory or Simple AD to manage cluster access. Make sure to check AWS Region availability for these services. To set up a cluster, specify an AWS ParallelCluster DirectoryService configuration. AWS Directory Service directories can be connected to multiple clusters. This allows for centralized management of identities across multiple environments and a unified login experience.

When you use AWS Directory Service for AWS ParallelCluster multiple user access, you can log in to the cluster with user credentials that are defined in the directory. These credentials consist of a user name and password. After you log in to the cluster for the first time, a user SSH key is automatically generated. You can use it to log in without a password.

You can create, delete, and modify a cluster's users or groups after your directory service is deployed. With AWS Directory Service, you can do this in the AWS Management Console or by using the *Active Directory Users and Computers* tool. This tool is accessible from any Amazon EC2

instance that's joined to your Active Directory. For more information, see Installing the Active Directory administration tools.

If you plan to use AWS ParallelCluster in a single subnet with no internet access, see AWS ParallelCluster in a single subnet with no internet access for additional requirements.

Topics

Create an Active Directory

Make sure that you create an Active Directory (AD) before you create your cluster. For information about how to choose the type of active directory for your cluster, see Which to choose in the *AWS Directory Service Administration Guide*.

If the directory is empty, add users with user names and passwords. For more information, see the documentation that's specific to AWS Directory Service for Microsoft Active Directory or Simple AD.

> **Note**
>
> AWS ParallelCluster requires every Active Directory user directory to be in the `/home/$user` directory.

Create a cluster with an AD domain

> **Warning**
>
> This introductory section describes how to set up AWS ParallelCluster with a Managed Active Directory (AD) server over the Lightweight Directory Access Protocol (LDAP). LDAP is an insecure protocol. For production systems, we strongly recommended the use of TLS certificates (LDAPS) as described in the Example AWS Managed Microsoft AD over LDAP(S) cluster configurations section that follows.

Configure your cluster to integrate with a directory by specifying the relevant information in the `DirectoryService` section of the cluster configuration file. For more information, see the `DirectoryService` configuration section.

You can use this following example to integrate your cluster with an AWS Managed Microsoft AD over the Lightweight Directory Access Protocol (LDAP).

Specific definitions that are required for an AWS Managed Microsoft AD over LDAP configuration:

- You must set the `ldap_auth_disable_tls_never_use_in_production` parameter to `True` under `DirectoryService` / `AdditionalSssdConfigs`.
- You can specify either controller hostnames or IP addresses for `DirectoryService` / `DomainAddr`.
- `DirectoryService` / `DomainReadOnlyUser` syntax must be as follows:

```
cn=ReadOnly,ou=Users,ou=CORP,dc=corp,dc=example,dc=com
```

Get your AWS Managed Microsoft AD configuration data:

```
$ aws ds describe-directories --directory-id "d-abcdef01234567890"
```

```
{
    "DirectoryDescriptions": [
        {
            "DirectoryId": "d-abcdef01234567890",
            "Name": "corp.example.com",
            "DnsIpAddrs": [
                "203.0.113.225",
                "192.0.2.254"
            ],
            "VpcSettings": {
                "VpcId": "vpc-021345abcdef6789",
                "SubnetIds": [
                    "subnet-1234567890abcdef0",
                    "subnet-abcdef01234567890"
                ],
                "AvailabilityZones": [
                    "region-idb",
```

```
                    "region-idd"
                ]
            }
        }
    ]
}
```

Cluster configuration for an AWS Managed Microsoft AD:

```
Region: region-id
Image:
  Os: alinux2
HeadNode:
  InstanceType: t2.micro
  Networking:
    SubnetId: subnet-1234567890abcdef0
  Ssh:
    KeyName: pcluster
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue1
      ComputeResources:
        - Name: t2micro
          InstanceType: t2.micro
          MinCount: 1
          MaxCount: 10
      Networking:
        SubnetIds:
          - subnet-abcdef01234567890
DirectoryService:
  DomainName: dc=corp,dc=example,dc=com
  DomainAddr: ldap://203.0.113.225,ldap://192.0.2.254
  PasswordSecretArn: arn:aws:secretsmanager:region-
id:123456789012:secret:MicrosoftAD.Admin.Password-1234
  DomainReadOnlyUser: cn=ReadOnly,ou=Users,ou=CORP,dc=corp,dc=example,dc=com
  AdditionalSssdConfigs:
    ldap_auth_disable_tls_never_use_in_production: True
```

To use this configuration for a Simple AD, change the `DomainReadOnlyUser` property value in the `DirectoryService` section:

```
DirectoryService:
```

```
  DomainName: dc=corp,dc=example,dc=com
  DomainAddr: ldap://203.0.113.225,ldap://192.0.2.254
  PasswordSecretArn: arn:aws:secretsmanager:region-
id:123456789012:secret:SimpleAD.Admin.Password-1234
  DomainReadOnlyUser: cn=ReadOnlyUser,cn=Users,dc=corp,dc=example,dc=com
  AdditionalSssdConfigs:
    ldap_auth_disable_tls_never_use_in_production: True
```

Considerations:

- We recommend that you use LDAP over TLS/SSL (or LDAPS) rather than LDAP alone. TLS/SSL ensures that the connection is encrypted.
- The `DirectoryService` / `DomainAddr` property value matches the entries in the `DnsIpAddrs` list from the `describe-directories` output.
- We recommend that your cluster use subnets that are located in the same Availability Zone that the `DirectoryService` / `DomainAddr` points to. If you use custom Dynamic Host Configuration Protocol (DHCP) configuration that's recommended for directory VPCs and your subnets *aren't* located in the `DirectoryService` / `DomainAddr` Availability Zone, cross traffic among Availability Zones is possible. The use of custom DHCP configurations *isn't* required to use the multi-user AD integration feature.
- The `DirectoryService` / `DomainReadOnlyUser` property value specifies a user that must be created in the directory. This user *isn't* created by default. We recommend that you *don't* give this user permission to modify directory data.
- The `DirectoryService` / `PasswordSecretArn` property value points to an AWS Secrets Manager secret that contains the password of the user that you specified for the `DirectoryService` / `DomainReadOnlyUser` property. If this user's password changes, update the secret value and update the cluster. To update the cluster for the new secret value, you must stop the compute fleet with the `pcluster update-compute-fleet` command. If you configured your cluster to use `LoginNodes`, stop the `LoginNodes` / `Pools` and update the cluster after setting `LoginNodes` / `Pools` / `Count` to 0. Then, run the following command from within the cluster head node.

  ```
  sudo /opt/parallelcluster/scripts/directory_service/
update_directory_service_password.sh
  ```

For another example, see also Integrating Active Directory.

Log in to a cluster integrated with an AD domain

If you enabled the Active Delivery (AD) domain integration feature, authentication by password is enabled on the cluster head node. The home directory of an AD user is created at the first user login to the head node or the first time a sudo-user switches to the AD user on the head node.

Password authentication isn't enabled for cluster compute nodes. AD users must log in to compute nodes with SSH keys.

By default, SSH keys are set up in the AD user `/${HOME}/.ssh` directory at the first SSH login to the head node. This behavior can be disabled by setting `DirectoryService` / `GenerateSshKeysForUsers` boolean property to `false` in the cluster configuration. By default, `DirectoryService` / `GenerateSshKeysForUsers` is set to `true`.

If an AWS ParallelCluster application requires passwordless SSH between cluster nodes, make sure that the SSH keys are correctly set up in the user's home directory.

AWS Managed Microsoft AD passwords expire after 42 days. For more information, see Manage password policies for AWS Managed Microsoft AD in the *AWS Directory Service Administration Guide*. If your password expires, it must be reset to restore cluster access. For more information, see How to reset a user password and expired passwords.

> **Note**
>
> If the AD integration feature doesn't work as expected, the SSSD logs can provide useful diagnostic information for troubleshooting the issue. These logs are located in the `/var/log/sssd` directory on cluster nodes. By default, they're also stored in a cluster's Amazon CloudWatch log group.
> For more information, see Troubleshooting multi-user integration with Active Directory.

Running MPI jobs

As suggested in SchedMD, bootstrap MPI jobs using Slurm as the MPI bootstrapping method. For more information, refer to the official Slurm documentation or the official documentation for your MPI library.

For example, in the IntelMPI official documentation, you learn that when running a StarCCM job, you must set Slurm as process orchestrator by exporting the environment variable `I_MPI_HYDRA_BOOTSTRAP=slurm`.

Note

Known issue
In the case where your MPI application relies on SSH as mechanism to spawn MPI jobs, it's possible to incur in a known bug in Slurm that causes the wrong resolution of the directory user name to "nobody".
Either configure your application to use Slurm as the MPI bootstrapping method or refer to Known issues with username resolution in the Troubleshooting section for further details and possible workarounds.

Example AWS Managed Microsoft AD over LDAP(S) cluster configurations

AWS ParallelCluster supports multiple user access by integrating with an AWS Directory Service over the Lightweight Directory Access Protocol (LDAP), or LDAP over TLS/SSL (LDAPS).

The following examples show how to create cluster configurations to integrate with an AWS Managed Microsoft AD over LDAP(S).

AWS Managed Microsoft AD over LDAPS with certificate verification

You can use this example to integrate your cluster with an AWS Managed Microsoft AD over LDAPS, with certificate verification.

Specific definitions for an AWS Managed Microsoft AD over LDAPS with certificates configuration:

- `DirectoryService` / `LdapTlsReqCert` must be set to `hard` (default) for LDAPS with certificate verification.
- `DirectoryService` / `LdapTlsCaCert` must specify the path to your certificate of authority (CA) certificate.

  The CA certificate is a certificate bundle that contains the certificates of the entire CA chain that issued certificates for the AD domain controllers.

  Your CA certificate and certificates must be installed on the cluster nodes.
- Controllers hostnames must be specified for `DirectoryService` / `DomainAddr`, not IP addresses.

- `DirectoryService` / `DomainReadOnlyUser` syntax must be as follows:

```
cn=ReadOnly,ou=Users,ou=CORP,dc=corp,dc=example,dc=com
```

Example cluster configuration file for using AD over LDAPS:

```
Region: region-id
Image:
  Os: alinux2
HeadNode:
  InstanceType: t2.micro
  Networking:
    SubnetId: subnet-1234567890abcdef0
  Ssh:
    KeyName: pcluster
  Iam:
    AdditionalIamPolicies:
      - Policy: arn:aws:iam::aws:policy/AmazonS3ReadOnlyAccess
  CustomActions:
    OnNodeConfigured:
      Script: s3://&example-s3-bucket;/scripts/pcluster-dub-msad-ldaps.post.sh
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue1
      ComputeResources:
        - Name: t2micro
          InstanceType: t2.micro
          MinCount: 1
          MaxCount: 10
      Networking:
        SubnetIds:
          - subnet-abcdef01234567890
      Iam:
        AdditionalIamPolicies:
          - Policy: arn:aws:iam::aws:policy/AmazonS3ReadOnlyAccess
      CustomActions:
        OnNodeConfigured:
          Script: s3://&example-s3-bucket;/scripts/pcluster-dub-msad-ldaps.post.sh
DirectoryService:
  DomainName: dc=corp,dc=example,dc=com
```

```
  DomainAddr: ldaps://win-abcdef01234567890.corp.example.com,ldaps://win-
abcdef01234567890.corp.example.com
  PasswordSecretArn: arn:aws:secretsmanager:region-
id:123456789012:secret:MicrosoftAD.Admin.Password-1234
  DomainReadOnlyUser: cn=ReadOnly,ou=Users,ou=CORP,dc=corp,dc=example,dc=com
  LdapTlsCaCert: /etc/openldap/cacerts/corp.example.com.bundleca.cer
  LdapTlsReqCert: hard
```

Add certificates and configure domain controllers in post install script:

```
*#!/bin/bash*
set -e

AD_CERTIFICATE_S3_URI="s3://amzn-s3-demo-bucket/bundle/corp.example.com.bundleca.cer"
AD_CERTIFICATE_LOCAL="/etc/openldap/cacerts/corp.example.com.bundleca.cer"

AD_HOSTNAME_1="win-abcdef01234567890.corp.example.com"
AD_IP_1="192.0.2.254"

AD_HOSTNAME_2="win-abcdef01234567890.corp.example.com"
AD_IP_2="203.0.113.225"

# Download CA certificate
mkdir -p $(dirname "${AD_CERTIFICATE_LOCAL}")
aws s3 cp "${AD_CERTIFICATE_S3_URI}" "${AD_CERTIFICATE_LOCAL}"
chmod 644 "${AD_CERTIFICATE_LOCAL}"

# Configure domain controllers reachability
echo "${AD_IP_1} ${AD_HOSTNAME_1}" >> /etc/hosts
echo "${AD_IP_2} ${AD_HOSTNAME_2}" >> /etc/hosts
```

You can retrieve the domain controllers hostnames from instances joined to the domain as shown in the following examples.

From Windows instance

```
$ nslookup 192.0.2.254
```

```
Server:  corp.example.com
Address:  192.0.2.254
```

```
Name:    win-abcdef01234567890.corp.example.com
Address:  192.0.2.254
```

From Linux instance

```
$ nslookup 192.0.2.254
```

```
192.0.2.254.in-addr.arpa   name = corp.example.com
192.0.2.254.in-addr.arpa   name = win-abcdef01234567890.corp.example.com
```

AWS Managed Microsoft AD over LDAPS without certificate verification

You can use this example to integrate your cluster with an AWS Managed Microsoft AD over LDAPS, without certificate verification.

Specific definitions for an AWS Managed Microsoft AD over LDAPS without certificate verification configuration:

- `DirectoryService` / `LdapTlsReqCert` must be set to `never`.
- Either controller hostnames or IP addresses can be specified for `DirectoryService` / `DomainAddr`.
- `DirectoryService` / `DomainReadOnlyUser` syntax must be as follows:

  ```
  cn=ReadOnly,ou=Users,ou=CORP,dc=corp,dc=example,dc=com
  ```

Example cluster configuration file for using AWS Managed Microsoft AD over LDAPS without certificate verification:

```
Region: region-id
Image:
  Os: alinux2
HeadNode:
  InstanceType: t2.micro
  Networking:
    SubnetId: subnet-1234567890abcdef0
  Ssh:
    KeyName: pcluster
Scheduling:
```

```
  Scheduler: slurm
  SlurmQueues:
    - Name: queue1
      ComputeResources:
        - Name: t2micro
          InstanceType: t2.micro
          MinCount: 1
          MaxCount: 10
      Networking:
        SubnetIds:
          - subnet-abcdef01234567890
DirectoryService:
  DomainName: dc=corp,dc=example,dc=com
  DomainAddr: ldaps://203.0.113.225,ldaps://192.0.2.254
  PasswordSecretArn: arn:aws:secretsmanager:region-
id:123456789012:secret:MicrosoftAD.Admin.Password-1234
  DomainReadOnlyUser: cn=ReadOnly,ou=Users,ou=CORP,dc=corp,dc=example,dc=com
  LdapTlsReqCert: never
```

Best practices

The following sections provide best practices for using AWS ParallelCluster, which includes network performance and budget alerts. If you encounter problems even though you follow these best practices, see AWS ParallelCluster troubleshooting for possible solutions.

Best practices: head node instance type selection

Even though the head node doesn't run a job, its functions and its sizing are crucial to the overall performance of the cluster. When you choose the instance type to use for your head node, consider the following characteristics:

Cluster size: The head node orchestrates the scaling logic of the cluster and is responsible of attaching new nodes to the scheduler. To scale up and down a cluster that has a large number nodes, provide the head node some extra compute capacity.

Shared file systems: When you use shared file systems, choose an instance type with enough network bandwidth, and enough Amazon EBS bandwidth, to handle your workflows. Ensure that the head node is able to both expose sufficient NFS server directories for the cluster and handle the artifacts that need to be shared between the compute nodes and head node.

Best practices: network performance

Network performance is critical for high performance computing (HPC) applications. Without reliable network performance, these applications can't perform as expected. To optimize network performance, consider the following best practices.

- **Placement group:** If you're using Slurm, consider configuring each Slurm queue to use a cluster placement group . A cluster's *placement group* is a logical grouping of instances within a single Availability Zone. For more information, see placement groups in the *Amazon EC2 User Guide*. You can specify a `PlacementGroup` in the queue's `Networking` section, each compute resource is assigned to the queue's placement group. When specifying a `PlacementGroup` in the compute resource's `Networking` section, that specific compute resource is assigned to that placement group. The compute resource placement group specification overrides the queue specification for the compute resource. For more information, see `SlurmQueues` / `Networking` / `PlacementGroup` and `SlurmQueues` / `ComputeResources` / `Networking` / `PlacementGroup`.

  ```
  Networking:
    PlacementGroup:
      Enabled: true
      Id: your-placement-group-name
  ```

  Alternatively, have AWS ParallelCluster create a placement group for you.

  ```
  Networking:
    PlacementGroup:
      Enabled: true
  ```

  Starting with AWS ParallelCluster version 3.3.0, placement group creation and management is modified. When you specify the placement group to be enabled, without a `name` or `Id`, in the queue, each compute resource is assigned its own managed placement group, instead of one managed group for the entire queue. This helps to reduce insufficient capacity errors. If you need to have one placement group for the entire queue, you can use a named placement group.

  `SlurmQueues` / `Networking` / `PlacementGroup` / `Name` was added as a preferred alternative to `SlurmQueues` / `Networking` / `PlacementGroup` / `Id`.

  For more information, see `Networking`.

- **Enhanced networking:** Consider choosing an instance type that supports enhanced networking. This recommendation applies to all current generation instances. For more information, see enhanced networking on Linux in the *Amazon EC2 User Guide*.
- **Elastic Fabric Adapter:** To support high levels of scalable instance to instance communication, consider choosing EFA network interfaces for your network. The EFA's custom-built operating system (OS) bypass hardware enhances instance to instance communications with the on-demand elasticity and flexibility of the AWS Cloud. You can configure each Slurm queue `ComputeResource` to use `Efa`. For more information about using EFA with AWS ParallelCluster, see Elastic Fabric Adapter.

  ```
  ComputeResources:
    - Name: your-compute-resource-name
      Efa:
        Enabled: true
  ```

  For more information about EFA, see Elastic Fabric Adapter in the *Amazon EC2 User Guide for Linux Instances*.
- **Instance bandwidth:** The bandwidth scales with instance size. For information about the different instance types, see Amazon EBS–optimized instances and Amazon EBS volume types in the *Amazon EC2 User Guide*.

Best practices: budget alerts

To manage resource costs in AWS ParallelCluster, we recommend that you use AWS Budgets actions to create a budget. You can also create defined budget threshold alerts for selected AWS resources. For more information, see Configuring a budget action in the *AWS Budgets User Guide*. Similarly, you can also use Amazon CloudWatch to create a billing alarm. For more information, see Creating a billing alarm to monitor your estimated AWS charges.

Best practices: moving a cluster to a new AWS ParallelCluster minor or patch version

Currently each AWS ParallelCluster minor version is self-contained along with its `pcluster` CLI. To move a cluster to a new minor or patch version, you must re-create the cluster using the new version's CLI.

To optimize the process of moving a cluster to a new minor or patch version, we recommend that you do the following:

- Save personal data in external volumes that are created outside the cluster, such as Amazon EFS and FSx for Lustre. By doing this, you can easily move the data from one cluster to another in the future.
- Create shared storage systems using the following types. You can create these systems using the AWS CLI or AWS Management Console.
  - `SharedStorage` / `EbsSettings` / `VolumeId`
  - `SharedStorage` / `EfsSettings` / `FileSystemId`
  - `SharedStorage` / `FsxLustreSettings` / `FileSystemId`

  Define a file system or volume in a cluster configuration as existing file system or volume. This way, they're preserved when you delete the cluster and can be attached to a new cluster.

  We recommend that you use Amazon EFS or FSx for Lustre file systems. Both of these systems can be attached to multiple clusters at the same time. Moreover, you can attach either of these systems to a new cluster before you delete your existing cluster.
- Use custom bootstrap actions to customize your instances rather than using a custom AMI. If instead, you use a custom AMI, then you need to delete and recreate that AMI for each new version release.
- We recommend that you apply the preceding recommendations in the following sequence:
  1. Update the existing cluster configuration to use existing file system definitions.
  2. Verify the `pcluster` version and update it if needed.
  3. Create and test the new cluster. When you test the new cluster, check the following:
     - Make sure that your data is available in the new cluster.
     - Make sure that your application works in the new cluster.
  4. After your new cluster is fully tested and operational and you no longer need the existing cluster, delete it.

Moving from AWS ParallelCluster 2.x to 3.x

The following sections describe what happens when you move from AWS ParallelCluster 2.x to 3.x, including the changes from one version to the other.

Custom Bootstrap actions

With AWS ParallelCluster 3, you can specify different custom bootstrap actions scripts for the head node and compute nodes using `OnNodeStart` (`pre_install` in AWS ParallelCluster version 2) and `OnNodeConfigured` (`post_install` in AWS ParallelCluster version 2) parameters in the `HeadNode` and `Scheduling` / `SlurmQueues` sections. For more information, see Custom bootstrap actions.

Custom bootstrap actions scripts that are developed for AWS ParallelCluster 2 must be adapted to be used in AWS ParallelCluster 3:

- We don't recommend using `/etc/parallelcluster/cfnconfig` and `cfn_node_type` to differentiate between head and compute nodes. Instead, we recommend that you specify two different scripts in the `HeadNode` and `Scheduling` / `SlurmQueues`.
- If you prefer to continue loading `/etc/parallelcluster/cfnconfig` for use in your bootstrap actions script, note the value of `cfn_node_type` is changed from "MasterServer" to "HeadNode" (see: Inclusive language).
- On AWS ParallelCluster 2, the first input argument to bootstrap action scripts was the S3 URL to the script and was reserved. In AWS ParallelCluster 3, only the arguments configured in the configuration are passed to the scripts.

> **Warning**
>
> Using internal variables provided through the `/etc/parallelcluster/cfnconfig` file isn't officially supported. This file might be removed as part of a future release.

AWS ParallelCluster 2.x and 3.x use different configuration file syntax

AWS ParallelCluster 3.x configuration uses YAML syntax. The full reference can be found at Configuration files.

In addition to requiring a YAML file format, a number of configuration sections, settings, and parameter values have been updated in AWS ParallelCluster 3.x. In this section, we note key changes to the AWS ParallelCluster configuration along with side-by-side examples illustrating these differences across each version of AWS ParallelCluster.

Example of multiple scheduler queues configuration with hyperthreading enabled and disabled

AWS ParallelCluster 2:

```
[cluster default]
queue_settings = ht-enabled, ht-disabled
...

[queue ht-enabled]
compute_resource_settings = ht-enabled-i1
disable_hyperthreading = false

[queue ht-disabled]
compute_resource_settings = ht-disabled-i1
disable_hyperthreading = true

[compute_resource ht-enabled-i1]
instance_type = c5n.18xlarge
[compute_resource ht-disabled-i1]
instance_type = c5.xlarge
```

AWS ParallelCluster 3:

```
...
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: ht-enabled
      Networking:
        SubnetIds:
          - compute_subnet_id
      ComputeResources:
        - Name: ht-enabled-i1
          DisableSimultaneousMultithreading: true
          InstanceType: c5n.18xlarge
    - Name: ht-disabled
      Networking:
        SubnetIds:
          - compute_subnet_id
      ComputeResources:
        - Name: ht-disabled-i1
          DisableSimultaneousMultithreading: false
          InstanceType: c5.xlarge
```

Example of new FSx for Lustre file-system configuration

AWS ParallelCluster 2:

```
[cluster default]
fsx_settings = fsx
...

[fsx fsx]
shared_dir = /shared-fsx
storage_capacity = 1200
imported_file_chunk_size = 1024
import_path = s3://amzn-s3-demo-bucket
export_path = s3://amzn-s3-demo-bucket/export_dir
weekly_maintenance_start_time = 3:02:30
deployment_type = PERSISTENT_1
data_compression_type = LZ4
```

AWS ParallelCluster 3:

```
...
SharedStorage:
  - Name: fsx
    MountDir: /shared-fsx
    StorageType: FsxLustre
    FsxLustreSettings:
      StorageCapacity: 1200
      ImportedFileChunkSize: 1024
      ImportPath: s3://amzn-s3-demo-bucket
      ExportPath: s3://amzn-s3-demo-bucket/export_dir
      WeeklyMaintenanceStartTime: "3:02:30"
      DeploymentType: PERSISTENT_1
      DataCompressionType: LZ4
```

Example of a cluster configuration mounting an existing FSx for Lustre file-system

AWS ParallelCluster 2:

```
[cluster default]
fsx_settings = fsx
...
```

```
[fsx fsx]
shared_dir = /shared-fsx
fsx_fs_id = fsx_fs_id
```

AWS ParallelCluster 3:

```
...
SharedStorage:
  - Name: fsx
    MountDir: /shared-fsx
    StorageType: FsxLustre
    FsxLustreSettings:
      FileSystemId: fsx_fs_id
```

Example of a cluster with the Intel HPC Platform Specification software stack

AWS ParallelCluster 2:

```
[cluster default]
enable_intel_hpc_platform = true
...
```

AWS ParallelCluster 3:

```
...
AdditionalPackages:
  IntelSoftware:
    IntelHpcPlatform: true
```

Notes:

- The installation of Intel HPC Platform Specification software is subject to the terms and conditions of the applicable Intel End User License Agreement.

Example of custom IAM configurations including: instance profile, instance role, additional policies for instances and the role for the lambda functions associated with the cluster

AWS ParallelCluster 2:

```
[cluster default]
```

```
additional_iam_policies = arn:aws:iam::aws:policy/
AmazonS3ReadOnlyAccess,arn:aws:iam::aws:policy/AmazonDynamoDBReadOnlyAccess
ec2_iam_role = ec2_iam_role
iam_lambda_role = lambda_iam_role
...
```

AWS ParallelCluster 3:

```
...
Iam:
  Roles:
    CustomLambdaResources: lambda_iam_role
HeadNode:
  ...
  Iam:
    InstanceRole: ec2_iam_role
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue1
      ...
      Iam:
        InstanceProfile: iam_instance_profile
    - Name: queue2
      ...
      Iam:
        AdditionalIamPolicies:
          - Policy: arn:aws:iam::aws:policy/AmazonS3ReadOnlyAccess
          - Policy: arn:aws:iam::aws:policy/AmazonDynamoDBReadOnlyAccess
```

Notes:

- For AWS ParallelCluster 2, the IAM settings are applied to all the instances of a cluster and `additional_iam_policies` can't be used in conjunction with `ec2_iam_role`.
- For AWS ParallelCluster 3, you can have different IAM settings for head and compute nodes and even specify different IAM settings for each compute queue.
- For AWS ParallelCluster 3, you can use an IAM instance profile as an alternative to an IAM role. `InstanceProfile`, `InstanceRole` or `AdditionalIamPolicies` can't be configured together.

Example of custom bootstrap actions

AWS ParallelCluster 2:

```
[cluster default]
s3_read_resource = arn:aws:s3:::amzn-s3-demo-bucket/*
pre_install = s3://amzn-s3-demo-bucket/scripts/pre_install.sh
pre_install_args = 'R curl wget'
post_install = s3://amzn-s3-demo-bucket/scripts/post_install.sh
post_install_args = "R curl wget"
...
```

AWS ParallelCluster 3:

```
...
HeadNode:
  ...
  CustomActions:
    OnNodeStart:
      Script: s3://amzn-s3-demo-bucket/scripts/pre_install.sh
      Args:
        - R
        - curl
        - wget
    OnNodeConfigured:
      Script: s3://amzn-s3-demo-bucket/scripts/post_install.sh
      Args: ['R', 'curl', 'wget']
  Iam:
    S3Access:
      - BucketName: amzn-s3-demo-bucket
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue1
      ...
      CustomActions:
        OnNodeStart:
          Script: s3://amzn-s3-demo-bucket/scripts/pre_install.sh
          Args: ['R', 'curl', 'wget']
        OnNodeConfigured:
          Script: s3://amzn-s3-demo-bucket/scripts/post_install.sh
          Args: ['R', 'curl', 'wget']
      Iam:
        S3Access:
          - BucketName: amzn-s3-demo-bucket
```

Example of a cluster with read and write access to the S3 bucket resources

AWS ParallelCluster 2:

```
[cluster default]
s3_read_resource = arn:aws:s3:::amzn-s3-demo-bucket/read_only/*
s3_read_write_resource = arn:aws:s3:::amzn-s3-demo-bucket/read_and_write/*
...
```

AWS ParallelCluster 3:

```
...
HeadNode:
  ...
  Iam:
    S3Access:
      - BucketName: amzn-s3-demo-bucket
        KeyName: read_only/
        EnableWriteAccess: False
      - BucketName: amzn-s3-demo-bucket
        KeyName: read_and_write/
        EnableWriteAccess: True
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue1
      ...
      Iam:
        S3Access:
          - BucketName: amzn-s3-demo-bucket
            KeyName: read_only/
            EnableWriteAccess: False
          - BucketName: amzn-s3-demo-bucket
            KeyName: read_and_write/
            EnableWriteAccess: True
```

Inclusive language

AWS ParallelCluster 3 uses the words "head node" in places where "master" was used in AWS ParallelCluster 2. This includes the following:

- Variable exported in the AWS Batch job environment changed: from `MASTER_IP` to `PCLUSTER_HEAD_NODE_IP`.
- All AWS CloudFormation outputs changed from `Master*` to `HeadNode*`.
- All NodeType and tags changed from `Master` to `HeadNode`.

Scheduler support

AWS ParallelCluster 3.x doesn't support Son of Grid Engine (SGE) and Torque schedulers.

The AWS Batch commands `awsbhosts`, `awsbkill`, `awsbout`, `awsbqueues`, `awsbstat`, and `awsbsub` are distributed as a separate `aws-parallelcluster-awsbatch-cli` PyPI package. This package is installed by AWS ParallelCluster on the head node. You can still use these AWS Batch commands from the cluster's head node. However, if you wish to use AWS Batch commands from a location other than the head node, you must first install the `aws-parallelcluster-awsbatch-cli` PyPI package.

AWS ParallelCluster CLI

The AWS ParallelCluster command line interface (CLI) has been changed. The new syntax is described in AWS ParallelCluster CLI commands. The output format for the CLI is a JSON string.

Configuring a new cluster

The `pcluster configure` command includes different parameters in AWS ParallelCluster 3 as compared to AWS ParallelCluster 2. For more information, see `pcluster configure`.

Note also that the configuration file syntax has changed from AWS ParallelCluster 2. For a full reference of the cluster configuration settings, see Cluster configuration file.

Creating a new cluster

AWS ParallelCluster 2's `pcluster create` command has been replaced by the `pcluster create-cluster` command.

Note the default behavior in AWS ParallelCluster 2.x, without the `-nw` option, is to wait on cluster creation events, while AWS ParallelCluster 3.x command returns immediately. The progress of the cluster creation can be monitored using `pcluster describe-cluster`.

An AWS ParallelCluster 3 configuration file contains a single cluster definition, so the `-t` parameter is no more needed.

The following is an example configuration file.

```
# AWS ParallelCluster v2
$ pcluster create \
    -r REGION \
    -c V2_CONFIG_FILE \
    -nw \
    -t CLUSTER_TEMPLATE \
    CLUSTER_NAME

# AWS ParallelCluster v3
$ pcluster create-cluster \
    --region REGION \
    --cluster-configuration V3_CONFIG_FILE \
    --cluster-name CLUSTER_NAME
```

Listing clusters

The `pcluster list` AWS ParallelCluster 2.x command must be replaced with `pcluster list-clusters` command.

Note: You need AWS ParallelCluster v2 CLI to list clusters created with 2.x versions of AWS ParallelCluster. See Install AWS ParallelCluster in a virtual environment (recommended) for how to install multiple versions of AWS ParallelCluster using virtual environments.

```
# AWS ParallelCluster v2
$ pcluster list -r REGION

# AWS ParallelCluster v3
$ pcluster list-clusters --region REGION
```

Starting and Stopping a cluster

The `pcluster start` and `pcluster stop` AWS ParallelCluster 2.x commands must be replaced with `pcluster update-compute-fleet` commands.

Starting a compute fleet:

```
# AWS ParallelCluster v2
$ pcluster start \
    -r REGION \
    CLUSTER_NAME
```

```
# AWS ParallelCluster v3 - Slurm fleets
$ pcluster update-compute-fleet \
    --region REGION \
    --cluster-name CLUSTER_NAME \
    --status START_REQUESTED

# AWS ParallelCluster v3 - AWS Batch fleets
$ pcluster update-compute-fleet \
    --region REGION \
    --cluster-name CLUSTER_NAME \
    --status ENABLED
```

Stopping a compute fleet:

```
# AWS ParallelCluster v2
$ pcluster stop \
    -r REGION \
    CLUSTER_NAME

# AWS ParallelCluster v3 - Slurm fleets
$ pcluster update-compute-fleet \
    --region REGION \
    --cluster-name CLUSTER_NAME \
    --status STOP_REQUESTED

# AWS ParallelCluster v3 - AWS Batch fleets
$ pcluster update-compute-fleet \
    --region REGION \
    --cluster-name CLUSTER_NAME \
    --status DISABLED
```

Connecting to a cluster

The `pcluster ssh` AWS ParallelCluster 2.x command has different parameters names in AWS ParallelCluster 3.x. See `pcluster ssh`.

Connecting to a cluster:

```
# AWS ParallelCluster v2
$ pcluster ssh \
    -r REGION \
    CLUSTER_NAME \
```

```
    -i ~/.ssh/id_rsa

# AWS ParallelCluster v3
$ pcluster ssh \
    --region REGION \
    --cluster-name CLUSTER_NAME \
    -i ~/.ssh/id_rsa
```

IMDS configuration update

Starting with version 3.0.0, AWS ParallelCluster introduced support for restricting access to the head node's IMDS (and the instance profile credentials) to a subset of superusers, by default. For more information, see `Imds` properties.

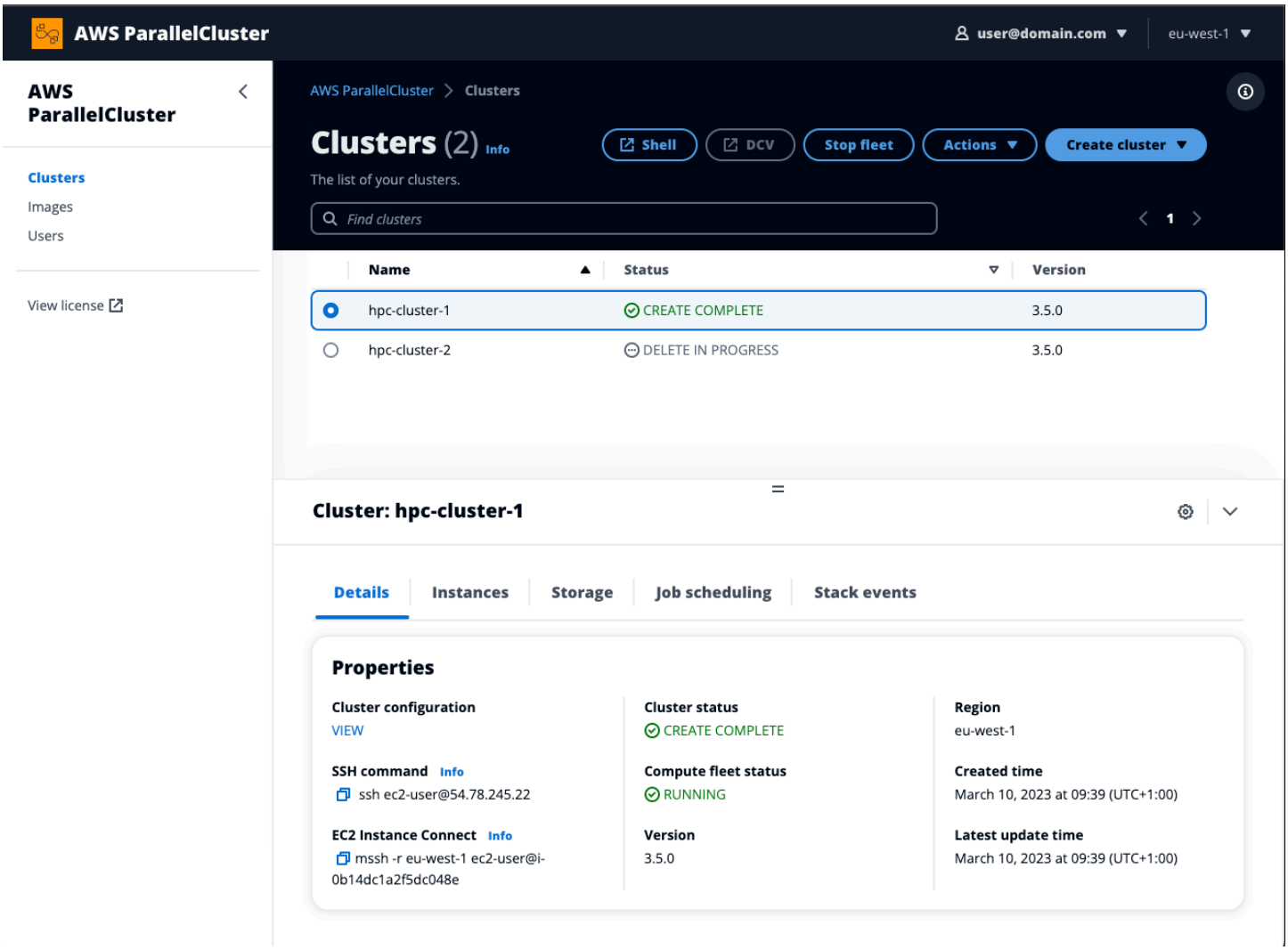
Using AWS ParallelCluster

Topics

AWS ParallelCluster UI

The AWS ParallelCluster UI (PCUI) is a web-based user interface that serves as a dashboard for creating, monitoring, and managing clusters. You install and access the PCUI in your AWS account. The PCUI is added with AWS ParallelCluster version 3.5.0.

To install the PCUI and get started, see Installing the PCUI and Configure and create a cluster with the PCUI.



The PCUI supports the following features:

- Displays the following:
  - The list of clusters you've created in your AWS account with AWS ParallelCluster.
  - The available status and details for your listed clusters.
  - CloudFormation stack event and AWS ParallelCluster logs that you can use for monitoring.

  - The status of jobs that are running on your clusters.
  - The list of custom images that you can use to build clusters.
  - The list of official images that the UI uses to create clusters.
  - The list of users that have access to the PCUI. You can add and remove users.
- Provides step-by-step guidance for creating and editing (updating) a cluster and selecting supported cluster features to add, edit, or remove. Inaccessible input fields can't be changed for the cluster configuration being edited. You have the option to perform a dry run validation of your cluster configuration before cluster deployment.
- Features direct shell links to access the head node in the **Clusters** view. Choose **Add SSM session** during the step-by-step guidance to add the direct shell access, and the **SSM Managed Instance Core** policy on the head node.

Consider the following when using the PCUI to create and manage your clusters:

- You can only create and edit clusters or build images with the same AWS ParallelCluster version that was used to create the PCUI. Earlier version clusters or images can only be viewed. If you manage multiple versions of clusters and images, we recommend that you create an PCUI instance to support each version.
- The PCUI is designed to mirror the `pcluster` CLI functionality. There are some differences. If you align with the step-by-step guidance, then you are using all of the supported features. Before deployment, you have the option to edit the cluster or image configuration manually. If you do this, we recommend that you validate the configuration by choosing **Dry run** to verify that your edits are fully supported.

> **Note**
>
> PCUI doesn't support AWS Batch.

AWS Lambda VPC configuration in AWS ParallelCluster

AWS ParallelCluster uses AWS Lambda to perform operations during the lifecycle of the cluster. An AWS Lambda function always runs in a VPC owned by the Lambda service. This Lambda function can also be connected to private subnets in a virtual private cloud (VPC) to access private resources.

Note

Lambda functions can't connect directly to a VPC with dedicated instance tenancy. To connect to resources in a dedicated VPC, peer the dedicated VPC to a second VPC with a default tenancy that can connect to a dedicated VPC.
For more information, see Dedicated Instances in the *Amazon EC2 User Guide for Linux Instances* and How do I connect a Lambda function to a dedicated VPC? from the *AWS Knowledge Center*.

Lambda functions that are created by AWS ParallelCluster can be connected to a private VPC. These Lambda functions need to access AWS services. You can provide access through the internet or VPC endpoints by using the following methods.

- **Internet access**

  To access the internet and AWS services, a Lambda function requires network address translation (NAT). Route outbound traffic from your private subnet to a NAT gateway in a public subnet.

- **VPC endpoints**

  Several AWS services offer VPC endpoints. You can use VPC endpoints to connect to AWS services from a VPC that doesn't have internet access. To view the list of AWS ParallelCluster VPC endpoints, see Networking.

Note

Every combination of subnets and security groups must provide access to AWS services using one these methods. Subnets and security groups must be in the same VPC.

For more information, see VPC endpoints in the *Amazon Virtual Private Cloud User Guide* and Internet and service access for VPC-connected functions in the *AWS Lambda Developer Guide*.

To configure the use of Lambda functions and VPCs, see `DeploymentSettings` / `LambdaFunctionsVpcConfig` for clusters or `DeploymentSettings` / `LambdaFunctionsVpcConfig` for images.

AWS Identity and Access Management permissions in AWS ParallelCluster

AWS ParallelCluster uses IAM permissions to control access to resources when creating and managing clusters.

To create and manage clusters in an AWS account, AWS ParallelCluster requires permissions at two levels:

- Permissions that the `pcluster` user requires to invoke the `pcluster` CLI commands for creating and managing clusters.
- Permissions that the cluster resources require to perform cluster actions.

AWS ParallelCluster uses an Amazon EC2 instance profile and role to provide cluster resource permissions. To manage cluster resource permissions, AWS ParallelCluster also requires permissions to IAM resources. For more information, see AWS ParallelCluster user example policies for managing IAM resources.

`pcluster` users require IAM permissions to use the `pcluster` CLI to create and manage a cluster and its resources. These permissions are included in IAM policies that can be added to a user or role. For more information on IAM roles, see Creating a user role in the *AWS Identity and Access Management User Guide*.

You can also use AWS ParallelCluster configuration parameters to manage IAM permissions.

The following sections contain the required permissions with examples.

To use the example policies, replace *<REGION>*, *<AWS ACCOUNT ID>*, and similar strings with the appropriate values.

The following example policies include Amazon Resource Names (ARNs) for the resources. If you're working in the AWS GovCloud (US) or AWS China partitions, the ARNs must be changed. Specifically, they must be changed from "arn:aws" to "arn:aws-us-gov" for the AWS GovCloud (US) partition or "arn:aws-cn" for the AWS China partition. For more information, see Amazon Resource Names (ARNs) in AWS GovCloud (US) Regions in the *AWS GovCloud (US) User Guide* and ARNs for AWS services in China in *Getting Started with AWS services in China*.

You can track changes to the example policies in AWS ParallelCluster documentation on GitHub.

Topics

AWS ParallelCluster Amazon EC2 instance roles

When you create a cluster with the default configuration settings, AWS ParallelCluster uses Amazon EC2 instance profiles to automatically create a default cluster Amazon EC2 instance role that provides the permissions required to create and manage the cluster and its resources.

Alternatives to using the default AWS ParallelCluster instance role

In place of the default AWS ParallelCluster instance role, you can use the `InstanceRole` cluster configuration setting to specify your own existing IAM role for EC2. For more information, see AWS ParallelCluster configuration parameters to manage IAM permissions. Typically, you specify existing IAM roles to fully control the permissions granted to EC2.

If your intent is to add extra policies to the default instance role, we recommend that you pass the additional IAM policies by using the `AdditionalIamPolicies` configuration setting instead of `InstanceProfile` or `InstanceRole` settings. You can update `AdditionalIamPolicies` when you update your cluster, however, you can't update the `InstanceRole` when you update your cluster.

AWS ParallelCluster example `pcluster` user policies

The following examples show the user policies required to create and manage AWS ParallelCluster and its resources by using the `pcluster` CLI. You can attach policies to a user or role.

Topics

Base AWS ParallelCluster `pcluster` user policy

The following policy shows the permissions required to run AWS ParallelCluster `pcluster` commands.

The last action listed in the policy is included to provide validation of any secrets specified in the cluster configuration. For example, an AWS Secrets Manager secret is used to configure the `DirectoryService` integration. In this case, a cluster is created only if a valid secret exists in the `PasswordSecretArn`. If this action is omitted, secret validation is skipped. To improve your security posture, we recommend that you scope down this policy statement by adding only the secrets specified in your cluster configuration.

> **Note**
>
> If existing Amazon EFS file systems are the only file systems used in your cluster, you can scope down the example Amazon EFS policy statements to the specific file systems referenced in the `SharedStorage` section of the cluster configuration file.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "ec2:Describe*"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "EC2Read"
        },
        {
            "Action": [
                "ec2:AllocateAddress",
                "ec2:AssociateAddress",
                "ec2:AttachNetworkInterface",
                "ec2:AuthorizeSecurityGroupEgress",
                "ec2:AuthorizeSecurityGroupIngress",
                "ec2:CreateFleet",
                "ec2:CreateLaunchTemplate",
```

```
                "ec2:CreateLaunchTemplateVersion",
                "ec2:CreateNetworkInterface",
                "ec2:CreatePlacementGroup",
                "ec2:CreateSecurityGroup",
                "ec2:CreateSnapshot",
                "ec2:CreateTags",
                "ec2:DeleteTags",
                "ec2:CreateVolume",
                "ec2:DeleteLaunchTemplate",
                "ec2:DeleteNetworkInterface",
                "ec2:DeletePlacementGroup",
                "ec2:DeleteSecurityGroup",
                "ec2:DeleteVolume",
                "ec2:DisassociateAddress",
                "ec2:ModifyLaunchTemplate",
                "ec2:ModifyNetworkInterfaceAttribute",
                "ec2:ModifyVolume",
                "ec2:ModifyVolumeAttribute",
                "ec2:ReleaseAddress",
                "ec2:RevokeSecurityGroupEgress",
                "ec2:RevokeSecurityGroupIngress",
                "ec2:RunInstances",
                "ec2:TerminateInstances"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "EC2Write"
        },
        {
            "Action": [
                "dynamodb:DescribeTable",
                "dynamodb:ListTagsOfResource",
                "dynamodb:CreateTable",
                "dynamodb:DeleteTable",
                "dynamodb:GetItem",
                "dynamodb:PutItem",
                "dynamodb:UpdateItem",
                "dynamodb:Query",
                "dynamodb:TagResource"
            ],
            "Resource": "arn:aws:dynamodb:*:111122223333:table/parallelcluster-
*",
            "Effect": "Allow",
            "Sid": "DynamoDB"
```

```
        },
        {
            "Action": [
                "route53:ChangeResourceRecordSets",
                "route53:ChangeTagsForResource",
                "route53:CreateHostedZone",
                "route53:DeleteHostedZone",
                "route53:GetChange",
                "route53:GetHostedZone",
                "route53:ListResourceRecordSets",
                "route53:ListQueryLoggingConfigs"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "Route53HostedZones"
        },
        {
            "Action": [
                "cloudformation:*"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "CloudFormation"
        },
        {
            "Action": [
                "cloudwatch:PutDashboard",
                "cloudwatch:ListDashboards",
                "cloudwatch:DeleteDashboards",
                "cloudwatch:GetDashboard",
                "cloudwatch:PutMetricAlarm",
                "cloudwatch:DeleteAlarms",
                "cloudwatch:DescribeAlarms",
                "cloudwatch:PutCompositeAlarm"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "CloudWatch"
        },
        {
            "Action": [
                "iam:GetRole",
                "iam:GetRolePolicy",
                "iam:GetPolicy",
```

```
                "iam:SimulatePrincipalPolicy",
                "iam:GetInstanceProfile"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/*",
                "arn:aws:iam::111122223333:policy/*",
                "arn:aws:iam::aws:policy/*",
                "arn:aws:iam::111122223333:instance-profile/*"
            ],
            "Effect": "Allow",
            "Sid": "IamRead"
        },
        {
            "Action": [
                "iam:CreateInstanceProfile",
                "iam:DeleteInstanceProfile",
                "iam:AddRoleToInstanceProfile",
                "iam:RemoveRoleFromInstanceProfile"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:instance-profile/parallelcluster/*"
            ],
            "Effect": "Allow",
            "Sid": "IamInstanceProfile"
        },
        {
            "Condition": {
                "StringEqualsIfExists": {
                    "iam:PassedToService": [
                        "lambda.amazonaws.com",
                        "ec2.amazonaws.com",
                        "spotfleet.amazonaws.com"
                    ]
                }
            },
            "Action": [
                "iam:PassRole"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/parallelcluster/*"
            ],
            "Effect": "Allow",
            "Sid": "IamPassRole"
        },
```

```
        {
            "Action": [
                "lambda:CreateFunction",
                "lambda:DeleteFunction",
                "lambda:GetFunctionConfiguration",
                "lambda:GetFunction",
                "lambda:InvokeFunction",
                "lambda:AddPermission",
                "lambda:RemovePermission",
                "lambda:UpdateFunctionConfiguration",
                "lambda:TagResource",
                "lambda:ListTags",
                "lambda:UntagResource"
            ],
            "Resource": [
                "arn:aws:lambda:*:111122223333:function:parallelcluster-*",
                "arn:aws:lambda:*:111122223333:function:pcluster-*"
            ],
            "Effect": "Allow",
            "Sid": "Lambda"
        },
        {
            "Action": [
                "s3:*"
            ],
            "Resource": [
                "arn:aws:s3:::parallelcluster-*",
                "arn:aws:s3:::aws-parallelcluster-*"
            ],
            "Effect": "Allow",
            "Sid": "S3ResourcesBucket"
        },
        {
            "Action": [
                "s3:Get*",
                "s3:List*"
            ],
            "Resource": "arn:aws:s3:::*-aws-parallelcluster*",
            "Effect": "Allow",
            "Sid": "S3ParallelClusterReadOnly"
        },
        {
            "Action": [
                "elasticfilesystem:*"
```

```
            ],
            "Resource": [
                "arn:aws:elasticfilesystem:*:111122223333:*"
            ],
            "Effect": "Allow",
            "Sid": "EFS"
        },
        {
            "Action": [
                "logs:DeleteLogGroup",
                "logs:PutRetentionPolicy",
                "logs:DescribeLogGroups",
                "logs:CreateLogGroup",
                "logs:TagResource",
                "logs:UntagResource",
                "logs:FilterLogEvents",
                "logs:GetLogEvents",
                "logs:CreateExportTask",
                "logs:DescribeLogStreams",
                "logs:DescribeExportTasks",
                "logs:DescribeMetricFilters",
                "logs:PutMetricFilter",
                "logs:DeleteMetricFilter"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "CloudWatchLogs"
        },
        {
            "Action": [
                "resource-groups:ListGroupResources"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "ResourceGroupRead"
        },
        {
            "Sid": "AllowDescribingFileCache",
            "Effect": "Allow",
            "Action": [
                "fsx:DescribeFileCaches"
            ],
            "Resource": "*"
        },
```

```
        {
            "Action": "secretsmanager:DescribeSecret",
            "Resource": "arn:aws:secretsmanager:us-
east-1:111122223333:secret:<SECRET NAME>",
            "Effect": "Allow"
        }
    ]
}
```

Additional AWS ParallelCluster `pcluster` user policy when using AWS Batch scheduler

In case you need to create and manage a cluster with AWS Batch scheduler, the following additional policy is required.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Condition": {
                "StringEqualsIfExists": {
                    "iam:PassedToService": [
                        "ecs-tasks.amazonaws.com",
                        "batch.amazonaws.com",
                        "codebuild.amazonaws.com"
                    ]
                }
            },
            "Action": [
                "iam:PassRole"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/parallelcluster/*"
            ],
            "Effect": "Allow",
            "Sid": "IamPassRole"
        },
        {
            "Condition": {
```

```
                "StringEquals": {
                    "iam:AWSServiceName": [
                        "batch.amazonaws.com"
                    ]
                }
            },
            "Action": [
                "iam:CreateServiceLinkedRole",
                "iam:DeleteServiceLinkedRole"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/aws-service-role/
batch.amazonaws.com/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "codebuild:*"
            ],
            "Resource": "arn:aws:codebuild:*:111122223333:project/pcluster-*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "ecr:*"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "ECR"
        },
        {
            "Action": [
                "batch:*"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "Batch"
        },
        {
            "Action": [
                "events:*"
            ],
            "Resource": "*",
```

```
            "Effect": "Allow",
            "Sid": "AmazonCloudWatchEvents"
        },
        {
            "Action": [
                "ecs:DescribeContainerInstances",
                "ecs:ListContainerInstances"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "ECS"
        }
    ]
}
```

Additional AWS ParallelCluster `pcluster` user policy when using Amazon FSx for Lustre

In case you need to create and manage a cluster with Amazon FSx for Lustre, the following additional policy is required.

> **Note**
>
> If existing Amazon FSx file systems are the only file systems used in your cluster, you can scope down the example Amazon FSx policy statements to the specific file systems referenced in the SharedStorage section of the cluster configuration file.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Condition": {
                "StringEquals": {
                    "iam:AWSServiceName": [
                        "fsx.amazonaws.com",
                        "s3.data-source.lustre.fsx.amazonaws.com"
                    ]
```

```
                }
            },
            "Action": [
                "iam:CreateServiceLinkedRole",
                "iam:DeleteServiceLinkedRole"
            ],
            "Resource": "*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "fsx:*"
            ],
            "Resource": [
                "arn:aws:fsx:*:111122223333:*"
            ],
            "Effect": "Allow",
            "Sid": "FSx"
        },
        {
            "Action": [
                "iam:CreateServiceLinkedRole",
                "iam:AttachRolePolicy",
                "iam:PutRolePolicy"
            ],
            "Resource": "arn:aws:iam::111122223333:role/aws-service-role/s3.data-
source.lustre.fsx.amazonaws.com/*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "s3:Get*",
                "s3:List*",
                "s3:PutObject"
            ],
            "Resource": "arn:aws:s3:::amzn-s3-demo-bucket",
            "Effect": "Allow"
        }
    ]
}
```

AWS ParallelCluster image build `pcluster` user policy

Users that intend to create custom Amazon EC2 images with AWS ParallelCluster must have the following set of permissions.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "ec2:DescribeSecurityGroups",
                "ec2:DescribeImages",
                "ec2:DescribeInstanceTypeOfferings",
                "ec2:DescribeInstanceTypes",
                "ec2:DeregisterImage",
                "ec2:DeleteSnapshot"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "EC2"
        },
        {
            "Action": [
                "iam:CreateInstanceProfile",
                "iam:AddRoleToInstanceProfile",
                "iam:GetRole",
                "iam:GetRolePolicy",
                "iam:GetInstanceProfile",
                "iam:RemoveRoleFromInstanceProfile"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:instance-profile/parallelcluster/*",
                "arn:aws:iam::111122223333:instance-profile/
ParallelClusterImage*",
                "arn:aws:iam::111122223333:role/parallelcluster/*"
            ],
            "Effect": "Allow",
            "Sid": "IAM"
        },
        {
            "Condition": {
```

```
                "StringEquals": {
                    "iam:PassedToService": [
                        "lambda.amazonaws.com",
                        "ec2.amazonaws.com"
                    ]
                }
            },
            "Action": [
                "iam:PassRole"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:instance-profile/parallelcluster/*",
                "arn:aws:iam::111122223333:role/parallelcluster/*"
            ],
            "Effect": "Allow",
            "Sid": "IAMPassRole"
        },
        {
            "Action": [
                "logs:GetLogEvents",
                "logs:CreateLogGroup",
                "logs:TagResource",
                "logs:UntagResource",
                "logs:DeleteLogGroup"
            ],
            "Resource": [
                "arn:aws:logs:*:111122223333:log-group:/aws/imagebuilder/
ParallelClusterImage-*",
                "arn:aws:logs:*:111122223333:log-group:/aws/lambda/
ParallelClusterImage-*"
            ],
            "Effect": "Allow",
            "Sid": "CloudWatch"
        },
        {
            "Action": [
                "cloudformation:DescribeStacks",
                "cloudformation:CreateStack",
                "cloudformation:DeleteStack"
            ],
            "Resource": [
                "arn:aws:cloudformation:*:111122223333:stack/*"
            ],
            "Effect": "Allow",
```

```
            "Sid": "CloudFormation"
        },
        {
            "Action": [
                "lambda:CreateFunction",
                "lambda:GetFunction",
                "lambda:AddPermission",
                "lambda:RemovePermission",
                "lambda:DeleteFunction",
                "lambda:TagResource",
                "lambda:ListTags",
                "lambda:UntagResource"
            ],
            "Resource": [
                "arn:aws:lambda:*:111122223333:function:ParallelClusterImage-*"
            ],
            "Effect": "Allow",
            "Sid": "Lambda"
        },
        {
            "Action": [
                "imagebuilder:Get*"
            ],
            "Resource": "*",
            "Effect": "Allow",
            "Sid": "ImageBuilderGet"
        },
        {
            "Action": [
                "imagebuilder:CreateImage",
                "imagebuilder:TagResource",
                "imagebuilder:CreateImageRecipe",
                "imagebuilder:CreateComponent",
                "imagebuilder:CreateDistributionConfiguration",
                "imagebuilder:CreateInfrastructureConfiguration",
                "imagebuilder:DeleteImage",
                "imagebuilder:DeleteComponent",
                "imagebuilder:DeleteImageRecipe",
                "imagebuilder:DeleteInfrastructureConfiguration",
                "imagebuilder:DeleteDistributionConfiguration"
            ],
            "Resource": [
                "arn:aws:imagebuilder:*:111122223333:image/parallelclusterimage-
*",
```

```
                "arn:aws:imagebuilder:*:111122223333:image-recipe/
parallelclusterimage-*",
                "arn:aws:imagebuilder:*:111122223333:component/
parallelclusterimage-*",
                "arn:aws:imagebuilder:*:111122223333:distribution-configuration/
parallelclusterimage-*",
                "arn:aws:imagebuilder:*:111122223333:infrastructure-
configuration/parallelclusterimage-*"
            ],
            "Effect": "Allow",
            "Sid": "ImageBuilder"
        },
        {
            "Action": [
                "s3:CreateBucket",
                "s3:ListBucket",
                "s3:ListBucketVersions"
            ],
            "Resource": [
                "arn:aws:s3:::parallelcluster-*"
            ],
            "Effect": "Allow",
            "Sid": "S3Bucket"
        },
        {
            "Action": [
                "sns:GetTopicAttributes",
                "sns:TagResource",
                "sns:CreateTopic",
                "sns:Subscribe",
                "sns:Publish",
                "SNS:DeleteTopic",
                "SNS:Unsubscribe"
            ],
            "Resource": [
                "arn:aws:sns:*:111122223333:ParallelClusterImage-*"
            ],
            "Effect": "Allow",
            "Sid": "SNS"
        },
        {
            "Action": [
                "s3:PutObject",
                "s3:GetObject",
```

```
                "s3:GetObjectVersion",
                "s3:DeleteObject",
                "s3:DeleteObjectVersion"
            ],
            "Resource": [
                "arn:aws:s3:::parallelcluster-*/*"
            ],
            "Effect": "Allow",
            "Sid": "S3Objects"
        },
        {
            "Action": "iam:CreateServiceLinkedRole",
            "Effect": "Allow",
            "Resource": "arn:aws:iam::*:role/aws-service-role/
imagebuilder.amazonaws.com/AWSServiceRoleForImageBuilder",
            "Condition": {
                "StringLike": {
                    "iam:AWSServiceName": "imagebuilder.amazonaws.com"
                }
            }
        }
    ]
}
```

AWS ParallelCluster user example policies for managing IAM resources

When using AWS ParallelCluster to create clusters or custom AMIs, IAM policies must be provided that contain permissions to grant the required set of permissions to AWS ParallelCluster components. These IAM resources can be either automatically created by AWS ParallelCluster or be provided as input when creating a cluster or a custom image.

You can use the following modes to provide the AWS ParallelCluster user with the permissions required to access IAM resources by using additional IAM policies in the configuration.

Topics

- Privileged IAM access mode
- Restricted IAM access mode
- PermissionsBoundary mode

Privileged IAM access mode

With this mode, AWS ParallelCluster automatically creates all necessary IAM resources. These IAM policies are scoped down to enable access to cluster resources only.

To enable Privileged IAM access mode, add the following policy to the user role.

Note

If you configure HeadNode / Iam / AdditionalPolicies or Scheduling / SlurmQueues / Iam / AdditionalPolicies parameters, you must provide the AWS ParallelCluster user with permission to attach and detach role policies for each additional policy as shown in the following policy. Add the additional policy ARNs to the condition for attaching and detaching role policies.

Warning

This mode enables the user to have IAM Administrator privileges in the AWS account

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "iam:CreateServiceLinkedRole",
                "iam:DeleteRole",
                "iam:TagRole"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/parallelcluster/*"
            ],
            "Effect": "Allow",
            "Sid": "IamRole"
        },
        {
            "Action": [
```

```
                "iam:CreateRole"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/parallelcluster/*"
            ],
            "Effect": "Allow",
            "Sid": "IamCreateRole"
        },
        {
            "Action": [
                "iam:PutRolePolicy",
                "iam:DeleteRolePolicy"
            ],
            "Resource": "arn:aws:iam::111122223333:role/parallelcluster/*",
            "Effect": "Allow",
            "Sid": "IamInlinePolicy"
        },
        {
            "Condition": {
                "ArnLike": {
                    "iam:PolicyARN": [
                        "arn:aws:iam::111122223333:policy/parallelcluster*",
                        "arn:aws:iam::111122223333:policy/parallelcluster/*",
                        "arn:aws:iam::aws:policy/CloudWatchAgentServerPolicy",
                        "arn:aws:iam::aws:policy/AmazonSSMManagedInstanceCore",
                        "arn:aws:iam::aws:policy/AWSBatchFullAccess",
                        "arn:aws:iam::aws:policy/AmazonS3ReadOnlyAccess",
                        "arn:aws:iam::aws:policy/service-role/
AWSBatchServiceRole",
                        "arn:aws:iam::aws:policy/service-role/
AmazonEC2ContainerServiceforEC2Role",
                        "arn:aws:iam::aws:policy/service-role/
AmazonECSTaskExecutionRolePolicy",
                        "arn:aws:iam::aws:policy/service-role/
AmazonEC2SpotFleetTaggingRole",
                        "arn:aws:iam::aws:policy/
EC2InstanceProfileForImageBuilder",
                        "arn:aws:iam::aws:policy/service-role/
AWSLambdaBasicExecutionRole"
                    ]
                }
            },
            "Action": [
                "iam:AttachRolePolicy",
```

```
                "iam:DetachRolePolicy"
            ],
            "Resource": "arn:aws:iam::111122223333:role/parallelcluster/*",
            "Effect": "Allow",
            "Sid": "IamPolicy"
        }
    ]
}
```

Restricted IAM access mode

When no additional IAM policies are granted to the user, IAM roles required by clusters or custom image build need to be manually created by an administrator and passed as part of the cluster configuration.

When creating a cluster the following parameters are required:

- Iam / Roles / LambdaFunctionsRole
- HeadNode / Iam / InstanceRole | InstanceProfile
- Scheduling / SlurmQueues / Iam / InstanceRole | InstanceProfile

When building a custom image the following parameters are required:

- Build / Iam / InstanceRole | InstanceProfile
- Build / Iam / CleanupLambdaRole

The IAM roles passed as part of the above listed parameters must be created on the `/parallelcluster/` path prefix. If this isn't possible, the user policy needs to be updated to grant `iam:PassRole` permission on the specific custom roles, as in the following example.

```
{
   "Condition": {
       "StringEqualsIfExists": {
           "iam:PassedToService": [
               "ecs-tasks.amazonaws.com",
               "lambda.amazonaws.com",
               "ec2.amazonaws.com",
               "spotfleet.amazonaws.com",
               "batch.amazonaws.com",
```

```
                "codebuild.amazonaws.com"
            ]
        }
    },
    "Action": [
        "iam:PassRole"
    ],
    "Resource": [
        <list all custom IAM roles>
    ],
    "Effect": "Allow",
    "Sid": "IamPassRole"
}
```

Warning

Currently this mode does not allow the management of AWS Batch clusters because not all IAM roles can be passed in the cluster configuration.

`PermissionsBoundary` mode

This mode delegates to AWS ParallelCluster the creation of IAM roles that are bound to the configured IAM permissions boundary. For more information on IAM permissions boundaries, see Permissions boundaries for IAM entities in the *IAM User Guide*.

The following policy needs to be added to the user role.

In the policy, replace *<permissions-boundary-arn>* with the IAM policy ARN to be enforced as permissions boundary.

Warning

If you configure the `HeadNode` / `Iam` / `AdditionalPolicies` or `Scheduling` / `SlurmQueues` / `Iam` / `AdditionalPolicies` parameters, you must grant the user permission to attach and detach role policies for each additional policy as shown in the following policy. Add the additional policy ARNs to the condition for attaching and detaching role policies.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "iam:CreateServiceLinkedRole",
                "iam:DeleteRole",
                "iam:TagRole"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/parallelcluster/*"
            ],
            "Effect": "Allow",
            "Sid": "IamRole"
        },
        {
            "Condition": {
                "StringEquals": {
                    "iam:PermissionsBoundary": [
                        "<permissions-boundary-arn>"
                    ]
                }
            },
            "Action": [
                "iam:CreateRole"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/parallelcluster/*"
            ],
            "Effect": "Allow",
            "Sid": "IamCreateRole"
        },
        {
            "Condition": {
                "StringEquals": {
                    "iam:PermissionsBoundary": [
                        "<permissions-boundary-arn>"
                    ]
                }
            },
            "Action": [
```

```
                "iam:PutRolePolicy",
                "iam:DeleteRolePolicy"
            ],
            "Resource": "arn:aws:iam::111122223333:role/parallelcluster/*",
            "Effect": "Allow",
            "Sid": "IamInlinePolicy"
        },
        {
            "Condition": {
                "StringEquals": {
                    "iam:PermissionsBoundary": [
                        "<permissions-boundary-arn>"
                    ]
                },
                "ArnLike": {
                    "iam:PolicyARN": [
                        "arn:aws:iam::111122223333:policy/parallelcluster*",
                        "arn:aws:iam::111122223333:policy/parallelcluster/*",
                        "arn:aws:iam::aws:policy/CloudWatchAgentServerPolicy",
                        "arn:aws:iam::aws:policy/AmazonSSMManagedInstanceCore",
                        "arn:aws:iam::aws:policy/AWSBatchFullAccess",
                        "arn:aws:iam::aws:policy/AmazonS3ReadOnlyAccess",
                        "arn:aws:iam::aws:policy/service-role/
AWSBatchServiceRole",
                        "arn:aws:iam::aws:policy/service-role/
AmazonEC2ContainerServiceforEC2Role",
                        "arn:aws:iam::aws:policy/service-role/
AmazonECSTaskExecutionRolePolicy",
                        "arn:aws:iam::aws:policy/service-role/
AmazonEC2SpotFleetTaggingRole",
                        "arn:aws:iam::aws:policy/
EC2InstanceProfileForImageBuilder",
                        "arn:aws:iam::aws:policy/service-role/
AWSLambdaBasicExecutionRole"
                    ]
                }
            },
            "Action": [
                "iam:AttachRolePolicy",
                "iam:DetachRolePolicy"
            ],
            "Resource": "arn:aws:iam::111122223333:role/parallelcluster/*",
            "Effect": "Allow",
            "Sid": "IamPolicy"
```

```
        }
    ]
}
```

When this mode is enabled, you must specify the permissions boundary ARN in the `Iam` / `PermissionsBoundary` configuration parameter when creating or updating a cluster and in the `Build` / `Iam` / `PermissionBoundary` parameter when building a custom image.

AWS ParallelCluster configuration parameters to manage IAM permissions

AWS ParallelCluster exposes a series of configuration options to customize and manage the IAM permissions and roles that are used in a cluster or during the custom AMI creation process.

Topics

Cluster configuration

Topics

Head node IAM role

`HeadNode` / `Iam` / `InstanceRole` | `InstanceProfile`

With this option, you override the default IAM role that's assigned to the head node of the cluster. For additional details, please refer to the `InstanceProfile` reference.

Here is the minimal set of policies to be used as part of this role when the scheduler is Slurm:

- `arn:aws:iam::aws:policy/CloudWatchAgentServerPolicy` managed IAM policy. For more information, see Create IAM roles and users for use with the CloudWatch agent in the *Amazon CloudWatch User Guide*.
- `arn:aws:iam::aws:policy/AmazonSSMManagedInstanceCore` managed IAM policy. For more information, see AWS managed policies for AWS Systems Manager in the *AWS Systems Manager User Guide*.
- Additional IAM policy:

  JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "s3:GetObject",
                "s3:GetObjectVersion"
            ],
            "Resource": [
                "arn:aws:s3:::us-east-1-aws-parallelcluster/*",
                "arn:aws:s3:::dcv-license.us-east-1/*",
                "arn:aws:s3:::parallelcluster-*-v1-do-not-delete/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "dynamodb:GetItem",
                "dynamodb:PutItem",
                "dynamodb:UpdateItem",
                "dynamodb:BatchWriteItem",
                "dynamodb:BatchGetItem"
            ],
            "Resource": "arn:aws:dynamodb:us-east-1:111122223333:table/
parallelcluster-*",
            "Effect": "Allow"
        },
        {
            "Condition": {
                "StringEquals": {
                    "ec2:ResourceTag/parallelcluster:node-type": "Compute"
                }
            },
```

```
            "Action": "ec2:TerminateInstances",
            "Resource": "*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "ec2:RunInstances",
                "ec2:CreateFleet"
            ],
            "Resource": "*",
            "Effect": "Allow"
        },
        {
            "Condition": {
                "StringEquals": {
                    "iam:PassedToService": [
                        "ec2.amazonaws.com"
                    ]
                }
            },
            "Action": [
                "iam:PassRole"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/parallelcluster/*",
                "arn:aws:iam::111122223333:instance-profile/parallelcluster/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "ec2:DescribeInstances",
                "ec2:DescribeInstanceStatus",
                "ec2:DescribeVolumes",
                "ec2:DescribeInstanceAttribute",
                "ec2:DescribeCapacityReservations"
            ],
            "Resource": "*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "ec2:CreateTags",
                "ec2:AttachVolume"
```

```
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1:111122223333:instance/*",
                "arn:aws:ec2:us-east-1:111122223333:volume/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "cloudformation:DescribeStacks",
                "cloudformation:DescribeStackResource",
                "cloudformation:SignalResource"
            ],
            "Resource": "*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "route53:ChangeResourceRecordSets"
            ],
            "Resource": "*",
            "Effect": "Allow"
        },
        {
            "Action": "secretsmanager:GetSecretValue",
            "Resource": "arn:aws:secretsmanager:us-
east-1:111122223333:secret:<SECRET_ID>",
            "Effect": "Allow"
        }
    ]
}
```

Note that in case `Scheduling` / `SlurmQueues` / `Iam` / `InstanceRole` is used to override the compute IAM role, the head node policy reported above needs to include such role in the `Resource` section of the `iam:PassRole` permission.

Here is the minimal set of policies to be used as part of this role when the scheduler is AWS Batch:

- `arn:aws:iam::aws:policy/CloudWatchAgentServerPolicy` managed IAM policy. For more information, see Create IAM roles and users for use with the CloudWatch agent in the *Amazon CloudWatch User Guide*.

- `arn:aws:iam::aws:policy/AmazonSSMManagedInstanceCore` managed IAM policy. For more information, see AWS managed policies for AWS Systems Manager in the *AWS Systems Manager User Guide*.
- Additional IAM policy:

  JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "s3:GetObject",
                "s3:PutObject",
                "s3:GetObjectVersion"
            ],
            "Resource": [
                "arn:aws:s3:::parallelcluster-*-v1-do-not-delete/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": "s3:GetObject",
            "Resource": [
                "arn:aws:s3:::dcv-license.us-east-1/*",
                "arn:aws:s3:::us-east-1-aws-parallelcluster/*"
            ],
            "Effect": "Allow"
        },
        {
            "Condition": {
                "StringEquals": {
                    "iam:PassedToService": [
                        "batch.amazonaws.com"
                    ]
                }
            },
            "Action": [
                "iam:PassRole"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/parallelcluster/*",
                "arn:aws:iam::111122223333:instance-profile/parallelcluster/*"
```

```
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "batch:DescribeJobQueues",
                "batch:DescribeJobs",
                "batch:ListJobs",
                "batch:DescribeComputeEnvironments"
            ],
            "Resource": "*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "batch:SubmitJob",
                "batch:TerminateJob",
                "logs:GetLogEvents",
                "ecs:ListContainerInstances",
                "ecs:DescribeContainerInstances"
            ],
            "Resource": [
                "arn:aws:logs:us-east-1:111122223333:log-group:/aws/batch/
job:log-stream:PclusterJobDefinition*",
                "arn:aws:ecs:us-east-1:111122223333:container-instance/
AWSBatch-PclusterComputeEnviron*",
                "arn:aws:ecs:us-east-1:111122223333:cluster/AWSBatch-
Pcluster*",
                "arn:aws:batch:us-east-1:111122223333:job-queue/
PclusterJobQueue*",
                "arn:aws:batch:us-east-1:111122223333:job-definition/
PclusterJobDefinition*:*",
                "arn:aws:batch:us-east-1:111122223333:job/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "ec2:DescribeInstances",
                "ec2:DescribeInstanceStatus",
                "ec2:DescribeVolumes",
                "ec2:DescribeInstanceAttribute"
            ],
            "Resource": "*",
```

```
            "Effect": "Allow"
        },
        {
            "Action": [
                "ec2:CreateTags",
                "ec2:AttachVolume"
            ],
            "Resource": [
                "arn:aws:ec2:us-east-1:111122223333:instance/*",
                "arn:aws:ec2:us-east-1:111122223333:volume/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": [
                "cloudformation:DescribeStackResource",
                "cloudformation:DescribeStacks",
                "cloudformation:SignalResource"
            ],
            "Resource": "*",
            "Effect": "Allow"
        },
        {
            "Action": "secretsmanager:GetSecretValue",
            "Resource": "arn:aws:secretsmanager:us-
east-1:111122223333:secret:<SECRET_ID>",
            "Effect": "Allow"
        }
    ]
}
```

Amazon S3 access

HeadNode / Iam / S3Access or Scheduling / SlurmQueues / S3Access

In these configuration sections, you can customize the Amazon S3 access by granting additional Amazon S3 policies to the IAM roles associated with the head node or compute nodes of the cluster when such roles are created by AWS ParallelCluster. For more information, see the reference documentation for each of the configuration parameter.

This parameter can be only used when the user is configured with Privileged IAM access mode or PermissionsBoundary mode.

Additional IAM policies

`HeadNode` / `Iam` / `AdditionalIamPolicies` or `SlurmQueues` / `Iam` / `AdditionalIamPolicies`

Use this option to attach additional managed IAM policies to the IAM roles associated with the head node or compute nodes of the cluster when such roles are created by AWS ParallelCluster.

> **Warning**
>
> To use this option, make sure the AWS ParallelCluster user is granted `iam:AttachRolePolicy` and `iam:DetachRolePolicy` permissions for the IAM policies that need to be attached.

AWS Lambda functions role

`Iam` / `Roles` / `LambdaFunctionsRole`

This option overrides the role attached to all AWS Lambda functions that are used during the cluster creation process. AWS Lambda needs to be configured as the principal allowed to assume the role.

> **Note**
>
> If `DeploymentSettings` / `LambdaFunctionsVpcConfig` is set, the `LambdaFunctionsRole` must include the AWS Lambda role permission to set the VPC configuration.

Here is the minimal set of policies to be used as part of this role:

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "route53:ListResourceRecordSets",
```

```
                "route53:ChangeResourceRecordSets"
            ],
            "Resource": "arn:aws:route53:::hostedzone/*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "logs:CreateLogStream",
                "logs:PutLogEvents"
            ],
            "Effect": "Allow",
            "Resource": "arn:aws:logs:us-east-1:111122223333:log-group:/aws/
lambda/pcluster-*"
        },
        {
            "Action": "ec2:DescribeInstances",
            "Effect": "Allow",
            "Resource": "*"
        },
        {
            "Action": "ec2:TerminateInstances",
            "Condition": {
                "StringEquals": {
                    "ec2:ResourceTag/parallelcluster:node-type": "Compute"
                }
            },
            "Effect": "Allow",
            "Resource": "*"
        },
        {
            "Action": [
                "s3:DeleteObject",
                "s3:DeleteObjectVersion",
                "s3:ListBucket",
                "s3:ListBucketVersions"
            ],
            "Effect": "Allow",
            "Resource": [
                "arn:aws:s3:::parallelcluster-*-v1-do-not-delete",
                "arn:aws:s3:::parallelcluster-*-v1-do-not-delete/*"
            ]
        }
    ]
```

```
}
```

Compute nodes IAM role

Scheduling / SlurmQueues / Iam / InstanceRole | InstanceProfile

This option allows to override the IAM role that is assigned to the compute nodes of the cluster. For more information, see InstanceProfile.

Here is the minimal set of policies to be used as part of this role:

- arn:aws:iam::aws:policy/CloudWatchAgentServerPolicy managed IAM policy. For more information, see Create IAM roles and users for use with the CloudWatch agent in the *Amazon CloudWatch User Guide*.
- arn:aws:iam::aws:policy/AmazonSSMManagedInstanceCore managed IAM policy. For more information, see AWS managed policies for AWS Systems Manager in the *AWS Systems Manager User Guide*.
- Additional IAM policy:

  JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "dynamodb:Query",
                "dynamodb:UpdateItem",
                "dynamodb:PutItem",
                "dynamodb:GetItem"
            ],
            "Resource": "arn:aws:dynamodb:us-east-1:111122223333:table/
parallelcluster-*",
            "Effect": "Allow"
        },
        {
            "Action": "s3:GetObject",
            "Resource": [
                "arn:aws:s3:::us-east-1-aws-parallelcluster/*"
            ],
            "Effect": "Allow"
        },
```

```
        {
            "Action": "ec2:DescribeInstanceAttribute",
            "Resource": "*",
            "Effect": "Allow"
        },
        {
            "Action": "cloudformation:DescribeStackResource",
            "Resource": [
                "arn:aws:cloudformation:us-east-1:111122223333:stack/*/*"
            ],
            "Effect": "Allow"
        }
    ]
}
```

Permissions boundary

Iam / PermissionsBoundary

This parameter forces AWS ParallelCluster to attach the given IAM policy as a `PermissionsBoundary` to all IAM roles that are created as part of a cluster deployment.

See PermissionsBoundary mode for the list of policies required by the user when this setting is defined.

Custom Image configuration

Topics

- Instance role for EC2 Image Builder
- AWS Lambda cleanup role
- Additional IAM policies
- Permissions boundary

Instance role for EC2 Image Builder

Build / Iam / InstanceRole | InstanceProfile

With this option you override the IAM role that is assigned to the Amazon EC2 instance launched by EC2 Image Builder to create a custom AMI.

Here is the minimal set of policies to be used as part of this role:

- `arn:aws:iam::aws:policy/AmazonSSMManagedInstanceCore` managed IAM policy. For more information, see AWS managed policies for AWS Systems Manager in the *AWS Systems Manager User Guide*.
- `arn:aws:iam::aws:policy/EC2InstanceProfileForImageBuilder` managed IAM policy. For more information, see `EC2InstanceProfileForImageBuilder` policy in the *Image Builder User Guide*.
- Additional IAM policy:

  JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "ec2:CreateTags",
                "ec2:ModifyImageAttribute"
            ],
            "Resource": "arn:aws:ec2:us-east-1::image/*",
            "Effect": "Allow"
        }
    ]
}
```

AWS Lambda cleanup role

`Build` / `Iam` / `CleanupLambdaRole`

This option overrides the role attached to all AWS Lambda functions that are used during the custom image build process. AWS Lambda needs to be configured as the principal allowed to assume the role.

Note

If `DeploymentSettings` / `LambdaFunctionsVpcConfig` is set, the `CleanupLambdaRole` must include the AWS Lambda role permission to set the VPC configuration.

Here is the minimal set of policies to be used as part of this role:

- `arn:aws:iam::aws:policy/service-role/AWSLambdaBasicExecutionRole` managed IAM policy. For more information, see AWS managed policies for Lambda features in the *AWS Lambda Developer Guide*.
- Additional IAM policy:

  JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "iam:DetachRolePolicy",
                "iam:DeleteRole",
                "iam:DeleteRolePolicy"
            ],
            "Resource": "arn:aws:iam::111122223333:role/parallelcluster/*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "iam:DeleteInstanceProfile",
                "iam:RemoveRoleFromInstanceProfile"
            ],
            "Resource": "arn:aws:iam::111122223333:instance-profile/
parallelcluster/*",
            "Effect": "Allow"
        },
        {
            "Action": "imagebuilder:DeleteInfrastructureConfiguration",
            "Resource": "arn:aws:imagebuilder:us-
east-1:111122223333:infrastructure-configuration/parallelclusterimage-*",
            "Effect": "Allow"
```

```
        },
        {
            "Action": [
                "imagebuilder:DeleteComponent"
            ],
            "Resource": [
                "arn:aws:imagebuilder:us-east-1:111122223333:component/
parallelclusterimage-*/*"
            ],
            "Effect": "Allow"
        },
        {
            "Action": "imagebuilder:DeleteImageRecipe",
            "Resource": "arn:aws:imagebuilder:us-east-1:111122223333:image-
recipe/parallelclusterimage-*/*",
            "Effect": "Allow"
        },
        {
            "Action": "imagebuilder:DeleteDistributionConfiguration",
            "Resource": "arn:aws:imagebuilder:us-
east-1:111122223333:distribution-configuration/parallelclusterimage-*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "imagebuilder:DeleteImage",
                "imagebuilder:GetImage",
                "imagebuilder:CancelImageCreation"
            ],
            "Resource": "arn:aws:imagebuilder:us-east-1:111122223333:image/
parallelclusterimage-*/*",
            "Effect": "Allow"
        },
        {
            "Action": "cloudformation:DeleteStack",
            "Resource": "arn:aws:cloudformation:us-east-1:111122223333:stack/*/
*",
            "Effect": "Allow"
        },
        {
            "Action": "ec2:CreateTags",
            "Resource": "arn:aws:ec2:us-east-1::image/*",
            "Effect": "Allow"
        },
```

```
        {
            "Action": "tag:TagResources",
            "Resource": "*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "lambda:DeleteFunction",
                "lambda:RemovePermission"
            ],
            "Resource": "arn:aws:lambda:us-
east-1:111122223333:function:ParallelClusterImage-*",
            "Effect": "Allow"
        },
        {
            "Action": "logs:DeleteLogGroup",
            "Resource": "arn:aws:logs:us-east-1:111122223333:log-group:/aws/
lambda/ParallelClusterImage-*:*",
            "Effect": "Allow"
        },
        {
            "Action": [
                "SNS:GetTopicAttributes",
                "SNS:DeleteTopic",
                "SNS:GetSubscriptionAttributes",
                "SNS:Unsubscribe"
            ],
            "Resource": "arn:aws:sns:us-
east-1:111122223333:ParallelClusterImage-*",
            "Effect": "Allow"
        }
    ]
}
```

Additional IAM policies

Build / Iam / AdditionalIamPolicies

You use this option to attach additional managed IAM policies to the role associated with the Amazon EC2 instance used by EC2 Image Builder to produce the custom AMI.

Warning

To use this option, make sure the AWS ParallelClusteruser is granted `iam:AttachRolePolicy` and `iam:DetachRolePolicy` permissions for the IAM policies that need to be attached.

Permissions boundary

`Build` / `Iam` / `PermissionsBoundary`

This parameter forces AWS ParallelCluster to attach the given IAM policy as a `PermissionsBoundary` to all IAM roles that are created as part of custom AMI build.

See `PermissionsBoundary mode` for the list of policies required to use such functionality.

Network configurations

AWS ParallelCluster uses Amazon Virtual Private Cloud (VPC) for networking. VPC provides a flexible and configurable networking platform where you can deploy clusters.

The VPC must have `DNS Resolution = yes`, `DNS Hostnames = yes` and DHCP options with the correct domain name for the Region. The default DHCP Option Set already specifies the required *AmazonProvidedDNS*. If specifying more than one domain name server, see DHCP options sets in the *Amazon VPC User Guide*.

AWS ParallelCluster supports the following high-level configurations:

- One subnet for both head and compute nodes.
- Two subnets, with the head node in one public subnet, and compute nodes in a private subnet. The subnets can be either new or existing ones.

All of these configurations can operate with or without public IP addressing. AWS ParallelCluster can also be deployed to use an HTTP proxy for all AWS requests. The combinations of these configurations result in many deployment scenarios. For example, you can configure a single public subnet with all access over the internet. Or, you can configure a fully private network using AWS Direct Connect and HTTP proxy for all traffic.

Starting from AWS ParallelCluster 3.0.0 it is possible to configure different `SecurityGroups`, `AdditionalSecurityGroups` and `PlacementGroup` settings for each queue. For more information, see `HeadNode` / `Networking` and `SlurmQueues` / `Networking` and `AwsBatchQueues` / `Networking`.

For illustrations of some networking scenarios, see the following architecture diagrams.

Topics

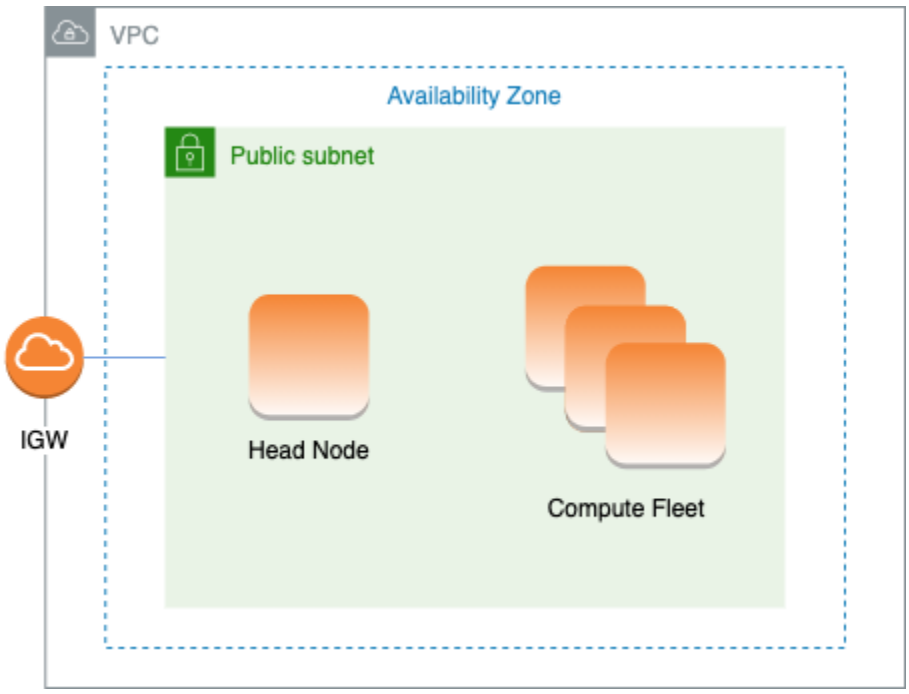
AWS ParallelCluster in a single public subnet

In this configuration, all instances of the cluster must be assigned a public IP in order to get internet access. To achieve this, do the following:

- Make sure the head node is assigned a public IP address by either turning on the "Enable auto-assign public IPv4 address" setting for the subnet used in `HeadNode` / `Networking` / `SubnetId` or by assigning an Elastic IP in `HeadNode` / `Networking` / `ElasticIp`.
- Make sure the compute nodes are assigned a public IP address by either turning on the "Enable auto-assign public IPv4 address" setting for the subnet used in `Scheduling` / `SlurmQueues` / `Networking` / `SubnetIds` or by setting `AssignPublicIp`: true in `Scheduling` / `SlurmQueues` / `Networking`.
- If you define a p4d instance type, or another instance type that has multiple network interfaces or a network interface card to the head node, you must set `HeadNode` / `Networking` / `ElasticIp` to `true` to provide public access. AWS public IPs can only be assigned to instances launched with a single network interface. For this case, we recommend that you use a NAT gateway to provide public access to the cluster compute nodes. For more information on IP addresses, see Assign a public IPv4 address during instance launch in the *Amazon EC2 User Guide for Linux Instances*.
- You can't define a p4d or hp6id instance type, or another instance type that has multiple network interfaces or a network interface card to compute nodes because AWS public IPs can only be assigned to instances launched with a single network interface. For more information

on IP addresses, see Assign a public IPv4 address during instance launch in the *Amazon EC2 User Guide for Linux Instances*.

For more information, see Enabling internet access in *Amazon VPC User Guide*.



The configuration for this architecture requires the following settings:

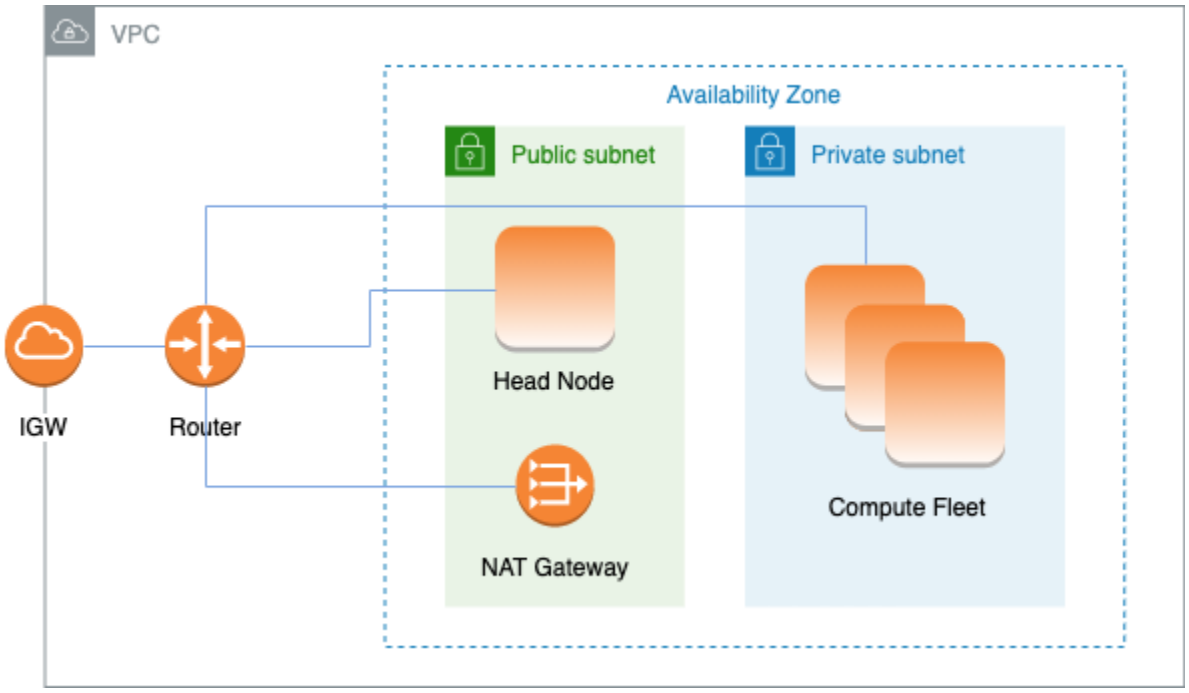
```
# Note that all values are only provided as examples
HeadNode:
  ...
  Networking:
    SubnetId: subnet-12345678 # subnet with internet gateway
    #ElasticIp: true | false | eip-12345678
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - ...
      Networking:
        SubnetIds:
          - subnet-12345678 # subnet with internet gateway
        #AssignPublicIp: true
```

AWS ParallelCluster using two subnets

In this configuration, only the head node of the cluster is required to have a public IP assigned. You can achieve this by either turning on the "Enable auto-assign public IPv4 address" setting for the subnet used in `HeadNode` / `Networking` / `SubnetId` or by assigning an Elastic IP in `HeadNode` / `Networking` / `ElasticIp`.

If you define a p4d instance type or another instance type that has multiple network interfaces or a network interface card to the head node, you must set `HeadNode` / `Networking` / `ElasticIp` to `true` to provide public access. AWS public IPs can only be assigned to instances launched with a single network interface. For more information on IP addresses, see Assign a public IPv4 address during instance launch in the *Amazon EC2 User Guide for Linux Instances*.

This configuration requires a NAT gateway or an internal proxy in the subnet used for the queues, to give internet access to the compute instances.



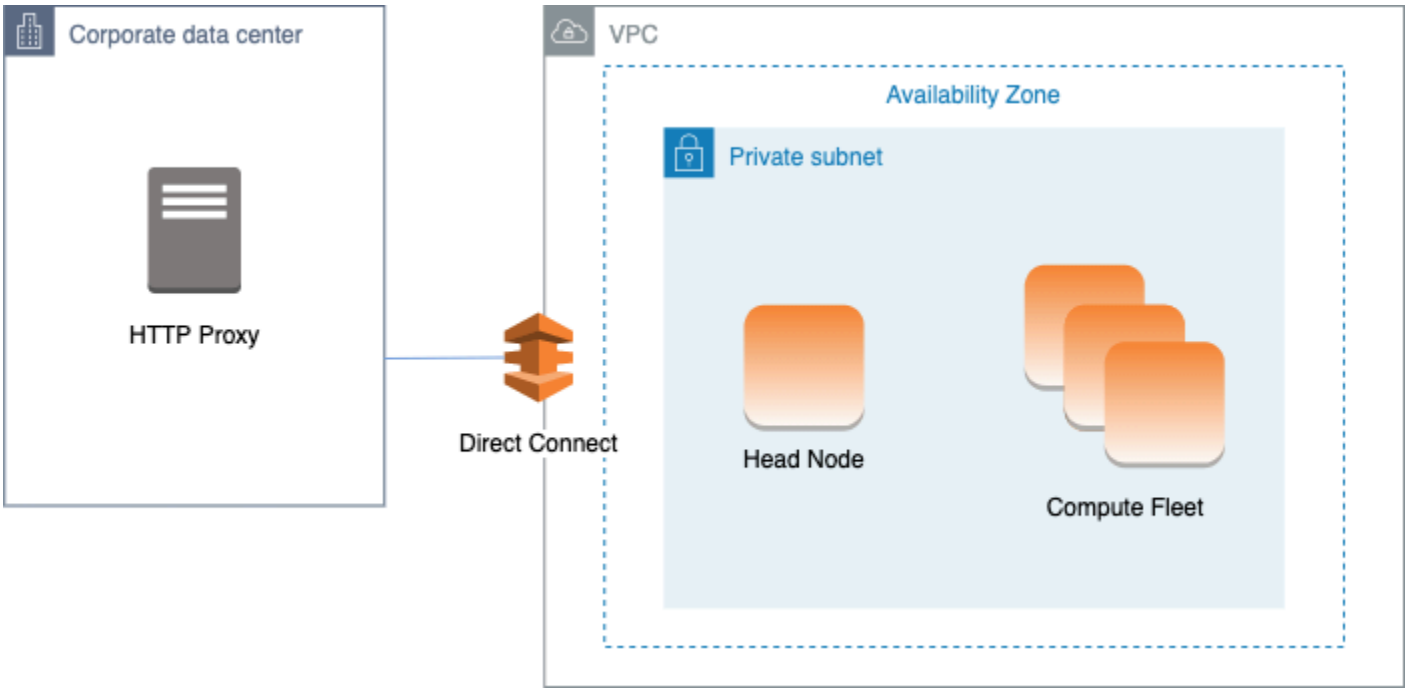
The configuration to use an existing private subnet for compute instances requires the following settings:

```
# Note that all values are only provided as examples
HeadNode:
  ...
  Networking:
```

```
    SubnetId: subnet-12345678 # subnet with internet gateway
    #ElasticIp: true | false | eip-12345678
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - ...
      Networking:
        SubnetIds:
          - subnet-23456789 # subnet with NAT gateway
        #AssignPublicIp: false
```

AWS ParallelCluster in a single private subnet connected using AWS Direct Connect

When `Scheduling` / `SlurmQueues` / `Networking` / `AssignPublicIp` is set to `false`, the subnets must be correctly set up to use the Proxy for all traffic. Web access is required for both head and compute nodes.



The configuration for this architecture requires the following settings:

```
# Note that all values are only provided as examples
HeadNode:
  ...
```

```
  Networking:
    SubnetId: subnet-34567890 # subnet with proxy
    Proxy:
      HttpProxyAddress: http://proxy-address:port
  Ssh:
    KeyName: ec2-key-name
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - ...
      Networking:
        SubnetIds:
          - subnet-34567890 # subnet with proxy
        AssignPublicIp: false
        Proxy:
          HttpProxyAddress: http://proxy-address:port
```

AWS ParallelCluster with AWS Batch scheduler

When you use `awsbatch` as the scheduler type, AWS ParallelCluster creates an AWS Batch managed compute environment. The AWS Batch environment manages Amazon Elastic Container Service (Amazon ECS) container instances. These instances are launched in the subnet configured in the `AwsBatchQueues` / `Networking` / `SubnetIds` parameter. For AWS Batch to function correctly, Amazon ECS container instances need external network access to communicate with the Amazon ECS service endpoint. This translates into the following scenarios:

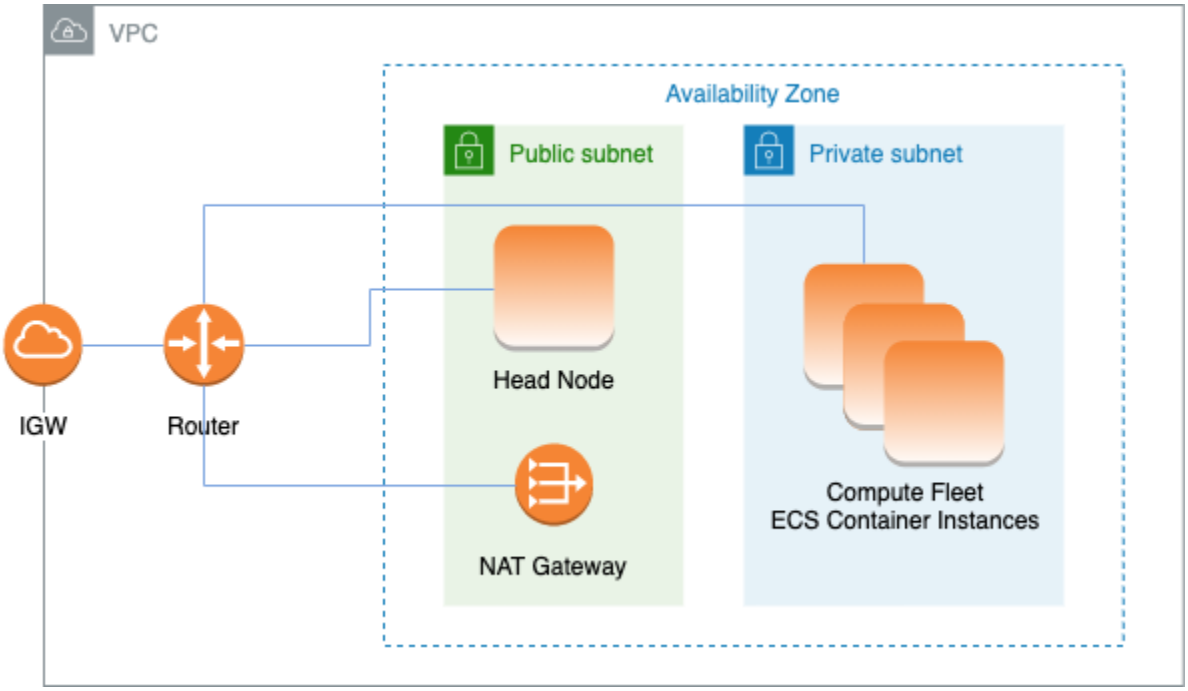
- The Subnet ID specified for the queue uses a NAT gateway to access the internet. We recommended this approach.
- Instances launched in the queue subnet have public IP addresses and can reach the internet through an Internet Gateway.

Additionally, if you're interested in multi-node parallel jobs (from the AWS Batch docs):

AWS Batch multi-node parallel jobs use the Amazon ECS `awsvpc` network mode. This gives your multi-node parallel job containers the same networking properties as Amazon EC2 instances. Each multi-node parallel job container gets its own elastic network interface, a primary private IP address, and an internal DNS hostname. The network interface is created in the same Amazon VPC subnet as its host compute resource. Any security groups that are applied to your compute resources are also applied to it.

When using Amazon ECS Task Networking, the `awsvpc` network mode doesn't provide elastic network interfaces with public IP addresses for tasks that use the Amazon EC2 launch type. To access the internet, tasks that use the Amazon EC2 launch type must be launched in a private subnet that's configured to use a NAT gateway.

You must configure a NAT gateway in order to enable the cluster to run multi-node parallel jobs.



All the previous configuration and considerations are valid for AWS Batch, too. The following is an example of a AWS Batch networking configuration.

```
# Note that all values are only provided as examples
HeadNode:
  ...
  Networking:
    SubnetId: subnet-12345678 # subnet with internet gateway, NAT gateway or proxy
    #ElasticIp: true | false | eip-12345678
    #Proxy:
      #HttpProxyAddress: http://proxy-address:port
  Ssh:
    KeyName: ec2-key-name
Scheduling:
  Scheduler: awsbatch
  AwsBatchQueues:
    - ...
      Networking:
```

```
        SubnetIds:
          - subnet-23456789 # subnet with internet gateway, NAT gateway or proxy
        #AssignPublicIp: true | false
```

In the `Scheduling` / `AwsBatchQueues` / `Networking` section, the `SubnetIds` is a list type but, currently, only one subnet is supported.

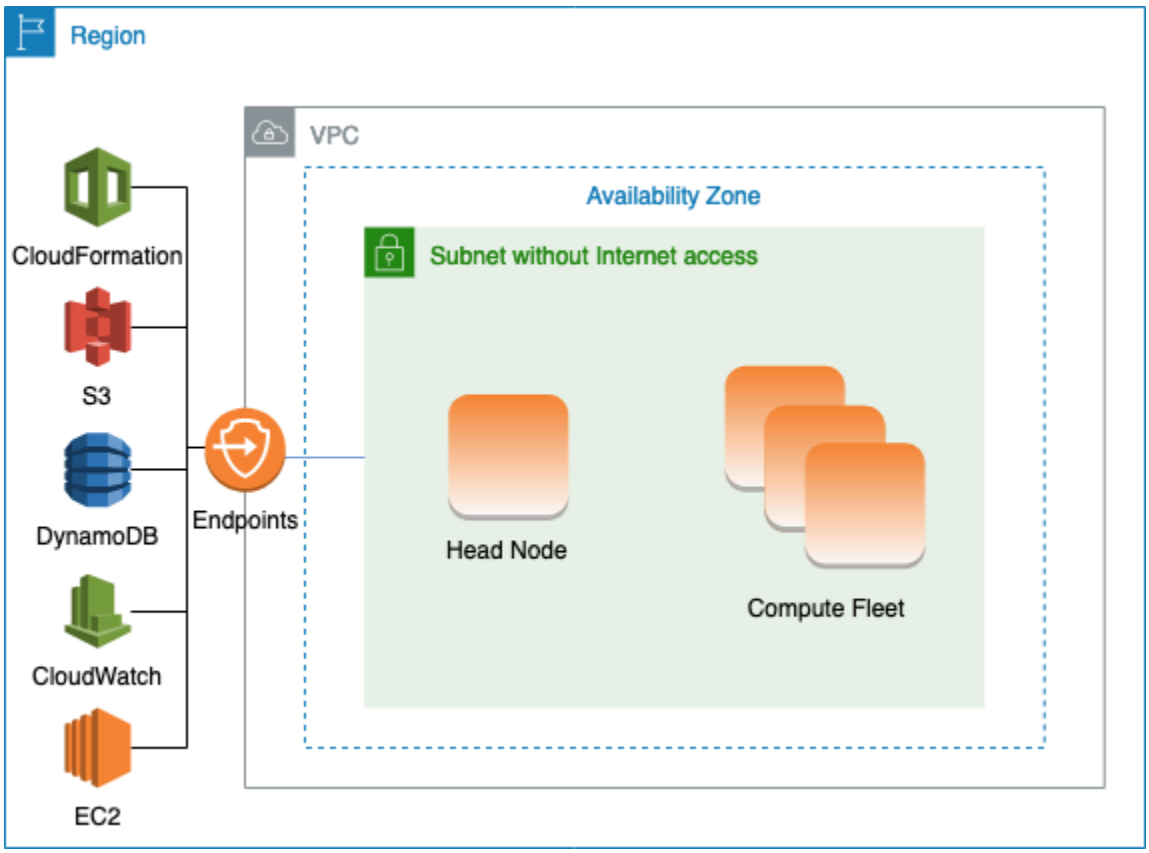
For more information, see the following topics:

- AWS Batch managed compute environments
- AWS Batch multi-node parallel jobs
- Amazon ECS task networking with the awsvpc network mode

AWS ParallelCluster in a single subnet with no internet access

A subnet without internet access doesn't allow inbound or outbound connections to the internet. This AWS ParallelCluster configuration can help customers concerned with security further enhance the security of their AWS ParallelCluster resources. AWS ParallelCluster nodes are built from AWS ParallelCluster AMIs that include all of the software that's required to run a cluster with no internet access. This way, AWS ParallelCluster can create and manage clusters with nodes that don't have internet access.

In this section, you learn about how to configure the cluster. You also learn about limitations in running clusters without internet access.



Configuring VPC endpoints

To ensure the proper functioning of the cluster, the cluster nodes must be able to interact with a number of AWS Services.

Create and configure the following VPC endpoints so that cluster nodes can interact with the AWS Services, without internet access:

Commercial and AWS GovCloud (US) partitions

Service	Service name	Type
Amazon CloudWatch	com.amazonaws.*region-id*.logs	Interface
AWS CloudFormation	com.amazonaws.*region-id*.cloudformation	Interface
Amazon EC2	com.amazonaws.*region-id*.ec2	Interface

Service	Service name	Type
Amazon S3	com.amazonaws.*region-id* .s3	Gateway
Amazon DynamoDB	com.amazonaws.*region-id* .dynamodb	Gateway
AWS Secrets Manager**	com.amazonaws.*region-id* .secretsmanager	Interface
AWS Elastic Load Balancing ***	com.amazonaws.*region-id* .elasticloadbalancing	Interface
AWS Auto Scaling***	com.amazonaws.*region-id* .autoscaling	Interface

China partition

Service	Service name	Type
Amazon CloudWatch	com.amazonaws.*region-id* .logs	Interface
AWS CloudFormation	cn.com.amazonaws.*region-id* .cloudformation	Interface
Amazon EC2	cn.com.amazonaws.*region-id* .ec2	Interface
Amazon S3	com.amazonaws.*region-id* .s3	Gateway
Amazon DynamoDB	com.amazonaws.*region-id* .dynamodb	Gateway
AWS Secrets Manager**	com.amazonaws.*region-id* .secretsmanager	Interface

Service	Service name	Type
AWS Elastic Load Balancing ***	com.amazonaws.*region-id* .elasticloadbalancing	Interface
AWS Auto Scaling***	cn.com.amazonaws.*region-id* .autoscaling	Interface

** This endpoint is only required when `DirectoryService` is enabled, otherwise it is optional.

*** These endpoints are only required when LoginNodes is enabled, otherwise they are optional.

All instances in the VPC must have proper security groups to communicate with the endpoints. You can do this by adding security groups to `AdditionalSecurityGroups` under the `HeadNode` and `AdditionalSecurityGroups` under the `SlurmQueues` configurations. For example, if the VPC endpoints are created without explicitly specifying a security group, the default security group is associated with the endpoints. By adding the default security group to `AdditionalSecurityGroups`, you enable the communication between the cluster and the endpoints.

Note

When you use IAM policies to restrict access to VPC endpoints, you must add the following to the Amazon S3 VPC endpoint:

```
PolicyDocument:
  Version: 2012-10-17
  Statement:
    - Effect: Allow
      Principal: "*"
      Action:
        - "s3:PutObject"
      Resource:
        - !Sub "arn:${AWS::Partition}:s3:::cloudformation-waitcondition-
${AWS::Region}/*"
```

Disable Route 53 and use Amazon EC2 hostnames

When creating a Slurm cluster, AWS ParallelCluster creates a private Route 53 hosted zone that is used to resolve the custom compute node hostnames, such as `{queue_name}-{st|dy}-{compute_resource}-{N}`. Because Route 53 doesn't support VPC endpoints, this feature must be disabled. Additionally, AWS ParallelCluster must be configured to use the default Amazon EC2 hostnames, such as `ip-1-2-3-4`. Apply the following settings to your cluster configuration:

```
...
Scheduling:
  ...
  SlurmSettings:
    Dns:
      DisableManagedDns: true
      UseEc2Hostnames: true
```

Warning

For clusters created with `SlurmSettings` / `Dns` / `DisableManagedDns` and `UseEc2Hostnames` set to `true`, the Slurm `NodeName` isn't resolved by the DNS. Use the Slurm `NodeHostName` instead.

Note

This note isn't relevant starting with AWS ParallelCluster version 3.3.0.
For AWS ParallelCluster supported versions prior to 3.3.0:
When `UseEc2Hostnames` is set to `true`, the Slurm configuration file is set with the AWS ParallelCluster `prolog` and `epilog` scripts:

- `prolog` runs to add nodes info to `/etc/hosts` on compute nodes when each job is allocated.
- `epilog` runs to clean contents written by `prolog`.

To add custom `prolog` or `epilog` scripts, add them to the `/opt/slurm/etc/pcluster/prolog.d/` or `/opt/slurm/etc/pcluster/epilog.d/` folders respectively.

Cluster configuration

Learn how to configure your cluster to run in a subnet with no connection to the internet.

The configuration for this architecture requires the following settings:

```
# Note that all values are only provided as examples
...
HeadNode:
  ...
  Networking:
    SubnetId: subnet-1234567890abcdef0 # the VPC of the subnet needs to have VPC
 endpoints
    AdditionalSecurityGroups:
      - sg-abcdef01234567890 # optional, the security group that enables the
 communication between the cluster and the VPC endpoints
LoginNodes: # optional, if enabled, requires creation and configuration of VPC
 endpoints for AWS Elastic Load Balancing (ELB) and Auto Scaling services
  Pools:
    - ...
      Networking:
        SubnetIds:
          - subnet-1234567890abcdef0 # the VPC of the subnet needs to have VPC
 endpoints attached
        AdditionalSecurityGroups:
          - sg-1abcdef01234567890 # optional, the security group that enables the
 communication between the cluster and the VPC endpoints
Scheduling:
  Scheduler: Slurm # Cluster in a subnet without internet access is supported only when
 the scheduler is Slurm.
  SlurmSettings:
    Dns:
      DisableManagedDns: true
      UseEc2Hostnames: true
  SlurmQueues:
    - ...
      Networking:
        SubnetIds:
          - subnet-1234567890abcdef0 # the VPC of the subnet needs to have VPC
 endpoints attached
        AdditionalSecurityGroups:
          - sg-1abcdef01234567890 # optional, the security group that enables the
 communication between the cluster and the VPC endpoints
```

- SubnetId(s): The subnet without internet access.

To enable communication between AWS ParallelCluster and AWS Services, the VPC of the subnet must have the VPC endpoints attached. Before you create your cluster, verify that auto-assign public IPv4 address is disabled in the subnet to ensure that the `pcluster` commands have access to the cluster.

- `AdditionalSecurityGroups`: The security group that enables the communication between the cluster and the VPC endpoints.

  Optional:

  - If the VPC endpoints are created without explicitly specifying a security group, the default security group of the VPC is associated. Therefore, provide the default security group to `AdditionalSecurityGroups`.
  - If custom security groups are used when creating the cluster and/or the VPC endpoints, `AdditionalSecurityGroups` is unnecessary as long as the custom security groups enable communication between the cluster and the VPC endpoints.

- `Scheduler`: The cluster scheduler.

  `slurm` is the only valid value. Only the Slurm scheduler supports a cluster in a subnet without internet access.

- `SlurmSettings`: The Slurm settings.

  See the preceding section *Disable Route53 and use Amazon EC2 hostnames*.

Limitations

- *Connecting to the head node over SSH or Amazon DCV:* When connecting to a cluster, make sure the client of the connection can reach the head node of the cluster through its private IP address. If the client isn't in the same VPC as the head node, use a proxy instance in a public subnet of the VPC. This requirement applies to both SSH and DCV connections. The public IP of a head node isn't accessible if the subnet doesn't have internet access. The `pcluster ssh` and `dcv-connect` commands use the public IP if it exists or the private IP. Before you create your cluster, verify that auto-assign public IPv4 address is disabled in the subnet to ensure that the `pcluster` commands have access to the cluster.

  The following example shows how you can connect to a DCV session running in the head node of your cluster. You connect through a proxy Amazon EC2 instance. The instance functions as a Amazon DCV server for your PC and as the client for the head node in the private subnet.

Connect over DCV through a proxy instance in a public subnet:

1. Create an Amazon EC2 instance in a public subnet, which is in the same VPC as the cluster's subnet.
2. Ensure that the Amazon DCV client and server are installed on your Amazon EC2 instance.
3. Attach an AWS ParallelCluster User Policy to the proxy Amazon EC2 instance. For more information, see AWS ParallelCluster example `pcluster` user policies.
4. Install AWS ParallelCluster on the proxy Amazon EC2 instance.
5. Connect over DCV to the proxy Amazon EC2 instance.
6. Use the `pcluster dcv-connect` command on the proxy instance to connect to the cluster inside the subnet without internet access.

- *Interacting with other AWS services:* Only services strictly required by AWS ParallelCluster are listed above. If your cluster must interact with other services, create the corresponding VPC endpoints.

Login nodes provisioned by AWS ParallelCluster

Starting from version 3.7.0, AWS ParallelCluster cluster administrators can provision login nodes that can be used to provide access to users to run jobs vs directly accessing the cluster head node. Cluster users with appropriate permissions can use Active Directory or their ssh credential to login, submit and manage their jobs. As a result, cluster management can be improved and the chances of depleting the resources of the head node required by Slurm to manage the cluster can be minimized. Logged in users will also have access to all shared storage of the cluster mounted on login nodes. If a login node needs to be stopped, users that are logged in will be notified in advance through the active shell session they are using.

Login nodes are specified as pools where a pool defines a group of login nodes that have the same resource configuration. All the login nodes in a pool are configured to be part of a network load balancer that enables distributing sessions across login nodes in a round-robin fashion. The present implementation allows users to specify multiple login node pools.

Security for login nodes

Login nodes inherit the `AllowedIPs` settings `AllowedIps` from the head node, unless AllowedIps is specified for the login node pool. In this manner, cluster administrators can restrict the security

posture of the cluster by specifying the source CIDR or a prefix list from where SSH connections are allowed on either the head node or a pool of login nodes.

In the present implementation the access to the head node is not automatically restricted when enabling login nodes. If needed, a cluster administrator can restrict this access updating the head nodes ssh configuration using standard Linux commands. This can be also be accomplished by specifying custom Security Groups on the head node by using the `AdditionalSecurityGroups` setting in the head node section of the ParallelCluster YAML file to deny connections from unauthorized users.

Networking for login nodes

Login nodes are provisioned with a single connection address to the network load balancer configured for the pool of login nodes. The connectivity settings of the address are based on the type of subnet specified in the Login nodes Pool configuration.

- If the subnet is private, the address will be private and, in order to grant access to the login nodes, the cluster administrator must provision a bastion host.
- If the subnet is public, the address will be public

All connection requests are managed by the Network Load Balancer using round-robin routing.

Storage for login nodes

All shared storage configured on the cluster using ParallelCluster including managed storage will be mounted on all the login nodes.

Retrieve login nodes information

To retrieve the address of the single connection provisioned to access the login nodes, the cluster administrator can run the `describe-cluster` command. The command will also provide more information about the status of the login nodes.

Login nodes are a new node type supported by ParallelCluster that can be specified with the `describe-cluster-instances` command when querying the status of a specific node type.

The availability of a single connection address to the Login nodes pool doesn't prevent direct access to a specific login node. However, it is not recommended to use the direct connection to avoid warnings from the ssh client. The ssh client stores host identifiers locally for each

target address. Since the host identifier is specific per pool, use of different IPs and/or the single connection address may have the same host identifier associated with different target addresses: this may cause warning from the ssh client since the same host identifier is associated multiple targets.

Imds properties for login nodes

Access to the login node's IMDS (and the instance profile credentials) is restricted to root user, cluster administrative user (`pc-cluster-admin`by default) and operating system specific default user (`ec2-user`on Amazon Linux 2 and RedHat, and `ubuntu`on Ubuntu 18.04.

To restrict IMDS access, AWS ParallelCluster manages a chain of`iptables`.

> **Note**
>
> Any customization of`iptables`or`ip6tables`rules can interfere with the mechanism used to restrict IMDS access on the login node.See also `Imds property setting`.

Login Nodes lifecycle

Currently, there is no dedicated command to stop and start the login nodes in a pool. In order to stop the login nodes in a pool the cluster administrator has to update the cluster configuration specifying zero on the count of login nodes (`Count: 0`) and then run an `pcluster.update-cluster-v3` command.

> **Note**
>
> Logged in users are notified about the termination of the specific instance and about the related gracetime period. During the gracetime period no new connections will be allowed except for the ones from the cluster default user. The message shown is customizable by the cluster administrator from the head node or from a login node editing the file `/opt/parallelcluster/shared_login_nodes/loginmgtd_config.json`. This termination message is not visible when you are connected using the AWS Systems Manager Session Manager Session Manager.

In order to start the login nodes pool the cluster administrator has to restore the previous `Count` value in the cluster configuration and then run an `update-cluster` command.

Permissions required to run the login nodes pool

In order to manage the login nodes pool the cluster administrator must have the following additional permissions:

```
        - Action:
          - iam:CreateServiceLinkedRole
          - autoscaling:DeleteAutoScalingGroup
          - autoscaling:DeleteLifecycleHook
          - autoscaling:Describe*
          - autoscaling:PutLifecycleHook
          - autoscaling:UpdateAutoScalingGroup
          - elasticloadbalancing:CreateListener
          - elasticloadbalancing:CreateTargetGroup
          - elasticloadbalancing:DeleteListener
          - elasticloadbalancing:DeleteLoadBalancer
          - elasticloadbalancing:DeleteTargetGroup
          - elasticloadbalancing:Describe*
          - elasticloadbalancing:ModifyLoadBalancerAttributes
        Resource: '*'
        Condition:
          ForAllValues:StringEquals:
            aws:TagKeys: [ "parallelcluster:cluster-name" ]
        - Action:
          - autoscaling:CreateAutoScalingGroup
          - elasticloadbalancing:AddTags
          - elasticloadbalancing:CreateLoadBalancer
        Resource: '*'
        Effect: Allow
```

Custom bootstrap actions

If you define the `HeadNode` / `CustomActions` / `OnNodeStart` configuration settings, AWS ParallelCluster runs arbitrary code immediately after the node starts. If you define the `HeadNode` / `CustomActions` / `OnNodeConfigured` configuration settings, AWS ParallelCluster runs the code after the node configuration is correctly completed.

Starting with AWS ParallelCluster version 3.4.0, the code can be run after the head node update, if you define the `HeadNode` / `CustomActions` / `OnNodeUpdated` configuration settings.

In most cases, this code is stored in Amazon Simple Storage Service (Amazon S3) and accessed through an HTTPS connection. The code is run as `root` and can be in any script language that's supported by the cluster OS. Often the code is in *Bash* or *Python*.

Note

Starting with AWS ParallelCluster version 3.7.0, the cluster `Imds` / `ImdsSupport` default setting is `v2.0`.
When you create a new cluster to upgrade to version 3.7.0 and later versions, either update your custom bootstrap action scripts to be compatible with IMDSv2 or set `Imds` / `ImdsSupport` to `v1.0` in your cluster configuration file.

Warning

It is your responsibility to configure the custom scripts and arguments as described in the Shared responsibility model. Verify that your custom bootstrap scripts and arguments are from sources that you trust to have full access to your cluster nodes.

Warning

AWS ParallelCluster doesn't support the use of internal variables that are provided through the `/etc/parallelcluster/cfnconfig` file. This file might be removed as part of a future release.

`OnNodeStart` actions are called before any node deployment bootstrap action is started, such as configuring NAT, Amazon Elastic Block Store (Amazon EBS) or the scheduler. `OnNodeStart` bootstrap actions may include modifying storage, adding extra users, and adding packages.

Note

If you configure `DirectoryService` and a `HeadNode` / `CustomActions` / `OnNodeStart` script for your cluster, AWS ParallelCluster configures `DirectoryService` and restarts the `sssd`, before it runs the `OnNodeStart` script.

`OnNodeConfigured` actions are called after the node bootstrap processes are complete. `OnNodeConfigured` actions serve the last actions to occur before an instance is considered fully configured and complete. Some `OnNodeConfigured` actions include changing scheduler settings, modifying storage, and modifying packages. You can pass arguments to scripts by specifying them during configuration.

`OnNodeUpdated` actions are called after the head node update is completed and the scheduler and shared storage are aligned with the latest cluster configuration changes.

When `OnNodeStart` or `OnNodeConfigured` custom actions succeed, success is indicated with exit code zero (0). Any other exit code indicates the instance bootstrap failed.

When `OnNodeUpdated` custom actions succeed, success is signaled with exit code zero (0). Any other exit code indicates the update failed.

> **Note**
>
> If you configure `OnNodeUpdated`, you must manually restore the `OnNodeUpdated` actions to the previous state on update failures.
> If an `OnNodeUpdated` custom action fails, the update rolls back to the previous state. However, the `OnNodeUpdated` action is only run at update time and not at stack rollback time.

You can specify different scripts for the head node and for each queue, in the `HeadNode` / `CustomActions` and i `Scheduling` / `SlurmQueues` / `CustomActions` configuration sections. `OnNodeUpdated` can only be configured in the `HeadNode` section.

> **Note**
>
> Before AWS ParallelCluster version 3.0, it was not possible to specify different scripts for head and compute nodes. Please refer to Moving from AWS ParallelCluster 2.x to 3.x.

Topics

- Configuration settings to define actions and arguments
- Arguments
- Example cluster with custom bootstrap actions

- Example of how to update a custom bootstrap script for IMDSv2
- Example of how to update a configuration for IMDSv1

Configuration settings to define actions and arguments

The following configuration settings are used to define `HeadNode` / `CustomActions` / `OnNodeStart` & `OnNodeConfigured` & `OnNodeUpdated` and `Scheduling` / `CustomActions` / `OnNodeStart` & `OnNodeConfigured` actions and arguments.

```
HeadNode:
  [...]
  CustomActions:
    OnNodeStart:
      # Script URL. This is run before any of the bootstrap scripts are run
      Script: s3://amzn-s3-demo-bucket/on-node-start.sh
      Args:
        - arg1
    OnNodeConfigured:
      # Script URL. This is run after all the bootstrap scripts are run
      Script: s3://amzn-s3-demo-bucket/on-node-configured.sh
      Args:
        - arg1
    OnNodeUpdated:
      # Script URL. This is run after the head node update is completed.
      Script: s3://amzn-s3-demo-bucket/on-node-updated.sh
      Args:
        - arg1
  # Bucket permissions
  Iam:
    S3Access:
      - BucketName: bucket_name
        EnableWriteAccess: false
Scheduling:
  Scheduler: slurm
   [...]
  SlurmQueues:
    - Name: queue1
      [...]
      CustomActions:
        OnNodeStart:
          Script: s3://amzn-s3-demo-bucket/on-node-start.sh
          Args:
```

```
            - arg1
        OnNodeConfigured:
          Script: s3://amzn-s3-demo-bucket/on-node-configured.sh
          Args:
            - arg1
      Iam:
        S3Access:
          - BucketName: bucket_name
            EnableWriteAccess: false
```

Using the Sequence setting (added in AWS ParallelCluster version 3.6.0):

```
HeadNode:
  [...]
  CustomActions:
    OnNodeStart:
      # Script URLs. The scripts are run in the same order as listed in the
 configuration, before any of the bootstrap scripts are run.
      Sequence:
        - Script: s3://amzn-s3-demo-bucket/on-node-start1.sh
          Args:
            - arg1
        - Script: s3://amzn-s3-demo-bucket/on-node-start2.sh
          Args:
            - arg1
        [...]
    OnNodeConfigured:
      # Script URLs. The scripts are run in the same order as listed in the
 configuration, after all the bootstrap scripts are run.
      Sequence:
        - Script: s3://amzn-s3-demo-bucket/on-node-configured1.sh
          Args:
            - arg1
        - Script: s3://amzn-s3-demo-bucket/on-node-configured2.sh
          Args:
            - arg1
        [...]
    OnNodeUpdated:
      # Script URLs. The scripts are run in the same order as listed in the
 configuration, after the head node update is completed.
      Sequence:
        - Script: s3://amzn-s3-demo-bucket/on-node-updated1.sh
          Args:
```

```
            - arg1
        - Script: s3://amzn-s3-demo-bucket/on-node-updated2.sh
          Args:
            - arg1
        [...]
  # Bucket permissions
  Iam:
    S3Access:
      - BucketName: bucket_name
        EnableWriteAccess: false
Scheduling:
  Scheduler: slurm
   [...]
  SlurmQueues:
    - Name: queue1
      [...]
      CustomActions:
        OnNodeStart:
          # Script URLs. The scripts are run in the same order as listed in the
 configuration, before any of the bootstrap scripts are run
          Sequence:
            - Script: s3://amzn-s3-demo-bucket/on-node-start1.sh
              Args:
                - arg1
            - Script: s3://amzn-s3-demo-bucket/on-node-start2.sh
              Args:
                - arg1
            [...]
        OnNodeConfigured:
          # Script URLs. The scripts are run in the same order as listed in the
 configuration, after all the bootstrap scripts are run
          Sequence:
            - Script: s3://amzn-s3-demo-bucket/on-node-configured1.sh
              Args:
                - arg1
            - Script: s3://amzn-s3-demo-bucket/on-node-configured2.sh
              Args:
                - arg1
            [...]
      Iam:
        S3Access:
          - BucketName: bucket_name
            EnableWriteAccess: false
```

The Sequence setting is added starting with AWS ParallelCluster version 3.6.0. When you specify Sequence, you can list multiple scripts for a custom action. AWS ParallelCluster continues to support configuring a custom action with a single script, without including Sequence.

AWS ParallelCluster doesn't support including both a single script and Sequence for the same custom action. For example, AWS ParallelCluster fails if you specify the following configuration.

```
[...]
  CustomActions:
    OnNodeStart:
      # Script URL. This is run before any of the bootstrap scripts are run
      Script: s3://amzn-s3-demo-bucket/on-node-start.sh
          Args:
            - arg1
      # Script URLs. The scripts are run in the same order as listed in the
 configuration, before any of the bootstrap scripts are run.
      Sequence:
        - Script: s3://amzn-s3-demo-bucket/on-node-start1.sh
          Args:
            - arg1
        - Script: s3://amzn-s3-demo-bucket/on-node-start2.sh
          Args:
            - arg1
[...]
```

Arguments

In AWS ParallelCluster 2.x the $1 arguments was a reserved one, to store the URL of the custom script. If you want to re-use the custom bootstrap scripts created for AWS ParallelCluster 2.x with AWS ParallelCluster 3.x you need to adapt them by considering the shift of the arguments. Please refer to Moving from AWS ParallelCluster 2.x to 3.x.

Example cluster with custom bootstrap actions

The following steps create a simple script to be executed after the node is configured, that installs the `R`, `curl` and `wget` packages in the nodes of the cluster.

1. Create a script.

```
#!/bin/bash
  echo "The script has $# arguments"
```

```
  for arg in "$@"
  do
      echo "arg: ${arg}"
  done
  yum -y install "${@:1}"
```

2. Upload the script with the correct permissions to Amazon S3. If public read permissions aren't appropriate for you, use HeadNode / Iam / S3Access and Scheduling / SlurmQueues configuration sections. For more information, see Working with Amazon S3.

```
$ aws s3 cp --acl public-read /path/to/myscript.sh s3://amzn-s3-demo-bucket/myscript.sh
```

> **Important**
>
> If the script was edited on Windows, line endings must be changed from CRLF to LF before the script is uploaded to Amazon S3.

3. Update the AWS ParallelCluster configuration to include the new OnNodeConfigured action.

```
CustomActions:
  OnNodeConfigured:
    Script: https://<amzn-s3-demo-bucket>.s3.<region>.amazonaws.com/myscript.sh
    Args:
      - "R"
      - "curl"
      - "wget"
```

If the bucket doesn't have public-read permission, use s3 as the URL protocol.

```
CustomActions:
  OnNodeConfigured:
    Script: s3://amzn-s3-demo-bucket/myscript.sh
    Args:
      - "R"
      - "curl"
      - "wget"
```

4. Launch the cluster.

```
$ pcluster create-cluster --cluster-name mycluster \
```

```
  --region <region> --cluster-configuration config-file.yaml
```

5. Verify the output.

- If you added custom actions to the `HeadNode` configuration, log in to the head node and check the `cfn-init.log` file located at `/var/log/cfn-init.log` by running the following command:

```
$ less /var/log/cfn-init.log
  2021-09-03 10:43:54,588 [DEBUG] Command run
  postinstall output: The script has 3 arguments
  arg: R
  arg: curl
  arg: wget
  Loaded plugins: dkms-build-requires, priorities, update-motd, upgrade-helper
  Package R-3.4.1-1.52.amzn1.x86_64 already installed and latest version
  Package curl-7.61.1-7.91.amzn1.x86_64 already installed and latest version
  Package wget-1.18-4.29.amzn1.x86_64 already installed and latest version
  Nothing to do
```

- If you added custom actions to the `SlurmQueues` setting, check the `cloud-init.log` located at `/var/log/cloud-init.log` in a compute node. Use CloudWatch to view these logs.

You can view both of these logs in the Amazon CloudWatch console. For more information, see Integration with Amazon CloudWatch Logs.

Example of how to update a custom bootstrap script for IMDSv2

In the following example, we update a custom bootstrap action script that was used with IMDSv1 for use with IMDSv2. The IMDSv1 script retrieves Amazon EC2 instance AMI ID metadata.

```
#!/bin/bash
AMI_ID=$(curl http://169.254.169.254/latest/meta-data/ami-id)
echo $AMI_ID >> /home/ami_id.txt
```

The following shows the custom bootstrap action script modified to be compatible with IMDSv2.

```
#!/bin/bash
AMI_ID=$(TOKEN=`curl -X PUT "http://169.254.169.254/latest/api/token" -H "X-aws-ec2-
metadata-token-ttl-seconds: 21600"` \
```

```
         && curl -H "X-aws-ec2-metadata-token: $TOKEN" -v http://169.254.169.254/
latest/meta-data/ami-id)
echo $AMI_ID >> /home/ami_id.txt
```

For more information, see Retrieve instance metadata in the *Amazon EC2 User Guide for Linux Instances*.

Example of how to update a configuration for IMDSv1

The following is an example of a cluster configuration that supports IMDSv1 when using AWS ParallelCluster versions 3.7.0 and older.

```
Region: us-east-1
Imds:
  ImdsSupport: v1.0
Image:
  Os: alinux2
HeadNode:
  InstanceType: t2.micro
  Networking:
    SubnetId: subnet-abcdef01234567890
  Ssh:
    KeyName: key-name
  CustomActions:
    OnNodeConfigured:
      Script: Script-path
Scheduling:
  Scheduler: slurm
  SlurmQueues:
  - Name: queue1
    CustomActions:
      OnNodeConfigured:
        Script: Script-path
    ComputeResources:
    - Name: t2micro
      Instances:
      - InstanceType: t2.micro
      MinCount: 11
    Networking:
      SubnetIds:
      - subnet-abcdef01234567890
```

Working with Amazon S3

You can configure AWS ParallelCluster's access to Amazon S3 through the HeadNode / Iam / S3Access and Scheduling / SlurmQueues / - Name / Iam / S3Access parameters in the AWS ParallelCluster configuration.

Examples

The following example configures read-only access to all objects in *firstbucket/read_only/* and read/write access to all objects in *secondbucket/read_and_write/*.

```
...
HeadNode:
  ...
  Iam:
    S3Access:
      - BucketName: firstbucket
        KeyName: read_only/*
        EnableWriteAccess: false
      - BucketName: secondbucket
        KeyName: read_and_write/*
        EnableWriteAccess: true
...
```

The next example configures read-only access to all objects in folder *read_only/* in any bucket (*) in the account.

```
...
HeadNode:
  ...
  Iam:
    S3Access:
      - BucketName: *
        KeyName: read_only/*
        EnableWriteAccess: false
...
```

The final example configures read_only access to all buckets and objects in the account.

```
...
```

```
HeadNode:
  ...
  Iam:
    S3Access:
      - BucketName: *
...
```

Working with Spot Instances

AWS ParallelCluster uses Spot Instances if you have set `SlurmQueues` / `CapacityType` or `AwsBatchQueues` / `CapacityType` to SP0T in the cluster configuration file. Spot Instances are more cost effective than On-Demand Instances, but they might be interrupted. It might help to take advantage of *Spot Instance interruption notices*, which provide a two minute warning before Amazon EC2 must stop or terminate your Spot Instance. For more information, see Spot Instance interruptions in *Amazon EC2 User Guide*. To learn how `AwsBatchQueues` works with Spot Instances, see Compute Resources in the *AWS Batch User Guide*.

The AWS ParallelCluster configured scheduler assigns jobs to compute resources in queues with spot instances in the same way it assigns jobs to compute resources in queues with on-demand instances.

When using Spot Instances, an AWSServiceRoleForEC2Spot service-linked role must exist in your account. To create this role in your account using the AWS CLI, run the following command:

```
$ aws iam create-service-linked-role --aws-service-name spot.amazonaws.com
```

For more information, see Service-linked role for Spot Instance requests in the *Amazon EC2 User Guide*.

The following sections describe three scenarios in which Spot Instances can be interrupted when using `SlurmQueues`.

Scenario 1: Spot Instance with no running jobs is interrupted

When this interruption occurs, AWS ParallelCluster tries to replace the instance if the scheduler queue has pending jobs that require additional instances, or if the number of active instances is lower than the `SlurmQueues` / `ComputeResources` / `MinCount`. If AWS ParallelCluster can't provision new instances, then a request for new instances is periodically repeated.

Scenario 2: Spot Instance running single node jobs is interrupted

The job fails with a state code of `NODE_FAIL`, and the job is requeued (unless `--no-requeue` is specified when the job is submitted). If the node is a static node, it's replaced. If the node is a dynamic node, the node is terminated and reset. For more information about `sbatch`, including the `--no-requeue` parameter, see sbatch in the *Slurm documentation*.

Scenario 3: Spot Instance running multi-node jobs is interrupted

The job fails with a state code of `NODE_FAIL`, and the job is requeued (unless `--no-requeue` was specified when the job was submitted). If the node is a static node, it's replaced. If the node is a dynamic node, the node is terminated and reset. Other nodes that were running the terminated jobs might be allocated to other pending jobs, or scaled down after the configured `SlurmSettings` / `ScaledownIdletime` time has passed.

For more information about Spot Instances, see Spot Instances in the *Amazon EC2 User Guide*.

Schedulers supported by AWS ParallelCluster

AWS ParallelCluster supports Slurm and AWS Batch schedulers, which are set using the `Scheduler` setting. The following topics will describe each scheduler and how to use them.

Topics

Slurm Workload Manager (`slurm`)

Cluster capacity size and update

The capacity of the cluster is defined by the number of compute nodes the cluster can scale. Compute nodes are backed by Amazon EC2 instances defined within compute resources in the AWS ParallelCluster configuration (`Scheduling/SlurmQueues/ComputeResources`), and are organized into queues (`Scheduling/SlurmQueues`) that map 1:1 to Slurm partitions.

Within a compute resource it's possible to configure the minimum number of compute nodes (instances) that must always be kept running in the cluster (`MinCount`), and the maximum number of instances the compute resource can scale to (`MaxCount3`).

At cluster creation time, or upon a cluster update, AWS ParallelCluster launches as many Amazon EC2 instances as configured in `MinCount` for each compute resource (`Scheduling/SlurmQueues/ ComputeResources`) defined in the cluster. The instances launched to cover the minimal amount of nodes for a compute resources in the cluster are called ***static nodes***. Once started, static nodes are meant to be persistent in the cluster and they are not terminated by the system, unless a particular event or condition occurs. Such events include, for example, the failure of Slurm or Amazon EC2 health checks and the change of the Slurm node status to DRAIN or DOWN.

The Amazon EC2 instances, in the range of 1 to `'MaxCount - MinCount'` (`MaxCount` *minus* `MinCount`), launched on-demand to deal with the increased load of the cluster, are referred to as ***dynamic nodes***. Their nature is ephemeral, they are launched to serve pending jobs and are terminated once they stay idle for a period of time defined by `Scheduling/SlurmSettings/ScaledownIdletime` in the cluster configuration (default: 10 minutes).

Static nodes and dynamic node comply to the following naming schema:

- Static nodes `<Queue/Name>-st-<ComputeResource/Name>-<num>` where `<num> = 1..ComputeResource/MinCount`
- Dynamic nodes `<Queue/Name>-dy-<ComputeResource/Name>-<num>` where `<num> = 1..(ComputeResource/MaxCount - ComputeResource/MinCount)`

For example given the following AWS ParallelCluster configuration:

```
Scheduling:
    Scheduler: Slurm
    SlurmQueues:
        - Name: queue1
            ComputeResources:
                - Name: c5xlarge
                    Instances:
                        - InstanceType: c5.xlarge
                        MinCount: 100
                        MaxCount: 150
```

The following nodes will be defined in Slurm

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```

When a compute resource has `MinCount == MaxCount`, all the corresponding compute nodes will be static and all the instances will be launched at cluster creation/update time and kept up and running. For example:

```
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue1
      ComputeResources:
        - Name: c5xlarge
          Instances:
            - InstanceType: c5.xlarge
          MinCount: 100
          MaxCount: 100
```

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```

Cluster capacity update

The update of the cluster capacity includes adding or removing queues, compute resources or changing the `MinCount/MaxCount` of a compute resource. Starting from AWS ParallelCluster version 3.9.0, reducing the size of a queue requires the compute fleet to be stopped or QueueUpdateStrategy set to TERMINATE for before a cluster update to take place. It's not required to stop the compute fleet or to set QueueUpdateStrategy to TERMINATE when:

- Adding new queues to Scheduling/`SlurmQueues`
- Adding new compute resources `Scheduling/SlurmQueues/`ComputeResources to a queue

- Increasing the `MaxCount` of a compute resource
- Increasing MinCount of a compute resource and increasing MaxCount of the same compute resource of at least the same amount

Considerations and limitations

This section is meant to outline any important factors, constraints, or limitations that should be taken into account when resizing the cluster capacity.

- When removing a queue from `Scheduling/SlurmQueues` all the compute nodes with name `<Queue/Name>-*` , both static and dynamic, will be removed from the Slurm configuration and the corresponding Amazon EC2 instances will be terminated.
- When removing a compute resource `Scheduling/SlurmQueues/ComputeResources` from a queue, all the compute nodes with name `<Queue/Name>-*-<ComputeResource/Name>-*` , both static and dynamic, will be removed from the Slurm configuration and the corresponding Amazon EC2 instances will be terminated.

When changing the `MinCount` parameter of a compute resource we can distinguish two different scenarios, if `MaxCount` is kept equal to `MinCount` (static capacity only), and if `MaxCount` is greater than `MinCount` (mixed static and dynamic capacity).

Capacity changes with static nodes only

- If `MinCount == MaxCount` , when increasing `MinCount` (and `MaxCount`), the cluster will be configured by extending the number of static nodes to the new value of `MinCount` `<Queue/Name>-st-<ComputeResource/Name>-<new_MinCount>` and the system will keep trying to launch Amazon EC2 instances to fulfill the new required static capacity.
- If `MinCount == MaxCount` , when decreasing `MinCount` (and `MaxCount`) of the amount N, the cluster will be configured by removing the last N static nodes `<Queue/Name>-st-<ComputeResource/Name>-<old_MinCount - N>...<old_MinCount>]` and the system will terminate the corresponding Amazon EC2 instances.
  - Initial state `MinCount = MaxCount = 100`
  -
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
```

```
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```

- Update `-30` on `MinCount` and `MaxCount: MinCount = MaxCount = 70`
-

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     70   idle queue1-st-c5xlarge-[1-70]
```

Capacity changes with mixed nodes

If `MinCount < MaxCount`, when increasing `MinCount` by an amount N (assuming `MaxCount` will be kept unchanged), the cluster will be configured by extending the number static nodes to the new value of `MinCount` (`old_MinCount + N`): `<Queue/Name>-st-<ComputeResource/Name>-<old_MinCount + N>` and the system will keep trying to launch Amazon EC2 instances to fulfill the new required static capacity. Moreover, to honor the `MaxCount` capacity of the compute resource, the cluster configuration is updated by *removing the last N dynamic nodes*: `<Queue/Name>-dy-<ComputeResource/Name>-[<MaxCount - old_MinCount - N>...<MaxCount - old_MinCount>]` and the system will terminate the corresponding Amazon EC2 instances.

- Initial state: `MinCount = 100; MaxCount = 150`
-

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```

- Update `+30` to `MinCount : MinCount = 130 (MaxCount = 150)`
-

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     20  idle~ queue1-dy-c5xlarge-[1-20]
queue1*      up   infinite    130   idle queue1-st-c5xlarge-[1-130]
```

If `MinCount < MaxCount`, when increasing `MinCount` and `MaxCount` of the same amount N, the cluster will be configured by extending the number static nodes to the new value of `MinCount` (`old_MinCount + N`): `<Queue/Name>-st-<ComputeResource/Name>-<old_MinCount + N>` and the system will keep trying to launch Amazon EC2 instances to fulfill the new required static capacity. Moreover, no changes will be done on the number of dynamic nodes to honor the new

`MaxCount` value.

- Initial state: `MinCount = 100; MaxCount = 150`
-
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```
- Update +30 to `MinCount : MinCount = 130 (MaxCount = 180)`
-
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     20  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite    130   idle queue1-st-c5xlarge-[1-130]
```

If `MinCount < MaxCount`, when decreasing `MinCount` of the amount N (assuming `MaxCount` will be kept unchanged), the cluster will be configured by removing the last N static nodes static nodes `<Queue/Name>-st-<ComputeResource/Name>-[<old_MinCount - N>...<old_MinCount>`and the system will terminate the corresponding Amazon EC2 instances. Moreover, to honor the `MaxCount` capacity of the compute resource, the cluster configuration is updated by extending the number of the dynamic nodes to fill the gap `MaxCount - new_MinCount: <Queue/Name>-dy-<ComputeResource/Name>-[1..<MazCount - new_MinCount>]` In this case, since those are dynamic nodes, no new Amazon EC2 instances will be launched unless the scheduler has jobs in pending on the new nodes.

- Initial state: `MinCount = 100; MaxCount = 150`
-
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```
- Update -30 on `MinCount : MinCount = 70 (MaxCount = 120)`
-
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     80  idle~ queue1-dy-c5xlarge-[1-80]
queue1*      up   infinite     70   idle queue1-st-c5xlarge-[1-70]
```

If `MinCount < MaxCount`, when decreasing `MinCount` and `MaxCount` of the same amount N, the cluster will be configured by removing the last N static nodes `<Queue/Name>-st-<ComputeResource/Name>-<old_MinCount - N>...<oldMinCount>]` and the system will terminate the corresponding Amazon EC2 instances.

Moreover, no changes will be done on the number of dynamic nodes to honor the new `MaxCount` value.

- Initial state: `MinCount = 100; MaxCount = 150`
-
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```
- Update -30 on `MinCount : MinCount = 70 (MaxCount = 120)`
-
```
$ sinfo
```

```
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     80  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite     70   idle queue1-st-c5xlarge-[1-70]
```

If `MinCount < MaxCount`, when decreasing `MaxCount` of the amount N (assuming `MinCount` will be kept unchanged), the cluster will be configured by removing the last N dynamic nodes `<Queue/Name>-dy-<ComputeResource/Name>-<old_MaxCount - N...<oldMaxCount>]` and the system will terminate the corresponding Amazon EC2 instances in the case they were running.No impact is expected on the static nodes.

- Initial state: `MinCount = 100; MaxCount = 150`
-
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```
- Update -30 on `MaxCount : MinCount = 100 (MaxCount = 120)`
-
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     20  idle~ queue1-dy-c5xlarge-[1-20]
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```

Impacts on the Jobs

In all the cases where nodes are removed and Amazon EC2 instances terminated, a sbatch job running on the removed nodes will be re-queued, unless there are no other nodes satisfying the job requirements. In this last case, the job fails with status NODE_FAIL and disappears from the queue and it must be re-submitted manually.

If you are planning to perform a cluster resize update, you can prevent jobs to go running in the nodes that are going to be removed during the planned update. This is possible by setting the nodes to be removed in maintenance. Please be aware that setting a node in maintenance would not impact jobs that are eventually already running in the node.

Suppose that with the planned cluster resize update you are going to remove the node `qeueu-st-computeresource-[9-10]`. You can create a Slurm reservation with the following command

```
sudo -i scontrol create reservation ReservationName=maint_for_update user=root
 starttime=now duration=infinite flags=maint,ignore_jobs nodes=qeueu-st-
computeresource-[9-10]
```

This will create a Slurm reservation named `maint_for_update` on the nodes `qeueu-st-computeresource-[9-10]`. From the time when the reservation is created, no more jobs can go running into the nodes `qeueu-st-computeresource-[9-10]`. Please be aware that the reservation will not prevent jobs to be eventually allocated on the nodes `qeueu-st-computeresource-[9-10]`.

After the cluster resize update, if the Slurm reservation was set only on nodes that were removed during the resize update, the maintenance reservation will be automatically deleted. If instead you had created a Slurm reservation on nodes that are still present after the cluster resize update, we may want to remove the maintenance reservation on the nodes after the resize update is performed, by using the following command

```
sudo -i scontrol delete ReservationName=maint_for_update
```

For additional details on Slurm reservation, see the official SchedMD doc here.

Cluster update process on capacity changes

Upon a scheduler configuration change, the following steps are executed during the cluster update process:

- Stop AWS ParallelCluster `clustermgtd (supervisorctl stop clustermgtd)`
- Generate updated Slurm partitions configuration from AWS ParallelCluster configuration
- Restart `slurmctld` (done through Chef service recipe)
- Check `slurmctld` status `(systemctl is-active --quiet slurmctld.service)`

- Reload Slurm configuration (`scontrol reconfigure`)
- Start `clustermgtd` (`supervisorctl start clustermgtd`)

For information about Slurm, see https://slurm.schedmd.com. For downloads, see https://github.com/SchedMD/slurm/tags. For the source code, see https://github.com/SchedMD/slurm.

Supported cluster and SLURM versions

The following table lists the AWS ParallelCluster and Slurm versions that AWS supports.

AWS ParallelCluster version(s)	Supported Slurm version
3.13.0	24.05.07
3.12.0	23.11.10
3.11.0	23.11.10
3.9.2, 3.9.3, 3.10.0	23.11.7
3.9.0, 3.9.1	23.11.4
3.8.0	23.02.7
3.7.2	23.02.6
3.7.1	23.02.5
3.7.0	23.02.4
3.6.0, 3.6.1	23.02.2
3.5.0, 3.5.1	22.05.8
3.4.0, 3.4.1	22.05.7
3.3.0, 3.3.1	22.05.5
3.1.4, 3.1.5, 3.2.0, 3.2.1	21.08.8-2
3.1.2, 3.1.3	21.08.6

AWS ParallelCluster version(s)	Supported Slurm version
3.1.1	21.08.5
3.0.0	20.11.8

Topics

Configuration of multiple queues

With AWS ParallelCluster version 3, you can configure multiple queues by setting the `Scheduler` to `slurm` and specifying more than one queue for `SlurmQueues` in the configuration file. In this mode, different instance types coexist in the compute nodes that are specified in the `ComputeResources` section of the configuration file. `ComputeResources` with different instance types are scaled up or down as needed for the `SlurmQueues`.

Multiple *queues* within a single cluster are generally preferred over multiple clusters when the workloads share the same underlying infrastructure and resources (like shared storage, networking, or login nodes). If workloads have similar compute, storage, and networking needs, using multiple queues within a single cluster is more efficient because it allows for resource sharing and avoids unnecessary duplication. This approach simplifies management and reduces overhead, while still allowing for efficient job scheduling and resource allocation. On the other hand, multiple

clusters should be used when there are strong security, data, or operational isolation requirements between workloads. For example, if you need to manage and operate workloads independently, with different schedules, update cycles, or access policies, multiple clusters are more appropriate.

Cluster queue and compute resource quotas

Resource	Quota
Slurm queues	50 queues per cluster
Compute resources	50 compute resources per queue 50 compute resources per cluster

Node Counts

Each compute resource in ComputeResources for a queue must have a unique Name, InstanceType, MinCount, and MaxCount. MinCount and MaxCount have default values that define the range of instances for a compute resource in ComputeResources for a queue. You can also specify your own values for MinCount and MaxCount. Each compute resource in ComputeResources is composed of static nodes numbered from 1 to the value of MinCount and dynamic nodes numbered from the value of MinCount to the value of MaxCount.

Example Configuration

The following is an example of a Scheduling section for a cluster configuration file. In this configuration there are two queues named queue1 and queue2 and each of the queues has ComputeResources with a specified MaxCount.

```
Scheduling:
  Scheduler: slurm
  SlurmQueues:
  - Name: queue1
    ComputeResources:
    - InstanceType: c5.xlarge
      MaxCount: 5
      Name: c5xlarge
    - InstanceType: c4.xlarge
      MaxCount: 5
      Name: c4xlarge
  - Name: queue2
```

```
    ComputeResources:
    - InstanceType: c5.xlarge
      MaxCount: 5
      Name: c5xlarge
```

Hostnames

The instances that are launched into the compute fleet are dynamically assigned. Hostnames are generated for each node. By default AWS ParallelCluster will use the following format of the hostname :

`$HOSTNAME=$QUEUE-$STATDYN-$COMPUTE_RESOURCE-$NODENUM`

- `$QUEUE` is the name of the queue. For example, if the `SlurmQueues` section has an entry with the `Name` set to "`queue-name`" then "`$QUEUE`" is "`queue-name`".
- `$STATDYN` is `st` for static nodes or `dy` for dynamic nodes.
- `$COMPUTE_RESOURCE` is the `Name` of the `ComputeResources` compute resource corresponding to this node.
- `$NODENUM` is the number of the node. `$NODENUM` is between one (1) and the value of `MinCount` for static nodes and between one (1) and `MaxCount`-`MinCount` for dynamic nodes.

From the example configuration file above, a given node from `queue1` and compute resource `c5xlarge` has a hostname: `queue1-dy-c5xlarge-1`.

Both hostnames and fully-qualified domain names (FQDN) are created using Amazon Route 53 hosted zones. The FQDN is `$HOSTNAME.$CLUSTERNAME.pcluster`, where `$CLUSTERNAME` is the name of the cluster.

Note that the same format will be used for the Slurm node names as well.

Users can choose to use the default Amazon EC2 hostname of the instance powering the compute node instead of the default host name format used by AWS ParallelCluster. This can be done by setting the `UseEc2Hostnames` parameter to be true. However, Slurm node names will continue to use the default AWS ParallelCluster format.

Slurm guide for multiple queue mode

Here you can learn how AWS ParallelCluster and Slurm manage queue (partition) nodes and how you can monitor the queue and node states.

Overview

The scaling architecture is based on Slurm's Cloud Scheduling Guide and power saving plugin. For more information about the power saving plugin, see Slurm Power Saving Guide. In the architecture, resources that can potentially be made available for a cluster are typically predefined in the Slurm configuration as cloud nodes.

Cloud node lifecycle

Throughout their lifecycle, cloud nodes enter several if not all of the following states: `POWER_SAVING`, `POWER_UP` (`pow_up`), `ALLOCATED` (`alloc`), and `POWER_DOWN` (`pow_dn`). In some cases, a cloud node might enter the `OFFLINE` state. The following list details several aspects of these states in the cloud node lifecycle.

- **A node in a `POWER_SAVING` state** appears with a ~ suffix (for example `idle~`) in `sinfo`. In this state, no EC2 instances are backing the node. However, Slurm can still allocate jobs to the node.
- **A node transitioning to a `POWER_UP` state** appears with a # suffix (for example `idle#`) in `sinfo`. A node automatically transitions to a `POWER_UP` state, when Slurm allocates a job to a node in a `POWER_SAVING` state.

  Alternatively, you can transition the nodes to the `POWER_UP` state manually as an `su` root user with the command:

  ```
  $ scontrol update nodename=nodename state=power_up
  ```

  In this stage, the `ResumeProgram` is invoked, EC2 instances are launched and configured, and the node transitions to the `POWER_UP` state.
- **A node that is currently available for use** appears without a suffix (for example `idle`) in `sinfo`. After the node is set up and has joined the cluster, it becomes available to run jobs. In this stage, the node is properly configured and ready for use.

  As a general rule, we recommend that the number of Amazon EC2 instances be the same as the number of available nodes. In most cases, static nodes are available after the cluster is created.
- **A node that is transitioning to a `POWER_DOWN` state** appears with a `%` suffix (for example `idle%`) in `sinfo`. Dynamic nodes automatically enter the `POWER_DOWN` state after `ScaledownIdletime`. In contrast, static nodes in most cases aren't powered down. However, you can place the nodes in the `POWER_DOWN` state manually as an `su` root user with the command:

```
$ scontrol update nodename=nodename state=down reason="manual draining"
```

In this state, the instances associated with a node are terminated, and the node is set back to the POWER_SAVING state and available for use after ScaledownIdletime.

The ScaledownIdletime setting is saved to the Slurm configuration SuspendTimeout setting.

- **A node that is offline** appears with a * suffix (for example down*) in sinfo. A node goes offline if the Slurm controller can't contact the node or if the static nodes are disabled and the backing instances are terminated.

Consider the node states shown in the following sinfo example.

```
$ sinfo
  PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
  efa          up   infinite      4  idle~ efa-dy-efacompute1-[1-4]
  efa          up   infinite      1   idle efa-st-efacompute1-1
  gpu          up   infinite      1  idle% gpu-dy-gpucompute1-1
  gpu          up   infinite      9  idle~ gpu-dy-gpucompute1-[2-10]
  ondemand     up   infinite      2   mix# ondemand-dy-ondemandcompute1-[1-2]
  ondemand     up   infinite     18  idle~ ondemand-dy-ondemandcompute1-
[3-10],ondemand-dy-ondemandcompute2-[1-10]
  spot*        up   infinite     13  idle~ spot-dy-spotcompute1-[1-10],spot-dy-
spotcompute2-[1-3]
  spot*        up   infinite      2   idle spot-st-spotcompute2-[1-2]
```

The spot-st-spotcompute2-[1-2] and efa-st-efacompute1-1 nodes already have backing instances set up and are available for use. The ondemand-dy-ondemandcompute1-[1-2] nodes are in the POWER_UP state and should be available within a few minutes. The gpu-dy-gpucompute1-1 node is in the POWER_DOWN state, and it transitions into POWER_SAVING state after ScaledownIdletime (defaults to 10 minutes).

All of the other nodes are in POWER_SAVING state with no EC2 instances backing them.

Working with an available node

An available node is backed by an Amazon EC2 instance. By default, the node name can be used to directly SSH into the instance (for example ssh efa-st-efacompute1-1). The private IP address of the instance can be retrieved using the command:

```
$ scontrol show nodes nodename
```

Check for IP address in the returned `NodeAddr` field.

For nodes that aren't available, the `NodeAddr` field shouldn't point to a running Amazon EC2 instance. Rather, it should be the same as the node name.

Job states and submission

Jobs submitted in most cases are immediately allocated to nodes in the system, or placed in pending if all the nodes are allocated.

If nodes allocated for a job include any nodes in a `POWER_SAVING` state, the job starts out with a `CF`, or `CONFIGURING` state. At this time, the job waits for the nodes in the `POWER_SAVING` state to transition to the `POWER_UP` state and become available.

After all nodes allocated for a job are available, the job enters the `RUNNING` (`R`) state.

By default, all jobs are submitted to the default queue (known as a partition in Slurm). This is signified by a `*` suffix after the queue name. You can select a queue using the `-p` job submission option.

All nodes are configured with the following features, which can be used in job submission commands:

- An instance type (for example `c5.xlarge`)
- A node type (This is either `dynamic` or `static`.)

You can see the features for a particular node by using the command:

```
$ scontrol show nodes nodename
```

In the return, check the `AvailableFeatures` list.

Consider the initial state of the cluster, which you can view by running the `sinfo` command.

```
$ sinfo
  PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
  efa          up   infinite      4  idle~ efa-dy-efacompute1-[1-4]
```

```
  efa          up   infinite      1   idle efa-st-efacompute1-1
  gpu          up   infinite     10  idle~ gpu-dy-gpucompute1-[1-10]
  ondemand     up   infinite     20  idle~ ondemand-dy-ondemandcompute1-
[1-10],ondemand-dy-ondemandcompute2-[1-10]
  spot*        up   infinite     13  idle~ spot-dy-spotcompute1-[1-10],spot-dy-
spotcompute2-[1-3]
  spot*        up   infinite      2   idle spot-st-spotcompute2-[1-2]
```

Note that `spot` is the default queue. It is indicated by the * suffix.

Submit a job to one static node in the default queue (`spot`).

```
$ sbatch --wrap "sleep 300" -N 1 -C static
```

Submit a job to one dynamic node in the EFA queue.

```
$ sbatch --wrap "sleep 300" -p efa -C dynamic
```

Submit a job to eight (8) `c5.2xlarge` nodes and two (2) `t2.xlarge` nodes in the ondemand queue.

```
$ sbatch --wrap "sleep 300" -p ondemand -N 10 -C "[c5.2xlarge*8&t2.xlarge*2]"
```

Submit a job to one GPU node in the gpu queue.

```
$ sbatch --wrap "sleep 300" -p gpu -G 1
```

Consider the state of the jobs using the `squeue` command.

```
$ squeue
 JOBID PARTITION    NAME   USER   ST       TIME  NODES NODELIST(REASON)
  12   ondemand     wrap   ubuntu CF       0:36     10 ondemand-dy-ondemandcompute1-
[1-8],ondemand-dy-ondemandcompute2-[1-2]
  13        gpu     wrap   ubuntu CF       0:05      1 gpu-dy-gpucompute1-1
   7       spot     wrap   ubuntu  R       2:48      1 spot-st-spotcompute2-1
   8        efa     wrap   ubuntu  R       0:39      1 efa-dy-efacompute1-1
```

Jobs 7 and 8 (in the `spot` and `efa` queues) are already running (R). Jobs 12 and 13 are still configuring (CF), probably waiting for instances to become available.

```
# Nodes states corresponds to state of running jobs
$ sinfo
 PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
 efa          up   infinite      3  idle~ efa-dy-efacompute1-[2-4]
 efa          up   infinite      1    mix efa-dy-efacompute1-1
 efa          up   infinite      1   idle efa-st-efacompute1-1
 gpu          up   infinite      1   mix~ gpu-dy-gpucompute1-1
 gpu          up   infinite      9  idle~ gpu-dy-gpucompute1-[2-10]
 ondemand     up   infinite     10   mix# ondemand-dy-ondemandcompute1-[1-8],ondemand-
dy-ondemandcompute2-[1-2]
 ondemand     up   infinite     10  idle~ ondemand-dy-ondemandcompute1-[9-10],ondemand-
dy-ondemandcompute2-[3-10]
 spot*        up   infinite     13  idle~ spot-dy-spotcompute1-[1-10],spot-dy-
spotcompute2-[1-3]
 spot*        up   infinite      1    mix spot-st-spotcompute2-1
 spot*        up   infinite      1   idle spot-st-spotcompute2-2
```

Node state and features

In most cases, node states are fully managed by AWS ParallelCluster according to the specific processes in the cloud node lifecycle described earlier in this topic.

However, AWS ParallelCluster also replaces or terminates unhealthy nodes in `DOWN` and `DRAINED` states and nodes that have unhealthy backing instances. For more information, see `clustermgtd`.

Partition states

AWS ParallelCluster supports the following partition states. A Slurm partition is a queue in AWS ParallelCluster.

- UP: Indicates that the partition is in an active state. This is the default state of a partition. In this state, all nodes in the partition are active and available for use.
- `INACTIVE`: Indicates that the partition is in the inactive state. In this state, all instances backing nodes of an inactive partition are terminated. New instances aren't launched for nodes in an inactive partition.

pcluster update-compute-fleet

- **Stopping the compute fleet** - When the following command is executed, all partitions transition to the `INACTIVE` state, and AWS ParallelCluster processes keep the partitions in the `INACTIVE` state.

```
$ pcluster update-compute-fleet --cluster-name testSlurm \
   --region eu-west-1 --status STOP_REQUESTED
```

- **Starting the compute fleet** - When the following command is executed, all partitions initially transition to the UP state. However, AWS ParallelCluster processes don't keep the partition in an UP state. You need to change partition states manually. All static nodes become available after a few minutes. Note that setting a partition to UP doesn't power up any dynamic capacity.

```
$ pcluster update-compute-fleet --cluster-name testSlurm \
   --region eu-west-1 --status START_REQUESTED
```

When `update-compute-fleet` is run, you can check the state of the cluster by running the `pcluster describe-compute-fleet` command and checking the `Status`. The following lists possible states:

- `STOP_REQUESTED`: The stop compute fleet request is sent to the cluster.
- `STOPPING`: The `pcluster` process is currently stopping the compute fleet.
- `STOPPED`: The `pcluster` process finished the stopping process, all partitions are in `INACTIVE` state, and all compute instances are terminated.
- `START_REQUESTED`: The start compute fleet request is sent to the cluster.
- `STARTING`: The `pcluster` process is currently starting the cluster.
- `RUNNING`: The `pcluster` process finished the starting process, all partitions are in the UP state, and static nodes are available after a few minutes.
- `PROTECTED`: This status indicates that some partitions have consistent bootstrap failures. Affected partitions are inactive. Please investigate the issue and then run `update-compute-fleet` to re-enable the fleet.

Manual control of queues

In some cases, you might want to have some manual control over the nodes or queue (known as a partition in Slurm) in a cluster. You can manage nodes in a cluster through the following common procedures using the `scontrol` command.

- **Power up dynamic nodes in `POWER_SAVING` state**

Run the command as an `su` root user:

```
$ scontrol update nodename=nodename state=power_up
```

You can also submit a placeholder `sleep 1` job requesting a certain number of nodes and then rely on Slurm to power up the required number of nodes.

- **Power down dynamic nodes before `ScaledownIdletime`**

  We recommend that you set dynamic nodes to `DOWN` as an `su` root user with the command:

  ```
  $ scontrol update nodename=nodename state=down reason="manually draining"
  ```

  AWS ParallelCluster automatically terminates and resets the downed dynamic nodes.

  In general, we don't recommend that you set nodes to `POWER_DOWN` directly using the `scontrol update nodename=`*`nodename`* `state=power_down` command. This is because AWS ParallelCluster automatically handles the power down process.

- **Disable a queue (partition) or stop all static nodes in specific partition**

  Set a specific queue to `INACTIVE` as an `su` root user with the command:

  ```
  $ scontrol update partition=queuename state=inactive
  ```

  Doing this terminates all instances backing nodes in the partition.

- **Enable a queue (partition)**

  Set a specific queue to `UP` an `su` root user with the command:

  ```
  $ scontrol update partition=queuename state=up
  ```

Scaling behavior and adjustments

Here is an example of the normal scaling workflow:

- The scheduler receives a job that requires two nodes.

- The scheduler transitions two nodes to a POWER_UP state, and calls ResumeProgram with the node names (for example queue1-dy-spotcompute1-[1-2]).
- ResumeProgram launches two Amazon EC2 instances and assigns the private IP addresses and hostnames of queue1-dy-spotcompute1-[1-2], waiting for ResumeTimeout (the default period is 30 minutes before resetting the nodes.
- Instances are configured and join the cluster. A job starts running on instances.
- The job completes and stops running.
- After the configured SuspendTime has elapsed (which is set to ScaledownIdletime), the scheduler sets the instances to the POWER_SAVING state. The scheduler then sets queue1-dy-spotcompute1-[1-2] to the POWER_DOWN state and calls SuspendProgram with the node names.
- SuspendProgram is called for two nodes. Nodes remain in the POWER_DOWN state, for example, by remaining idle% for a SuspendTimeout (the default period is 120 seconds (2 minutes)). After clustermgtd detects that nodes are powering down, it terminates the backing instances. Then, it transitions queue1-dy-spotcompute1-[1-2] to the idle state and resets the private IP address and hostname so it is ready to power up for future jobs.

If things go wrong and an instance for a particular node can't be launched for some reason, then the following happens:

- The scheduler receives a job that requires two nodes.
- The scheduler transitions two cloud bursting nodes to the POWER_UP state and calls ResumeProgram with the nodenames, (for example queue1-dy-spotcompute1-[1-2]).
- ResumeProgram launches only one (1) Amazon EC2 instance and configures queue1-dy-spotcompute1-1, with one (1) instance, queue1-dy-spotcompute1-2, failing to launch.
- queue1-dy-spotcompute1-1 isn't impacted and comes online after reaching the POWER_UP state.
- queue1-dy-spotcompute1-2 transitions to the POWER_DOWN state, and the job is requeued automatically because Slurm detects a node failure.
- queue1-dy-spotcompute1-2 becomes available after SuspendTimeout (the default is 120 seconds (2 minutes)). In the meantime, the job is requeued and can start running on another node.
- The above process repeats until the job can run on an available node without a failure occurring.

There are two timing parameters that can be adjusted if needed:

- **`ResumeTimeout` (the default is 30 minutes)**: `ResumeTimeout` controls the time Slurm waits before transitioning the node to the down state.
  - It might be useful to extend `ResumeTimeout` if your pre/post installation process takes nearly that long.
  - `ResumeTimeout` is also the maximum time that AWS ParallelCluster waits before replacing or resetting a node if there is an issue. Compute nodes self-terminate if any error occurs during launch or setup. AWS ParallelCluster processes replace a node upon detection of a terminated instance.
- **`SuspendTimeout` (the default is 120 seconds (2 minutes))**: `SuspendTimeout` controls how quickly nodes get placed back into the system and are ready for use again.
  - A shorter `SuspendTimeout` means that nodes are reset more quickly, and Slurm can try to launch instances more frequently.
  - A longer `SuspendTimeout` means that failed nodes are reset more slowly. In the meantime, Slurm tries to use other nodes. If `SuspendTimeout` is more than a few minutes, Slurm tries to cycle through all nodes in the system. A longer `SuspendTimeout` might be beneficial for large-scale systems (over1,000 nodes) to reduce stress on Slurm when it tries to frequently re-queue failing jobs.
  - Note that `SuspendTimeout` doesn't refer to the time AWS ParallelCluster waits to terminate a backing instance for a node. Backing instances for `POWER_DOWN` nodes are immediately terminated. The terminate process usually is finished in a few minutes. However, during this time, the node remains in the `POWER_DOWN` state and isn't available for the scheduler's use.

Logs for the architecture

The following list contains the key logs. The log stream name used with Amazon CloudWatch Logs has the format *`{hostname}`*`.`*`{instance_id}`*`.`*`{logIdentifier}`*, where *`logIdentifier`* follows the log names.

- `ResumeProgram`: `/var/log/parallelcluster/slurm_resume.log` (`slurm_resume`)
- `SuspendProgram`: `/var/log/parallelcluster/slurm_suspend.log` (`slurm_suspend`)
- `clustermgtd`: `/var/log/parallelcluster/clustermgtd.log` (`clustermgtd`)
- `computemgtd`: `/var/log/parallelcluster/computemgtd.log` (`computemgtd`)
- `slurmctld`: `/var/log/slurmctld.log` (`slurmctld`)

- `slurmd`: `/var/log/slurmd.log` (`slurmd`)

Common issues and how to debug:

Nodes that failed to launch, power up, or join the cluster

- Dynamic nodes:
  - Check the `ResumeProgram` log to see if `ResumeProgram` was called with the node. If not, check the `slurmctld` log to determine if Slurm tried to call `ResumeProgram` with the node. Note that incorrect permissions on `ResumeProgram` might cause it to fail silently.
  - If `ResumeProgram` is called, check to see if an instance was launched for the node. If the instance didn't launch, there should be clear error message as to why the instance failed to launch.
  - If an instance was launched, there may have been some problem during the bootstrap process. Find the corresponding private IP address and instance ID from the `ResumeProgram` log and look at corresponding bootstrap logs for the specific instance in CloudWatch Logs.
- Static nodes:
  - Check the `clustermgtd` log to see if instances were launched for the node. If instances didn't launch, there should be clear errors on why the instances failed to launch.
  - If an instance was launched, there is some problem with the bootstrap process. Find the corresponding private IP and instance ID from the `clustermgtd` log and look at corresponding bootstrap logs for the specific instance in CloudWatch Logs.

Nodes replaced or terminated unexpectedly, and node failures

- Nodes replaced/terminated unexpectedly:
  - In most cases, `clustermgtd` handles all node maintenance actions. To check if `clustermgtd` replaced or terminated a node, check the `clustermgtd` log.
  - If `clustermgtd` replaced or terminated the node, there should be a message indicating the reason for the action. If the reason is scheduler related (for example, the node was `DOWN`), check in the `slurmctld` log for more details. If the reason is Amazon EC2 related, use tools such as Amazon CloudWatch or the Amazon EC2 console, CLI, or SDKs, to check status or logs for that instance. For example, you can check if the instance had scheduled events or failed Amazon EC2 health status checks.

  - If `clustermgtd` didn't terminate the node, check if `computemgtd` terminated the node or if EC2 terminated the instance to reclaim a Spot Instance.
- Node failures:
  - In most cases, jobs are automatically requeued if a node failed. Look in the `slurmctld` log to see why a job or a node failed and assess the situation from there.

Failure when replacing or terminating instances, failure when powering down nodes

- In general, `clustermgtd` handles all expected instance termination actions. Look in the `clustermgtd` log to see why it failed to replace or terminate a node.
- For dynamic nodes failing `ScaledownIdletime`, look in the `SuspendProgram` log to see if `slurmctld` processes made calls with the specific node as argument. Note `SuspendProgram` doesn't actually perform any specific action. Rather, it only logs when it's called. All instance termination and `NodeAddr` resets are completed by `clustermgtd`. Slurm transitions nodes to `IDLE` after `SuspendTimeout`.

Other issues:

- AWS ParallelCluster doesn't make job allocation or scaling decisions. It only tries to launch, terminate, and maintain resources according to Slurm's instructions.

  For issues regarding job allocations, node allocation and scaling decision, look at the `slurmctld` log for errors.

Slurm cluster protected mode

When a cluster runs with protected mode enabled, AWS ParallelCluster monitors and tracks compute node bootstrap failures as the compute nodes are being launched. It does this to detect whether these failures are occurring continuously.

If the following is detected in a queue (partition), the cluster enters protected status:

1. Consecutive compute node bootstrap failures occur continuously with no successful compute node launches.
2. The failure count reaches a predefined threshold.

After the cluster enters protected status, AWS ParallelCluster disables queues with failures at or above the predefined threshold.

Slurm cluster protected mode was added in AWS ParallelCluster version 3.0.0.

You can use protected mode to reduce the time and resources spent on compute node bootstrap failure cycling.

Protected mode parameter

`protected_failure_count`

`protected_failure_count` specifies the number of consecutive failures in a queue (partition) that activate cluster protected status.

The default `protected_failure_count` is 10 and protected mode is enabled.

If `protected_failure_count` is greater than zero, protected mode is enabled.

If `protected_failure_count` is less than or equal to zero, protected mode is disabled.

You can change the `protected_failure_count` value by adding the parameter in the `clustermgtd` config file that's located at `/etc/parallelcluster/slurm_plugin/parallelcluster_clustermgtd.conf` in the `HeadNode`.

You can update this parameter anytime and you don't need to stop the compute fleet to do so. If a launch succeeds in a queue before the failure count reaches `protected_failure_count`, the failure count is reset to zero.

Check cluster status in protected status

When a cluster is in protected status, you can check the compute fleet status and node states.

Compute fleet status

The status of the compute fleet is `PROTECTED` in a cluster running in protected status.

```
$ pcluster describe-compute-fleet --cluster-name <cluster-name> --region <region-id>
{
   "status": "PROTECTED",
   "lastStatusUpdatedTime": "2022-04-22T00:31:24.000Z"
```

```
}
```

Node status

To learn which queues (partitions) have bootstrap failures that have activated protected status, log in to the cluster and run the `sinfo` command. Partitions with bootstrap failures at or above `protected_failure_count` are in the `INACTIVE` state. Partitions without bootstrap failures at or above `protected_failure_count` are in the UP state and work as expected.

PROTECTED status doesn't impact running jobs. If jobs are running on a partition with bootstrap failures at or above `protected_failure_count`, the partition is set to `INACTIVE` after the running jobs complete.

Consider the node states shown in the following example.

```
$ sinfo
PARTITION AVAIL TIMELIMIT NODES STATE NODELIST
queue1* inact infinite 10 down% queue1-dy-c5xlarge-[1-10]
queue1* inact infinite 3490 idle~ queue1-dy-c5xlarge-[11-3500]
queue2 up infinite 10 idle~ queue2-dy-c5xlarge-[1-10]
```

Partition `queue1` is `INACTIVE` because 10 consecutive compute node bootstrap failures were detected.

Instances behind nodes `queue1-dy-c5xlarge-[1-10]` launched but failed to join the cluster because of an unhealthy status.

The cluster is in protected status.

Partition `queue2` isn't impacted by the bootstrap failures in `queue1`. It's in the UP state and can still run jobs.

How to deactivate protected status

After the bootstrap error has been resolved, you can run the following command to take the cluster out of protected status.

```
$ pcluster update-compute-fleet --cluster-name <cluster-name> \
  --region <region-id> \
  --status START_REQUESTED
```

Bootstrap failures that activate protected status

Bootstrap errors that activate protected status are subdivided into the following three types. To identify the type and issue, you can check if AWS ParallelCluster generated logs. If logs were generated, you can check them for error details. For more information, see Retrieving and preserving logs.

1. **Bootstrap error that causes an instance to self-terminate**.

   An instance fails early in the bootstrap process, such as an instance that self-terminates because of errors in the `SlurmQueues` \ `CustomActions` \ `OnNodeStart` | `OnNodeConfigured` script.

   For dynamic nodes, look for errors similar to the following:

   ```
   Node bootstrap error: Node ... is in power up state without valid backing instance
   ```

   For static nodes, look in the `clustermgtd` log (`/var/log/parallelcluster/clustermgtd`) for errors similar to the following:

   ```
   Node bootstrap error: Node ... is in power up state without valid backing instance
   ```

2. **Nodes `resume_timeout` or `node_replacement_timeout` expires**.

   An instance can't join the cluster within the `resume_timeout` (for dynamic nodes) or `node_replacement_timeout` (for static nodes). It doesn't self-terminate before the timeout. For example, networking isn't set up correctly for the cluster and the node is set to the `DOWN` state by Slurm after the timeout expires.

   For dynamic nodes, look for errors similar to the following:

   ```
   Node bootstrap error: Resume timeout expires for node
   ```

   For static nodes, look in the `clustermgtd` log (`/var/log/parallelcluster/clustermgtd`) for errors similar to the following:

   ```
   Node bootstrap error: Replacement timeout expires for node ... in replacement.
   ```

3. **Nodes fail health check**.

An instance behind the node fails an Amazon EC2 health check or scheduled event health check, and the nodes are treated as bootstrap failure nodes. In this case, the instance terminates for a reason outside the control of AWS ParallelCluster.

Look in the `clustermgtd` log (`/var/log/parallelcluster/clustermgtd`) for errors similar to the following:

```
Node bootstrap error: Node %s failed during bootstrap when performing health check.
```

4. **Compute nodes fail Slurm registration**.

   The registration of the `slurmd` daemon with the Slurm control daemon (`slurmctld`) fails and causes the compute node state to change to the `INVALID_REG` state. Incorrectly configured Slurm compute nodes can cause this error, such as computed nodes configured with `CustomSlurmSettings` compute node specification errors.

   Look in the `slurmctld` log file (`/var/log/slurmctld.log`) on the head node, or look in the `slurmd` log file (`/var/log/slurmd.log`) of the failed compute node for errors similar to the following:

   ```
   Setting node %s to INVAL with reason: ...
   ```

How to debug protected mode

If your cluster is in protected status, and if AWS ParallelCluster generated `clustermgtd` logs from the `HeadNode` and the `cloud-init-output` logs from problematic compute nodes, then you can check the logs for error details. For more information about how to retrieve logs, see Retrieving and preserving logs.

`clustermgtd` log(`/var/log/parallelcluster/clustermgtd`) on the head node

Log messages show which partitions have bootstrap failures and the corresponding bootstrap failure count.

```
[slurm_plugin.clustermgtd:_handle_protected_mode_process] - INFO - Partitions
bootstrap failure count: {'queue1': 2}, cluster will be set into protected mode if
 protected failure count reach threshold.
```

In the `clustermgtd` log, search for `Found the following bootstrap failure nodes` to find which node failed to bootstrap.

```
[slurm_plugin.clustermgtd:_handle_protected_mode_process] - WARNING -
Found the following bootstrap failure nodes: (x2)  ['queue1-st-
c5large-1(192.168.110.155)',  'broken-st-c5large-2(192.168.65.215)']
```

In the `clustermgtd` log, search for `Node bootstrap error` to find the reason for the failure.

```
[slurm_plugin.clustermgtd:_is_node_bootstrap_failure] - WARNING - Node bootstrap
 error:
Node broken-st-c5large-2(192.168.65.215) is currently in  replacement and no backing
 instance
```

`cloud-init-output` log(`/var/log/cloud-init-output.log`) on the compute nodes

After obtaining the bootstrap failure node private IP address in the `clustermgtd` log, you can find the corresponding compute node log by either logging into the compute node or by following the guidance in Retrieving and preserving logs to retrieve logs. In most cases, the `/var/log/cloud-init-output` log from the problematic node shows the step that caused the compute node bootstrap failure.

Slurm cluster fast insufficient capacity fail-over

Starting with AWS ParallelCluster version 3.2.0, clusters run with the fast insufficient capacity fail-over mode enabled by default. This minimizes the time spent retrying to queue a job when Amazon EC2 insufficient capacity errors are detected. This is particularly effective when you configure your queue with multiple compute resources that use different instance types.

Amazon EC2 detected insufficient capacity failures:

- `InsufficientInstanceCapacity`
- `InsufficientHostCapacity`
- `InsufficientReservedInstanceCapacity`
- `MaxSpotInstanceCountExceeded`
- `SpotMaxPriceTooLow`: Activated if the Spot request price is lower than the minimum required Spot request fulfillment price.
- `Unsupported`: Activated with the use of an instance type that isn't supported in a specific AWS Region.

In fast insufficient capacity failure-over mode, if an insufficient capacity error is detected when a job is assigned to a `SlurmQueues` / `compute resource`, AWS ParallelCluster does the following:

1. It sets the compute resource to a disabled (`DOWN`) state for a predefined period of time.
2. It uses `POWER_DOWN_FORCE` to cancel the compute resource failing node jobs and to suspend the failing node. It sets the failing node to the `IDLE` and `POWER_DOWN (!)` state, and then to `POWERING_DOWN (%)`.
3. It requeues the job to another compute resource.

The static and powered up nodes of the disabled compute resource aren't impacted. Jobs can complete on these nodes.

This cycle repeats until the job is successfully assigned to a compute resource node or nodes. For information about node states, see the Slurm guide for multiple queue mode.

If no compute resources are found to run the job, the job is set to the `PENDING` state until the predefined period of time elapses. In this case, you can modify the predefined period of time as described in the following section.

Insufficient capacity timeout parameter

`insufficient_capacity_timeout`

`insufficient_capacity_timeout` specifies the period of time (in seconds) that the compute resource is kept in the disabled (down) state when an insufficient capacity error is detected.

By default, `insufficient_capacity_timeout` is enabled.

The default `insufficient_capacity_timeout` is 600 seconds (10 minutes).

If the `insufficient_capacity_timeout` value is less than or equal to zero, fast insufficient capacity failure-over mode is disabled.

You can change the `insufficient_capacity_timeout` value by adding the parameter in the `clustermgtd` config file located at `/etc/parallelcluster/slurm_plugin/parallelcluster_clustermgtd.conf` in the `HeadNode`.

The parameter can be updated at any time without stopping the compute fleet.

For example:

- `insufficient_capacity_timeout=600`:

  If an insufficient capacity error is detected, the compute resource is set to a disabled (`DOWN`). After 10 minutes, its failed node is set to the `idle~` (`POWER_SAVING`) state.
- `insufficient_capacity_timeout=60`:

  If an insufficient capacity error is detected, the compute resource is in a disabled (`DOWN`). After 1 minute, its failed node is set to the `idle~` state.
- `insufficient_capacity_timeout=0`:

  Fast insufficient capacity failure-over mode is disabled. The compute resource isn't disabled.

> **Note**
>
> There might be a delay of up to one minute between the time when nodes fail with insufficient capacity errors and the time when the cluster management daemon detects the node failures. This is because the cluster management daemon checks for node insufficient capacity failures and sets the compute resources to the down state at one minute intervals.

Fast insufficient capacity fail-over mode status

When a cluster is in fast insufficient capacity fail-over mode, you can check its status and node states.

Node states

When a job is submitted to a compute resource dynamic node and an insufficient capacity error is detected, the node is placed in the down# state with reason.

```
(Code:InsufficientInstanceCapacity)Failure when resuming nodes.
```

Then powered off nodes (nodes in `idle~` state) are set to down~ with reason.

```
(Code:InsufficientInstanceCapacity)Temporarily disabling node due to insufficient
 capacity.
```

The job is requeued to other compute resources in the queue.

The compute resource static nodes and nodes that are UP aren't impacted by fast insufficient capacity fail-over mode.

Consider the node states shown in the following example.

```
$ sinfo
PARTITION AVAIL TIMELIMIT NODES STATE NODELIST
queue1*   up    infinite    30  idle~ queue1-dy-c-1-[1-15],queue1-dy-c-2-[1-15]
queue2    up    infinite    30  idle~ queue2-dy-c-1-[1-15],queue2-dy-c-2-[1-15]
```

We submit a job to queue1 that requires one node.

```
$ sinfo
PARTITION AVAIL TIMELIMIT NODES STATE NODELIST
queue1*   up   infinite  1   down# queue1-dy-c-1-1
queue1*   up   infinite  15  idle~ queue1-dy-c-2-[1-15]
queue1*   up   infinite  14  down~ queue1-dy-c-1-[2-15]
queue2    up   infinite  30  idle~ queue2-dy-c-1-[1-15],queue2-dy-c-2-[1-15]
```

Node `queue1-dy-c-1-1` is launched to run the job. However, the instance failed to launch due to an insufficient capacity error. Node `queue1-dy-c-1-1` is set to down. The powered off dynamic node within the compute resource (`queue2-dy-c-1`) is set to down.

You can check the node reason with `scontrol show nodes`.

```
$ scontrol show nodes queue1-dy-c-1-1
NodeName=broken-dy-c-2-1 Arch=x86_64 CoresPerSocket=1
CPUAlloc=0 CPUTot=96 CPULoad=0.00
...
ExtSensorsJoules=n/s ExtSensorsWatts=0 ExtSensorsTemp=n/s
Reason=(Code:InsufficientInstanceCapacity)Failure when resuming nodes
 [root@2022-03-10T22:17:50]

$ scontrol show nodes queue1-dy-c-1-2
NodeName=broken-dy-c-2-1 Arch=x86_64 CoresPerSocket=1
CPUAlloc=0 CPUTot=96 CPULoad=0.00
...
ExtSensorsJoules=n/s ExtSensorsWatts=0 ExtSensorsTemp=n/s
Reason=(Code:InsufficientInstanceCapacity)Temporarily disabling node due to
 insufficient capacity [root@2022-03-10T22:17:50]
```

The job is queued to another instance type within the queue compute resources.

After the `insufficient_capacity_timeout` elapses, nodes in the compute resource are reset to the `idle~` state.

```
$ sinfo
PARTITION AVAIL TIMELIMIT NODES STATE NODELIST
queue1*   up    infinite    30  idle~ queue1-dy-c-1-[1-15],queue1-dy-c-2-[1-15]
queue2    up    infinite    30  idle~ queue2-dy-c-1-[1-15],queue2-dy-c-2-[1-15]
```

After the `insufficient_capacity_timeout` elapses and nodes in the compute resource are reset to the `idle~` state, the Slurm scheduler gives the nodes lower priority. The scheduler keeps selecting nodes from other queue compute resources with higher weights unless one of the following occurs:

- A job's submission requirements match the recovered compute resource.
- No other compute resources are available because they are at capacity.
- `slurmctld` is restarted.
- The AWS ParallelCluster compute fleet is stopped and started to power down and power up all nodes.

Related logs

Logs related to insufficient capacity errors and fast insufficient capacity fail-over mode can be found in Slurm's `resume` log and `clustermgtd` log in the head node.

Slurm `resume` (`/var/log/parallelcluster/slurm_resume.log`)

Error messages when a node fails to launch because of insufficient capacity.

```
[slurm_plugin.instance_manager:_launch_ec2_instances] - ERROR - Failed RunInstances
 request: dcd0c252-90d4-44a7-9c79-ef740f7ecd87
[slurm_plugin.instance_manager:add_instances_for_nodes] - ERROR - Encountered
 exception when launching instances for nodes (x1) ['queue1-dy-c-1-1']: An error
 occurred
(InsufficientInstanceCapacity) when calling the RunInstances operation (reached max
 retries: 1): We currently do not have sufficient p4d.24xlarge capacity in the
Availability Zone you requested (us-west-2b). Our system will be working on
 provisioning additional capacity. You can currently get p4d.24xlarge capacity by
 not
specifying an Availability Zone in your request or choosing us-west-2a, us-west-2c.
```

Slurm `clustermgtd` (`/var/log/parallelcluster/clustermgtd`)

Compute resource c-1 in queue1 is disabled because of insufficient capacity.

```
[slurm_plugin.clustermgtd:_reset_timeout_expired_compute_resources] - INFO - The
 following compute resources are in down state
due to insufficient capacity: {'queue1': {'c-1':
 ComputeResourceFailureEvent(timestamp=datetime.datetime(2022, 4, 14, 23, 0, 4,
 769380, tzinfo=datetime.timezone.utc),
error_code='InsufficientInstanceCapacity')}}, compute resources are reset after
 insufficient capacity timeout (600 seconds) expired
```

After the insufficient capacity timeout expires, the compute resource is reset, nodes within the compute resources are set to `idle~`.

```
[root:_reset_insufficient_capacity_timeout_expired_nodes] - INFO - Reset the
 following compute resources because insufficient capacity
timeout expired: {'queue1': ['c-1']}
```

Slurm memory-based scheduling

Starting with version 3.2.0, AWS ParallelCluster supports Slurm memory-based scheduling with the `SlurmSettings` / `EnableMemoryBasedScheduling` cluster configuration parameter.

Note

Starting with AWS ParallelCluster version 3.7.0, `EnableMemoryBasedScheduling` can be enabled if you configure multiple instance types in Instances.
For AWS ParallelCluster versions 3.2.0 to 3.6.*x*, `EnableMemoryBasedScheduling` can't be enabled if you configure multiple instance types in Instances.

Warning

When you specify multiple instances types in a Slurm queue compute resource with `EnableMemoryBasedScheduling` enabled, the `RealMemory` value is the minimum amount of memory made available to all instance types. This might lead to significant

amounts of unused memory if you specify instance types with very different memory capacities.

With `EnableMemoryBasedScheduling: true`, the Slurm scheduler tracks the amount of memory that each job requires on each node. Then, the Slurm scheduler uses this information to schedule multiple jobs on the same compute node. The total amount of memory that jobs require on a node can't be larger than the available node memory. The scheduler prevents a job from using more memory than what was requested when the job was submitted.

With `EnableMemoryBasedScheduling: false`, jobs might compete for memory on a shared node and cause job failures and `out-of-memory` events.

Warning

Slurm uses a power of 2 notation for its labels, such as MB or GB. Read these labels as MiB and GiB, respectively.

Slurm configuration and memory-based scheduling

With `EnableMemoryBasedScheduling: true`, Slurm sets the following Slurm configuration parameters:

- `SelectTypeParameters=CR_CPU_Memory` in the `slurm.conf`. This option configures node memory to be a consumable resource in Slurm.
- `ConstrainRAMSpace=yes` in the Slurm `cgroup.conf`. With this option, a job's access to memory is limited to the amount of memory that the job requested when submitted.

Note

Several other Slurm configuration parameters can impact the behavior of the Slurm scheduler and resource manager when these two options are set. For more information, see the Slurm Documentation.

Slurm scheduler and memory-based scheduling

`EnableMemoryBasedScheduling: false` (default)

By default, `EnableMemoryBasedScheduling` is set to false. When false, Slurm doesn't include memory as a resource in its scheduling algorithm and doesn't track the memory that jobs use. Users can specify the `--mem MEM_PER_NODE` option to set the minimum amount of memory per node that a job requires. This forces the scheduler to choose nodes with a `RealMemory` value of at least `MEM_PER_NODE` when scheduling the job.

For example, suppose that a user submits two jobs with `--mem=5GB`. If requested resources such as CPUs or GPUs are available, the jobs can run at the same time on a node with 8 GiB of memory. The two jobs aren't scheduled on compute nodes with less than 5 GiB of `RealMemory`.

Warning

When memory-based scheduling is disabled, Slurm doesn't track the amount of memory that jobs use. Jobs that run on the same node might compete for memory resources and cause the other job to fail.
When memory-based scheduling is disabled, we recommend that users don't specify the `--mem-per-cpu` or `--mem-per-gpu` options. These options might cause behavior that differs from what's described in the Slurm Documentation.

`EnableMemoryBasedScheduling: true`

When `EnableMemoryBasedScheduling` is set to true, Slurm tracks the memory usage of each job and prevents jobs from using more memory than requested with the `--mem` submission options.

Using the previous example, a user submits two jobs with `--mem=5GB`. The jobs can't run at the same time on a node with 8 GiB of memory. This is because the total amount of memory that's required is greater than the memory that's available on the node.

With memory-based scheduling enabled, `--mem-per-cpu` and `--mem-per-gpu` behave consistently with what's described in the Slurm documentation. For example, a job is submitted with `--ntasks-per-node=2 -c 1 --mem-per-cpu=2GB`. In this case, Slurm assigns the job a total of 4 GiB for each node.

Warning

When memory-based scheduling is enabled, we recommend that users include a `--mem` specification when submitting a job. With the default Slurm configuration that's included with AWS ParallelCluster, if no memory option is included (`--mem`, `--mem-per-cpu`, or `--mem-per-gpu`), Slurm assigns entire memory of the allocated nodes to the job, even if it requests only a portion of the other resources, such as CPUs or GPUs. This effectively prevents node sharing until the job is finished because no memory is available to other jobs. This happens because Slurm sets the memory per node for the job to `DefMemPerNode` when no memory specifications are provided at job submission time. The default value for this parameter is 0 and specifies unlimited access to a node's memory. If multiple types of compute resources with different amounts of memory are available in the same queue, a job submitted without memory options might be assigned different amounts of memory on different nodes. This depends on which nodes the scheduler makes available to the job. Users can define a custom value for options, such as `DefMemPerNode` or `DefMemPerCPU`, at the cluster or partition level in the Slurm configuration files to prevent this behavior.

Slurm RealMemory and AWS ParallelCluster SchedulableMemory

With the Slurm configuration that's shipped with AWS ParallelCluster, Slurm interprets RealMemory to be the amount of memory per node that's available to jobs. Starting with version 3.2.0, by default, AWS ParallelCluster sets `RealMemory` to 95 percent of the memory listed in Amazon EC2 Instance Types and returned by the Amazon EC2 API DescribeInstanceTypes.

When memory-based scheduling is disabled, the Slurm scheduler uses `RealMemory` to filter nodes when users submit a job with `--mem` specified.

When memory-based scheduling is enabled, the Slurm scheduler interprets `RealMemory` to be the maximum amount of memory that's available to jobs that are running on the compute node.

The default setting might not be optimal for all instance types:

- This setting might be higher than the amount of memory that nodes can actually access. This can happen when compute nodes are small instance types.

- This setting might be lower than the amount of memory that nodes can actually access. This can happen when compute nodes are large instance types and can lead to a significant amount of unused memory.

You can use `SlurmQueues` / `ComputeResources` / `SchedulableMemory` to fine-tune the value of `RealMemory` configured by AWS ParallelCluster for compute nodes. To override the default, define a custom value for `SchedulableMemory` specifically for your cluster configuration.

To check a compute node's actual available memory, run the `/opt/slurm/sbin/slurmd -C` command on the node. This command returns the hardware configuration of the node, including the `RealMemory` value. For more information, see `slurmd -C`.

Make sure that the compute node's operating system processes have sufficient memory. To do this, limit the memory available to jobs by setting the `SchedulableMemory` value to lower than the `RealMemory` value that the `slurmd -C` command returned.

Multiple instance type allocation with Slurm

Starting with AWS ParallelCluster version 3.3.0, you can configure your cluster to allocate from a compute resource's set of defined instance types. Allocation can be based on Amazon EC2 fleet low cost or optimal capacity strategies.

This set of defined instance types must either all have the same number of vCPUs or, if multithreading is disabled, the same number of cores. Moreover, this set of instance types must have the same number of accelerators of the same manufacturers. If `Efa` / `Enabled` is set to `true`, the instances must have EFA supported. For more information and requirements, see `Scheduling` / `SlurmQueues` / `AllocationStrategy` and `ComputeResources` / `Instances`.

You can set `AllocationStrategy` to `lowest-price` or `capacity-optimized` depending on your `CapacityType` configuration.

In `Instances`, you can configure a set of instance types.

> **Note**
>
> Starting with AWS ParallelCluster version 3.7.0, `EnableMemoryBasedScheduling` can be enabled if you configure multiple instance types in `Instances`.
> For AWS ParallelCluster versions 3.2.0 to 3.6.*x*, `EnableMemoryBasedScheduling` can't be enabled if you configure multiple instance types in `Instances`.

The following examples show how you can query instance types for vCPUs, EFA support, and architecture.

Query InstanceTypes with 96 vCPUs and x86_64 architecture.

```
$ aws ec2 describe-instance-types --region region-id \
  --filters "Name=vcpu-info.default-vcpus,Values=96" "Name=processor-info.supported-
architecture,Values=x86_64" \
  --query "sort_by(InstanceTypes[*].
{InstanceType:InstanceType,MemoryMiB:MemoryInfo.SizeInMiB,CurrentGeneration:CurrentGeneration,V
 &InstanceType)" \
  --output table
```

Query InstanceTypes with 64 cores, EFA support, and arm64 architecture.

```
$ aws ec2 describe-instance-types --region region-id \
  --filters "Name=vcpu-info.default-cores,Values=64" "Name=processor-
info.supported-architecture,Values=arm64" "Name=network-info.efa-
supported,Values=true" --query "sort_by(InstanceTypes[*].
{InstanceType:InstanceType,MemoryMiB:MemoryInfo.SizeInMiB,CurrentGeneration:CurrentGeneration,V
 &InstanceType)" \
  --output table
```

The next example cluster configuration snippet shows how you can use these InstanceType and AllocationStrategy properties.

```
...
 Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue-1
      CapacityType: ONDEMAND
      AllocationStrategy: lowest-price
      ...
      ComputeResources:
        - Name: computeresource1
          Instances:
            - InstanceType: r6g.2xlarge
            - InstanceType: m6g.2xlarge
            - InstanceType: c6g.2xlarge
          MinCount: 0
          MaxCount: 500
```

```
        - Name: computeresource2
          Instances:
            - InstanceType: m6g.12xlarge
            - InstanceType: x2gd.12xlarge
          MinCount: 0
          MaxCount: 500
...
```

Cluster scaling for dynamic nodes

ParallelCluster supports Slurm's methods to dynamically scale clusters by using Slurm's power saver plugin. For more information, see the Cloud Scheduling Guide and the Slurm Power Saving Guide in the Slurm documentation. The following topics describe the Slurm strategies for each version.

Topics

- Slurm dynamic node allocation strategies in version 3.8.0
- Slurm dynamic node allocation strategies in version 3.7.x
- Slurm dynamic node allocation strategies in version 3.6.x and previous

Slurm dynamic node allocation strategies in version 3.8.0

Starting with ParallelCluster version 3.8.0, ParallelCluster uses **Job-level resume** or **job-level scaling** as the default dynamic node allocation strategy to scale the cluster: ParallelCluster scales up the cluster based on the requirements of each job, the number of nodes allocated to the job, and which nodes need to be resumed. ParallelCluster gets this information from the SLURM_RESUME_FILE environment variable.

The scaling for dynamic nodes is a two steps process, which involves the launch of the EC2 instances and the assignment of the launched Amazon EC2 instances to the Slurm nodes. Each of these two steps can be done using an **all-or-nothing** or **best-effort** logic.

For launch of the Amazon EC2 instances:

- **all-or-nothing** calls the launch Amazon EC2 API with minimum target equals to the total target capacity
- **best-effort** calls the launch Amazon EC2 API with minimum target equals to 1 and the total target capacity equals to the requested capacity

For assignment of the Amazon EC2 instances to Slurm nodes:

- **all-or-nothing** assigns Amazon EC2 instances to Slurm nodes only if it's possible to assign an Amazon EC2 instance to every requested node
- **best-effort** assigns Amazon EC2 instances to Slurm nodes even if all the requested nodes are not covered by Amazon EC2 instance capacity

  The possible combinations of the above strategies translates into the ParallelCluster launch strategies.

Example

<caption>The available **ParallelCluster launch strategies** that can be set into the ScalingStrategy cluster configuration to be used with **job-level scaling** are:</caption>

all-or-nothing scaling:

This strategy involves AWS ParallelCluster initiating an Amazon EC2 launch instance API call for each job, that requires all instances necessary for the requested compute nodes to be successfully launched. This ensures that the cluster scales only when the required capacity per job is available, avoiding idle instances left at the end of the scaling process.

The strategy uses an **all-or-nothing** logic for the launch of the Amazon EC2 instances for each job plus and **all-or-nothing** logic for the assignment of the Amazon EC2 instances to Slurm nodes.

The strategy groups launch requests into batches, one for each compute resource requested and up to 500 nodes each. For requests spanning multiple compute resources or exceeding 500 nodes, ParallelCluster sequentially processes multiple batches.

The failure of any single resource's batch results in the termination of all associated unused capacity, ensuring that no idle instances will be left at the end of the scaling process.

Limitations

- The time taken for scaling is directly proportional to the number of jobs submitted per execution of the Slurm resume program.
- The scaling operation is limited by the RunInstances resource account limit, set at 1000 instances by default. This limitation is in accordance with AWS's EC2 API throttling policies, for more details refer to Amazon EC2 API throttling documentation

- When you submit a job in a compute resource with a single instance type, in a queue that spans multiple Availability Zones, the **all-or-nothing** EC2 launch API call only succeeds if all of the capacity can be provided in a single Availability Zone.
- When you submit a job in a compute resource with multiple instance types, in a queue with a single Availability Zone, the **all-or-nothing** Amazon EC2 launch API call only succeeds if all of the capacity can be provided by a single instance type.
- When you submit a job in a compute resource with multiple instance types, in a queue spanning multiple Availability Zones, the **all-or-nothing** Amazon EC2 launch API call isn't supported and ParallelCluster performs **best-effort** scaling instead.

greedy-all-or-nothing scaling:

This variant of the all-or-nothing strategy still ensures that the cluster scales only when the required capacity per job is available, avoiding idle instances at the end of the scaling process, but it involves ParallelCluster initiating an Amazon EC2 launch instance API call that aims for a minimum target capacity of 1, attempting to maximize the number of nodes launched up to the requested capacity. The strategy uses a best-effort logic for the launch of the EC2 instances for all the jobs plus the **all-or-nothing** logic for the assignment of the Amazon EC2 instances to Slurm nodes for each job.

The strategy groups launch requests into batches, one for each compute resource requested and up to 500 nodes each. For requests spanning multiple compute resources or exceeding 500 nodes, ParellelCluster sequentially processes multiple batches.

It ensure that no idle instances will be left at the end of the scaling process, by maximizing the throughput at the cost of temporary over-scaling during the scaling process.

Limitations

- Temporary over-scaling is possible, leading to additional costs for instances that transition to a running state before scaling completion.
- The same instance limit as in the all-or-nothing strategy applies, subject to AWS's RunInstances resource account limit.

best-effort scaling:

This strategy calls Amazon EC2 launch instance API call by targeting a minimum capacity of 1 and aiming to achieve the total requested capacity at the cost of leaving idle instances after the scaling

process execution if not all the requested capacity is available. The strategy uses a best-effort logic for the launch of the Amazon EC2 instances for all the jobs plus the **best-effort** logic for the assignment of the Amazon EC2 instances to Slurm nodes for each job.

The strategy groups launch requests into batches, one for each compute resource requested and up to 500 nodes each. For requests spanning multiple compute resources or exceeding 500 nodes, ParallelCluster sequentially processes multiple batches.

This strategy allows for scaling far beyond the default 1000 instances limit over multiple scaling process executions, at the cost of having idle instances across the different scaling processes.

Limitations

- Possible idle running instances at the end of the scaling process, for the case when it's not possible to allocate all the nodes requested by the jobs.

The following is an example that shows how the scaling of dynamic nodes behave using the different **ParallelCluster launch strategies**. Suppose you have submitted two jobs requesting 20 nodes each, for a total of 40 nodes of the same type, but there are only 30 Amazon EC2 instances available to cover the requested capacity on EC2.

all-or-nothing scaling:

- For the first job, an **all-or-nothing** Amazon EC2 launch instance API is called, requesting 20 instances. A successful call has results in the launch of 20 instances
- **all-or-nothing** assignment of the 20 launched instances to Slurm nodes for the first job is successful
- Another **all-or-nothing** Amazon EC2 launch instance API is called, requesting 20 instances for the second job. The call is not successful, since there is only capacity for another 10 instances. No instances are launched at this time

greedy-all-or-nothing scaling:

- A **best-effort** Amazon EC2 launch instance API is called, requesting 40 instances, which is the total capacity requested by all the jobs. This results in the launch of 30 instances
- An **all-or-nothing** assignment of 20 of the launched instances to Slurm nodes for the first job is successful

- Another **all-or-nothing** assignment of the remaining launched instances to Slurm nodes for the second job is tried, but since there are only 10 available instances out of the total 20 requested by the job, the assignment is not successful
- The 10 unassigned launched instances are terminated

best-effort scaling:

- A **best-effort** Amazon EC2 launch instance API is called, requesting 40 instances, which is the total capacity requested by all the jobs. This results in the launch of 30 instances.
- A **best-effort** assignment of 20 of the launched instances to Slurm nodes for the first job is successful.
- Another **best-effort** assignment of the remaining 10 launched instances to Slurm nodes for the second job is successful, even if the total requested capacity was 20. But since the job was requesting the 20 nodes, and it was possible to assign Amazon EC2 instances to only 10 of them, the job cannot start and the instances are left running idle, until enough capacity is found to start the missing 10 instances at a later call of the scaling process, or the scheduler schedules the job on other, already running, compute nodes.

Slurm dynamic node allocation strategies in version 3.7.x

ParallelCluster uses 2 types of dynamic node allocation strategies to scale the cluster:

- **Allocation based on available requested node information:**
  - **All-nodes resume** or **node-list** scaling:

    ParallelCluster scales up the cluster based only on Slurm's requested node list names when Slurm's `ResumeProgram` runs. It allocates compute resources to nodes only by node name. The list of node names can span multiple jobs.
  - **Job-level resume** or **job-level** scaling:

    ParallelCluster scales up the cluster based on the requirements of each job, the current number of nodes that are allocated to the job, and which nodes need to be resumed. ParallelCluster gets this information from the `SLURM_RESUME_FILE` environment variable.
- **Allocation with an Amazon EC2 launch strategy:**
  - **Best-effort** scaling:

ParallelCluster scales up the cluster by using an Amazon EC2 launch instance API call with the minimum target capacity equal to 1, to launch some, but not necessarily all of instances needed to support the requested nodes.

- **All-or-nothing** scaling:

  ParallelCluster scales up the cluster by using an Amazon EC2 launch instance API call that only succeeds if all of the instances needed to support the requested nodes are launched. In this case, it calls the Amazon EC2 launch instance API with the minimum target capacity equal to the total requested capacity.

By default, ParallelCluster uses **node-list** scaling with a **best-effort** Amazon EC2 launch strategy to launch some, but not necessarily all of instances needed to support the requested nodes. It tries to provision as much capacity as possible to serve the submitted workload.

Starting with ParallelCluster version 3.7.0, ParallelCluster uses **job-level** scaling with an **all-or-nothing** EC2 launch strategy for jobs submitted in **exclusive mode**. When you submit a job in exclusive mode, the job has exclusive access to its allocated nodes. For more information, see EXCLUSIVE in the Slurm documentation.

To submit a job in exclusive mode:

- Pass the exclusive flag when submitting a Slurm job to the cluster. For example, `sbatch ... --exclusive`.

  OR

- Submit a job to a cluster queue that has been configured with `JobExclusiveAllocation` set to `true`.

When submitting a job in exclusive mode:

- ParallelCluster currently batches launch requests to include up to 500 nodes. If a job requests more than 500 nodes, ParallelCluster makes an **all-or-nothing** launch request for each set of 500 nodes and an additional launch request for the remainder of nodes.
- If node allocation is in a single compute resource, ParallelCluster makes an **all-or-nothing** launch request for each set of 500 nodes and an additional launch request for the remainder of nodes. If a launch request fails, ParallelCluster terminates the unused capacity created by all of the launch requests.

- If node allocation spans multiple compute resources, ParallelCluster needs to make an **all-or-nothing** launch request for each compute resource. These requests are also batched. If a launch request fails for one of the compute resources, ParallelCluster terminates the unused capacity created by all of the compute resource launch requests.

job-level scaling with **all-or-nothing** launch strategy known limitations:

- When you submit a job in a compute resource with a single instance type, in a queue that spans multiple Availability Zones, the **all-or-nothing** EC2 launch API call only succeeds if all of the capacity can be provided in a single Availability Zone.
- When you submit a job in a compute resource with multiple instance types, in a queue with a single Availability Zone, the **all-or-nothing** Amazon EC2 launch API call only succeeds if all of the capacity can be provided by a single instance type.
- When you submit a job in a compute resource with multiple instance types, in a queue spanning multiple Availability Zones, the **all-or-nothing** Amazon EC2 launch API call isn't supported and ParallelCluster performs **best-effort** scaling instead.

Slurm dynamic node allocation strategies in version 3.6.x and previous

AWS ParallelCluster uses only one type of dynamic node allocation strategy to scale the cluster:

- Allocation based on available requested node information:
  - **All-nodes resume** or **node-list** scaling: ParallelCluster scales up the cluster based only on Slurm's requested node list names when Slurm's`ResumeProgram` runs. It allocates compute resources to nodes only by node name. The list of node names can span multiple jobs.
- Allocation with an Amazon EC2 launch strategy:
  - **Best-effort** scaling: ParallelCluster scales up the cluster by using an Amazon EC2 launch instance API call with the minimum target capacity equal to 1, to launch some, but not necessarily all of instances needed to support the requested nodes.

ParallelCluster uses**node-list** scaling with a **best-effort** Amazon EC2 launch strategy to launch some, but not necessarily all of instances needed to support the requested nodes. It tries to provision as much capacity as possible to serve the submitted workload.

Limitations

- Possible idle running instances at the end of the scaling process, for the case when it's not possible to allocate all the nodes requested by the jobs.

Slurm accounting with AWS ParallelCluster

Starting with version 3.3.0, AWS ParallelCluster supports Slurm accounting with the cluster configuration parameter SlurmSettings / Database.

Starting with version 3.10.0, AWS ParallelCluster supports Slurm accounting with an external Slurmdbd with the cluster configuration parameter SlurmSettings / ExternalSlurmdbd. Using an external Slurmdbd is recommended if multiple clusters share the same database.

With Slurm accounting, you can integrate an external accounting database to do the following:

- Manage cluster users or groups of users and other entities. With this capability, you can use Slurm's more advanced features, such as resource limit enforcement, fair-share, and QOSs.
- Collect and save job data, such as the user that ran the job, the job's duration, and the resources it uses. You can view the saved data with the `sacct` utility.

> **Note**
>
> AWS ParallelCluster supports Slurm accounting for Slurm supported MySQL database servers.

Working with Slurm accounting using external Slurmdbd in AWS ParallelCluster v3.10.0 and later

Before you configure Slurm accounting, you must have an existing external Slurmdbd database server, which connects to an existing external database server.

To configure this, define the following:

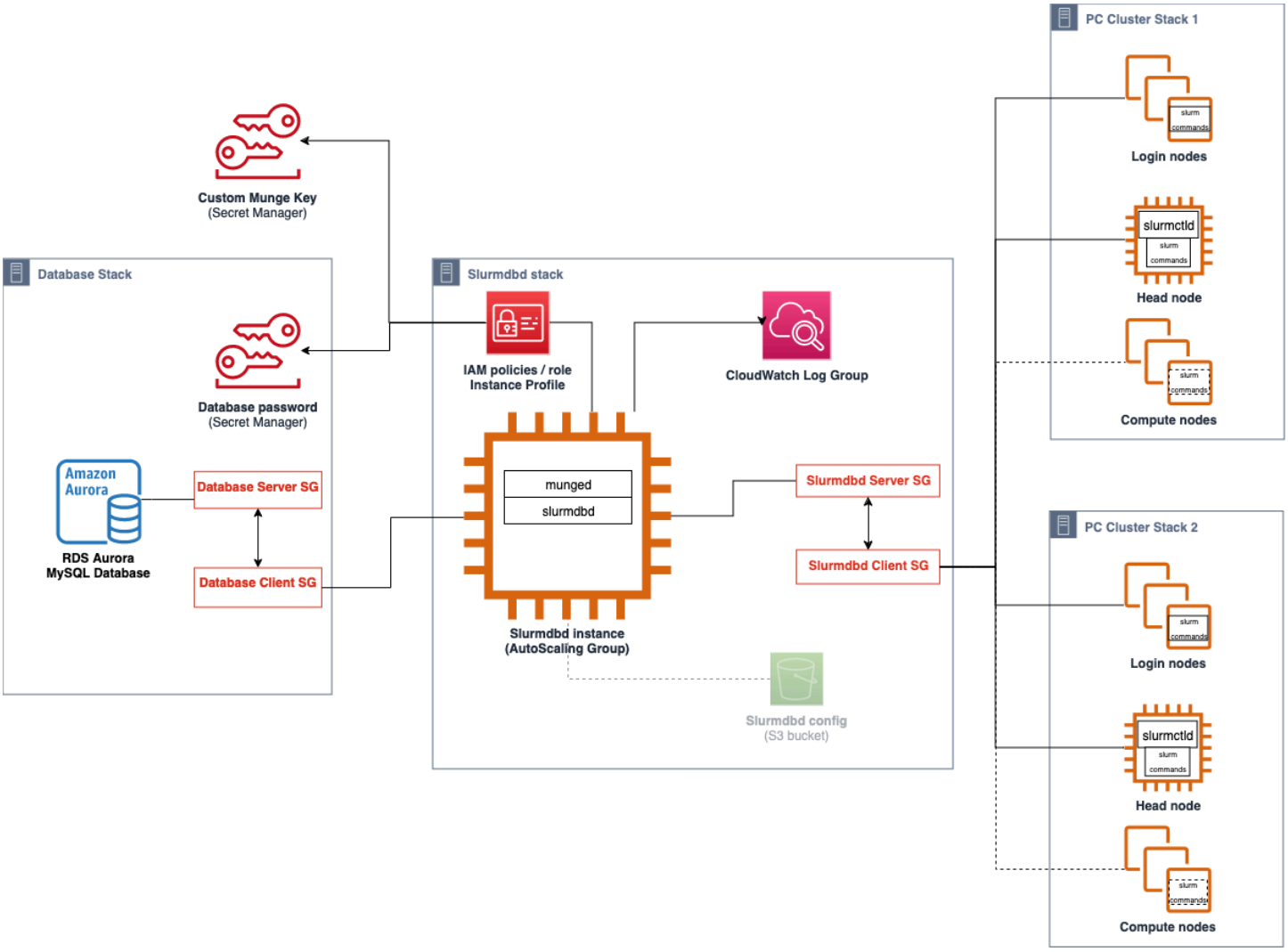
- The address of the external Slurmdbd server in ExternalSlurmdbd / Host. The server must exist and be reachable from the head node.
- The munge key to communicate with the external Slurmdbd server in MungeKeySecretArn.

To step through a tutorial, see Creating a cluster with an external Slurmdbd accounting.

Note

You are responsible to manage the Slurm database accounting entities.

The architecture of the AWS ParallelCluster external SlurmDB support feature enables multiple clusters sharing the same SlurmDB and the same database.



Warning

Traffic between AWS ParallelCluster and the external SlurmDB is not encrypted. It is recommended to run the cluster and the external SlurmDB in a trusted network.

Working with Slurm accounting using head node Slurmdbd in AWS ParallelCluster v3.3.0 and later

Before you configure Slurm accounting, you must have an existing external database server and database that uses `mysql` protocol.

To configure Slurm accounting with AWS ParallelCluster, you must define the following:

- The URI for the external database server in Database / Uri. The server must exist and be reachable from the head node.
- Credentials to access the external database that are defined in Database / PasswordSecretArn and Database / UserName. AWS ParallelCluster uses this information to configure accounting at the Slurm level and the `slurmdbd` service on the head node. `slurmdbd` is the daemon that manages communication between the cluster and the database server.

To step through a tutorial, see Creating a cluster with Slurm accounting.

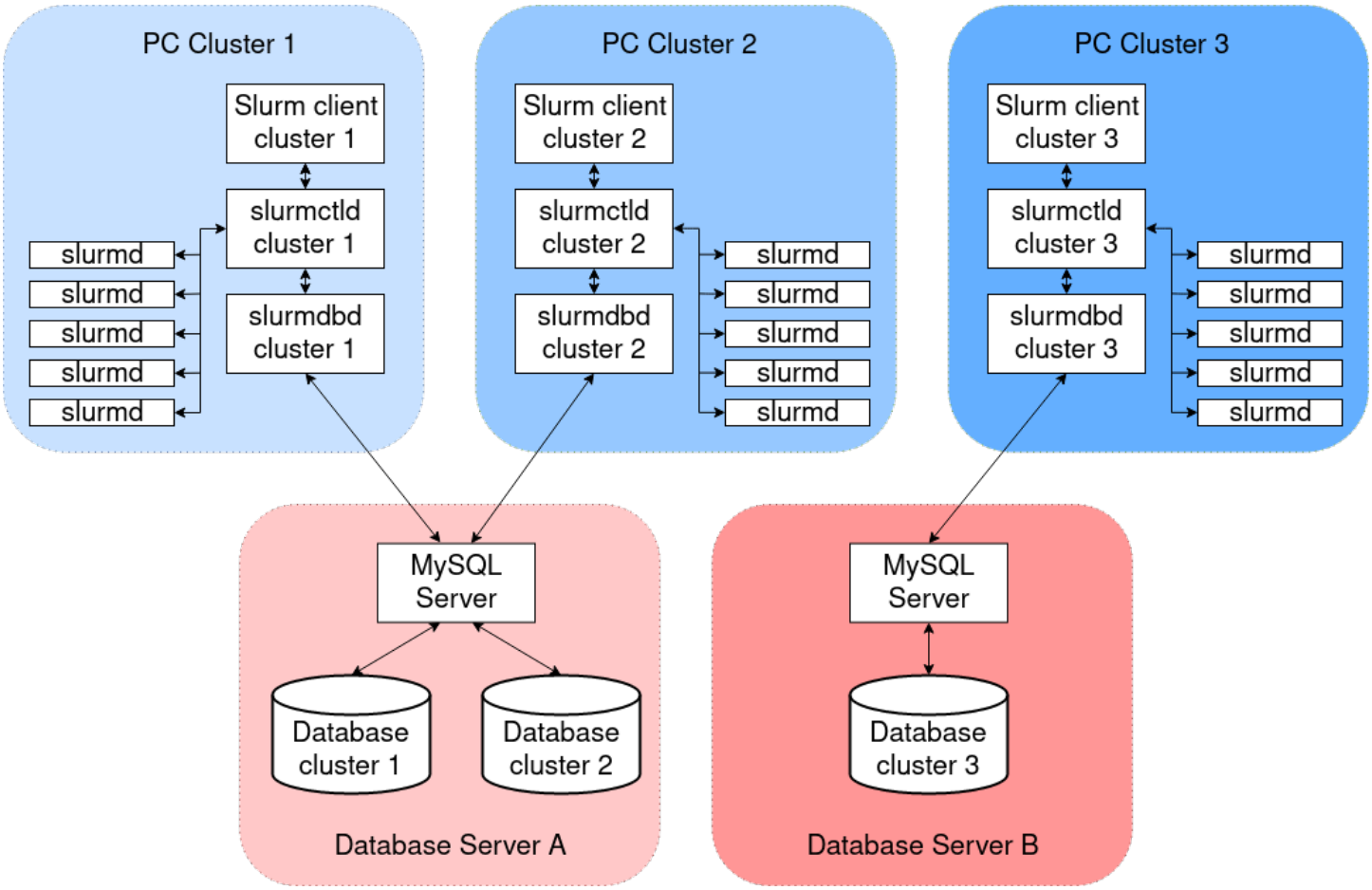
> **Note**
>
> AWS ParallelCluster performs a basic bootstrap of the Slurm accounting database by setting the default cluster user as database admin in the Slurm database. AWS ParallelCluster doesn't add any other user to the accounting database. The customer is responsible for managing the accounting entities in the Slurm database.

AWS ParallelCluster configures `slurmdbd` to ensure that a cluster has its own Slurm database on the database server. The same database server can be used across multiple clusters, but each cluster has its own separate database. AWS ParallelCluster uses the cluster name to define the name for the database in the `slurmdbd` configuration file `StorageLoc` parameter. Consider the following situation. A database that's present on the database server includes a cluster name that doesn't map to an active cluster name. In this case, you can create a new cluster with that cluster name to map to that database. Slurm reuses the database for the new cluster.

> **Warning**
>
> - We don't recommend setting up more than one cluster to use the same database at once. Doing so can lead to performance issues or even database deadlock situations.

- If Slurm accounting is enabled on the head node of a cluster, we recommend using an instance type with a powerful CPU, more memory, and higher network bandwidth. Slurm accounting can add strain on the head node of the cluster.

In the current architecture of the AWS ParallelCluster Slurm accounting feature, each cluster has its own instance of the `slurmdbd` daemon as shown in the following diagram example configurations.



If you're adding custom Slurm multi-cluster or federation functionalities to your cluster environment, all clusters must reference the same `slurmdbd` instance. For this alternative, we recommend that you enable AWS ParallelCluster Slurm accounting on one cluster and manually configure the other clusters to connect to the `slurmdbd` that are hosted on the first cluster.

If you're using AWS ParallelCluster versions prior to version 3.3.0, refer to the alternative method to implement Slurm accounting that's described in this HPC Blog Post.

Slurm accounting considerations

Database and cluster on different VPCs

To enable Slurm accounting, a database server is needed to serve as a backend for the read and write operations that the `slurmdbd` daemon performs. Before the cluster is created or updated to enable Slurm accounting, the head node must be able to reach the database server.

If you need to deploy the database server on a VPC other than the one that the cluster uses, consider the following:

- To enable communication between the `slurmdbd` on the cluster side and the database server, you must set up connectivity between the two VPCs. For more information, see VPC Peering in the *Amazon Virtual Private Cloud User Guide*.
- You must create the security group that you want to attach to the head node on the VPC of the cluster. After the two VPCs have been peered, cross-linking between the database side and the cluster side security groups is available. For more information, see Security Group Rules in the *Amazon Virtual Private Cloud User Guide*.

Configuring TLS encryption between `slurmdbd` and the database server

With the default Slurm accounting configuration that AWS ParallelCluster provides, `slurmdbd` establishes a TLS encrypted connection to the database server, if the server supports TLS encryption. AWS database services such as Amazon RDS and Amazon Aurora support TLS encryption by default.

You can require secure connections on the server side by setting the `require_secure_transport` parameter on the database server. This is configured in the provided CloudFormation template.

Following security best practice, we recommend that you also enable server identity verification on the `slurmdbd` client. To do this, configure the StorageParameters in the `slurmdbd.conf`. Upload the server CA certificate to the head node of the cluster. Next, set the SSL_CA option of `StorageParameters` in `slurmdbd.conf` to the path of the server CA certificate on the head node. Doing this enables server identity verification on the `slurmdbd` side. After you make these changes, restart the `slurmdbd` service to re-establish connectivity to the database server with identity verification enabled.

Updating the database credentials

To update the values for Database / UserName or PasswordSecretArn, you must first stop the compute fleet. Suppose that the secret value that's stored in the AWS Secrets Manager secret is changed and its ARN isn't changed. In this situation, the cluster doesn't automatically update the database password to the new value. To update the cluster for the new secret value, run the following command from the head node.

```
$ sudo /opt/parallelcluster/scripts/slurm/update_slurm_database_password.sh
```

> **Warning**
>
> To avoid losing accounting data, we recommend that you only change the database password when the compute fleet is stopped.

Database monitoring

We recommend that you enable the monitoring features of the AWS database services. For more information, see Amazon RDS monitoring or Amazon Aurora monitoring documentation.

Slurm configuration customization

Starting with AWS ParallelCluster version 3.6.0, you can customize the `slurm.conf` Slurm configuration in an AWS ParallelCluster cluster configuration.

In the cluster configuration, you can customize Slurm configuration parameters by using the following cluster configuration settings:

- Customize Slurm parameters for the entire cluster by using either the `SlurmSettings` / `CustomSlurmSettings` or the `CustomSlurmSettingsIncludeFile` parameter. AWS ParallelCluster fails if you specify both.
- Customize Slurm parameters for a queue by using `SlurmQueues` / `CustomSlurmSettings` (mapped to Slurm partitions).
- Customize Slurm parameters for a compute resource by using `SlurmQueues` / `ComputeResources` / `CustomSlurmSettings` (mapped to Slurm nodes).

Slurm configuration customization limits and considerations when using AWS ParallelCluster

- For `CustomSlurmSettings` and `CustomSlurmSettingsIncludeFile` settings, you can only specify and update `slurm.conf` parameters that are included in the Slurm version that's supported by the AWS ParallelCluster version that you are using to configure a cluster.
- If you specify custom Slurm configurations in any of the `CustomSlurmSettings` parameters, AWS ParallelCluster performs validation checks and prevents setting or updating Slurm configuration parameters that conflict with AWS ParallelCluster logic. The Slurm configuration parameters that are known to conflict with AWS ParallelCluster are identified in deny lists. The deny lists can change in future AWS ParallelCluster versions if other Slurm features are added. For more information, see Deny-listed Slurm configuration parameters for `CustomSlurmSettings`.
- AWS ParallelCluster only checks whether a parameter is in a deny list. AWS ParallelCluster doesn't validate your custom Slurm configuration parameter syntax or semantics. You are responsible for validating your custom Slurm configuration parameters. Invalid custom Slurm configuration parameters can cause Slurm daemon failures that can lead to cluster create and update failures.
- If you specify custom Slurm configurations in `CustomSlurmSettingsIncludeFile`, AWS ParallelCluster doesn't perform any validation.
- You can update `CustomSlurmSettings` and `CustomSlurmSettingsIncludeFile` without stopping and starting the compute fleet. In this case, AWS ParallelCluster restarts the `slurmctld` daemon and runs the `scontrol reconfigure` command.

  Some Slurm configuration parameters might require different operations before a change is registered in the entire cluster. For example, they might require a restart of all daemons in the cluster. You are responsible for verifying whether AWS ParallelCluster operations are sufficient for propagating your custom Slurm configuration parameter settings during updates. If you find that AWS ParallelCluster operations aren't sufficient, it's your responsibility to provide the additional actions required to propagate the updated settings as recommended in the Slurm documentation.

Deny-listed Slurm configuration parameters for `CustomSlurmSettings`

The following tables list the parameters with the AWS ParallelCluster versions that deny their use, starting with version 3.6.0. `CustomSlurmSettings` isn't supported for AWS ParallelCluster versions earlier than version 3.6.0.

Deny-listed parameters at cluster level:

Slurm parameter	Deny-listed in AWS ParallelCluster versions
CommunicationParameters	3.6.0
Epilog	3.6.0
GresTypes	3.6.0
LaunchParameters	3.6.0
Prolog	3.6.0
ReconfigFlags	3.6.0
ResumeFailProgram	3.6.0
ResumeProgram	3.6.0
ResumeTimeout	3.6.0
SlurmctldHost	3.6.0
SlurmctldLogFile	3.6.0
SlurmctldParameters	3.6.0
SlurmdLogfile	3.6.0
SlurmUser	3.6.0
SuspendExcNodes	3.6.0
SuspendProgram	3.6.0
SuspendTime	3.6.0
TaskPlugin	3.6.0
TreeWidth	3.6.0

Deny-listed parameters at cluster level when the native Slurm accounting integration is configured in the cluster configuration:

Slurm parameter	Deny-listed in AWS ParallelCluster versions
AccountingStorageType	3.6.0
AccountingStorageHost	3.6.0
AccountingStoragePort	3.6.0
AccountingStorageUser	3.6.0
JobAcctGatherType	3.6.0

Deny-listed parameters at the queue (partition) level for queues managed by AWS ParallelCluster:

Slurm parameter	Deny-listed in AWS ParallelCluster versions
Nodes	3.6.0
PartitionName	3.6.0
ResumeTimeout	3.6.0
State	3.6.0
SuspendTime	3.6.0

Deny-listed parameters at the compute resource (node) level for compute resource managed by AWS ParallelCluster:

Slurm parameter	Deny-listed in AWS ParallelCluster version and later versions
CPUs	3.6.0
Features	3.6.0

Slurm parameter	Deny-listed in AWS ParallelCluster version and later versions
Gres	3.6.0
NodeAddr	3.6.0
NodeHostname	3.6.0
NodeName	3.6.0
Weight	3.7.0

Slurm`prolog` and `epilog`

Starting with AWS ParallelCluster version 3.6.0, the Slurm configuration that's deployed with AWS ParallelCluster includes the `Prolog` and `Epilog` configuration parameters:

```
# PROLOG AND EPILOG
Prolog=/opt/slurm/etc/scripts/prolog.d/*
Epilog=/opt/slurm/etc/scripts/epilog.d/*
SchedulerParameters=nohold_on_prolog_fail
BatchStartTimeout=180
```

For more information, see the Prolog and Epilog Guide in the Slurm documentation.

AWS ParallelCluster includes the following prolog and epilog scripts:

- `90_plcuster_health_check_manager` (in the `Prolog` folder)
- `90_pcluster_noop` (in the `Epilog` folder)

> **Note**
>
> Both the `Prolog` and `Epilog` folder must contain at least one file.

You can use your own custom `prolog` or `epilog` scripts by adding them to the corresponding `Prolog` and `Epilog` folders.

> **Warning**
>
> Slurm runs every script in the folders, in reverse alphabetical order.

The run time duration of the `prolog` and `epilog` scripts impact the time needed to run a job. Update the `BatchStartTimeout` configuration setting when running multiple or long running `prolog` scripts. The default is 3 minutes.

If you are using custom `prolog` and `epilog` scripts, locate the scripts in the respective `Prolog` and `Epilog` folders. We recommend that you keep the `90_plcuster_health_check_manager` script that runs before every custom script. For more information, see Slurm configuration customization.

Cluster capacity size and update

The capacity of the cluster is defined by the number of compute nodes the cluster can scale. Compute nodes are backed by Amazon EC2 instances defined within compute resources in the AWS ParallelCluster configuration (`Scheduling/SlurmQueues/ ComputeResources`), and are organized into queues (`Scheduling/SlurmQueues`) that map 1:1 to Slurm partitions.

Within a compute resource it's possible to configure the minimum number of compute nodes (instances) that must always be kept running in the cluster (`MinCount`), and the maximum number of instances the compute resource can scale to (`MaxCount3`).

At cluster creation time, or upon a cluster update, AWS ParallelCluster launches as many Amazon EC2 instances as configured in `MinCount` for each compute resource (`Scheduling/SlurmQueues/ ComputeResources`) defined in the cluster. The instances launched to cover the minimal amount of nodes for a compute resources in the cluster are called ***static nodes***. Once started, static nodes are meant to be persistent in the cluster and they are not terminated by the system, unless a particular event or condition occurs. Such events include, for example, the failure of Slurm or Amazon EC2 health checks and the change of the Slurm node status to DRAIN or DOWN.

The Amazon EC2 instances, in the range of 1 to `'MaxCount - MinCount'` (`MaxCount` *minus* `MinCount`), launched on-demand to deal with the increased load of the cluster, are referred to as ***dynamic nodes***. Their nature is ephemeral, they are launched to serve pending jobs and are terminated once they stay idle for a period of time defined by `Scheduling/SlurmSettings/ScaledownIdletime` in the cluster configuration (default: 10 minutes).

Static nodes and dynamic node comply to the following naming schema:

- Static nodes `<Queue/Name>-st-<ComputeResource/Name>-<num>` where `<num> = 1..ComputeResource/MinCount`
- Dynamic nodes `<Queue/Name>-dy-<ComputeResource/Name>-<num>` where `<num> = 1..(ComputeResource/MaxCount - ComputeResource/MinCount)`

For example given the following AWS ParallelCluster configuration:

```
Scheduling:
    Scheduler: Slurm
    SlurmQueues:
        - Name: queue1
            ComputeResources:
                - Name: c5xlarge
                    Instances:
                        - InstanceType: c5.xlarge
                        MinCount: 100
                        MaxCount: 150
```

The following nodes will be defined in Slurm

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```

When a compute resource has `MinCount == MaxCount`, all the corresponding compute nodes will be static and all the instances will be launched at cluster creation/update time and kept up and running. For example:

```
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue1
      ComputeResources:
```

```
        - Name: c5xlarge
          Instances:
            - InstanceType: c5.xlarge
          MinCount: 100
          MaxCount: 100
```

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```

Cluster capacity update

The update of the cluster capacity includes adding or removing queues, compute resources or changing the `MinCount/MaxCount` of a compute resource. Starting from AWS ParallelCluster version 3.9.0, reducing the size of a queue requires the compute fleet to be stopped or QueueUpdateStrategy set to TERMINATE for before a cluster update to take place. It's not required to stop the compute fleet or to set QueueUpdateStrategy to TERMINATE when:

- Adding new queues to Scheduling/`SlurmQueues`
- Adding new compute resources `Scheduling/SlurmQueues/``ComputeResources` to a queue
- Increasing the `MaxCount` of a compute resource
- Increasing MinCount of a compute resource and increasing MaxCount of the same compute resource of at least the same amount

Considerations and limitations

This section is meant to outline any important factors, constraints, or limitations that should be taken into account when resizing the cluster capacity.

- When removing a queue from `Scheduling/``SlurmQueues` all the compute nodes with name `<Queue/Name>-*` , both static and dynamic, will be removed from the Slurm configuration and the corresponding Amazon EC2 instances will be terminated.
- When removing a compute resource `Scheduling/SlurmQueues/``ComputeResources` from a queue, all the compute nodes with name `<Queue/Name>-*-<ComputeResource/Name>-*` ,

both static and dynamic, will be removed from the Slurm configuration and the corresponding Amazon EC2 instances will be terminated.

When changing the `MinCount` parameter of a compute resource we can distinguish two different scenarios, if `MaxCount` is kept equal to `MinCount` (static capacity only), and if `MaxCount` is greater than `MinCount` (mixed static and dynamic capacity).

Capacity changes with static nodes only

- If `MinCount == MaxCount` , when increasing `MinCount` (and `MaxCount`), the cluster will be configured by extending the number of static nodes to the new value of `MinCount` `<Queue/Name>-st-<ComputeResource/Name>-<new_MinCount>` and the system will keep trying to launch Amazon EC2 instances to fulfill the new required static capacity.
- If `MinCount == MaxCount` , when decreasing `MinCount` (and `MaxCount`) of the amount N, the cluster will be configured by removing the last N static nodes `<Queue/Name>-st-<ComputeResource/Name>-<old_MinCount - N>...<old_MinCount>]` and the system will terminate the corresponding Amazon EC2 instances.
  - Initial state `MinCount = MaxCount = 100`
  -
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```
  - Update `-30` on `MinCount` and `MaxCount: MinCount = MaxCount = 70`
  -
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     70   idle queue1-st-c5xlarge-[1-70]
```

Capacity changes with mixed nodes

If `MinCount < MaxCount`, when increasing `MinCount` by an amount N (assuming `MaxCount` will be kept unchanged), the cluster will be configured by extending the number static nodes to

the new value of MinCount (old_MinCount + N): <Queue/Name>-st-<ComputeResource/Name>-<old_MinCount + N> and the system will keep trying to launch Amazon EC2 instances to fulfill the new required static capacity. Moreover, to honor the MaxCount capacity of the compute resource, the cluster configuration is updated by *removing the last N dynamic nodes*: <Queue/Name>-dy-<ComputeResource/Name>-[<MaxCount - old_MinCount - N>...<MaxCount - old_MinCount>] and the system will terminate the corresponding Amazon EC2 instances.

- Initial state: MinCount = 100; MaxCount = 150
- ```
  $ sinfo
  PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
  queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
  queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
  ```
- Update +30 to MinCount : MinCount = 130 (MaxCount = 150)
- ```
  $ sinfo
  PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
  queue1*      up   infinite     20  idle~ queue1-dy-c5xlarge-[1-20]
  queue1*      up   infinite    130   idle queue1-st-c5xlarge-[1-130]
  ```

If MinCount < MaxCount, when increasing MinCount and MaxCount of the same amount N, the cluster will be configured by extending the number static nodes to the new value of MinCount (old_MinCount + N): <Queue/Name>-st-<ComputeResource/Name>-<old_MinCount + N> and the system will keep trying to launch Amazon EC2 instances to fulfill the new required static capacity. Moreover, no changes will be done on the number of dynamic nodes to honor the new

MaxCount value.

- Initial state: MinCount = 100; MaxCount = 150
- ```
  $ sinfo
  PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
  queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
  ```
  ```

```
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```

- Update +30 to `MinCount : MinCount = 130 (MaxCount = 180)`
-

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     20  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite    130   idle queue1-st-c5xlarge-[1-130]
```

If `MinCount < MaxCount`, when decreasing `MinCount` of the amount N (assuming `MaxCount` will be kept unchanged), the cluster will be configured by removing the last N static nodes static nodes `<Queue/Name>-st-<ComputeResource/Name>-[<old_MinCount - N>...<old_MinCount>`and the system will terminate the corresponding Amazon EC2 instances. Moreover, to honor the `MaxCount` capacity of the compute resource, the cluster configuration is updated by extending the number of the dynamic nodes to fill the gap `MaxCount - new_MinCount: <Queue/Name>-dy-<ComputeResource/Name>-[1..<MazCount - new_MinCount>]` In this case, since those are dynamic nodes, no new Amazon EC2 instances will be launched unless the scheduler has jobs in pending on the new nodes.

- Initial state: `MinCount = 100; MaxCount = 150`
-

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```

- Update -30 on `MinCount : MinCount = 70 (MaxCount = 120)`
-

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     80  idle~ queue1-dy-c5xlarge-[1-80]
```

```
queue1*      up   infinite     70   idle queue1-st-c5xlarge-[1-70]
```

If `MinCount < MaxCount`, when decreasing `MinCount` and `MaxCount` of the same amount N, the cluster will be configured by removing the last N static nodes `<Queue/Name>-st-<ComputeResource/Name>-<old_MinCount - N>...<oldMinCount>]` and the system will terminate the corresponding Amazon EC2 instances.

Moreover, no changes will be done on the number of dynamic nodes to honor the new `MaxCount` value.

- Initial state: `MinCount = 100; MaxCount = 150`
-
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
```
- Update -30 on `MinCount : MinCount = 70 (MaxCount = 120)`
-
```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     80  idle~ queue1-dy-c5xlarge-[1-50]
queue1*      up   infinite     70   idle queue1-st-c5xlarge-[1-70]
```

If `MinCount < MaxCount`, when decreasing `MaxCount` of the amount N (assuming `MinCount` will be kept unchanged), the cluster will be configured by removing the last N dynamic nodes `<Queue/Name>-dy-<ComputeResource/Name>-<old_MaxCount - N...<oldMaxCount>]` and the system will terminate the corresponding Amazon EC2 instances in the case they were running.No impact is expected on the static nodes.

- Initial state: `MinCount = 100; MaxCount = 150`

- ```
  $ sinfo
  PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
  queue1*      up   infinite     50  idle~ queue1-dy-c5xlarge-[1-50]
  queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
  ```
- Update -30 on `MaxCount : MinCount = 100 (MaxCount = 120)`
- ```
  $ sinfo
  PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
  queue1*      up   infinite     20  idle~ queue1-dy-c5xlarge-[1-20]
  queue1*      up   infinite    100   idle queue1-st-c5xlarge-[1-100]
  ```

Impacts on the Jobs

In all the cases where nodes are removed and Amazon EC2 instances terminated, a sbatch job running on the removed nodes will be re-queued, unless there are no other nodes satisfying the job requirements. In this last case the job fails with status NODE_FAIL and disappears from the queue, and it must be re-submitted manually.

If you are planning to perform a cluster resize update, you can prevent jobs to go running in the nodes that are going to be removed during the planned update. This is possible by setting the nodes to be removed in maintenance. Please be aware that setting a node in maintenance would not impact jobs that are eventually already running in the node.

Suppose that with the planned cluster resize update you are going to remove the node `qeueu-st-computeresource-[9-10]`. You can create a Slurm reservation with the following command

```
sudo -i scontrol create reservation ReservationName=maint_for_update user=root
 starttime=now duration=infinite flags=maint,ignore_jobs nodes=qeueu-st-
computeresource-[9-10]
```

This will create a Slurm reservation named `maint_for_update` on the nodes `qeueu-st-computeresource-[9-10]`. From the time when the reservation is created, no more jobs can go running into the nodes `qeueu-st-computeresource-[9-10]`. Please be aware that

the reservation will not prevent jobs to be eventually allocated on the nodes `qeueu-st-computeresource-[9-10]`.

After the cluster resize update, if the Slurm reservation was set only on nodes that were removed during the resize update, the maintenance reservation will be automatically deleted. If instead you had created a Slurm reservation on nodes that are still present after the cluster resize update, we may want to remove the maintenance reservation on the nodes after the resize update is performed, by using the following command

```
sudo -i scontrol delete ReservationName=maint_for_update
```

For additional details on Slurm reservation, see the official SchedMD doc here.

Cluster update process on capacity changes

Upon a scheduler configuration change, the following steps are executed during the cluster update process:

- Stop AWS ParallelCluster `clustermgtd (supervisorctl stop clustermgtd)`
- Generate updated Slurm partitions configuration from AWS ParallelCluster configuration
- Restart `slurmctld` (done through Chef service recipe)
- Check `slurmctld` status `(systemctl is-active --quiet slurmctld.service)`
- Reload Slurm configuration `(scontrol reconfigure)`
- Start `clustermgtd (supervisorctl start clustermgtd)`

Using AWS Batch (`awsbatch`) scheduler with AWS ParallelCluster

AWS ParallelCluster also supports AWS Batch schedulers. The following topics describe how to use AWS Batch. For information about AWS Batch, see AWS Batch. For documentation, see the AWS Batch User Guide.

AWS ParallelCluster CLI commands for AWS Batch

When you use the `awsbatch` scheduler, the AWS ParallelCluster CLI commands for AWS Batch are automatically installed in the AWS ParallelCluster head node. The CLI uses AWS Batch API operations and permits the following operations:

- Submit and manage jobs.
- Monitor jobs, queues, and hosts.
- Mirror traditional scheduler commands.

> **Important**
>
> AWS ParallelCluster doesn't support GPU jobs for AWS Batch. For more information, see GPU jobs.

This CLI is distributed as a separate package. For more information, see Scheduler support.

Topics

- awsbsub
- awsbstat
- awsbout
- awsbkill
- awsbqueues
- awsbhosts

awsbsub

Submits jobs to the job queue of the cluster.

```
awsbsub [-h] [-jn JOB_NAME] [-c CLUSTER] [-cf] [-w WORKING_DIR]
        [-pw PARENT_WORKING_DIR] [-if INPUT_FILE] [-p VCPUS] [-m MEMORY]
        [-e ENV] [-eb ENV_DENYLIST] [-r RETRY_ATTEMPTS] [-t TIMEOUT]
        [-n NODES] [-a ARRAY_SIZE] [-d DEPENDS_ON]
        [command] [arguments [arguments ...]]
```

> **Important**
>
> AWS ParallelCluster doesn't support GPU jobs for AWS Batch. For more information, see GPU jobs.

Positional Arguments

command

Submits the job (the command specified must be available on the compute instances) or the file name to be transferred. See also `--command-file`.

arguments

(Optional) Specifies arguments for the command or the command-file.

Named Arguments

-jn *JOB_NAME*, --job-name *JOB_NAME*

Names the job. The first character must be either a letter or number. The job name can contain letters (both uppercase and lowercase), numbers, hyphens, and underscores, and be up to 128 characters in length.

-c *CLUSTER*, --cluster *CLUSTER*

Specifies the cluster to use.

-cf, --command-file

Indicates that the command is a file to be transferred to the compute instances.

Default: False

-w *WORKING_DIR*, --working-dir *WORKING_DIR*

Specifies the folder to use as the job's working directory. If a working directory isn't specified, the job is run in the `job-<AWS_BATCH_JOB_ID>` subfolder of the user's home directory. You can use either this parameter or the `--parent-working-dir` parameter.

-pw *PARENT_WORKING_DIR*, --parent-working-dir *PARENT_WORKING_DIR*

Specifies the parent folder of the job's working directory. If a parent working directory isn't specified, it defaults to the user's home directory. A subfolder named `job-<AWS_BATCH_JOB_ID>` is created in the parent working directory. You can use either this parameter or the `--working-dir` parameter.

-if *INPUT_FILE*, --input-file *INPUT_FILE*

Specifies the file to be transferred to the compute instances, in the job's working directory. You can specify multiple input file parameters.

-p *VCPUS*, --vcpus *VCPUS*

Specifies the number of vCPUs to reserve for the container. When used together with –nodes, it identifies the number of vCPUs for each node.

Default: 1

-m *MEMORY*, --memory *MEMORY*

Specifies the hard limit of memory (in MiB) to provide for the job. If your job attempts to exceed the memory limit specified here, the job is ended.

Default: 128

-e *ENV*, --env *ENV*

Specifies a comma-separated list of environment variable names to export to the job environment. To export all environment variables, specify ‘all’. Note that a list of 'all' environment variables doesn't include those listed in the `–env-blacklist` parameter, or variables starting with the `PCLUSTER_*` or `AWS_*` prefix.

-eb *ENV_DENYLIST*, --env-blacklist *ENV_DENYLIST*

Specifies a comma-separated list of environment variable names to **not** export to the job environment. By default, `HOME`, `PWD`, `USER`, `PATH`, `LD_LIBRARY_PATH`, `TERM`, and `TERMCAP` are not exported.

-r *RETRY_ATTEMPTS*, --retry-attempts *RETRY_ATTEMPTS*

Specifies the number of times to move a job to the `RUNNABLE` status. You can specify between 1 and 10 attempts. If the value of attempts is greater than 1, the job is retried if it fails, until it has moved to a `RUNNABLE` status for the specified number of times.

Default: 1

-t *TIMEOUT*, --timeout *TIMEOUT*

Specifies the time duration in seconds (measured from the job attempt’s `startedAt` timestamp) after which AWS Batch terminates your job if it hasn't finished. The timeout value must be at least 60 seconds.

-n *NODES*, --nodes *NODES*

Specifies the number of nodes to reserve for the job. Specify a value for this parameter to enable multi-node parallel submission.

Note

When the Scheduler / AwsBatchQueues / CapacityType parameter is set to SPOT, multi-node parallel jobs *aren't* supported. Additionally, there must be an AWSServiceRoleForEC2Spot service-linked role in your account. You can create this role with the following AWS CLI command:

```
$ aws iam create-service-linked-role --aws-service-name spot.amazonaws.com
```

For more information, see Service-linked role for Spot Instance requests in the *Amazon Elastic Compute Cloud User Guide for Linux Instances*.

`-a ARRAY_SIZE, --array-size ARRAY_SIZE`

Indicates the size of the array. You can specify a value between 2 and 10,000. If you specify array properties for a job, it becomes an array job.

`-d DEPENDS_ON, --depends-on DEPENDS_ON`

Specifies a semicolon-separated list of dependencies for a job. A job can depend upon a maximum of 20 jobs. You can specify a SEQUENTIAL type dependency without specifying a job ID for array jobs. A sequential dependency allows each child array job to complete sequentially, starting at index 0. You can also specify an N_TO_N type dependency with a job ID for array jobs. An N_TO_N dependency means that each index child of this job must wait for the corresponding index child of each dependency to complete before it can begin. The syntax for this parameter is "jobId=*<string>*,type=*<string>*;...".

awsbstat

Shows the jobs that are submitted in the cluster's job queue.

```
awsbstat [-h] [-c CLUSTER] [-s STATUS] [-e] [-d] [job_ids [job_ids ...]]
```

Positional Arguments

job_ids

Specifies the space-separated list of job IDs to show in the output. If the job is a job array, all of the child jobs are displayed. If a single job is requested, it is shown in a detailed version.

Named Arguments

-c *CLUSTER*, --cluster *CLUSTER*

Indicates the cluster to use.

-s *STATUS*, --status *STATUS*

Specifies a comma-separated list of job statuses to include. The default job status is “active.”. Accepted values are: SUBMITTED, PENDING, RUNNABLE, STARTING, RUNNING, SUCCEEDED, FAILED, and ALL.

Default: “SUBMITTED,PENDING,RUNNABLE,STARTING,RUNNING”

-e, --expand-children

Expands jobs with children (both array and multi-node parallel).

Default: False

-d, --details

Shows jobs details.

Default: False

awsbout

Shows the output of a given job.

```
awsbout [-h] [-c CLUSTER] [-hd HEAD] [-t TAIL] [-s] [-sp STREAM_PERIOD] job_id
```

Positional Arguments

job_id

Specifies the job ID.

Named Arguments

`-c CLUSTER, --cluster CLUSTER`

Indicates the cluster to use.

`-hd HEAD, --head HEAD`

Gets the first *HEAD* lines of the job output.

`-t TAIL, --tail TAIL`

Gets the last <tail> lines of the job output.

`-s, --stream`

Gets the job output, and then waits for additional output to be produced. This argument can be used together with –tail to start from the latest <tail> lines of the job output.

Default: False

`-sp STREAM_PERIOD, --stream-period STREAM_PERIOD`

Sets the streaming period.

Default: 5

awsbkill

Cancels or terminates jobs submitted in the cluster.

```
awsbkill [-h] [-c CLUSTER] [-r REASON] job_ids [job_ids ... ]
```

Positional Arguments

`job_ids`

Specifies the space-separated list of job IDs to cancel or terminate.

Named Arguments

`-c CLUSTER, --cluster CLUSTER`

Indicates the name of the cluster to use.

`-r REASON, --reason REASON`

Indicates the message to attach to a job, explaining the reason for canceling it.

Default: “Terminated by the user”

awsbqueues

Shows the job queue that is associated with the cluster.

```
awsbqueues [-h] [-c CLUSTER] [-d] [job_queues [job_queues ... ]]
```

Positional arguments

`job_queues`

Specifies the space-separated list of queue names to show. If a single queue is requested, it is shown in a detailed version.

Named arguments

`-c CLUSTER, --cluster CLUSTER`

Specifies the name of the cluster to use.

`-d, --details`

Indicates whether to show the details of the queues.

Default: False

awsbhosts

Shows the hosts that belong to the cluster’s compute environment.

```
awsbhosts [-h] [-c CLUSTER] [-d] [instance_ids [instance_ids ... ]]
```

Positional Arguments

instance_ids

Specifies a space-separated list of instances IDs. If a single instance is requested, it is shown in a detailed version.

Named Arguments

-c *CLUSTER*, --cluster *CLUSTER*

Specifies the name of the cluster to use.

-d, --details

Indicates whether to show the details of the hosts.

Default: False

Shared storage

AWS ParallelCluster supports either using Amazon EBS, FSx for ONTAP, and FSx for OpenZFS shared storage volumes, Amazon EFS and FSx for Lustre shared storage file systems, or File Caches. We recommend that you follow the AWS well-architected framework reliability pillar guidance and back up your volumes and file systems.

Select a storage system that meets your HPC application I/O requirements. You can optimize each file system based on your specific use case. For more information, see storage options overview.

Amazon EBS volumes are attached to the head node and shared with compute nodes through NFS. This option can be cost effective, but performance depends on the head node resources as storage needs scale. This can become a bottleneck as more compute nodes are added to the cluster and the throughput demand increases.

Amazon EFS files systems scale as storage needs change. You can configure these file systems for a variety of use cases. Use Amazon EFS file systems to run parallelized and latency sensitive applications on your cluster.

FSx for Lustre file systems can process massive data sets at up to hundreds of gigabytes per second throughput, millions of IOPS, and sub-millisecond latencies. Use FSx for Lustre file systems for demanding high performance compute environments.

In the SharedStorage section, you can define either external or AWS ParallelCluster managed storage:

- **External storage** refers to an existing volume or file system that you manage. AWS ParallelCluster doesn't create or delete this storage.
- **Managed storage** refers to a volume or file system that AWS ParallelCluster created and can delete.

External storage

You can configure AWS ParallelCluster to attach external storage to the cluster when the cluster is created or updated. Similarly you can configure it to detach external storage from the cluster when the cluster is deleted or updated. Your data is preserved and you can use it for long-term permanent shared storage outside of the cluster lifecycle.

> **Note**
>
> Versions of AWS ParallelCluster prior to 3.8 do not allow for externally managed filesystems to be mounted at /home. Starting from version 3.8, AWS ParallelCluster allows you to use /home as a mount point for an external managed filesystem. You can mount an externally managed file system to /home by specifying /home as the value to the MountDir parameter under the SharedStorage section.
> Amazon File Cache is not suitable for use as the system /home directory and therefore is not supported at this time for mounting /home.
> When specifying a /home directory under the SharedStorage section the SharedStorageType configuration option will be overridden, meaning the settings under SharedStorage section will be used instead.
> When mounting an external filesystem to the /home directory AWS ParallelCluster copies the head node's /home contents to the external filesystem, without overwriting existing files on the external storage. This includes transferring the cluster's SSH key for the default user, if it is absent on the external filesystem. For more information refer to AWS ParallelCluster shared storage considerations.

AWS ParallelCluster managed storage

AWS ParallelCluster managed storage is dependent on the lifecycle of the cluster by default in the configuration. The `SharedStorage` `DeletionPolicy` configuration parameter is set to `Delete` by default.

By default, an AWS ParallelCluster managed file system or volume and its data are deleted if one of the following is true.

- You delete the cluster.
- You change the managed shared storage configuration `Name`.
- You remove the managed shared storage from the configuration.

Set `DeletionPolicy` to `Retain` to persist your managed shared file system or volume and data. We recommend that you backup your data regularly to avoid the loss of data. You can use AWS Backup to centrally manage backups for all of your storage options.

You can remove the life cycle dependency with configuration settings. For more information, see Convert AWS ParallelCluster managed storage to external storage.

For information on shared storage quotas, see Quotas for shared storage.

For more information about shared storage and switching to new AWS ParallelCluster versions, see Best practices: moving a cluster to a new AWS ParallelCluster minor or patch version.

You can configure AWS ParallelCluster to attach external storage to the cluster when the cluster is created or updated. Similarly, you can configure it to detach external storage from the cluster when the cluster is deleted or updated. Your data is preserved and you can use it for long term permanent shared storage solutions that are independent of the cluster lifecycle.

By default, managed storage is dependent on the lifecycle of the cluster. You can remove the dependency with configuration settings that are described in Convert AWS ParallelCluster managed storage to external storage.

With specific settings, you can optimize each of the supported storage solutions for your use cases.

For shared storage quotas, see Quotas for shared storage.

For more information about shared storage and switching to new AWS ParallelCluster versions, see Best practices: moving a cluster to a new AWS ParallelCluster minor or patch version.

The following topics describe how to configure shared storage for each storage service that AWS ParallelCluster supports.

Topics

Amazon Elastic Block Store

To use an existing external Amazon EBS volume for long term permanent storage that's independent of the cluster life cycle, specify `EbsSettings` / `VolumeId`.

If you don't specify `VolumeId`, by default, AWS ParallelCluster creates a managed EBS volume from `EbsSettings` when your cluster is created. AWS ParallelCluster also deletes the volume and data when the cluster is deleted or the volume is removed from the cluster configuration.

For an AWS ParallelCluster managed EBS volume, you can use `EbsSettings` / `DeletionPolicy` to instruct AWS ParallelCluster to `Delete`, `Retain`, or `Snapshot` the volume when either the cluster is deleted or when the volume is removed from the cluster configuration. By default, `DeletionPolicy` is set to `Delete`.

Warning

For AWS ParallelCluster managed shared storage, `DeletionPolicy` is set to `Delete` by default.
This means that, if one of the following is true, a managed volume and its data are deleted:

- You delete the cluster.
- You change the managed shared storage configuration `SharedStorage` / `Name`.
- You remove the managed shared storage from the configuration.

We recommend that you back up your shared storage with snapshots regularly to avoid the loss of data. For more information about Amazon EBS snapshots, see Amazon EBS snapshots in the *Amazon Elastic Compute Cloud User Guide for Linux Instances*. To learn

how to manage data backups across AWS services, see AWS Backup in the *AWS Backup Developer Guide*.

Amazon Elastic File System

To use an existing external Amazon EFS file system for long-term permanent storage outside of the cluster life cycle, specify `EfsSettings` / `FileSystemId`, by default, AWS ParallelCluster creates a managed Amazon EFS file system from `EfsSettings` when it creates the cluster. AWS ParallelCluster also deletes the file system and data when the cluster is deleted or when the file system is removed from the cluster configuration.

For an AWS ParallelCluster managed Amazon EFS file system, you can use the `EfsSettings` / `DeletionPolicy` to instruct AWS ParallelCluster to `Delete` or `Retain` either when the cluster is deleted or when the file system removed from the cluster configuration. By default, `DeletionPolicy` is set to `Delete`.

Warning

For AWS ParallelCluster managed shared storage, `DeletionPolicy` is set to `Delete` by default.
This means that, if one of the following is true, a managed file system and its data are deleted:

- You delete the cluster.
- You change the managed shared storage configuration `SharedStorage` / `Name`.
- You remove the managed shared storage from the configuration.

We recommend that you back up your shared storage regularly to avoid the loss of data. For more information about how to back up individual Amazon EFS volumes, see Backing up your Amazon EFS file systems in the *Amazon Elastic File System User Guide*. To learn how to manage data backups across AWS services, see AWS Backup in the *AWS Backup Developer Guide*.

Amazon FSx for Lustre

To use an existing external FSx for Lustre file system for long-term permanent storage outside of the cluster life cycle, specify `FsxLustreSettings` / `FileSystemId`.

If you don't specify `FsxLustreSettings` / `FileSystemId`, by default, AWS ParallelCluster creates a managed FSx for Lustre file system from `FsxLustreSettings` when it creates the cluster. AWS ParallelCluster also deletes the file system and data when the cluster is deleted or when the file system is removed from the cluster configuration.

For an AWS ParallelCluster managed FSx for Lustre file system, you can use the `FsxLustreSettings` / `DeletionPolicy` to instruct AWS ParallelCluster to `Delete` or `Retain` the file system when either the cluster is deleted or when the file system is removed from the cluster configuration. By default, `DeletionPolicy` is set to `Delete`.

Warning

For AWS ParallelCluster managed shared storage, `DeletionPolicy` is set to `Delete` by default.
This means that, if one of the following is true, a managed file system and its data are deleted:

- You delete the cluster.
- You change the managed shared storage configuration `SharedStorage` / `Name`.
- You remove the managed shared storage from the configuration.

We recommend that you back up your shared storage regularly to avoid the loss of data. You can define backups in your cluster with `SharedStorage` / `FsxLustreSettings` / `AutomaticBackupRetentionDays` and `DailyAutomaticBackupStartTime`. To learn how to manage data backups across AWS services, see AWS Backup in the *AWS Backup Developer Guide*.

Configure FSx for ONTAP, FSx for OpenZFS, and File Cache shared storage

For FSx for ONTAP, FSx for OpenZFS, and File Cache, you can use `FsxOntapSettings` / `VolumeId`, `FsxOpenZfsSettings` / `VolumeId`, and `FileCacheSettings` / `FileCacheId` to specify mounting an external existing volume or File Cache for your cluster.

AWS ParallelCluster managed shared storage isn't supported for FSx for ONTAP, FSx for OpenZFS, and File Cache.

Working with shared storage in AWS ParallelCluster

In the following sections, you will learn about working with AWS ParallelCluster and shared storage, including shared storage considerations and how to convert managed storage to external storage.

Topics

AWS ParallelCluster shared storage considerations

Consider the following when working with shared storage in AWS ParallelCluster.

- Back up your file system data with AWS Backup or another method to manage backups for all of your storage systems.
- To add shared storage, you add a shared storage section to your configuration file and create or update the cluster.
- To remove shared storage, you remove the shared storage section from your configuration file and update the cluster.
- To replace existing AWS ParallelCluster managed shared storage with new managed storage, change the value for `SharedStorage` / `Name` and update the cluster.

> **Warning**
>
> By default, the existing AWS ParallelCluster managed storage and data is deleted when you perform the cluster update with a new `Name` parameter. If you need to change

Name and retain the existing managed shared storage data, make sure you either set the DeletionPolicy to Retain or back up the data before you update the cluster.

- If you don't back up AWS ParallelCluster managed storage data and DeletionPolicy is Delete, your data is deleted when either your cluster is deleted or when your managed storage is removed from the cluster configuration and the cluster is updated.
- If you don't back up AWS ParallelCluster managed storage data and DeletionPolicy is Retain, your file system is detached before the cluster is deleted and can be re-attached to another cluster as an external file system. Your data is preserved.
- If AWS ParallelCluster managed storage is removed from the cluster configuration and DeletionPolicy is Retain, it can be re-attached to the cluster as an external file system with your cluster data preserved.
- Starting with AWS ParallelCluster version 3.4.0, you can enhance security for Amazon EFS file system mounts by configuring SharedStorage / EfsSettings / EncryptionInTransit and IamAuthorization settings.
- When mounting an external filesystem to the /home directory, AWS ParallelCluster copies the contents of the head node's /home directory to the external filesystem. It copies existing data in the /home directory without overwriting existing files or directories on the external storage. This includes the cluster's SSH key for the default user in case it does not already exist on the external filesystem. Consequently all other clusters that mount the same external filesystem to their respective /home directory will also have the same SSH key for their default user of the cluster.
- In a multi-cluster environment that mounts the same external filesystem to the /home directories of clusters, SSH keys that grant access to the compute nodes, created on the head node by AWS ParallelCluster, are generated only once when the first cluster mounts the external filesystem to /home. All other clusters use the same SSH key. As a result, anyone possessing the SSH key for the default user of these shared clusters can access any cluster. All compute nodes allow connections using the initially generated key.

Convert AWS ParallelCluster managed storage to external storage

Learn how to convert AWS ParallelCluster managed storage to external storage.

The procedures are based on the following example configuration file snippet.

```
...
```

```
  - MountDir: /fsx
    Name: fsx
    StorageType: FsxLustre
    FsxLustreSettings:
      StorageCapacity: 1200
      DeletionPolicy: Delete
...
```

Convert AWS ParallelCluster managed storage to external storage

1. Set the `DeletionPolicy` to `Retain` in the cluster configuration file.

```
...
   - MountDir: /fsx
     Name: fsx
     StorageType: FsxLustre
     FsxLustreSettings:
       StorageCapacity: 1200
       DeletionPolicy: Retain
...
```

2. To set the `DeletionPolicy` change, run the following command.

```
pcluster update-cluster -n cluster-name -c cluster-config.yaml
```

3. Remove the `SharedStorage` section from the cluster configuration file.

```
...
...
```

4. To change the managed `SharedStorage` to external `SharedStorage` and detach it from the cluster, run the following command.

```
pcluster update-cluster -n cluster-name -c cluster-config.yaml
```

5. Your shared storage is now external and detached from the cluster.
6. To attach your external file system to the original cluster or another cluster, follow these steps.

   a. Get the FSx for Lustre file system ID.

      i. To use the AWS CLI run the following command and find the file system with a name that includes the name of your original cluster and note the file system ID.

```
aws fsx describe-file-systems
```

ii. To use the AWS Management Console, log in and navigate to the https://console.aws.amazon.com/fsx/. In the list of file systems, find the file system with a name that includes the name of your original cluster and note the file system ID.

b. Update the file system security group rules to provide access to and from the file system and cluster subnets. You can find the file system security group name and ID in the Amazon FSx console.

   Add rules to the file system security group that allow inbound and outbound TCP traffic from and to the head node and the compute node IP CIDR ranges or prefixes. Specify TCP ports 988, 1021, 1022, and 1023 for the inbound and outbound TCP traffic.

   For more information, see `SharedStorage` / `FsxLustreSettings` / `FileSystemId` and Creating, configuring, and deleting security groups for Amazon EC2 in the *AWS Command Line Interface User Guide for Version 2*.

c. Add the `SharedStorage` section to the cluster configuration.

```
...
   - MountDir: /fsx
     Name: fsx-external
     StorageType: FsxLustre
     FsxLustreSettings:
       FileSystemId: fs-02e5b4b4abd62d51c
...
```

d. To add the external shared storage to the cluster, run the following command.

```
pcluster update-cluster -n cluster-name -c cluster-config.yaml
```

Quotas for shared storage

Configure cluster `SharedStorage` to mount existing shared file storage and create new shared file storage based on the quotas that are listed in the following table.

The mounted file storage quotas for each cluster

File shared storage type	AWS ParallelCluster managed storage	External storage	Quota net total
Amazon EBS	5	5	5
RAID	1	0	1
Amazon EFS	1	20	21
Amazon FSx †	1 FSx for Lustre	20	21

Note

This table of quotas is added in AWS ParallelCluster version 3.2.0.

† AWS ParallelCluster only supports mounting existing Amazon FSx for NetApp ONTAP, Amazon FSx for OpenZFS, and File Cache systems. It doesn't support the creation of new FSx for ONTAP, FSx for OpenZFS, and File Cache systems.

Note

If you use AWS Batch as a scheduler, FSx for Lustre is only available on the cluster head node.
File Caches don't support AWS Batch schedulers.

AWS ParallelCluster resources and tagging

With AWS ParallelCluster you can create tags to track and manage your AWS ParallelCluster resources. You define the tags that you want AWS CloudFormation to create and propagate to all cluster resources in the Tags section of the cluster configuration file. You can also use tags that AWS ParallelCluster automatically generates to track and manage your resources.

When you create a cluster, the cluster and its resources are tagged with the AWS ParallelCluster and AWS systems tags defined in this section.

AWS ParallelCluster applies tags to the cluster instances, volumes, and resources. To identify the cluster stack, AWS CloudFormation applies AWS system tags to the cluster instances. To identify the cluster Amazon EC2 launch templates, Amazon EC2 applies system tags to the instances. You can use these tags to view and manage your AWS ParallelCluster resources.

> **Warning**
>
> All AWS ParallelCluster tags are essential and must not be modified in order to avoid impacts to system functionality. Because of this, you can't modify AWS system tags.

The following is an example of an AWS system tag for an AWS ParallelCluster resource.

```
"aws:cloudformation:stack-name"="clustername"
```

The following is an example of an AWS ParallelCluster tag applied to a resource.

```
"parallelcluster:cluster-name"="clustername"
```

You can view these tags in the Amazon EC2 section of the AWS Management Console.

View tags

Complete the following steps to view tags in the Amazon EC2 section of the AWS Management Console.

View tags

1. Navigate the Amazon EC2 console at https://console.aws.amazon.com/ec2/.
2. To view all cluster tags, choose **Tags** in the navigation pane.
3. To view cluster tags by instance, choose **Instances** in the navigation pane.
4. Select a cluster instance.
5. Choose the **Manage tags** tab in the instance details and view the tags.
6. Choose the **Storage** tab in the instance details.
7. Select the **Volume ID**.
8. In **Volumes**, choose the volume.
9. Choose the **Tags** tab in the volume details and view the tags.

AWS ParallelCluster head node instance tags

Key	Tag value
parallelcluster:cluster-name	*clustername*
Name	HeadNode
aws:ec2launchtemplate:id	*lt-1234567890abcdef0*
aws:ec2launchtemplate:version	*1*
parallelcluster:node-type	HeadNode
aws:cloudformation:stack-name	*clustername*
aws:cloudformation:logical-id	HeadNode
aws:cloudformation:stack-id	arn:aws:cloudformation: *region-id* :*ACCOUNTID* :stack/*clustername* /*1234abcd-12ab-12ab-12ab-1234567890abcdef0*
parallelcluster:version	*3.13.2*

AWS ParallelCluster head node root volume tags

Tag key	Tag value
parallelcluster:cluster-name	*clustername*
parallelcluster:node-type	HeadNode
parallelcluster:version	*3.13.2*

AWS ParallelCluster compute node instance tags

Key	Tag value
parallelcluster:cluster-name	*clustername*

Key	Tag value
parallelcluster:compute-resource-name	*compute-resource-name*
aws:ec2launchtemplate:id	*lt-1234567890abcdef0*
aws:ec2launchtemplate:version	*1*
parallelcluster:node-type	Compute
parallelcluster:queue-name	*queue-name*
parallelcluster:version	*3.13.2*

AWS ParallelCluster compute node root volume tags

Tag key	Tag value
parallelcluster:cluster-name	*clustername*
parallelcluster:compute-resource-name	*compute-resource-name*
parallelcluster:node-type	Compute
parallelcluster:queue-name	*queue-name*
parallelcluster:version	*3.13.2*

PCUI tags

Tag key	Tag value
parallelcluster-ui	true

Monitoring AWS ParallelCluster and logs

Monitoring is an important part of maintaining the reliability, availability, and performance of AWS ParallelCluster and your other AWS solutions. AWS provides the following monitoring tools to watch AWS ParallelCluster, report when something is wrong, and take automatic actions when appropriate:

- *Amazon CloudWatch* monitors your AWS resources and the applications you run on AWS in real time. You can collect and track metrics, create customized dashboards, and set alarms that notify you or take actions when a specified metric reaches a threshold that you specify. For example, you can have CloudWatch track CPU usage or other metrics of your Amazon EC2 instances and automatically launch new instances when needed. For more information, see the Amazon CloudWatch User Guide.
- *Amazon CloudWatch Logs* enables you to monitor, store, and access your log files from Amazon EC2 instances, CloudTrail, and other sources. CloudWatch Logs can monitor information in the log files and notify you when certain thresholds are met. You can also archive your log data in highly durable storage. For more information, see the Amazon CloudWatch Logs User Guide.
- *AWS CloudTrail* captures API calls and related events made by or on behalf of your AWS account and delivers the log files to an Amazon S3 bucket that you specify. You can identify which users and accounts called AWS, the source IP address from which the calls were made, and when the calls occurred. For more information, see the AWS CloudTrail User Guide.
- *Amazon EventBridge* is a serverless event bus service that makes it easy to connect your applications with data from a variety of sources. EventBridge delivers a stream of real-time data from your own applications, Software-as-a-Service (SaaS) applications, and AWS services and routes that data to targets such as Lambda. This enables you to monitor events that happen in services, and build event-driven architectures. For more information, see the Amazon EventBridge User Guide.

Topics

- Amazon EC2 console output logs
- Retrieve PCUI and AWS ParallelCluster runtime logs
- Retrieving and preserving logs

Integration with Amazon CloudWatch Logs

For more information about CloudWatch Logs, see Amazon CloudWatch Logs User Guide. To configure CloudWatch Logs integration, see the `Monitoring` section. To learn how to append custom logs to the CloudWatch configuration using `append-config`, see Multiple CloudWatch agent configuration files in the *Amazon CloudWatch User Guide*.

Amazon CloudWatch Logs cluster logs

A log group is created for each cluster with a name, `/aws/parallelcluster/`*`cluster-name-<timestamp>`* (for example, `/aws/parallelcluster/testCluster-202202050215`). Each log (or set of logs if the path contains a `*`) on each node has a log stream named *`{hostname}`*`.`*`{instance_id}`*`.`*`{logIdentifier}`*. (For example `ip-172-31-10-46.i-02587cf29cc3048f3.nodewatcher`.) Log data is sent to CloudWatch by the CloudWatch agent, which runs as `root` on all cluster instances.

An Amazon CloudWatch dashboard is created when the cluster is created. This dashboard gives you the ability to review the logs stored in CloudWatch Logs. For more information, see Amazon CloudWatch dashboard.

This list contains the *`logIdentifier`* and path for the log streams available for platforms, schedulers, and nodes.

Log streams available for platforms, schedulers, and nodes

Platfor	Schedu s	Nodes	Log streams
amazon redhat ubuntu	awsbat slurm	HeadNc	dcv-authenticator: `/var/log/parallelcluster/pc luster_dcv_authenticator.log` dcv-ext-authenticator: `/var/log/parallelcluster/pc luster_dcv_connect.log` dcv-agent: `/var/log/dcv/agent.*.log`

Platfor	Schedu s	Nodes	Log streams
			dcv-xsession: /var/log/dcv/dcv-xsession.*.log dcv-server: /var/log/dcv/server.log dcv-session-launcher: /var/log/dcv/sessionlauncher.log Xdcv: /var/log/dcv/Xdcv.*.log cfn-init: /var/log/cfn-init.log chef-client: /var/log/chef-client.log
amazon redhat ubuntu	awsbat slurm	Comput eet HeadNc	cloud-init: /var/log/cloud-init.log supervisord: /var/log/supervisord.log
amazon redhat ubuntu	slurm	Comput eet	cloud-init-output: /var/log/cloud-init-output.log computemgtd: /var/log/parallelcluster/computemgtd slurmd: /var/log/slurmd.log slurm_prolog_epilog: /var/log/parallelcluster/sl urm_prolog_epilog.log

Platfor	Schedu s	Nodes	Log streams
amazon redhat ubuntu	slurm	HeadNc	sssd: /var/log/sssd/sssd.log sssd_domain_default: /var/log/sssd/sssd_default.log pam_ssh_key_generator: /var/log/parallelcluster/pam_ssh_key_generator.log clusterstatusmgtd: /var/log/parallelcluster/clusterstatusmgtd clustermgtd: /var/log/parallelcluster/clustermgtd compute_console_output: /var/log/parallelcluster/compute_console_output slurm_resume: /var/log/parallelcluster/slurm_resume.log slurm_suspend: /var/log/parallelcluster/slurm_suspend.log slurmctld: /var/log/slurmctld.log slurm_fleet_status_manager: /var/log/parallelcluster/slurm_fleet_status_manager.log
amazon redhat	awsbatc slurm	Comput eet HeadNc	system-messages: /var/log/messages
ubuntu	awsbatc slurm	Comput eet HeadNc	syslog: /var/log/syslog

Jobs in clusters that use AWS Batch store the output of jobs that reached a state of RUNNING, SUCCEEDED, or FAILED in CloudWatch Logs. The log group is /aws/batch/job, and the log stream name format is *jobDefinitionName*/default/*ecs_task_id*. By default, these logs are set not to expire, but you can modify the retention period. For more information, see Change log data retention in CloudWatch Logs in the *Amazon CloudWatch Logs User Guide*.

Amazon CloudWatch Logs build image logs

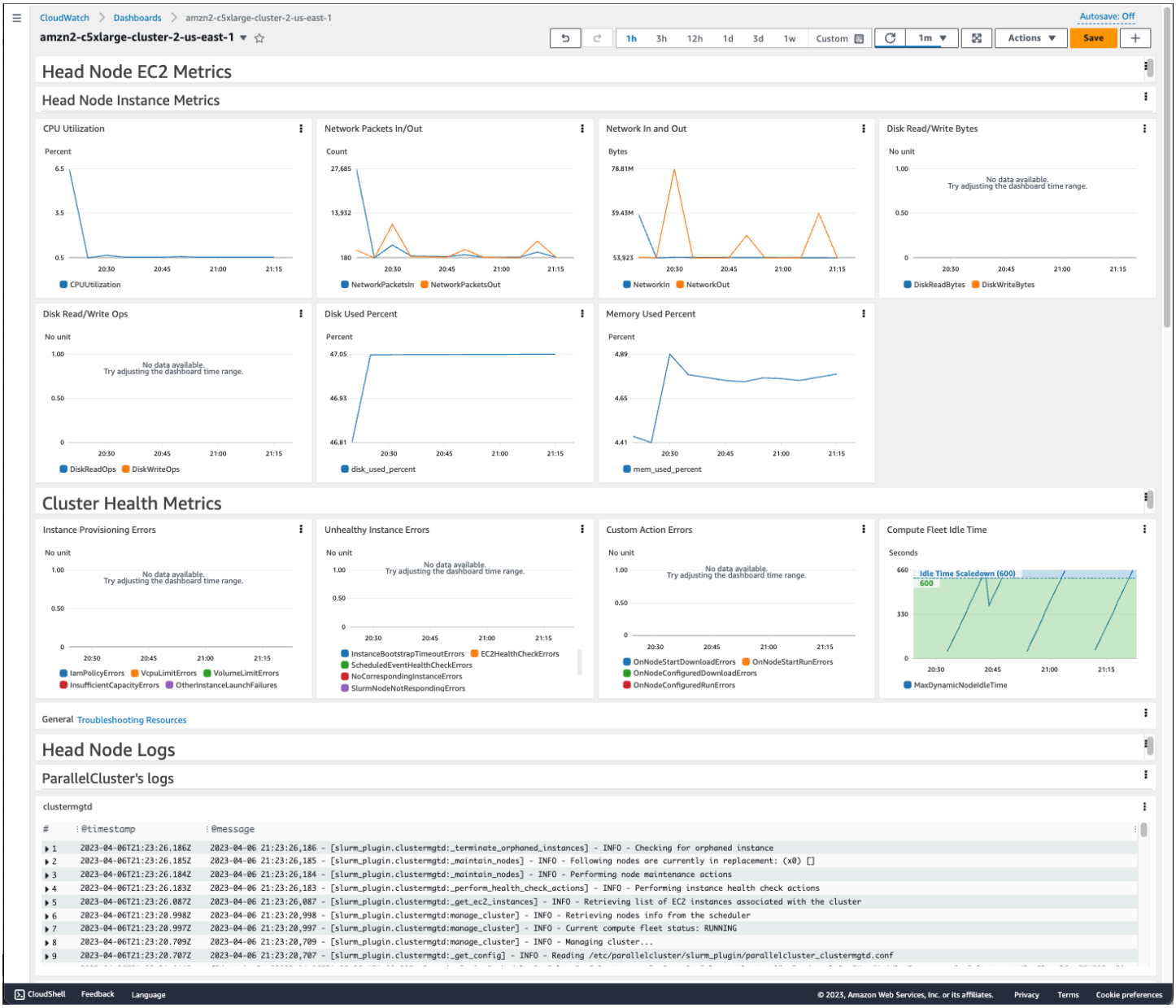
A log group is created for each custom build image with a name, /aws/imagebuilder/ParallelClusterImage-*<image-id>*. A unique log stream with name, *{pcluster-version}*/1 contains the output of the build image process.

You can access the logs by using the `pcluster` image commands. For more information, see AWS ParallelCluster AMI customization.

Amazon CloudWatch dashboard

An Amazon CloudWatch dashboard is created when a cluster is created. This makes it easier to monitor the nodes in your cluster, and to view the logs stored in Amazon CloudWatch Logs. The name of the dashboard is *ClusterName*-*Region*. *ClusterName* is the name of your cluster and *Region* is the AWS Region the cluster is in. You can access the dashboard in the console, or by opening https://console.aws.amazon.com/cloudwatch/home?region=*Region*#dashboards:name=*ClusterName*-*Region*.

The following image shows an example CloudWatch dashboard for a cluster.



Head Node Instance Metrics

The first section of the dashboard displays graphs of the head node Amazon EC2 metrics.

If your cluster has shared storage, the next section shows shared storage metrics.

Cluster Health Metrics

If your cluster uses Slurm for scheduling, the cluster health metric graphs show real-time cluster compute node errors. For more information, see Troubleshooting cluster health metrics. Cluster health metrics are added to the dashboard starting with AWS ParallelCluster version 3.6.0.

Head Node Logs

The final section lists head node logs grouped by AWS ParallelCluster's logs, Scheduler's logs, Amazon DCV integration logs, and System's logs.

For more information about Amazon CloudWatch dashboards, see Using Amazon CloudWatch dashboards in the *Amazon CloudWatch User Guide*.

If you don't want to create the Amazon CloudWatch dashboard, you can turn it off by setting `Monitoring` / `Dashboards` / `CloudWatch` / `Enabled` to `false`.

> **Note**
>
> If you disable the creation of the Amazon CloudWatch dashboard, you also disable the Amazon CloudWatch `disk_used_percent` and `memory_used_percent` alarms for your cluster. For more information, see Amazon CloudWatch alarms for cluster metrics.
> The `disk_used_percent` and `memory_used_percent` alarms are added starting with AWS ParallelCluster version 3.6.

Amazon CloudWatch alarms for cluster metrics

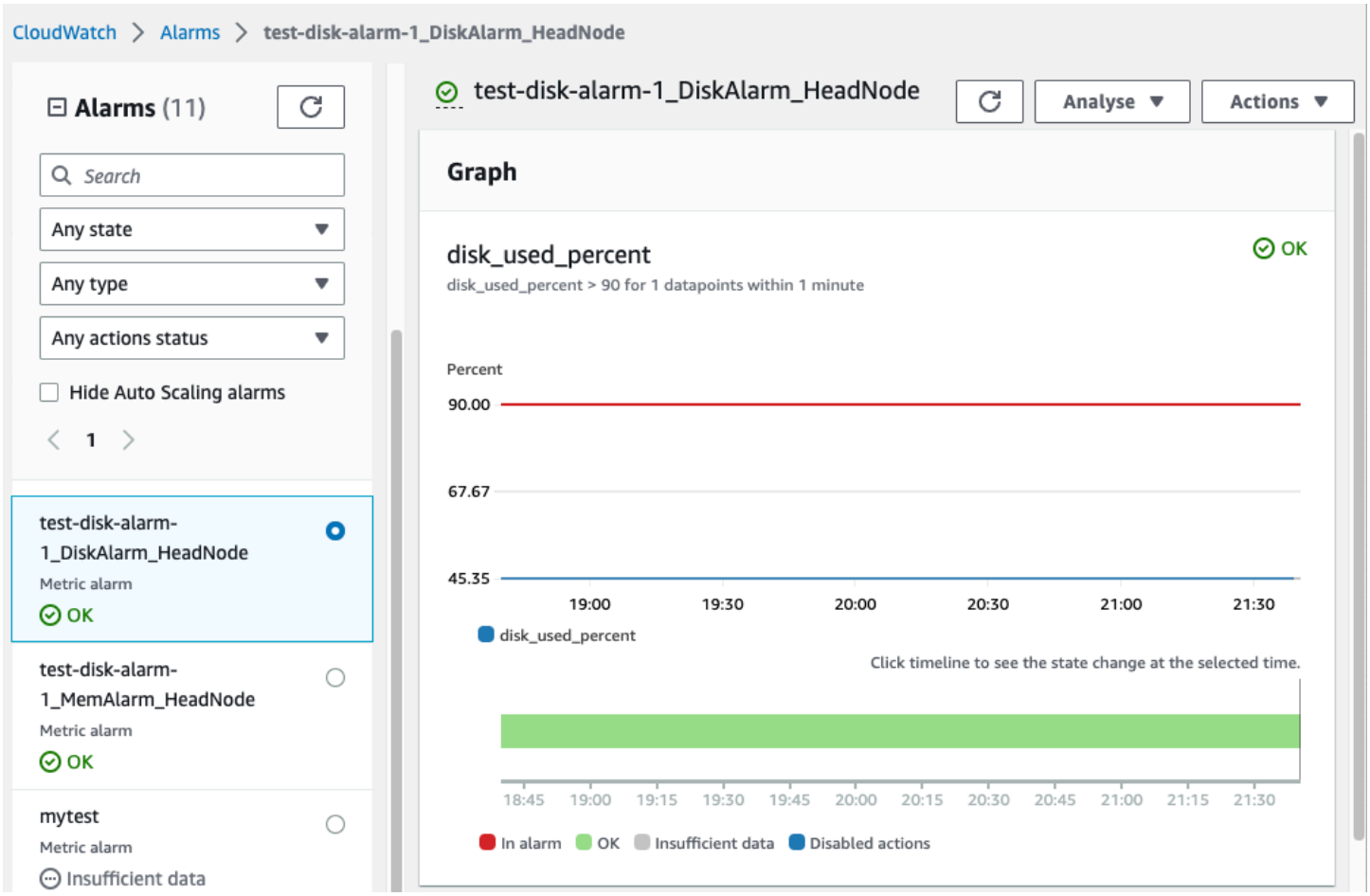
Starting with AWS ParallelCluster version 3.6, you can configure your cluster with Amazon CloudWatch alarms for monitoring the head node. One alarm monitors the root volume `disk_used_percent`. The other alarm monitors the `mem_used_percent` metric. For more information, see Metrics collected by the CloudWatch agent in the *Amazon CloudWatch User Guide*.

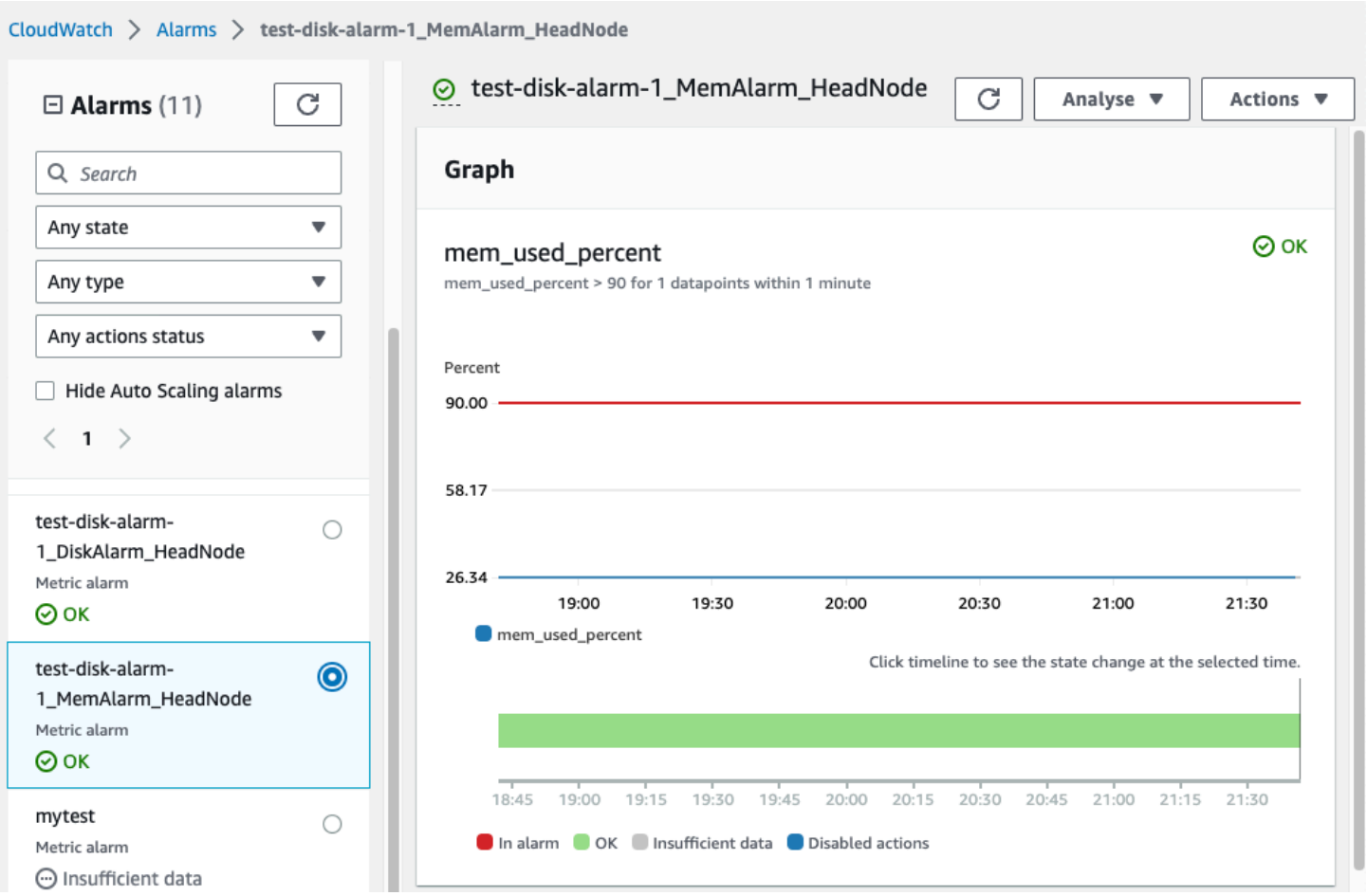
The alarms are named as follows:

- *`cluster-name`*`_DiskAlarm_HeadNode`
- *`cluster-name`*`_MemAlarm_HeadNode`

`cluster-name` is the name of your cluster.

Access the alarms in the CloudWatch console by choosing **Alarms** in the navigation pane. The following images show the disk usage alarm and memory usage alarm for a cluster.


CloudWatch > Alarms > test-disk-alarm-1_DiskAlarm_HeadNode
Alarms (11)
Search
Any state
Any type
Any actions status
Hide Auto Scaling alarms
1
test-disk-alarm-1_DiskAlarm_HeadNode
Metric alarm
OK
test-disk-alarm-1_MemAlarm_HeadNode
Metric alarm
OK
mytest
Metric alarm
Insufficient data
test-disk-alarm-1_DiskAlarm_HeadNode
Analyse
Actions
Graph
disk_used_percent
disk_used_percent > 90 for 1 datapoints within 1 minute
OK
Percent
90.00
67.67
45.35
19:00
19:30
20:00
20:30
21:00
21:30
disk_used_percent
Click timeline to see the state change at the selected time.
18:45
19:00
19:15
19:30
19:45
20:00
20:15
20:30
20:45
21:00
21:15
21:30
In alarm
OK
Insufficient data
Disabled actions



The disk usage alarm is in the `ALARM` state when the disk usage percentage is greater than 90% for 1 data point, within a 1 minute time period.

The memory usage alarm is in the `ALARM` state when the memory usage percentage is greater than 90% for 1 data point, within a 1 minute time period.

> **Note**
>
> AWS ParallelCluster doesn't configure alarm actions by default. For information about how to set up alarm actions, such as sending notifications, see Alarm actions. For more information about Amazon CloudWatch alarms, see Using Amazon CloudWatch alarms in the *Amazon CloudWatch User Guide*.

If you don't want to create these Amazon CloudWatch alarms, deactivate them by setting `Monitoring` / `Dashboards` / `CloudWatch` / `Enabled` to `false` in the cluster configuration.

This also disables the creation of the Amazon CloudWatch dashboard. For more information, see Amazon CloudWatch dashboard.

> **Note**
>
> If you deactivate the creation of the Amazon CloudWatch dashboard, you also deactivate the Amazon CloudWatch `disk_used_percent` and `memory_used_percent` alarms for your cluster.

AWS ParallelCluster configured log rotation

The AWS ParallelCluster log rotation configurations are located in `/etc/logrotate.d/parallelcluster_*_log_rotation` files. When a configured log rotates, the current log content is preserved in a single backup and the emptied log resumes logging.

Only 1 backup is maintained for each configured log.

AWS ParallelCluster configures a fast-growing log to rotate when it reaches 50 MB in size. Fast-growing logs are related to scaling and Slurm, including `/var/log/parallelcluster/clustermgtd`, `/var/log/parallelcluster/slurm_resume.log`, and `/var/log/slurmctld.log`.

AWS ParallelCluster configures a slow-growing log to rotate when it reaches 10 MB in size.

You can view earlier logs that are retained for the number of days defined in the cluster configuration `Logs` / `CloudWatch` / `RetentionInDays` setting with CloudFormation logging enabled. Check the `RetentionInDays` settings to see if the number of days needs to be increased for your use case.

AWS ParallelCluster configures and rotates the following logs:

Head node logs

```
/var/log/cloud-init.log
/var/log/supervisord.log
/var/log/cfn-init.log
/var/log/chef-client.log
/var/log/dcv/server.log
/var/log/dcv/sessionlauncher.log
```

```
/var/log/dcv/agent.*.log
/var/log/dcv/dcv-xsession.*.log
/var/log/dcv/Xdcv.*.log
/var/log/parallelcluster/pam_ssh_key_generator.log
/var/log/parallelcluster/clustermgtd
/var/log/parallelcluster/clusterstatusmgtd
/var/log/parallelcluster/slurm_fleet_status_manager.log
/var/log/parallelcluster/slurm_resume.log
/var/log/parallelcluster/slurm_suspend.log
/var/log/slurmctld.log
/var/log/slurmdbd.log
/var/log/parallelcluster/compute_console_output.log
```

Compute node logs

```
/var/log/cloud-init.log
/var/log/supervisord.log
/var/log/cloud-init-output.log
/var/log/parallelcluster/computemgtd
/var/log/slurmd.log
```

Login node logs

```
/var/log/cloud-init.log
/var/log/cloud-init.log
/var/log/cloud-init-output.log
/var/log/supervisord.log
/var/log/parallelcluster/pam_ssh_key_generator.log
```

`pcluster` CLI logs

The `pcluster` CLI writes logs of your commands to `pcluster.log.#` files in `/home/user/.parallelcluster/`.

For each command, the logs generally include the command with inputs, a copy of the CLI API version used to make the command, the response, and both info and error messages. For a create and build command, the logs also include the configuration file, configuration file validation operations, the CloudFormation template, and stack commands.

You can use these logs to verify errors, inputs, versions and `pcluster` CLI commands. They can also serve as a record of when commands were made.

Amazon EC2 console output logs

When AWS ParallelCluster detects that a static compute node instance terminates unexpectedly, it attempts to retrieve the Amazon EC2 console output from the terminated node instance after a period of time elapses. This way, if the compute node was unable to communicate with Amazon CloudWatch, useful troubleshooting information on why the node terminated might still be retrieved from the console output. This console output is recorded in the `/var/log/parallelcluster/compute_console_output` log on the head node. For more information about the Amazon EC2 console output, see Instance console output in the *Amazon EC2 User Guide for Linux Instances*.

By default, AWS ParallelCluster only retrieves the console output from a sample subset of terminated nodes. This prevents the cluster head node from being overwhelmed with multiple console output requests caused by large numbers of terminations. By default, AWS ParallelCluster waits 5 minutes between termination detection and console output retrieval to give Amazon EC2 time to retrieve the final console output from the nodes.

You can edit the sample size and wait time parameter values in the `/etc/parallelcluster/slurm_plugin/parallelcluster_clustermgtd.conf` file on the head node.

This feature is added in AWS ParallelCluster version 3.5.0.

Amazon EC2 console output parameters

You can edit the values of the following Amazon EC2 console output parameters in the `/etc/parallelcluster/slurm_plugin/parallelcluster_clustermgtd.conf` file on the head node.

`compute_console_logging_enabled`

To disable console output log collection, set `compute_console_logging_enabled` to `false`. The default is `true`.

You can update this parameter at any time, without stopping the compute fleet.

`compute_console_logging_max_sample_size`

`compute_console_logging_max_sample_size` sets the maximum number of compute nodes from which AWS ParallelCluster collects console outputs each time it detects an unexpected termination. If this value is less than 1, AWS ParallelCluster retrieves the console output from all terminated nodes. The default value is 1.

You can update this parameter at any time, without stopping the compute fleet.

`compute_console_wait_time`

`compute_console_wait_time` sets the time, in seconds, that AWS ParallelCluster waits between detecting a node failure and collecting the console output from that node. You can increase the wait time if you determine that Amazon EC2 needs more time to collect the final output from the terminated node. The default value is 300 seconds (5 minutes).

You can update this parameter at any time, without stopping the compute fleet.

Retrieve PCUI and AWS ParallelCluster runtime logs

Learn how to retrieve the PCUI and AWS ParallelCluster runtime logs for troubleshooting. To start, find the relevant PCUI and AWS ParallelCluster stack names. Use the stack name to locate the installation log groups. To finish, export the logs. These logs are specific to the AWS ParallelCluster runtime. For cluster logs, see Retrieving and preserving logs.

Prerequisites

- The AWS CLI is installed.
- You have credentials to run AWS CLI commands on the AWS account that the PCUI is on.
- You can access the Amazon CloudWatch console on the AWS account that the PCUI is on.

Step 1: Locate the stack names for the relevant stacks

In the following example, replace the red highlighted text with your actual values.

List the stacks, using the AWS Region where you installed the PCUI:

```
$ aws cloudformation list-stacks --region aws-region-id
```

Note the stack names for the following stacks:

- The name of the stack that deployed the PCUI on your account. You entered this name when you installed the PCUI; for example, `pcluster-ui`.
- The AWS ParallelCluster stack that is prefixed with the stack name you entered; for example, `pcluster-ui-ParallelClusterApi-ABCD1234EFGH`.

Step 2: Locate the log groups

List the log groups of the PCUI stack, as shown in the following example:

```
$ aws cloudformation describe-stack-resources \
   --region aws-region-id \
   --stack-name pcluster-ui \
   --query "StackResources[?ResourceType == 'AWS::Logs::LogGroup' &&
 (LogicalResourceId == 'ApiGatewayAccessLog' || LogicalResourceId ==
 'ParallelClusterUILambdaLogGroup')].PhysicalResourceId" \
   --output text
```

List the log groups of the AWS ParallelCluster API stack, as shown in the following example:

```
$ aws cloudformation describe-stack-resources \
   --region aws-region-id \
   --stack-name pcluster-ui-ParallelCluster-Api-ABCD1234EFGH \
   --query "StackResources[?ResourceType == 'AWS::Logs::LogGroup' && LogicalResourceId
 == 'ParallelClusterFunctionLogGroup'].PhysicalResourceId" \
   --output text
```

Note the lists of log groups for use in the next step.

Step 3: Export the logs

Use the following steps to gather and export the logs:

1. Log in to the AWS Management Console, and then navigate to the Amazon CloudWatch console on the AWS account that the PCUI is on.
2. Choose **Logs**, **Logs Insights** in the navigation pane.
3. Select all of the log groups listed in the previous step.
4. Choose a time range, such as 12 hours.
5. Run the following query:

   ```
   $ fields @timestamp, @message
   | sort @timestamp desc
   | limit 10000
   ```

6. Choose **Export results**, **Download table (JSON)**.

Retrieving and preserving logs

AWS ParallelCluster creates Amazon EC2 metrics for HeadNode and Compute instances and storage. You can view the metrics in the CloudWatch console **Custom Dashboards**. AWS ParallelCluster also creates cluster CloudWatch log streams in log groups. You can view these logs in the CloudWatch console **Custom Dashboards** or **Log groups**. The Monitoring cluster configuration section describes how you can modify the cluster CloudWatch logs and dashboard. For more information, see Integration with Amazon CloudWatch Logs and Amazon CloudWatch dashboard.

Logs are a useful resource for troubleshooting issues. For example, if you want to delete a failing cluster, it might be useful to first create an archive of the cluster logs. Follow the steps in Archive logs to create an archive.

Topics

Cluster logs unavailable in CloudWatch

If cluster logs aren't available in CloudWatch, check to make sure you haven't overwritten the AWS ParallelCluster CloudWatch log configuration when adding custom logs to the configuration.

To add custom logs to the CloudWatch configuration, make sure you append to the configuration rather than fetch and overwrite it. For more information on `fetch-config` and `append-config`, see Multiple CloudWatch agent configuration files in the *CloudWatch User Guide*.

To restore the AWS ParallelCluster CloudWatch log configuration, you can run the following commands inside an AWS ParallelCluster node:

```
$ PLATFORM="$(ohai platform | jq -r ".[]")"
LOG_GROUP_NAME="$(cat /etc/chef/dna.json | jq -r ".cluster.log_group_name")"
SCHEDULER="$(cat /etc/chef/dna.json | jq -r ".cluster.scheduler")"
NODE_ROLE="$(cat /etc/chef/dna.json | jq -r ".cluster.node_type")"
CONFIG_DATA_PATH="/usr/local/etc/cloudwatch_agent_config.json"
```

```
/opt/parallelcluster/pyenv/versions/cookbook_virtualenv/bin/python /usr/local/bin/
write_cloudwatch_agent_json.py --platform $PLATFORM --config $CONFIG_DATA_PATH --log-
group $LOG_GROUP_NAME --scheduler $SCHEDULER --node-role $NODE_ROLE
/opt/aws/amazon-cloudwatch-agent/bin/amazon-cloudwatch-agent-ctl -a fetch-config -m ec2
 -c file:/opt/aws/amazon-cloudwatch-agent/etc/amazon-cloudwatch-agent.json -s
```

Archive logs

You can archive the logs in Amazon S3 or in a local file (depending on the `--output-file` parameter).

> **Note**
>
> Starting with AWS ParallelCluster 3.12.0, you can export logs to the default AWS ParallelCluster bucket. In this case you don't need to set bucket permissions.

> **Note**
>
> Add permissions to the Amazon S3 bucket policy to grant CloudWatch access. For more information, see Set permissions on an Amazon S3 bucket in the *CloudWatch Logs User Guide*.

```
$ pcluster export-cluster-logs --cluster-name mycluster --region eu-west-1 \
  --bucket bucketname --bucket-prefix logs
{
  "url": "https://bucketname.s3.eu-west-1.amazonaws.com/export-log/mycluster-
logs-202109071136.tar.gz?..."
}

# use the --output-file parameter to save the logs locally
$ pcluster export-cluster-logs --cluster-name mycluster --region eu-west-1 \
  --bucket bucketname --bucket-prefix logs --output-file /tmp/archive.tar.gz
{
  "path": "/tmp/archive.tar.gz"
}
```

The archive contains the Amazon CloudWatch Logs streams and AWS CloudFormation stack events from the head node and compute nodes for the last 14 days, unless specified explicitly in the

configuration or in the parameters for the `export-cluster-logs` command. The time it takes for the command to finish depends on the number of nodes in the cluster and the number of log streams available in CloudWatch Logs. For more information about the available log streams, see Integration with Amazon CloudWatch Logs.

Preserved logs

Starting from version 3.0.0, AWS ParallelCluster preserves CloudWatch Logs by default when a cluster is deleted. If you want to delete a cluster and preserve its logs, make sure that `Monitoring` / `Logs` / `CloudWatch` / `DeletionPolicy` isn't set to `Delete` in the cluster configuration. Otherwise, change the value for this field to `Retain`, and run the `pcluster update-cluster` command. Then, run `pcluster delete-cluster --cluster-name` *`<cluster_name>`* to delete the cluster, but retain the log group that's stored in Amazon CloudWatch.

Terminated node logs

If a static compute node unexpectedly terminates and CloudWatch has no logs for it, check if AWS ParallelCluster has recorded the console output for that compute node on the head node in the `/var/log/parallelcluster/compute_console_output` log. For more information, see Key logs for debugging.

If the `/var/log/parallelcluster/compute_console_output` log isn't available or doesn't contain the output for the node, use the AWS CLI to retrieve the console output from the failed node. Log in to the cluster head node and get the failed node `instance-id` from the `/var/log/parallelcluster/slurm_resume.log` file.

Retrieve the console output by using the following command with the `instance-id`:

```
$ aws ec2 get-console-output --instance-id i-abcdef01234567890
```

If a dynamic compute node self-terminates after launching and CloudWatch has no logs for it, submit a job that activates a cluster scaling action. Wait for the instance to fail and retrieve the instance console log.

Log in to the cluster head node and get the compute node `instance-id` from the `/var/log/parallelcluster/slurm_resume.log` file.

To retrieve the instance console log, use the following command:

```
$ aws ec2 get-console-output --instance-id i-abcdef01234567890
```

The console output log can help you debug the root cause of a compute node failure when the compute node log isn't available.

AWS CloudFormation custom resource

Starting with AWS ParallelCluster version 3.6.0, you can use an AWS ParallelCluster CloudFormation custom resource in an AWS CloudFormation stack. The custom resource is an AWS ParallelCluster hosted stack. This way, you can use CloudFormation to configure and manage your clusters. For example, you can configure cluster external resources such as network, shared storage, and security group infrastructure in a CloudFormation stack. Furthermore, you can manage your cluster with a CloudFormation infrastructure as code pipeline.

Add an AWS ParallelCluster custom resource to your CloudFormation template by doing the following:

1. Add a custom resource provider stack that is owned and hosted by AWS ParallelCluster.
2. Reference the provider stack in your CloudFormation template as a custom resource.

The custom resource provider stack handles and responds to CloudFormation requests. For example, when you deploy your CloudFormation stack, you also configure and create a cluster. To update a cluster, you update your CloudFormation stack. You delete a cluster when you delete your stack. For more information about CloudFormation custom resources, see Custom resources in the *AWS CloudFormation User Guide*.

Warning

CloudFormation doesn't detect custom resource drift. Only use CloudFormation to update the cluster configuration and to delete a cluster.
You can use the `pcluster` CLI or the AWS ParallelCluster UI to monitor the state of the cluster or to update the compute fleet, but you must not use them to update the cluster configuration or to delete the cluster.

Note

We recommend that you add termination protection to your stack to avoid accidental removal.

Provider stack hosted by AWS ParallelCluster

The custom resource provider stack is formatted as shown in the following CloudFormation template snippet:

```
PclusterClusterProvider:
  Type: AWS::CloudFormation::Stack
  Properties:
    Parameters:
      CustomLambdaRole: # (Optional) RoleARN to override default
      AdditionalIamPolicies: # (Optional) comma-separated list of IAM policies to add
    TemplateURL: !Sub
      - https://${AWS::Region}-aws-parallelcluster.s3.${AWS::Region}.${AWS::URLSuffix}/
parallelcluster/${Version}/templates/custom_resource/cluster.yaml
      - { Version: 3.13.2 }
```

Properties:

Parameters:

CustomLambdaRole (optional):

A custom role with permissions to run the AWS Lambda that creates and manages the cluster. By default, the role uses the same policies defined by default in the AWS ParallelCluster documentation.

AdditionalIamPolicies (optional):

A comma-separated list of additional IAM Policy Amazon Resource Names (ARNs) to add to the role that the Lambda uses. This is only used if a `CustomLambdaRole` isn't specified and can be kept blank.

If you need additional policies for the head node, compute nodes, or for access to an Amazon S3 bucket, add them to the `CustomLambdaRole` or `AdditionalIamPolicy` property.

If you need to attach additional policies to the head node, you must also grant the necessary permissions to attach or detach those policies to the IAM role associated with the head node. Specifically, you'll need to attach the "iam:AttachRolePolicy" and "iam:DetachRolePolicy" permissions (or their equivalent in a managed policy) to the IAM role used by the head node. For more information, see AWS ParallelCluster user example policies for managing IAM resources.

For more information about the default policies, see AWS Identity and Access Management permissions in AWS ParallelCluster.

TemplateURL (required):

The AWS ParallelCluster custom resource file URL.

Outputs:

ServiceToken:

A value that can be used as a custom resource `ServiceToken` property. A custom resource `ServiceToken` specifies where AWS CloudFormation sends requests. This is a required input for a cluster resource that you include in your AWS CloudFormation template.

LogGroupArn:

The ARN of the CloudWatch `LogGroup` that the underlying resource logs to.

LambdaLayerArn:

The ARN of the Lambda layer that's used for running AWS ParallelCluster operations.

Cluster resource

The CloudFormation cluster resource is formatted as shown in the following CloudFormation template snippet:

```
PclusterCluster:
  Type: Custom::PclusterCluster
  Properties:
    ServiceToken: !GetAtt [ PclusterClusterProvider , Outputs.ServiceToken ]
    ClusterName: !Sub 'c-${AWS::StackName}' # Must be different from StackName
    ClusterConfiguration:
    # Your Cluster Configuration
```

Properties:

ServiceToken:

The AWS ParallelCluster provider stack `ServiceToken` output.

ClusterName:

The name of the cluster to be created and managed. The name must not match the CloudFormation stack's name. The name can't be changed after the cluster has been created.

ClusterConfiguration:

The cluster configuration YAML file, as described in Cluster configuration file. However, you can use the usual CloudFormation constructs, such as Intrinsic functions.

DeletionPolicy:

Defines whether to delete the cluster when the root stack is deleted. The default is `Delete`.

Retain:

Retain the cluster if the custom resource is deleted.

> **Note**
>
> To keep the retained cluster functioning, cluster-dependent resources, such as storage and networking, must have a deletion policy set to retain.

Delete:

Delete the cluster if the custom resource is deleted.

`Fn::GetAtt` return values:

The `Fn::GetAtt` intrinsic function returns a value for a specified attribute of a type. For more information about using the `Fn::GetAtt intrinsic` function, see Fn::GetAtt.

ClusterProperties:

The values from the `pcluster describe-cluster` operation.

validationMessages:

A string containing all the validation messages that occurred during the last create or update operation.

logGroupName:

The name of the log group that's used for logging Lambda cluster operations. The log events are retained for 90 days and the log group is retained after cluster deletion.

Example: `Fn::GetAtt`:

```
# Provide the public IP address of the head node as an output of a stack
Outputs:
  HeadNodeIp:
    Description: The public IP address of the head node
    Value: !GetAtt [ PclusterCluster, headNode.publicIpAddress ]
```

Example: Simple, complete CloudFormation template with an AWS ParallelCluster custom resource:

```
AWSTemplateFormatVersion: '2010-09-09'
Description: >
 AWS ParallelCluster CloudFormation Template

Parameters:
  HeadNodeSubnet:
    Description: Subnet where the HeadNode will run
    Type: AWS::EC2::Subnet::Id

  ComputeSubnet:
    Description: Subnet where the Compute Nodes will run
    Type: AWS::EC2::Subnet::Id

  KeyName:
    Description: KeyPair to login to the head node
    Type: AWS::EC2::KeyPair::KeyName

Resources:
  PclusterClusterProvider:
    Type: AWS::CloudFormation::Stack
    Properties:
      TemplateURL: !Sub
        - https://${AWS::Region}-aws-parallelcluster.s3.${AWS::Region}.
${AWS::URLSuffix}/parallelcluster/${Version}/templates/custom_resource/cluster.yaml
        - { Version: 3.13.2 }
```

```
  PclusterCluster:
    Type: Custom::PclusterCluster
    Properties:
      ServiceToken: !GetAtt [ PclusterClusterProvider , Outputs.ServiceToken ]
      ClusterName: !Sub 'c-${AWS::StackName}'
      ClusterConfiguration:
        Image:
          Os: alinux2
        HeadNode:
          InstanceType: t2.medium
          Networking:
            SubnetId: !Ref HeadNodeSubnet
          Ssh:
            KeyName: !Ref KeyName
        Scheduling:
          Scheduler: slurm
          SlurmQueues:
          - Name: queue0
            ComputeResources:
            - Name: queue0-cr0
              InstanceType: t2.micro
            Networking:
              SubnetIds:
              - !Ref ComputeSubnet

Outputs:
  HeadNodeIp:
    Description: The Public IP address of the HeadNode
    Value: !GetAtt [ PclusterCluster, headNode.publicIpAddress ]
  ValidationMessages:
    Description: Any warnings from cluster create or update operations.
    Value: !GetAtt PclusterCluster.validationMessages
```

To learn more about how to use the CloudFormation AWS ParallelCluster custom resource, see Creating a cluster with AWS CloudFormation.

Cluster operations

When a cluster custom resource is added to a CloudFormation stack, CloudFormation can perform the following cluster operations:

- CloudFormation creates a cluster in a new separate stack when it deploys a stack that includes the AWS ParallelCluster custom resource.

- If you update the cluster configuration defined in the stack, according to configuration update policies, CloudFormation updates the cluster. The AWS ParallelCluster custom resource provider doesn't stop the compute fleet before updating the cluster. We recommend that you use the `QueueUpdateStrategy` setting for cluster updates. This way, you can avoid making explicit `pcluster update-compute-fleet` calls before and after updates when using the AWS ParallelCluster custom resource.
- If you delete the stack, the cluster is deleted.

Troubleshooting stacks that include the AWS ParallelCluster custom resource

With an AWS ParallelCluster custom resource, CloudFormation deploys a cluster from a new, separate stack. You can monitor cluster creation by taking the following steps:

1. Navigate to CloudFormation in the AWS Management Console and choose **Stacks** in the navigation pane.
2. Choose the stack with the name that you defined for the cluster name.
3. If the stack state is `ROLLBACK_COMPLETE`, an error occurred during cluster creation.
4. Choose **Stack details**, and choose the **Events** tab.
5. Search **Events** on **Logical ID** for the name that you defined for the cluster name. It has a `Status reason` that gives a reason for an issue.
6. You can also choose the **Stacks** drop down menu, and then **Deleted** to see the list of deleted stacks. Select the stack with the cluster name and view **Events** for more details.
7. To view the output from the custom resource provider that manages the cluster, select the stack with the **Description** "AWS ParallelCluster Cluster Custom Resource." Choose the **Resources** tab, find the resource with **Logical ID** `PclusterCfnFunctionLogGroup`, and follow the given link. View the log streams that show the Lambda debug output.
8. To troubleshoot the cluster, see AWS ParallelCluster troubleshooting.

Elastic Fabric Adapter

Elastic Fabric Adapter (EFA) is a network device that has OS-bypass capabilities for low-latency network communications with other instances on the same subnet. EFA is exposed by using Libfabric, and can be used by applications using the Messaging Passing Interface (MPI).

To use EFA with AWS ParallelCluster and a Slurm scheduler, set `SlurmQueues` / `ComputeResources` / `Efa` / `Enabled` to `true`.

To view the list of Amazon EC2 instances that support EFA, see Supported instance types in the *Amazon EC2 User Guide for Linux Instances*.

We recommend that you run your EFA-enabled instances in a placement group. This way the instances are launched into a low-latency group in a single Availability Zone. For more information on how to configure placement groups with AWS ParallelCluster, see `SlurmQueues` / `Networking` / `PlacementGroup`.

For more information, see Elastic Fabric Adapter in the *Amazon EC2 User Guide* and Scale HPC workloads with elastic fabric adapter and AWS ParallelCluster in the *AWS Open Source Blog*.

Note

Elastic Fabric Adapter (EFA) isn't supported over different availability zones. For more information, see Scheduling / SlurmQueues / Networking / SubnetIds.

Note

By default, Ubuntu distributions enable ptrace (process trace) protection. ptrace protection is disabled so that Libfabric works properly. For more information, see Disable ptrace protection in the *Amazon EC2 User Guide*.

Enable Intel MPI

Intel MPI is available on the AWS ParallelCluster AMIs.

Note

To use Intel MPI, you must acknowledge and accept the terms of the Intel simplified software license.

By default, Open MPI is placed on the path. To enable Intel MPI instead of Open MPI, you must first load the Intel MPI module. Then, you need to install the latest version by using `module load`

`intelmpi`. The exact name of the module changes with every update. To see which modules are available, run `module avail`. The output is as follows.

```
$ module avail
----------------------------/usr/share/Modules/modulefiles
 --------------------------------
dot                             modules
libfabric-aws/1.16.0~amzn3.0    null
module-git                      openmpi/4.1.4
module-info                     use.own

----------------------------/opt/intel/mpi/2021.6.0/modulefiles
 ---------------------------
intelmpi
```

To load a module, run `module load` *`modulename`*. You can add this to the script used to run `mpirun`.

```
$ module load intelmpi
```

To see which modules are loaded, run `module list`.

```
$ module list
Currently Loaded Modulefiles:
  1) intelmpi
```

To verify that Intel MPI is enabled, run `mpirun --version`.

```
$ mpirun --version
Intel(R) MPI Library for Linux* OS, Version 2021.6 Build 20220227 (id: 28877f3f32)
Copyright 2003-2022, Intel Corporation.
```

After the Intel MPI module has been loaded, multiple paths are changed to use the Intel MPI tools. To run code that was compiled by the Intel MPI tools, load the Intel MPI module first.

> **Note**
>
> Intel MPI isn't compatible with AWS Graviton-based instances.

> **Note**
>
> Before AWS ParallelCluster version 2.5.0, Intel MPI wasn't available on the AWS ParallelCluster AMIs in the China (Beijing) and China (Ningxia) Regions.

AWS ParallelCluster API

What is AWS ParallelCluster API?

AWS ParallelCluster API is a serverless application that, once deployed to your AWS account, provides programmatic access to AWS ParallelCluster features through an API.

AWS ParallelCluster API is distributed as a self-contained AWS CloudFormation template that includes an Amazon API Gateway endpoint, that exposes AWS ParallelCluster features, and an AWS Lambda function, that takes care of processing the invoked features.

The following image shows a high-level architecture diagram of the AWS ParallelCluster API infrastructure.

AWS ParallelCluster API Documentation

The OpenAPI specification file that describes the AWS ParallelCluster API can be downloaded from:

```
https://<REGION>-aws-parallelcluster.s3.<REGION>.amazonaws.com/
parallelcluster/<VERSION>/api/ParallelCluster.openapi.yaml
```

Starting from the OpenAPI specification file, you can use one of the many available tools such as Swagger UI or Redoc to generate documentation for the AWS ParallelCluster API.

How to deploy AWS ParallelCluster API

To deploy AWS ParallelCluster API you need to be an Administrator of the AWS account.

The template used to deploy the API is available at the following URL:

```
https://<REGION>-aws-parallelcluster.s3.<REGION>.amazonaws.com/
parallelcluster/<VERSION>/api/parallelcluster-api.yaml
```

where *<REGION>* is the AWS Region where the API needs to be deployed to and *<VERSION>* is the AWS ParallelCluster version (e.g. 3.13.2).

AWS Lambda uses a Lambda layer interface with the AWS ParallelCluster Python library API to process the API invoked features.

> **Warning**
>
> Any user in the AWS account, that has privileged access to AWS Lambda or Amazon API Gateway services, automatically inherits permissions to administer AWS ParallelCluster API resources.

Deploy the AWS ParallelCluster API with AWS CLI

In this section, you will learn how to deploy with AWS CLI.

Configure AWS Credentials to be used with the CLI if you haven't already done so.

```
$ aws configure
```

Run the following commands to deploy the API:

```
$ REGION=<region>
$ API_STACK_NAME=<stack-name>  # This can be any name
$ VERSION=3.13.2
$ aws cloudformation create-stack \
    --region ${REGION} \
    --stack-name ${API_STACK_NAME} \
    --template-url https://${REGION}-aws-parallelcluster.s3.${REGION}.amazonaws.com/
parallelcluster/${VERSION}/api/parallelcluster-api.yaml \
    --capabilities CAPABILITY_NAMED_IAM CAPABILITY_AUTO_EXPAND
$ aws cloudformation wait stack-create-complete --stack-name ${API_STACK_NAME} --region
 ${REGION}
```

Customize your deployment

You can use the AWS CloudFormation parameters exposed by the template to customize the API deployment. To configure the value of a parameter when you deploy through the CLI, the following option can be used: `--parameters ParameterKey=KeyName,ParameterValue=Value`.

The following parameters are optional:

- **Region** - Use the `Region` parameter to specify whether the API is able to control resources in all AWS Regions (default) or in a single AWS Region. Set this value to the AWS Region the API is being deployed to in order to restrict access.
- **ParallelClusterFunctionRole** - This overrides the IAM role that gets assigned to the AWS Lambda function that implements AWS ParallelCluster features. The parameter accepts the ARN of an IAM role. This role needs to be configured to have AWS Lambda as the IAM principal. Also, since this role will replace the default role of the API Lambda function, it must have at least the default permissions required by the API as listed in AWS ParallelCluster example `pcluster` user policies.
- **ParallelClusterFunctionAdditionalPolicies** - ARN of the additional IAM policy to be attached to the AWS ParallelCluster API function role. Only one policy can be specified.
- **CustomDomainName, CustomDomainCertificate, CustomDomainHostedZoneId** - Use these parameters to set a custom domain for the Amazon API Gateway endpoint. `CustomDomainName` is the name of the domain to use, `CustomDomainCertificate` is the ARN of an AWS managed certificate for this domain name and `CustomDomainHostedZoneId` is the ID of the Amazon Route 53 hosted zone that you want to create records in.

  > **Warning**
  >
  > You can configure custom domain settings to enforce a minimum version of Transport Layer Security (TLS) for the API. For more information, see Choosing a minimum TLS version for a custom domain in API Gateway.

- **EnableIamAdminAccess** - By default the AWS Lambda function that processes AWS ParallelCluster API operations is configured with an IAM role that prevents any privileged IAM access (`EnableIamAdminAccess=false`). This makes the API unable to process operations that require the creation of IAM roles or policies. Because of this, the creation of clusters or custom images is successful only when IAM roles are provided as input as part of the resource configuration.

  When `EnableIamAdminAccess` is set to `true` the AWS ParallelCluster API is granted permissions to manage the creation of IAM roles required to deploy clusters or generate custom AMIs.

> **Warning**
>
> When this is set to true it grants IAM admin privileges to the AWS Lambda function that processes AWS ParallelCluster operations.

Refer to AWS ParallelCluster user example policies for managing IAM resources for additional details on the features that can be unlocked when you enable this mode.

- **PermissionsBoundaryPolicy** - This optional parameter accepts an existing IAM policy ARN that will be set as permissions boundary for all the IAM roles created by the PC API infrastructure and as a condition on the administrative IAM permissions so that only roles with this policy can be created by the PC API.

  Refer to `PermissionsBoundary` mode for additional details on the restrictions imposed by this mode.

- **CreateApiUserRole** - By default, the deployment of the AWS ParallelCluster API includes the creation of an IAM role which is set as the only role authorized to invoke the API. The Amazon API Gateway endpoint is configured with a resource based policy to grant invoke permission to the created user only. To change this, set `CreateApiUserRole=false` and then grant API access to selected IAM users. For more information, see Control access for invoking an API in the *Amazon API Gateway Developer Guide*.

  > **Warning**
  >
  > When `CreateApiUserRole=true` access to the API endpoint is not restricted by Amazon API Gateway resource policies, all IAM roles that have unconstrained `execute-api:Invoke` permission can access AWS ParallelCluster features. For more information, see Controlling access to an API with API Gateway resource policies in the *API Gateway Developer Guide*.

> **Warning**
>
> The `ParallelClusterApiUserRole` has permission to invoke all AWS ParallelCluster API operations. To restrict access to a subset of API resources, see the Control who can call an API Gateway API method with IAM policies in the *API Gateway Developer Guide*.

- **IAMRoleAndPolicyPrefix** - This optional parameter accepts a string containing a maximum of 10 characters that will be used as the prefix for both IAM roles and policies created as part of the PC API infrastructure.

Updating the API

In this section, you will learn how to use one of the two available options to update the API.

Upgrading to a newer AWS ParallelCluster version

Option 1: To remove the existing API, delete the corresponding AWS CloudFormation stack and deploy the new API as shown above.

Option 2: To update the existing API, run the following commands:

```
$ REGION=<region>
$ API_STACK_NAME=<stack-name>  # This needs to correspond to the existing API stack
 name
$ VERSION=3.13.2
$ aws cloudformation update-stack \
    --region ${REGION} \
    --stack-name ${API_STACK_NAME} \
    --template-url https://${REGION}-aws-parallelcluster.s3.${REGION}.amazonaws.com/
parallelcluster/${VERSION}/api/parallelcluster-api.yaml \
    --capabilities CAPABILITY_NAMED_IAM CAPABILITY_AUTO_EXPAND
$ aws cloudformation wait stack-update-complete --stack-name ${API_STACK_NAME} --region
 ${REGION}
```

Invoking AWS ParallelCluster API

The AWS ParallelCluster Amazon API Gateway endpoint is configured with `AWS_IAM` authorization type, and requires all requests to be SigV4 signed with valid IAM credentials (API reference: making http requests).

When deployed with default settings, API invoke permissions are only granted to the default IAM user created with the API.

To retrieve the ARN of the default IAM user, run:

```
$ REGION=<region>
$ API_STACK_NAME=<stack-name>
$ aws cloudformation describe-stacks --region ${REGION} --stack-name ${API_STACK_NAME}
 --query "Stacks[0].Outputs[?OutputKey=='ParallelClusterApiUserRole'].OutputValue" --
output text
```

To obtain temporary credentials for the default IAM user, run the STS AssumeRole command.

To retrieve the AWS ParallelCluster API endpoint run the following command:

```
$ REGION=<region>
$ API_STACK_NAME=<stack-name>
$ aws cloudformation describe-stacks --region ${REGION} --stack-name ${API_STACK_NAME}
 --query "Stacks[0].Outputs[?OutputKey=='ParallelClusterApiInvokeUrl'].OutputValue" --
output text
```

The AWS ParallelCluster API can be invoked by any HTTP client that complies with the OpenAPI specifications that can be found here:

```
https://<REGION>-aws-parallelcluster.s3.<REGION>.amazonaws.com/
parallelcluster/<VERSION>/api/ParallelCluster.openapi.yaml
```

Requests need to be SigV4 signed as documented here.

At this time, we do not offer any official API client implementation. However, you can use the OpenAPI Generator to easily generate API clients from the OpenAPI model. Once the client is generated, SigV4 signing needs to be added if not provided out of the box.

A reference implementation for a Python API client can be found in the AWS ParallelCluster repository. To learn more about how you can use the Python API client, see the Using the AWS ParallelCluster API tutorial.

To implement more advanced access control mechanisms, such as Amazon Cognito or Lambda Authorizers, or to further protect the API with AWS WAF or API keys, follow the Amazon API Gateway documentation.

Warning

An IAM user that is authorized to invoke the AWS ParallelCluster API, can indirectly control all AWS resources managed by AWS ParallelCluster in the AWS account. This includes the creation of AWS resources that the user can't control directly due to restrictions on the user IAM policy. For example, the creation of a AWS ParallelCluster cluster, depending on its configuration, might include the deployment of Amazon EC2 instances, Amazon Route 53, Amazon Elastic File System file systems, Amazon FSx file systems, IAM roles, and resources from other AWS services used by AWS ParallelCluster that the user might not have direct control over.

Warning

When you create a cluster with `AdditionalIamPolicies` specified in the configuration, the additional policies must match one of the following patterns:

```
- !Sub arn:${AWS::Partition}:iam::${AWS::AccountId}:policy/parallelcluster*
- !Sub arn:${AWS::Partition}:iam::${AWS::AccountId}:policy/parallelcluster/*
- !Sub arn:${AWS::Partition}:iam::aws:policy/CloudWatchAgentServerPolicy
- !Sub arn:${AWS::Partition}:iam::aws:policy/AmazonSSMManagedInstanceCore
- !Sub arn:${AWS::Partition}:iam::aws:policy/AWSBatchFullAccess
- !Sub arn:${AWS::Partition}:iam::aws:policy/AmazonS3ReadOnlyAccess
- !Sub arn:${AWS::Partition}:iam::aws:policy/service-role/AWSBatchServiceRole
- !Sub arn:${AWS::Partition}:iam::aws:policy/service-role/
AmazonEC2ContainerServiceforEC2Role
- !Sub arn:${AWS::Partition}:iam::aws:policy/service-role/
AmazonECSTaskExecutionRolePolicy
- !Sub arn:${AWS::Partition}:iam::aws:policy/service-role/
AmazonEC2SpotFleetTaggingRole
- !Sub arn:${AWS::Partition}:iam::aws:policy/EC2InstanceProfileForImageBuilder
- !Sub arn:${AWS::Partition}:iam::aws:policy/service-role/
AWSLambdaBasicExecutionRole
```

If you need other additional policies, you can do one of the following:

- Edit the `DefaultParallelClusterIamAdminPolicy` in:

```
https://<REGION>-aws-parallelcluster.s3.<REGION>.amazonaws.com/
parallelcluster/<VERSION>/api/parallelcluster-api.yaml
```

Add the policy in the `ArnLike/iam:PolicyARN` section.

- Don't specify policies for `AdditionalIamPolicies` in the configuration file and manually add policies to the AWS ParallelCluster Instance Role created within the cluster.

Accessing the API logs and metrics

API logs are published to Amazon CloudWatch with a retention of 30 days. To retrieve the LogGroup name associated with an API deployment, run the following command:

```
$ REGION=<region>
$ API_STACK_NAME=<stack-name>
$ aws cloudformation describe-stacks --region ${REGION} --
stack-name ${API_STACK_NAME} --query "Stacks[0].Outputs[?
OutputKey=='ParallelClusterLambdaLogGroup'].OutputValue" --output text
```

Lambda metrics, logs and AWS X-Ray trace logs can be also accessed through the Lambda console. To retrieve the ARN of the Lambda function associated with an API deployment run the following command:

```
$ REGION=<region>
$ API_STACK_NAME=<stack-name>
$ aws cloudformation describe-stacks --region ${REGION} --stack-name ${API_STACK_NAME}
 --query "Stacks[0].Outputs[?OutputKey=='ParallelClusterLambdaArn'].OutputValue" --
output text
```

AWS ParallelCluster for Terraform

Beginning with AWS ParallelCluster 3.8.0, you can deploy clusters and custom images using Terraform. To begin using this feature, see Terraform Provider for AWS ParallelCluster from the Terraform Registry.

Note

You must have ParallelCluster API deployed in your account to use the provider.

Use the following chart to determine the compatibility between the provider and the AWS ParallelCluster versions:

Provider version	AWS ParallelCluster version
1.0.0	3.8.0-3.10.1
1.1.0	3.11.0+

See examples of how to use the provider.

For an even smoother experience, use the official Terraform Module for AWS ParallelCluster from Terraform Registry. The module allows you to deploy:

1. ParallelCluster API
2. ParallelCluster clusters defined with YAML configuration file and HCL
3. Networking infrastructure required by a ParallelCluster cluster

See examples of how to use the module.

Connect to the head and login nodes through Amazon DCV

Amazon DCV is a remote visualization technology that enables users to securely connect to graphic intensive 3D applications that are hosted on a remote high-performance server. For more information, see Amazon DCV.

The Amazon DCV software is automatically installed on the head node and can be enabled by using the `Dcv` section from the `HeadNode`.

Starting from AWS ParallelCluster version 3.11, Amazon DCV can be enabled for a login node pool by using the Dcv section from the `LoginNodes` section / Pools configuration.

```
LoginNodes:
```

```
  Pools:
    Dcv:
      Enabled: true
```

AWS ParallelCluster sets /home/*<DEFAULT_AMI_USER>* as the DCV server storage folder. For more information about Amazon DCV configuration parameters, see `HeadNode` / `Dcv`.

To connect to the Amazon DCV session on the head node, use the `dcv-connect` command. To connect on a login node, use `dcv-connect` with the `--login-node-ip` parameter and pass in the public or private IP address of the login node you wish to connect to.

Amazon DCV HTTPS certificate

Amazon DCV automatically generates a self-signed certificate to secure traffic between the Amazon DCV client and Amazon DCV server.

To replace the default self-signed Amazon DCV certificate with another certificate, first connect to the head node. Then, copy both the certificate and key to the `/etc/dcv` folder before running the `pcluster dcv-connect` command.

For more information, see Changing the TLS certificate in the *Amazon DCV Administrator Guide*.

Licensing Amazon DCV

The Amazon DCV server doesn't require a license server when running on Amazon EC2 instances. However, the Amazon DCV server must periodically connect to an Amazon S3 bucket to determine if a valid license is available.

AWS ParallelCluster automatically adds the required permissions to the head node IAM policy. When using a custom IAM Instance Policy, use the permissions described in Amazon DCV on Amazon EC2 in the *Amazon DCV Administrator Guide*.

For troubleshooting tips, see Troubleshooting issues in Amazon DCV.

Using `pcluster update-cluster`

In AWS ParallelCluster 3.x, `pcluster update-cluster` analyzes the settings used to create the current cluster and the settings in the configuration file for issues. If any issues are discovered, they are reported, and the steps to take to fix the issues are displayed. For example, if the compute `InstanceType` is changed, the compute fleet must be stopped before an update can proceed.

This issue is reported when it is discovered. If no blocking issues are discovered, update process is started and the changes are reported.

You can use the `pcluster update-cluster --dryrun option` to see the changes before their run. For more information, see `pcluster update-cluster` examples.

For troubleshooting guidance, see AWS ParallelCluster troubleshooting.

Update Policy: definitions

Update policy: The login nodes in the cluster must be stopped for this setting to be changed for an update.

You can't change these settings while login nodes in the cluster are in use. Either you must revert the change, or you must stop the clusters login nodes. (You can stop the login nodes in the cluster by setting each pool's count equal to 0). After the cluster's login nodes are stopped, you can update the cluster (pcluster update-cluster) to activate the changes.

> **Note**
>
> This update policy is supported starting with AWS ParallelCluster version 3.7.0.

Update policy: Login node pools can be added, but removing a pool requires all login nodes in the cluster are stopped.

To remove a pool, you must stop all login nodes in the cluster. (You can stop login nodes in the cluster by setting each pool's Count equal to 0). After the cluster's login nodes have been stopped, you can update the cluster (`pcluster update-cluster`) to activate the changes.

> **Note**
>
> This update policy is supported starting with AWS ParallelCluster version 3.11.0.

Update policy: The login nodes in the pool must be stopped for this setting to be changed for an update.

You can't change these settings while login nodes in the pool are in use. Either you must revert the change, or you must stop the pool's login nodes. (You can stop the login nodes in the pool

by setting the pool's Count equal to 0). After the pool's login nodes have been stopped, you can update the cluster (`pcluster update-cluster`) to activate the changes.

> **Note**
>
> This update policy is supported starting with AWS ParallelCluster version 3.11.0.

Update policy: This setting can be changed during an update.

After changing this setting, the cluster can be updated using `pcluster update-cluster`.

Update policy: If this setting is changed, the update is not allowed.

After changing this setting, the cluster can't be updated. You must revert the settings for the original cluster and create a new cluster with the updated settings. You can delete the original cluster at a later date. To create the new cluster, use `pcluster create-cluster`. To delete the original cluster, use `pcluster delete-cluster`.

Update policy: This setting is not analyzed during an update.

These settings can be changed, and the cluster updated using `pcluster update-cluster`.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

These settings cannot be changed while the compute fleet exists. Either the change must be reverted or the compute fleet must be stopped (using `pcluster update-compute-fleet`). After the compute fleet is stopped you can update the cluster (`pcluster update-cluster`) to activate the changes. For example, if you are using a Slurm scheduler with `SlurmQueues` / `ComputeResources` / - `Name` / `MinCount` > 0, a compute fleet is started.

Update policy: The compute fleet and login nodes must be stopped for this setting to be changed for an update.

These settings cannot be changed while the compute fleet exists or if the login nodes are in use. Either the change must be reverted or the compute fleet and login nodes must be stopped (The compute fleet can be stopped using `pcluster update-compute-fleet`). After the compute fleet and login nodes have been stopped, you can update the cluster (`pcluster update-cluster`) to activate the changes.

Update policy: This setting can't be decreased during an update.

These settings can be changed, but they cannot be decreased. If these settings must be decreased, you must revert the settings for the original cluster and create a new cluster with the

updated settings. You can delete the original cluster at a later date. To create the new cluster, use `pcluster create-cluster`. To delete the original cluster, use `pcluster delete-cluster`.

Update policy: If this setting is changed, the update is not allowed. If you force the update, the new value will be ignored and the old value will be used.

After changing this setting, the cluster can't be updated. You must revert the settings for the original cluster and create a new cluster with the updated settings. You can delete the original cluster at a later date. To create the new cluster, use `pcluster create-cluster`. To delete the original cluster, use `pcluster delete-cluster`.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

These settings can be changed. Either the compute fleet must be stopped (using `pcluster update-compute-fleet`) or `QueueUpdateStrategy` must be set. After the compute fleet is stopped or `QueueUpdateStrategy` is set, you can update the cluster (`pcluster update-cluster`) to activate the changes.

> **Note**
>
> This update policy is supported starting with AWS ParallelCluster version 3.2.0.

Update policy: For this list values setting, a new value can be added during an update or the compute fleet must be stopped when removing an existing value.

A new value for these settings can be added during an update. After adding a new value to the list, the cluster can be updated using (`pcluster update-cluster`).

To remove an existing value from the list, the compute fleet must be stopped (using `pcluster update-compute-fleet`).

For example, if you are using a Slurm scheduler and adding a new instance type to Instances/InstanceType, you can update the cluster without stopping the compute fleet. To remove an existing instance type from Instances/InstanceType, the compute fleet must be stopped first (using pcluster update-compute-fleet).

> **Note**
>
> This update policy is supported starting with AWS ParallelCluster version 3.2.0.

Update policy: Reducing the size of a queue requires the compute fleet to be stopped or QueueUpdateStrategy must be set to TERMINATE for this setting to be changed for an update.

These settings can be changed, but if the change would reduce the size of the queue, the compute fleet must be stopped (using pcluster update-compute-fleet) or QueueUpdateStrategy must be set to TERMINATE. After the compute fleet is stopped or QueueUpdateStrategy is set to TERMINATE, you can update the cluster (pcluster update-cluster to activate the changes.

The TERMINATE set when resizing the capacity of the cluster, will only terminates the nodes from the back of the node list, and will leave untouched all the other nodes of the same partition.

For example, if cluster initial capacity is `MinCount = 5` and `MaxCount = 10`, the nodes are `st-[1-5]; dy-[1-5]`. When resizing the cluster to `MinCount = 3` and `MaxCount = 5`, the new cluster capacity will be composed by the nodes `st-[1-3]; dy-[1-2]`, which will not be touched during the update. Only the nodes `st-[4-5]; dy-[3-5]` are going to be terminated during the update.

The following changes are supported and don't require the compute fleet to be stopped nor the QueueUpdateStrategy set to TERMINATE:

- A new SlurmQueue is added
- A new ComputeResource is added
- MaxCount is increased
- MinCount is increased and MaxCount is increased of at least the same amount

Note: This update policy is supported starting with AWS ParallelCluster version 3.9.0.

Update policy: For this list values setting, the compute fleet must be stopped or `QueueUpdateStrategy` must be set to add a new value; the compute fleet must be stopped when removing an existing value.

A new value for these settings can be added during an update. Either the compute fleet must be stopped (using `pcluster update-compute-fleet`) or `QueueUpdateStrategy` must be

set. After the compute fleet is stopped or QueueUpdateStrategy is set, you can update the cluster (pcluster update-cluster) to activate the changes.

To remove an existing value from the list, the compute fleet must be stopped (using pcluster update-compute-fleet).

> **Note**
>
> This update policy is supported starting with AWS ParallelCluster version 3.3.0.

Update policy: All compute nodes must be stopped for a managed placement group deletion. The compute fleet must be stopped or QueueUpdateStrategy must be set for this setting to be changed for an update.

The compute fleet must be stopped (using pcluster update-compute-fleet) in order to remove a managed placement group. If you run a cluster update to remove a managed placement group before stopping the compute fleet, an invalid configuration message is returned and the update doesn't proceed. Stopping the compute fleet guarantees no instances are running.

pcluster update-cluster examples

These settings can be changed, but if the change would reduce the size of the queue, the compute fleet must be stopped (using pcluster update-compute-fleet) or QueueUpdateStrategy must be set to TERMINATE. After the compute fleet is stopped or QueueUpdateStrategy is set to TERMINATE, you can update the cluster (pcluster update-cluster to activate the changes.

- This example demonstrates an update with some allowed changes and the update is started directly.

```
$ pcluster update-cluster --cluster-name cluster_name --cluster-config
 ~/.parallelcluster/test_cluster --region us-east-1
{
  "cluster": {
    "clusterName": cluster_name,
    "cloudformationStackStatus": "UPDATE_IN_PROGRESS",
    "cloudformationStackArn": stack_arn,
    "region": "us-east-1",
```

```
    "version": "3.13.2",
    "clusterStatus": "UPDATE_IN_PROGRESS"
  },
  "changeSet": [
    {
      "parameter": "HeadNode.Networking.AdditionalSecurityGroups",
      "requestedValue": [
        "sg-0cd61884c4ad11234"
      ],
      "currentValue": [
        "sg-0cd61884c4ad16341"
      ]
    }
  ]
}
```

- This example demonstrates a dryrun update with some allowed changes. Dryrun is useful to report the change set without starting the update.

```
$ pcluster update-cluster --cluster-name cluster_name --cluster-config
 ~/.parallelcluster/test_cluster --region us-east-1 --dryrun true
{
  "message": "Request would have succeeded, but DryRun flag is set.",
  "changeSet": [
    {
      "parameter": "HeadNode.Networking.AdditionalSecurityGroups",
      "requestedValue": [
        "sg-0cd61884c4ad11234"
      ],
      "currentValue": [
        "sg-0cd61884c4ad16341"
      ]
    }
  ]
}
```

- This example demonstrates an update with some changes that block the update.

```
$ pcluster update-cluster --cluster-name cluster_name --cluster-config
 ~/.parallelcluster/test_cluster --region us-east-1
{
  "message": "Update failure",
  "updateValidationErrors": [
```

```
    {
      "parameter": "HeadNode.Ssh.KeyName",
      "requestedValue": "mykey_2",
      "message": "Update actions are not currently supported for the 'KeyName'
 parameter. Restore 'KeyName' value to 'jenkinsjun'. If you need this change, please
 consider creating a new cluster instead of updating the existing one.",
      "currentValue": "mykey_1"
    },
    {
      "parameter": "Scheduling.SlurmQueues[queue1].ComputeResources[queue1-
t2micro].InstanceType",
      "requestedValue": "c4.xlarge",
      "message": "All compute nodes must be stopped. Stop the compute fleet with the
 pcluster update-compute-fleet command",
      "currentValue": "t2.micro"
    },
    {
      "parameter": "SharedStorage[ebs1].MountDir",
      "requestedValue": "/my/very/very/long/shared_dir",
      "message": "Update actions are not currently supported for the 'MountDir'
 parameter. Restore 'MountDir' value to '/shared'. If you need this change, please
 consider creating a new cluster instead of updating the existing one.",
      "currentValue": "/shared"
    }
  ],
  "changeSet": [
    {
      "parameter": "HeadNode.Networking.AdditionalSecurityGroups",
      "requestedValue": [
        "sg-0cd61884c4ad11234"
      ],
      "currentValue": [
        "sg-0cd61884c4ad16341"
      ]
    },
    {
      "parameter": "HeadNode.Ssh.KeyName",
      "requestedValue": "mykey_2",
      "currentValue": "mykey_1"
    },
    {
      "parameter": "Scheduling.SlurmQueues[queue1].ComputeResources[queue1-
t2micro].InstanceType",
      "requestedValue": "c4.xlarge",
```

```
      "currentValue": "t2.micro"
    },
    {
      "parameter": "SharedStorage[ebs1].MountDir",
      "requestedValue": "/my/very/very/long/shared_dir",
      "currentValue": "/shared"
    }
  ]
}
```

AWS ParallelCluster AMI customization

There are scenarios where building a custom AMI for AWS ParallelCluster is necessary. This section covers what to consider when building a custom AWS ParallelCluster AMI.

You can build a custom AWS ParallelCluster AMI using one of the following methods:

1. Create a build image configuration file, and then use the `pcluster` CLI to build the image with EC2 Image Builder. This process is automated, repeatable, and supports monitoring. For more information, see the `pcluster` image commands.
2. Create an instance from an AWS ParallelCluster AMI, then log in to it and make manual modifications. Last, use Amazon EC2 to create a new AMI from the modified instance. This process takes less time. However, it isn't automated or repeatable, and it doesn't support use of the `pcluster` CLI image monitoring commands.

For more information about these methods, see Building a custom AWS ParallelCluster AMI.

AWS ParallelCluster AMI customization considerations

No matter how you create your custom image, we recommend that you perform preliminary validation tests and include provisions to monitor the status of the image being created.

To build a custom AMI using `pcluster`, you create a build image configuration file with a `Build` and `Image` section that EC2 Image Builder uses to build your custom image. The `Build` section specifies what Image Builder needs to build the image. This includes the `ParentImage` (base image), and `Components`. An Image Builder component defines a sequence of steps that are required to customize an instance before an image is created or to test an instance that was

launched by the created image. For more information, see Create a custom component with Image Builder in the *EC2 Image Builder User Guide*.

When called from pcluster `build-image` to create a custom image, Image Builder uses the build image configuration with the AWS ParallelCluster cookbook to bootstrap AWS ParallelCluster on your `ParentImage`. Image Builder downloads components, runs build and validate phases, creates the AMI, launches an instance from the AMI, and runs tests. When the process completes, Image Builder then produces a new image or a stop message.

Perform custom component validation tests

Before you include an Image Builder component in a configuration, test and validate it using one of the following methods. Because the Image Builder process can take up to 1 hour, we recommend that you test the components beforehand. This can save you a considerable amount of time.

Script case

Test the script in a running instance, outside the build image process, and verify that the script exits with exit code 0.

Amazon Resource Name (ARN) case

Test the component document in a running instance, outside the build image process. For a list of requirements, see Component manager in the *Image Builder User Guide*.

After successful validation, add the component to your build image configuration

After you verified that the custom component is working, add it to the Build image configuration file.

Monitor the Image Builder process with `pcluster` commands to aid in debugging

`describe-image`

Use this command to monitor the build image status.

`list-image-log-streams`

Use this command to get the IDs of log streams that you can use to retrieve log events with `get-image-log-events`.

`get-image-log-events`

Use this command to get the log stream of build image process events.

For example, you can tail build image events using the following command.

```
$ watch -n 1 'pcluster get-image-log-events -i <image-id> \
        --log-stream-name/1 <pcluster-version> \
        --query "events[*].message" | tail -n 50'
```

`get-image-stack-events`

Use this command to retrieve image stack events for the stack that Image Builder creates.

`export-image-logs`

Use this command save image logs.

For more information about AWS ParallelCluster logs and Amazon CloudWatch, see Amazon CloudWatch Logs build image logs and Amazon CloudWatch dashboard.

Other considerations

New AWS ParallelCluster releases and custom AMIs

If you build and use a custom AMI, you must repeat the steps that you used to create your custom AMI with each new AWS ParallelCluster release.

Custom bootstrap actions

Review the Custom bootstrap actions section to determine if the modifications you want to make can be scripted and supported with future AWS ParallelCluster releases.

Using custom AMIs

You can specify custom AMIs in the cluster configuration in the `Image` / `CustomAmi` and `Scheduling` / `SlurmQueues` / - `Name` / `Image` / `CustomAmi` sections.

To troubleshoot custom AMI validation warnings, see Troubleshooting custom AMI issues.

Launch instances with On-Demand Capacity Reservations (ODCR)

With On-Demand Capacity Reservations (ODCR), you can reserve capacity for your cluster Amazon EC2 instances in a specific Availability Zone. This way, you can create and manage Capacity Reservations independently from the billing accounts that Savings Plans or regional Reserved Instances offer.

You can configure `open` or `targeted` ODCR. *Open* ODCR cover any instances that match the ODCR attributes. These attributes are instance type, platform, and Availability Zone. You must explicitly define *Targeted* ODCR in the cluster configuration. To determine whether an ODCR is open or `targeted`, run the AWS CLI Amazon EC2 `describe-capacity-reservation` command.

You can also create an ODCR in a cluster placement group that's called a cluster placement group on-demand capacity reservation (CPG ODCR).

Multiple ODCRs can be grouped in a resource group. This can be defined in the cluster configuration file. For more information about resource groups, see What are resource groups? in the *Resource Groups and Tags User Guide*.

Using ODCR with AWS ParallelCluster

AWS ParallelCluster supports open ODCR. When using an open ODCR, you don't need to specify anything in AWS ParallelCluster. Instances are automatically selected for the cluster. You can specify an existing placement group or have AWS ParallelCluster create a new one for you.

ODCR in the cluster configuration

Starting with AWS ParallelCluster version 3.3.0, you can define ODCRs in the cluster configuration file, with no need to specify Amazon EC2 run-instances overrides.

You start by creating capacity reservations and resource groups using the methods described in the linked documentation for each. You must use the AWS CLI methods to create capacity reservation groups. If you use the AWS Management Console, you can only create Tag based or Stack based resource groups. Tag based and Stack based resource groups aren't supported by AWS ParallelCluster or the AWS CLI when launching instances with capacity reservations.

After the capacity reservations and resource groups are created, specify them in `SlurmQueues` / `CapacityReservationTarget` or `SlurmQueues` / `ComputeResources` /

CapacityReservationTarget as shown in the following example cluster configuration. Replace *values* highlighted in red with your valid values.

```
Image:
  Os: os
HeadNode:
  InstanceType: head_node_instance
  Networking:
    SubnetId: public_subnet_id
  Ssh:
    KeyName: key_name
Scheduling:
  Scheduler: scheduler
  SlurmQueues:
    - Name: queue1
      Networking:
        SubnetIds:
          - private_subnet_id
      ComputeResources:
        - Name: cr1
          Instances:
            - InstanceType: instance
          MaxCount: max_queue_size
          MinCount: max_queue_size
          Efa:
            Enabled: true
          CapacityReservationTarget:
            CapacityReservationResourceGroupArn: capacity_reservation_arn
```

OBSOLETE / NOT RECOMMENDED - Targeted ODCR with Amazon EC2 instance overrides

⚠ Warning

- Starting with AWS ParallelCluster version 3.3.0, we don't recommend this method. This section remains as a reference for implementations using prior versions.
- This method is not compatible with Multiple instance type allocation with Slurm.

Support for `targeted` ODCRs is added in AWS ParallelCluster 3.1.1. In this release, a mechanism was introduced that overrides EC2 `RunInstances` parameters and passes information about the reservation to use for each configured compute resource in AWS ParallelCluster. This mechanism

is compatible with `targeted` ODCR. However, when you use `targeted` ODCR, you must specify the `run-instances` override configuration. *Targeted* ODCRs must be explicitly defined in the AWS CLI Amazon EC2 `run-instances` command. To determine whether an ODCR is `open` or `targeted` run the AWS CLI Amazon EC2 command `describe-capacity-reservation`.

Multiple ODCRs can be grouped in a resource group. This can be used in the run-instances override to target multiple ODCRs at the same time.

If you're using a `targeted` ODCR, you can specify a placement group. However, you also need to specify a `run-instances` override configuration.

Suppose that AWS created a `targeted` ODCR for you or you have a specific set of Reserved Instances. Then, you can't specify a placement group. The rules that are configured by AWS might conflict with the placement group setting. So, if a placement group is required for your application, use a CPG ODCR. In either case, you must also specify the `run-instances` override configuration.

If you're using a CPG ODCR, you must specify the `run-instances` override configuration and you must specify the same placement group in the cluster configuration.

Using Reserved Instances with AWS ParallelCluster

Reserved instances are different than Capacity Reservations (ODCR). There are 2 types of reserved instances. A *Regional* Reserved Instance doesn't reserve capacity. A *zonal* Reserved Instance reserves capacity in the specified Availability Zone.

If you have Regional Reserved Instances, there's no capacity reservation and you can get Insufficient Capacity Errors. If you have zonal Reserved Instances, you have capacity reservation, but there are no `run-instances` API parameters that you can use to specify them.

Reserved instances are supported by any AWS ParallelCluster version. You don't have to specify anything in AWS ParallelCluster and the instances are automatically selected.

When using zonal Reserved Instances, you can avoid potential Insufficient Capacity Errors by omitting the placement group specification in the cluster configuration.

OBSOLETE / NOT RECOMMENDED - Using `RunInstances` customization in AWS ParallelCluster 3 for `targeted` On-Demand Capacity Reservations (ODCR)

Warning

- Starting with AWS ParallelCluster version 3.3.0, we don't recommend this method. This section remains as a reference for implementations using prior versions.
- This method is not compatible with Multiple instance type allocation with Slurm.

You can override Amazon EC2 `RunInstances` parameters for each compute resource that's configured in a cluster queue. To do so, create the `/opt/slurm/etc/pcluster/run_instances_overrides.json` file on the head node of the cluster with the following code snippet content:

- `${queue_name}` is the name of the queue that you want to apply overrides to.
- `${compute_resource_name}` is the compute resource that you want to apply overrides to.
- `${overrides}` is an arbitrary JSON object that contains a list of `RunInstances` overrides to use for the specific combination of queue and instance type. The overrides syntax needs to follow the same specifications that are documented in a run_instances boto3 call.

```
{
    "${queue_name}": {
        "${compute_resource_name}": {
            ${overrides}
        },
        ...
    },
    ...
}
```

For example, the following JSON configures the ODCR group `group_arn` to be used for `p4d.24xlarge` instances that are configured in `my-queue` and `my-compute-resource`.

```
{
    "my-queue": {
        "my-compute-resource": {
            "CapacityReservationSpecification": {
```

```
                "CapacityReservationTarget": {
                    "CapacityReservationResourceGroupArn": "group_arn"
                }
            }
        }
    }
}
```

After this JSON file is generated, the AWS ParallelCluster daemons that are responsible for cluster scaling automatically use the override configuration for instance launches. To confirm that the specified parameters are being used for instance provisioning, look at the following log files:

- `/var/log/parallelcluster/clustermgtd` (for static capacity)
- `/var/log/parallelcluster/slurm_resume.log` (for dynamic capacity)

If the parameters are correct, you'll find a log entry that contains the following:

```
Found RunInstances parameters override. Launching instances with: <parameters_list>
```

OBSOLETE / NOT RECOMMENDED - Create a cluster with `targeted` On-Demand Capacity Reservations (ODCR)

> **Warning**
>
> - Starting with AWS ParallelCluster version 3.3.0, we don't recommend this method. This section remains as a reference for implementations using prior versions.
> - This method is not compatible with Multiple instance type allocation with Slurm.

1. **Create a resource group, to group capacity.**

```
$ aws resource-groups create-group --name EC2CRGroup \
    --configuration '{"Type":"AWS::EC2::CapacityReservationPool"}'
 '{"Type":"AWS::ResourceGroups::Generic", "Parameters": [{"Name": "allowed-
resource-types", "Values": ["AWS::EC2::CapacityReservation"]}]}'
```

Note

A resource group doesn't support resources that are shared by other accounts.
If the target ODCR is shared by another account, you don't need to create a resource group. Use `CapacityReservationId` instead of a resource group in step 3.

```
#!/bin/bash
set -e

# Override run_instance attributes
cat > /opt/slurm/etc/pcluster/run_instances_overrides.json << EOF
{
    "my-queue": {
        "my-compute-resource": {
            "CapacityReservationSpecification": {
                "CapacityReservationTarget": {
                    "CapacityReservationId": "cr-abcdef01234567890"
                }
            }
        }
    }
}
EOF
```

Add capacity reservations to the resource group. Every time that you create a new ODCR, add it to the Group Reservation. Replace *ACCOUNT_ID* with your account ID, *PLACEHOLDER_CAPACITY_RESERVATION* with your capacity reservation ID, and *REGION_ID* with your AWS Region ID (for example, us-east-1).

```
$ aws resource-groups group-resources --region REGION_ID --group EC2CRGroup \
    --resource-arns arn:aws:ec2:REGION_ID:ACCOUNT_ID:capacity-
reservation/PLACEHOLDER_CAPACITY_RESERVATION
```

Create a policy document on your local computer. Replace *ACCOUNT_ID* with your account ID and *REGION_ID* with your AWS Region ID (for example, us-east-1).

```
cat > policy.json << EOF
{
```

```
    "Version": "2012-10-17",
    "Statement": [
        {
            "Sid": "RunInstancesInCapacityReservation",
            "Effect": "Allow",
            "Action": "ec2:RunInstances",
            "Resource": [
                "arn:aws:ec2:REGION_ID:ACCOUNT_ID:capacity-reservation/*",
                "arn:aws:resource-groups:REGION_ID:ACCOUNT_ID:group/*"
            ]
        }
    ]
}
EOF
```

2. **Create the IAM policy on your AWS account using the json file that you created.**

```
$ aws iam create-policy --policy-name RunInstancesCapacityReservation --policy-
document file://policy.json
```

3. **Create the following post install script locally on the instance and name it `postinstall.sh`.**

   Replace *ACCOUNT_ID* with your AWS account ID, and *REGION_ID* with your AWS Region ID (for example, us-east-1).

```
#!/bin/bash
set -e

# Override run_instance attributes
cat > /opt/slurm/etc/pcluster/run_instances_overrides.json << EOF
{
    "my-queue": {
        "my-compute-resource": {
            "CapacityReservationSpecification": {
                "CapacityReservationTarget": {
                    "CapacityReservationResourceGroupArn": "arn:aws:resource-
groups:REGION_ID:ACCOUNT_ID:group/EC2CRGroup"
                }
            }
        }
    }
}
```

```
EOF
```

Upload the file to an Amazon S3 bucket. Replace *amzn-s3-demo-bucket* with your specific S3 bucket name.

```
$ aws s3 mb s3://amzn-s3-demo-bucket
aws s3 cp postinstall.sh s3://amzn-s3-demo-bucket/postinstall.sh
```

4. **Create the local cluster configuration, replacing the placeholders with your own values.**

```
Region: REGION_ID
Image:
  Os: alinux2
HeadNode:
  InstanceType: c5.2xlarge
  Ssh:
    KeyName: YOUR_SSH_KEY
  Iam:
    S3Access:
      - BucketName: amzn-s3-demo-bucket
    AdditionalIamPolicies:
      - Policy: arn:aws:iam::ACCOUNT_ID:policy/RunInstancesCapacityReservation
  ## This post-install script is executed after the node is configured.
  ## It is used to install scripts at boot time and specific configurations
  ## In the script below we are overriding the calls to RunInstance to force
  ## the provisioning of our my-queue partition to go through
  ## the On-Demand Capacity Reservation
  CustomActions:
    OnNodeConfigured:
      Script: s3://amzn-s3-demo-bucket/postinstall.sh
  Networking:
    SubnetId: YOUR_PUBLIC_SUBNET_IN_TARGET_AZ

Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: my-queue
      ComputeResources:
        - MinCount: 0
          MaxCount: 100
          InstanceType: p4d.24xlarge
          Name: my-compute-resource
          Efa:
```

```
            Enabled: true
      Networking:
        ## PlacementGroup:
        ##   Enabled: true ## Keep PG disabled if using targeted ODCR
        SubnetIds:
          - YOUR_PRIVATE_SUBNET_IN_TARGET_AZ
```

5. **Create the cluster.**

   Use the following command to create the cluster. Replace *cluster-config.yaml* with your configuration file name, *cluster-dl* with your cluster name, and *REGION_ID* with your Region ID (for example, us-east-1).

   ```
   $ pcluster create-cluster --cluster-configuration cluster-config.yaml --cluster-
   name cluster-dl --region REGION_ID
   ```

   After the cluster is created, the post-install script runs in the head node. The script creates the `run_instances_overrides.json` file and overrides the calls to `RunInstances` to force the provisioning of the partition to go through the On-Demand Capacity Reservation.

   The AWS ParallelCluster daemons that are responsible for cluster scaling automatically use this configuration for new instances that are launched. To confirm that the specified parameters are being used to provision instances, you can look at the following log files:

   - `/var/log/parallelcluster/clustermgtd` (for static capacity - `MinCount > 0`)
   - `/var/log/parallelcluster/slurm_resume.log` (for dynamic capacity)

   If the parameters are correct, you'll find a log entry contains the following.

   ```
   Found RunInstances parameters override. Launching instances with: <parameters_list>
   ```

Updating `RunInstances` overrides

You can update the generated JSON configuration at any time without stopping the compute fleet. After the changes are applied, all new instances launch with the updated configuration. If you need to apply the updated configuration to running nodes, recycle the nodes by forcing an instance termination and wait for AWS ParallelCluster to replace those nodes. You can do this by

terminating the instance from the Amazon EC2 console or AWS CLI, or by setting the Slurm nodes in a DOWN or DRAIN state.

Use the following command to set the Slurm node to DOWN or DRAIN.

```
$ scontrol update nodename=my-queue-dy-my-compute-resource-1 state=down
 reason=your_reason
scontrol update nodename=my-queue-dy-my-compute-resource-1 state=drain
 reason=your_reason
```

Launch instances with Capacity Blocks (CB)

AWS ParallelCluster supports On-Demand Capacity Reservations (ODCR) and Capacity Blocks (CB) for Machine Learning. Unlike ODCR, CB can have a future start time and is time-bound. For more information about launching with ODCR, see Launch instances with On-Demand Capacity Reservations (ODCR).

Using CB with AWS ParallelCluster

To configure your new or existing clusters to use a CB, you first need to have a valid CB in your AWS account. You can use the AWS Management Console, AWS Command Line Interface, or SDK to find and purchase an available CB by following official documentation. Once you have a valid CB, you can set CB Amazon Resource Name (ARN) and related parameters in your AWS ParallelCluster configuration file. For more information, see Find and purchase Capacity Blocks (CB)

CB in the cluster configuration

To use a CB for a specific queue you need to use the CapacityReservationId parameter. Configure it to an existing CB ID. You can obtain the CB ARN from the AWS Management Console, AWS CLI, or SDK that you used to create the CB.

You have to set CapacityType = CAPACITY_BLOCK for the queue where you want to use the CB. Set it to the InstanceType of the compute resource (the same Amazon Elastic Compute Cloud instance type of the CB).

When CapacityReservationId is specified at compute resource level, InstanceType is optional because it will be automatically retrieved from the reservation.

When using `CapacityType = CAPACITY_BLOCK`, `MaxCount` must be equal to `MinCount` and greater than 0, because all the instances that are part of the CB reservation are managed as static nodes.

At the cluster creation time, the head node waits for all the static nodes to be ready before signaling the success of cluster creation. However, when using `CapacityType = CAPACITY_BLOCK`, the nodes that are part of the compute resources associated to won't be considered for this check. The cluster will be created even if not all the configured are active.

The following configuration file snippet shows the required parameters to enable in the AWS ParallelCluster configuration file.

```
SlurmQueues:
 - Name: string
   CapacityType: CAPACITY_BLOCK
   ComputeResources:
   - Name: string
     InstanceType: String (EC2 Instance type of the CB)
     MinCount: integer (<= total capacity of the CB)
     MaxCount: integer (equal to MinCount)
     CapacityReservationTarget:
        CapacityReservationId: String (CB id)
```

How AWS ParallelCluster uses Capacity Blocks (CB)

AWS ParallelCluster manages static nodes associated with in a peculiar way. AWS ParallelCluster creates a cluster even if the CB is not yet active, and instances are launched automatically once the CB is active.

The Slurm nodes that correspond to compute resources, associated with , and are not yet active, are kept in maintenance until they reach the CB start time. Slurm nodes remain in a reservation/maintenance state and are associated with the slurm admin user. This means they can accept jobs, but the jobs remain in `pending` until the reservation is removed.

AWS ParallelCluster automatically updates Slurm reservations and puts the related CB nodes in maintenance (corresponding to the CB state). When the CB is active, the Slurm reservation is removed, nodes start, and become available for the pending jobs or for new job submissions.

When the CB end time is reached, nodes will be moved back to a reservation/maintenance state. It's up to users to resubmit/requeue the jobs to a new queue/compute-resource when CB is no longer active and instances are terminated.

AMI patching and Amazon EC2 instance replacement

To ensure that all dynamically launched cluster compute nodes behave in a consistent manner, AWS ParallelCluster disables cluster instance automatic OS updates. Additionally, a specific set of AWS ParallelCluster AMIs are built for each version of AWS ParallelCluster and its associated CLI. This specific set of AMIs remain unchanged and they are only supported by the AWS ParallelCluster version they were built for. AWS ParallelCluster AMIs for released versions aren't updated.

However, due to emergent security issues, customers might want to add patches to these AMIs and then update their clusters with the patched AMI. This aligns with the AWS ParallelCluster Shared Responsibility Model.

To view the specific set of AWS ParallelCluster AMIs supported by the AWS ParallelCluster CLI version you are currently using, run:

```
$ pcluster version
$ pcluster list-official-images
```

The AWS ParallelCluster head node is a static instance and you can manually update it. Restart and reboot of the head node is fully supported starting with AWS ParallelCluster version 3.0.0.

If your instances have ephemeral instance stores, you must remember to save instance store data before manual updates. For more information, see the `HeadNode` / `LocalStorage` / `EphemeralVolume` cluster configuration and Instance types with instance store volumes in the *Amazon EC2 User Guide for Linux Instances*.

The compute nodes are ephemeral instances. By default you can only access them from the head node. Starting with AWS ParallelCluster version 3.0.0, you can update the AMI associated with compute instances by modifying the `Scheduling` / `SlurmQueues` / `Image` / `CustomAmi` parameter and running the `pcluster update-cluster` command, after stopping the compute fleet with `pcluster update-compute-fleet`:

```
$ pcluster update-compute-fleet-status --status STOP_REQUESTED
```

It's possible to automate the creation of an updated custom AMI for the compute nodes by using one of the following methods:

- Use the `pcluster build-image` command with an updated `Build` / `ParentImage`.
- Run the build with `Build` / `UpdateOsPackages` / `Enabled`:`true`.

Head node instance update or replacement

In some circumstances, you might be required to restart or reboot the head node. For example, this is required when you manually update the OS, or when there's an AWS instance scheduled retirement that imposes a head node instance restart.

If your instance doesn't have ephemeral drives, you can stop and start it again at any time. In the case of a scheduled retirement, starting the stopped instance migrates it to use the new hardware.

Similarly, you can manually stop and start an instance that doesn't have instance stores. For this case and for other cases of instances without ephemeral volumes, continue to Stop and start a cluster's head node.

If your instance has ephemeral drives and its been stopped, the data in the instance store is lost. You can determine if the instance type used for the head node has instance stores from the table found in Instance store volumes.

Save data from ephemeral drives

Starting with AWS ParallelCluster version 3.0.0, the head node restart and reboot is fully supported for every instance type. However, if instances have an ephemeral drive, its data is lost. Follow the next steps to preserve your data before a head node restart or reboot.

To check if you have data that needs to be preserved, view the content in the `EphemeralVolume` / `MountDir` folder (`/scratch` by default).

You can transfer the data to the root volume or the shared storage systems attached to the cluster, such as Amazon FSx, Amazon EFS, or Amazon EBS. Note that the data transfer to remote storage can incur additional costs.

After saving the data, continue to Stop and start a cluster's head node.

Stop and start a cluster's head node

1. Verify there aren't any running jobs in the cluster.

   When using a Slurm scheduler:

   - If the `sbatch --no-requeue` option isn't specified, running jobs are requeued.
   - If the `--no-requeue` option is specified, running jobs fail.

2. Request a cluster compute fleet stop:

```
$ pcluster update-compute-fleet --cluster-name cluster-name --status STOP_REQUESTED
{
  "status": "STOP_REQUESTED",
  ...
}
```

3. Wait until the compute fleet status is STOPPED:

```
$ pcluster update-compute-fleet --cluster-name cluster-name --status STOP_REQUESTED
{
  "status": "STOPPED",
  ...
}
```

4. For manual updates with an OS reboot or instance restart, you can use the AWS Management Console or AWS CLI. The following is an example of using the AWS CLI.

```
# Retrieve head node instance id
$ pcluster describe-cluster --cluster-name cluster-name --status STOP_REQUESTED
{
  "headNode": {
  "instanceId": "i-1234567890abcdef0",
  ...
},
  ...
}
# stop and start the instance
$ aws ec2 stop-instances --instance-ids 1234567890abcdef0
{
  "StoppingInstances": [
    {
      "CurrentState": {
        "Name": "stopping"
        ...
      },
      "InstanceId": "i-1234567890abcdef0",
      "PreviousState": {
        "Name": "running"
        ...
      }
    }
```

```
  ]
}
$ aws ec2 start-instances --instance-ids 1234567890abcdef0
{
  "StartingInstances": [
    {
      "CurrentState": {
        "Name": "pending"
        ...
      },
      "InstanceId": "i-1234567890abcdef0",
      "PreviousState": {
        "Name": "stopped"
        ...
      }
    }
  ]
}
```

5. Start the cluster compute fleet:

```
$ pcluster update-compute-fleet --cluster-name cluster-name --status
 START_REQUESTED
{
  "status": "START_REQUESTED",
  ...
}
```

Operating systems

AWS ParallelCluster supports Amazon Linux 2, Amazon Linux 2023, Ubuntu24.04, Ubuntu 22.04, Ubuntu 20.04, Red Hat Enterprise Linux 8 (RHEL8), Rocky 8, Red Hat Enterprise Linux 9 (RHEL9), and Rocky 9. AWS ParallelCluster offers pre-built AMIs for select operating systems, for more details on AMIs provided by AWS ParallelCluster refer to Image section.

Operating system considerations

Ubuntu 22.04 & Ubuntu 24.04

Ubuntu 22.04 & Ubuntu 24.04 require more secure keys for ssh and do not support RSA keys by default. Please generate an ed25519 key and use that for cluster creation.

Ubuntu 22.04 cannot be updated to the latest kernel because there is no FSx client for that kernel.

RHEL 8

RedHat Enterprise Linux 8.7 (rhel8) is added starting in AWS ParallelCluster version 3.6.0. If you configure your cluster to use rhel8, the on-demand cost for any instance type is higher than when you configure your cluster to use other supported operation systems.

For more information about pricing, see On-Demand Pricing and How is Red Hat Enterprise Linux on Amazon Elastic Compute Cloud offered and priced?.

Rocky 8

AWS ParallelCluster 3.8.0 supports Rocky Linux 8, but pre-built Rocky Linux 8 AMIs (for x86 and ARM architectures) are not available. AWS ParallelCluster 3.8.0 supports creating clusters with Rocky Linux 8 using custom AMIs using the CustomAmi property. For more information about building custom AMIs, refer to AWS ParallelCluster AMI customization.

To build your custom AMI from a base Rocky Linux 8 AMI, you can consider subscribing to the Rocky Linux 8 AMIs available on the AWS Marketplace. Make sure to review the pricing and subscription costs for Rocky Linux 8 AMIs on the AWS Marketplace. Alternatively you can also use the official Rocky Linux 8 AMIsas your base AMI.

Rocky9

AWS ParallelCluster 3.9.0 supports Rocky Linux 9, but pre-built Rocky Linux 9 AMIs (for x86 and ARM architectures) are not available. AWS ParallelCluster 3.9.0 supports creating clusters with Rocky Linux 9 using custom AMIs using the CustomAmi property. For more information about building custom AMIs, refer to AWS ParallelCluster AMI customization. To build your custom AMI from a base Rocky Linux 9 AMI, you can also use the official Rocky Linux 9 AMIs as your base AMI. Custom Rocky Linux 9 AMI build may fail if the base AMI does not have the latest kernel. To upgrade the kernel before building the AMI:

- Launch an instance using a rocky9 AMI id from here: https://rockylinux.org/cloud-images/
- ssh into the instance and run the following command:`sudo yum -y update`
- Create an image from the instance to use as `ParentImage`

Reference for AWS ParallelCluster

Topics

AWS ParallelCluster CLI commands

`pcluster` is the primary AWS ParallelCluster CLI command. You use `pcluster` to launch and manage HPC clusters in the AWS Cloud and to create and manage custom AMI images.

`pcluster3-config-converter` is used to convert cluster configurations in AWS ParallelCluster version 2 format into AWS ParallelCluster version 3 format.

```
pcluster [-h] ( build-image | configure |
               create-cluster | dcv-connect |
               delete-cluster | delete-cluster-instances | delete-image |
               describe-cluster | describe-cluster-instances |
               describe-compute-fleet | describe-image |
               export-cluster-logs | export-image-logs |
               get-cluster-log-events | get-cluster-stack-events |
               get-image-log-events | get-image-stack-events |
               list-cluster-log-streams | list-clusters |
               list-images | list-image-log-streams | list-official-images |
               ssh | update-cluster |
               update-compute-fleet | version ) ...
pcluster3-config-converter [-h] [-t CLUSTER_TEMPLATE]
               [-c CONFIG_FILE]
               [--force-convert]
               [-o OUTPUT_FILE]
```

Topics

pcluster

`pcluster` is the primary AWS ParallelCluster CLI command. You use `pcluster` to launch and manage HPC clusters in the AWS Cloud.

`pcluster` writes logs of your commands to `pcluster.log.#` files in `/home/user/.parallelcluster/`. For more information, see pcluster CLI logs.

To use `pcluster`, you must have an IAM role with the permissions required to run it.

```
pcluster [-h]
```

Arguments

`pcluster` *command*

Possible choices: build-image | configure | create-cluster | dcv-connect | delete-cluster | delete-cluster-instances | delete-image | describe-cluster | describe-cluster-instances | describe-compute-fleet | describe-image | export-cluster-logs | export-image-logs | get-cluster-log-events | get-cluster-stack-events | get-image-log-events | get-image-stack-events | list-clusters | list-cluster-log-streams | list-images | list-image-log-streams | list-official-images | ssh | update-cluster | update-compute-fleet | version

Sub-commands:

Topics

- pcluster build-image
- pcluster configure
- pcluster create-cluster
- pcluster dcv-connect
- pcluster delete-cluster
- pcluster delete-cluster-instances
- pcluster delete-image
- pcluster describe-cluster

- pcluster describe-cluster-instances
- pcluster describe-compute-fleet
- pcluster describe-image
- pcluster export-cluster-logs
- pcluster export-image-logs
- pcluster get-cluster-log-events
- pcluster get-cluster-stack-events
- pcluster get-image-log-events
- pcluster get-image-stack-events
- pcluster list-clusters
- pcluster list-cluster-log-streams
- pcluster list-images
- pcluster list-image-log-streams
- pcluster list-official-images
- pcluster ssh
- pcluster update-cluster
- pcluster update-compute-fleet
- pcluster version

pcluster build-image

Create a custom AWS ParallelCluster image in the specified Region.

```
pcluster build-image [-h]
                  --image-configuration IMAGE_CONFIGURATION
                  --image-id IMAGE_ID
                 [--debug]
                 [--dryrun DRYRUN]
                 [--query QUERY]
                 [--region REGION]
                 [--rollback-on-failure ROLLBACK_ON_FAILURE]
                 [--suppress-validators SUPPRESS_VALIDATORS [SUPPRESS_VALIDATORS ...]]
                 [--validation-failure-level {INFO,WARNING,ERROR}]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster build-image`.

`--image-configuration, -c` *`IMAGE_CONFIGURATION`*

Specifies the image configuration file as a YAML document.

`--image-id, -i` *`IMAGE_ID`*

Specifies the id of the image that will be built.

`--debug`

Turn on debug logging.

`--dryrun` *`DRYRUN`*

When `true`, the command performs validation without creating any resources. You can use this to validate the image configuration. (Defaults to `false`.)

`--query` *`QUERY`*

JMESPath query to perform on output.

`--region, -r` *`REGION`*

Specifies the AWS Region to use. The AWS Region must be specified, using the Region setting in the image configuration file, the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

`--rollback-on-failure` *`ROLLBACK_ON_FAILURE`*

When `true`, automatically initiates an image stack rollback on failure. (Defaults to `false`.)

`--suppress-validators` *`SUPPRESS_VALIDATORS`* `[`*`SUPPRESS_VALIDATORS ...`*`]`

Identifies one or more config validators to suppress.

Format: (`ALL|type:[A-Za-z0-9]+`)

`--validation-failure-level {INFO,WARNING,ERROR}`

Specifies the minimum validation level that will cause the creation to fail. (Defaults to ERROR.)

Example using AWS ParallelCluster version 3.1.2:

```
$ pcluster build-image --image-configuration image-config.yaml --image-id custom-
alinux2-image
{
  "image": {
    "imageId": "custom-alinux2-image",
    "imageBuildStatus": "BUILD_IN_PROGRESS",
    "cloudformationStackStatus": "CREATE_IN_PROGRESS",
    "cloudformationStackArn": "arn:aws:cloudformation:us-east-1:123456789012:stack/
custom-alinux2-image/1234abcd-56ef-78gh-90ij-abcd1234efgh",
    "region": "us-east-1",
    "version": "3.1.2"
  }
}
```

> **Warning**
>
> `pcluster build-image` uses the default VPC. If the default VPC has been deleted, perhaps by using AWS Control Tower or AWS Landing Zone, then the subnet ID must be specified in the image configuration file. For more information, see SubnetId.

`pcluster configure`

Begins an interactive configuration wizard for AWS ParallelCluster version 3. For more information, see Configure and create a cluster with the AWS ParallelCluster command line interface.

```
pcluster configure [-h]
                   --config CONFIG
                  [--debug]
                  [--region REGION]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster configure`.

`--config CONFIG`

Path to output the generated config file.

--debug

Turn on debug logging.

--region, -r *REGION*

Specifies the AWS Region to use. The Region must be specified, using the Region setting in the image configuration file, the AWS_DEFAULT_REGION environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

pcluster create-cluster

Creates an AWS ParallelCluster cluster.

```
pcluster create-cluster [-h]
                 --cluster-configuration CLUSTER_CONFIGURATION
                 --cluster-name CLUSTER_NAME
                [--debug]
                [--dryrun DRYRUN]
                [--query QUERY]
                [--region REGION]
                [--rollback-on-failure ROLLBACK_ON_FAILURE]
                [--suppress-validators SUPPRESS_VALIDATORS [SUPPRESS_VALIDATORS ...]]

                [--validation-failure-level {INFO,WARNING,ERROR}]
```

Named arguments

-h, --help

Shows the help text for `pcluster create-cluster`.

--cluster-configuration, -c *CLUSTER_CONFIGURATION*

Specifies the YAML cluster configuration file.

--cluster-name, -n *CLUSTER_NAME*

Specifies the name of the cluster to be created.

The name must start with an alphabetical character. The name can have up to 60 characters. If Slurm accounting is enabled, the name can have up to 40 characters.

Valid characters: a-z, A-Z, 0-9, and - (hyphen).

--debug

Enables debug logging.

--dryrun *DRYRUN*

When `true`, the command performs validation without creating any resources. You can use this to validate the cluster configuration. (Defaults to `false`.)

--query *QUERY*

Specifies the JMESPath query to perform on the output.

--region, -r *REGION*

Specifies the AWS Region to use. The AWS Region must be specified, using the `Region` setting in the cluster configuration file, the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

--rollback-on-failure *ROLLBACK_ON_FAILURE*

When `true`, automatically initiates a cluster stack rollback on failures. (Defaults to `true`.)

--suppress-validators *SUPPRESS_VALIDATORS* [*SUPPRESS_VALIDATORS* ...]

Identifies one or more config validators to suppress.

Format: (ALL|type:`[A-Za-z0-9]`+)

--validation-failure-level {INFO,WARNING,ERROR}

Specifies the minimum validation level that will cause the creation to fail. (Defaults to ERROR.)

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster create-cluster -c cluster-config.yaml -n cluster-v3
{
  "cluster": {
    "clusterName": "cluster-v3",
    "cloudformationStackStatus": "CREATE_IN_PROGRESS",
    "cloudformationStackArn": "arn:aws:cloudformation:us-east-1:123456789012:stack/
cluster-v3/1234abcd-56ef-78gh-90ij-abcd1234efgh",
    "region": "us-east-1",
```

```
    "version": "3.1.4",
    "clusterStatus": "CREATE_IN_PROGRESS"
  }
}
```

pcluster dcv-connect

Permits to connect to the head node through an interactive session by using Amazon DCV.

```
pcluster dcv-connect [-h]
                 --cluster-name CLUSTER_NAME
                [--debug]
                [--key-path KEY_PATH]
                [--login-node-ip LOGIN_NODE_IP]
                [--region REGION]
                [--show-url]
```

Named arguments

-h, --help

Shows the help text for `pcluster dcv-connect`.

--cluster-name, -n *CLUSTER_NAME*

Specifies the name of the cluster.

--debug

Enables debug logging.

--key-path *KEY_PATH*

Specifies the path of the SSH key to use for the connection.

--login-node-ip

Specifies the public or private IP address of a login node in the cluster. Using this argument allows connection to a login node in the cluster with DCV enabled.

> **Note**
>
> This argument is added in AWS ParallelCluster version 3.11.0.

`--region, -r` *`REGION`*

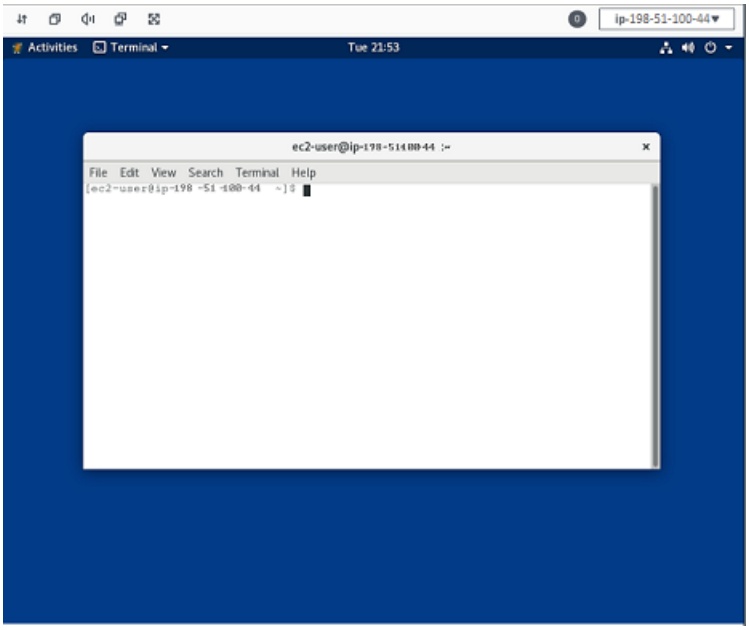
Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

`--show-url`

Prints the URL that would be used for the DCV connection and exits.

Example using AWS ParallelCluster version 3.11

```
$ pcluster dcv-connect -n cluster-3Dcv --login-node-ip 198.51.100.44 -r us-east-1 --
key-path /home/user/.ssh/key.pem
```



`pcluster delete-cluster`

Starts the deletion of a cluster.

```
pcluster delete-cluster [-h]
                 --cluster-name CLUSTER_NAME
                [--debug]
                [--query QUERY]
                [--region REGION]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster delete-cluster`.

`--cluster-name, -n` *`CLUSTER_NAME`*

Specifies the name of the cluster.

`--debug`

Enables debug logging.

`--query` *`QUERY`*

Specifies the JMESPath query to perform on the output.

`--region, -r` *`REGION`*

Specifies the AWS Region to use. The Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster delete-cluster -n cluster-v3
   {
  "cluster": {
    "clusterName": "cluster-v3",
    "cloudformationStackStatus": "DELETE_IN_PROGRESS",
    "cloudformationStackArn": "arn:aws:cloudformation:us-east-1:123456789012:stack/
cluster-v3/1234abcd-56ef-78gh-90ij-abcd1234efgh",
    "region": "us-east-1",
    "version": "3.1.4",
    "clusterStatus": "DELETE_IN_PROGRESS"
  }
}
```

`pcluster delete-cluster-instances`

Initiate the forced termination of all cluster compute nodes. This does not work with AWS Batch clusters.

```
pcluster delete-cluster-instances [-h]
                 --cluster-name CLUSTER_NAME
                [--debug]
                [--force FORCE]
                [--query QUERY]
                [--region REGION]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster delete-cluster-instances`.

`--cluster-name, -n CLUSTER_NAME`

Specifies the name of the cluster.

`--debug`

Enables debug logging.

`--force FORCE`

When `true`, forces the deletion by ignoring validation errors. (Defaults to `false`.)

`--query QUERY`

Specifies the JMESPath query to perform on the output.

`--region, -r REGION`

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

```
$ pcluster delete-cluster-instances -n cluster-v3
```

pcluster delete-image

Starts the deletion of the custom AWS ParallelCluster image.

```
pcluster delete-image [-h]
                 --image-id IMAGE_ID
                [--debug]
```

```
                [--force FORCE]
                [--query QUERY]
                [--region REGION]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster delete-image`.

`--image-id, -i IMAGE_ID`

Specifies the id of the image that will be deleted.

`--debug`

Enables debug logging.

`--force FORCE`

When `true`, forces the deletion in case there are instances using the AMI or if the AMI is shared. (Defaults to `false`.)

`--query QUERY`

Specifies the JMESPath query to perform on the output.

`--region, -r REGION`

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster delete-image --image-id custom-alinux2-image
{
  "image": {
    "imageId": "custom-alinux2-image",
    "imageBuildStatus": "DELETE_IN_PROGRESS",
    "region": "us-east-1",
    "version": "3.1.4"
  }
}
```

pcluster describe-cluster

Get detailed information about a cluster.

```
pcluster describe-cluster [-h]
                 --cluster-name CLUSTER_NAME
                [--debug]
                [--query QUERY]
                [--region REGION]
```

Named arguments

-h, --help

Shows the help text for `pcluster describe-cluster`.

--cluster-name, -n *CLUSTER_NAME*

Specifies the name of the cluster.

--debug

Enables debug logging.

--query *QUERY*

Specifies the JMESPath query to perform on the output.

--region, -r *REGION*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Examples using AWS ParallelCluster version 3.1.4:

Describe cluster details:

```
$ pcluster describe-cluster -n cluster-v3
{
  "creationTime": "2022-07-12T17:19:16.101Z",
  "headNode": {
    "launchTime": "2022-07-12T17:22:21.000Z",
    "instanceId": "i-1234567890abcdef0",
    "publicIpAddress": "198.51.100.44",
```

```
    "instanceType": "t2.micro",
    "state": "running",
    "privateIpAddress": "192.0.2.0.196"
  },
  "loginNodes": [
      {
          "status": "active",
          "poolName": "pool1",
          "address": "cluster-v3-eMr9BYRKZVDa-e5bb34f40b24f51d.elb.us-
east-1.amazonaws.com",
          "scheme": "internet-facing",
          "healthyNodes": 1,
          "unhealthyNodes": 0
      },
      {
          "status": "active",
          "poolName": "pool2",
          "address": "cluster-v3-PaQ7GgC27sic-aba10c890247b36b.elb.us-
east-1.amazonaws.com",
          "scheme": "internet-facing",
          "healthyNodes": 1,
          "unhealthyNodes": 0
      }
  ],
  "version": "3.1.4",
  "clusterConfiguration": {
    "url": "https://parallelcluster-e5ca74255d6c3886-v1-do-not-delete..."
  },
  "tags": [
    {
      "value": "3.11",
      "key": "parallelcluster:version"
    }
  ],
  "cloudFormationStackStatus": "CREATE_COMPLETE",
  "clusterName": "cluster-v3",
  "computeFleetStatus": "RUNNING",
  "cloudformationStackArn": "arn:aws:cloudformation:us-east-1:123456789012:stack/
cluster-v3/1234abcd-56ef-78gh-90ij-abcd1234efgh",
  "lastUpdatedTime": "2022-07-12T17:19:16.101Z",
  "region": "us-east-1",
  "clusterStatus": "CREATE_COMPLETE"
}
```

Use `describe-cluster` to retrieve the cluster configuration:

```
$ curl -o - $(pcluster describe-cluster -n cluster-v3 --query clusterConfiguration.url
 | xargs echo)
Region: us-east-1
Image:
  Os: alinux2
HeadNode:
  InstanceType: t2.micro
  Networking:
    SubnetId: subnet-abcdef01234567890
  Ssh:
    KeyName: adpc
  Iam:
    S3Access:
      - BucketName: cluster-v3-bucket
        KeyName: logs
        EnableWriteAccess: true
Scheduling:
  Scheduler: slurm
  SlurmQueues:
  - Name: queue1
    ComputeResources:
    - Name: t2micro
      InstanceType: t2.micro
      MinCount: 0
      MaxCount: 10
    Networking:
      SubnetIds:
      - subnet-021345abcdef6789
```

pcluster describe-cluster-instances

Describe the instances in a cluster.

```
pcluster describe-cluster-instances [-h]
                 --cluster-name CLUSTER_NAME
                [--debug]
                [--next-token NEXT_TOKEN]
                [--node-type {HeadNode,ComputeNode,LoginNode}]
                [--query QUERY]
                [--queue-name QUEUE_NAME]
```

```
                [--region REGION]
```

Named arguments

-h, --help

Shows the help text for `pcluster describe-cluster-instances`.

--cluster-name, -n *CLUSTER_NAME*

Specifies the name of the cluster.

--debug

Enables debug logging.

--next-token *NEXT_TOKEN*

The token for the next set of results.

--node-type {HeadNode,ComputeNode,LoginNode}

Specifies the node types to list. Supported values are `HeadNode`, `ComputeNode` and `LoginNode`. If this parameter is not specified, the `HeadNode`, `ComputeNode` and `LoginNodea` instances are described.

--query *QUERY*

Specifies the JMESPath query to perform on the output.

--queue-name *QUEUE_NAME*

Specifies the name of the queue to list. If this parameter is not specified, instances in all queues are described.

--region, -r *REGION*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster describe-cluster-instances -n cluster-v3
```

```
{
  "instances": [
    {
      "launchTime": "2022-07-12T17:22:21.000Z",
      "instanceId": "i-1234567890abcdef0",
      "publicIpAddress": "198.51.100.44",
      "instanceType": "t2.micro",
      "state": "running",
      "nodeType": "HeadNode",
      "privateIpAddress": "192.0.2.0.196"
    },
    {
      "launchTime": "2022-07-12T17:37:42.000Z",
      "instanceId": "i-021345abcdef6789",
      "queueName": "queue1",
      "publicIpAddress": "198.51.100.44",
      "instanceType": "t2.micro",
      "state": "pending",
      "nodeType": "ComputeNode",
      "privateIpAddress": "192.0.2.0.196"
    },
    {
        "launchTime": "2022-07-12T17:37:42.000Z",
        "instanceId": "i-021345abcdef6789",
        "poolName": "pool1",
        "publicIpAddress": "198.51.100.44",
        "instanceType": "t2.micro",
        "state": "pending",
        "nodeType": "loginNode",
        "privateIpAddress": "192.0.2.0.196"
    }
  ]
}
```

pcluster describe-compute-fleet

Describe the status of the compute fleet.

```
pcluster describe-compute-fleet [-h]
                 --cluster-name CLUSTER_NAME
                [--debug]
                [--query QUERY]
                [--region REGION]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster describe-compute-fleet`.

`--cluster-name, -n` *`CLUSTER_NAME`*

Specifies the name of the cluster.

`--debug`

Enables debug logging.

`--query` *`QUERY`*

Specifies the JMESPath query to perform on the output.

`--region, -r` *`REGION`*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster describe-compute-fleet -n pcluster-v3
{
  "status": "RUNNING",
  "lastStatusUpdatedTime": "2022-07-12T17:24:26.000Z"
}
```

pcluster describe-image

Get detailed information about an image.

```
pcluster describe-image [-h]
                 --image-id IMAGE_ID
                [--debug]
                [--query QUERY]
                [--region REGION]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster describe-image`.

`--image-id, -i` *`IMAGE_ID`*

Specifies the ID of the image.

`--debug`

Enables debug logging.

`--query` *`QUERY`*

Specifies the JMESPath query to perform on the output.

`--region, -r` *`REGION`*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.2:

```
$ pcluster describe-image --image-id custom-alinux2-image
{
  "imageConfiguration: {
    "url": "https://parallelcluster-1234abcd5678-v1-do-not-delete.../configs/image-
config.yaml"
  },
  "imageId": "custom-alinux2-image",
  "creationTime": "2022-04-05T20:23:07.000Z"
  "imageBuildStatus": "BUILD_COMPLETE",
  "region": "us-east-1",
  "ec2AmiInfo": {
    "amiName": "custom-alinux2-image 2022-04-05T19-55-22.518Z",
    "amiId": "ami-1234abcd5678efgh",
    "description": "AWS ParallelCluster AMI for alinux2,
 kernel-4.14.268-205.500.amzn2.x86_64, lustre-2.10.8-5.amzn2.x86_64,
 efa-1.14.2-1.amzn2.x86_64, dcv-2021.3.11591-1.el7.x86_64, slurm-21-08-6-1",
    "state": "AVAILABLE",
  "tags": [
      {
```

```
        "value": "arn:aws:imagebuilder:us-east-1:123456789012:image/
parallelclusterimage-custom-alinux2-image/3.1.2/1",
        "key": "Ec2ImageBuilderArn"
      },
      {
        "value": "parallelcluster-1234abcd5678efgh-v1-do-not-delete",
        "key": "parallelcluster:amzn-s3-demo-bucket"
      },
      {
        "value": "custom-alinux2-image",
        "key": "parallelcluster:image_name"
      },
      {
        "value": "available",
        "key": "parallelcluster:build_status"
      },
      {
        "value": "s3://amzn-s3-demo-bucket/parallelcluster/3.1.2/images/custom-alinux2-
image-1234abcd5678efgh/configs/image-config.yaml",
        "key": "parallelcluster:build_config"
      },
      {
        "value": "Amazon EC2 Image Builder",
        "key": "CreatedBy"
      },
      {
        "value": "arn:aws:logs:us-east-1:123456789012:log-group:/aws/imagebuilder/
ParallelClusterImage-custom-alinux2-image",
        "key": "parallelcluster:build_log"
      },
      {
        "value": "4.14.268-205.500.amzn2.x86_64",
        "key": "parallelcluster:kernel_version"
      },
      {
        "value": "arn:aws:imagebuilder:us-east-1:444455556666:image/amazon-linux-2-
x86/2022.3.16/1",
        "key": "parallelcluster:parent_image"
      },
      {
        "value": "3.1.2",
        "key": "parallelcluster:version"
      },
      {
```

```
        "value": "0.5.14",
        "key": "parallelcluster:munge_version"
      },
      {
        "value": "21-08-6-1",
        "key": "parallelcluster:slurm_version"
      },
      {
        "value": "2021.3.11591-1.el7.x86_64",
        "key": "parallelcluster:dcv_version"
      },
      {
        "value": "alinux2-image",
        "key": "parallelcluster:image_id"
      },
      {
        "value": "3.2.3",
        "key": "parallelcluster:pmix_version"
      },
      {
        "value": "parallelcluster/3.13.2/images/alinux2-image-abcd1234efgh56781234",
        "key": "parallelcluster:s3_image_dir"
      },
      {
        "value": "1.14.2-1.amzn2.x86_64",
        "key": "parallelcluster:efa_version"
      },
      {
        "value": "alinux2",
        "key": "parallelcluster:os"
      },
      {
        "value": "aws-parallelcluster-cookbook-3.1.2",
        "key": "parallelcluster:bootstrap_file"
      },
      {
        "value": "1.8.23-10.amzn2.1.x86_64",
        "key": "parallelcluster:sudo_version"
      },
      {
        "value": "2.10.8-5.amzn2.x86_64",
        "key": "parallelcluster:lustre_version"
      }
    ],
```

```
    "architecture": "x86_64"
  },
  "version": "3.1.2"
}
```

pcluster export-cluster-logs

Export the logs of the cluster to a local `tar.gz` archive by passing through an Amazon S3 Bucket.

```
pcluster export-cluster-logs [-h]
                 --cluster-name CLUSTER_NAME
                [--bucket BUCKET_NAME]
                [--bucket-prefix BUCKET_PREFIX]
                [--debug]
                [--end-time END_TIME]
                [--filters FILTER [FILTER ...]]
                [--keep-s3-objects KEEP_S3_OBJECTS]
                [--output-file OUTPUT_FILE]
                [--region REGION]
                [--start-time START_TIME]
```

Note

The `export-cluster-logs` command waits for CloudWatch Logs to complete the export of logs, so it is expected to experience a period of time without any output.

Named arguments

`-h, --help`

Shows the help text for `pcluster export-cluster-logs`.

`--bucket BUCKET_NAME`

Specifies the name of the Amazon S3 bucket to export cluster logs data to. It must be in the same Region as the cluster.

Note

- You must add permissions to the Amazon S3 bucket policy to grant CloudWatch access. For more information, see Set permissions on an Amazon S3 bucket in the *CloudWatch Logs User Guide*.
- Starting with AWS ParallelCluster version **3.12.0**, the `--bucket` option is **optional**. If the option is not specified, either the AWS ParallelCluster regional default bucket (`parallelcluster-hash-v1-DO-NOT-DELETE`) will be used, or if the Amazon S3 bucket pointed to by `CustomS3Bucket` is specified in the cluster configuration, that will be used. If you do not specify the `--bucket` option and you use the default AWS ParallelCluster bucket, you cannot export logs to the `parallelcluster/` folder, because it is a protected folder reserved for internal use.

Important

If the AWS ParallelCluster default bucket is used, pcluster will take care of configuring the bucket policy. If you customized the bucket policy and then upgrade to AWS ParallelCluster version **3.12.0**, the bucket policy will be overridden and you will need to reapply your changes.

`--cluster-name, -n` *`CLUSTER_NAME`*

Specifies the name of the cluster.

`--bucket-prefix` *`BUCKET_PREFIX`*

Specifies the path in the Amazon S3 bucket where exported logs data is to be stored.

By default, the bucket-prefix is:

```
cluster-name-logs-202209061743.tar.gz
```

`202209061743` is an example of the time in `%Y%m%d%H%M` format.

> **Note**
>
> Starting with AWS ParallelCluster version **3.12.0**, if you don't specify the `--bucket` option and you use the default AWS ParallelCluster bucket, you cannot export logs to the `parallelcluster/` folder, because it is a protected folder reserved for internal use.

`--debug`

Enables debug logging.

`--end-time` *`END_TIME`*

Specifies the end of the time range to collect log events, expressed in ISO 8601 format (YYYY-MM-DDThh:mm:ssZ, for example `2021-01-01T20:00:00Z`'). Events with a timestamp equal to or later than this time are not included. Time elements (e.g. minutes and seconds) may be omitted. The default value is the current time.

`--filters` *`FILTER`* **`[`***`FILTER`* **`...]`**

Specifies filters for the log. Format: `Name=a,Values=1 Name=b,Values=2,3`. Supported filters are:

`private-dns-name`

Specifies the short form of the private DNS name of the instance (e.g. `ip-10-0-0-101`).

`node-type`

Specifies the node type, the only accepted value for this filter is `HeadNode`.

`--keep-s3-objects` *`KEEP_S3_OBJECTS`*

If `true`, the exported objects exports to Amazon S3 are kept. (Defaults to `false`.)

`--output-file` *`OUTPUT_FILE`*

Specifies the file path to save the log archive to. If this is provided, then the logs are saved locally. Otherwise they are uploaded to Amazon S3 with the URL returned in the output. Default is to upload to Amazon S3.

`--region, -r REGION`

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

`--start-time START_TIME`

Specifies the start of the time range, expressed in ISO 8601 format (`YYYY-MM-DDThh:mm:ssZ`, for example `2021-01-01T20:00:00Z`). Log events with a timestamp equal to this time or later than this time are included. If not specified, the default is the time the cluster was created.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster export-cluster-logs --bucket cluster-v3-bucket -n cluster-v3
{
  "url": "https://cluster-v3-bucket..."
}
```

Cannot retrieve the logs?

If you cannot retrieve the logs using the `export-cluster-logs` command, then do one of the following:

- Retrieve the logs manually from the CloudWatch log group of the cluster.
- If the log group is empty, SSH into cluster nodes and retrieve the logs listed in Troubleshooting node initialization issues.
- If cluster nodes are not accessible because the cluster failed to create, then recreate the cluster with option `--rollback-on-failure false` and retrieve the logs from the nodes.

`pcluster export-image-logs`

Export the logs of the image builder stack to a local `tar.gz` archive by passing through an Amazon S3 Bucket.

```
pcluster export-image-logs [-h]
                 --image-id IMAGE_ID
                [--bucket BUCKET]
                [--bucket-prefix BUCKET_PREFIX]
```

```
                [--debug]
                [--end-time END_TIME]
                [--keep-s3-objects KEEP_S3_OBJECTS]
                [--output-file OUTPUT_FILE]
                [--region REGION]
                [--start-time START_TIME]
```

Note

The `export-image-logs` command waits for CloudWatch Logs to complete the export of logs, so it is expected to experience a period of time without any output.

Named arguments

`-h, --help`

Shows the help text for `pcluster export-image-logs`.

`--bucket BUCKET_NAME`

Specifies the Amazon S3 bucket name to export image build logs to. It must be in the same Region as the image.

Note

- You must add permissions to the Amazon S3 bucket policy to grant CloudWatch access. For more information, see Set permissions on an Amazon S3 bucket in the *CloudWatch Logs User Guide*.
- Starting with AWS ParallelCluster version **3.12.0**, the `--bucket` option is **optional**. If the option is not specified, either the AWS ParallelCluster regional default bucket (`parallelcluster-hash-v1-DO-NOT-DELETE`) will be used, or if the `CustomS3Bucket` is specified in the image configuration, that will be used.

Important

If the AWS ParallelCluster default bucket is used, pcluster will take care of configuring the bucket policy. If you customize the bucket policy before you upgrade to AWS

ParallelCluster version **3.12.0**, the bucket policy will be overridden and you will need to reapply the changes.

`--image-id, -i` *`IMAGE_ID`*

The image ID whose logs will be exported.

`--bucket-prefix` *`BUCKET_PREFIX`*

Specifies the path in the Amazon S3 bucket where exported logs data is to be stored.

By default, the bucket-prefix is:

```
ami-id-logs-202209061743.tar.gz
```

`202209061743` is the current time in `%Y%m%d%H%M` format.

Note

Starting with AWS ParallelCluster version **3.12.0**, if you don't specify the `--bucket` option and use the default AWS ParallelCluster bucket, you cannot export logs to the `parallelcluster/` folder, because it is a protected folder reserved for internal use.

`--debug`

Enables debug logging.

`--end-time` *`END_TIME`*

Specifies the end of the time range to collect log events, expressed in ISO 8601 format (YYYY-MM-DDThh:mm:ssZ, for example `2021-01-01T20:00:00Z`'). Events with a timestamp equal to or later than this time are not included. Time elements (e.g. minutes and seconds) may be omitted. The default value is the current time.

`--keep-s3-objects` *`KEEP_S3_OBJECTS`*

If `true`, the exported objects exports to Amazon S3 are kept. (Defaults to `false`.)

--output-file *OUTPUT_FILE*

Specifies the file path to save the log archive to. If this is provided, then the logs are saved locally. Otherwise they are uploaded to Amazon S3 with the URL returned in the output. Default is to upload to Amazon S3.

--region, -r *REGION*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

--start-time *START_TIME*

Specifies the start of the time range, expressed in ISO 8601 format (`YYYY-MM-DDThh:mm:ssZ`, for example `2021-01-01T20:00:00Z`). Log events with a timestamp equal to this time or later than this time are included. If not specified, the default is the time the cluster was created.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster export-image-logs --bucket image-v3-bucket --image-id ami-1234abcd5678efgh
{
  "url": "https://image-v3-bucket..."
}
```

pcluster get-cluster-log-events

Retrieve the events associated with a log stream.

```
pcluster get-cluster-log-events [-h]
                 --cluster-name CLUSTER_NAME
                 --log-stream-name LOG_STREAM_NAME
                [--debug]
                [--end-time END_TIME]
                [--limit LIMIT]
                [--next-token NEXT_TOKEN]
                [--query QUERY]
                [--region REGION]
                [--start-from-head START_FROM_HEAD]
                [--start-time START_TIME]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster get-cluster-log-events`.

`--cluster-name, -n` ***`CLUSTER_NAME`***

Specifies the name of the cluster.

`--log-stream-name` ***`LOG_STREAM_NAME`***

Specifies the name of the log stream. You can use the `list-cluster-log-streams` command to retrieve a log stream associated with an event or events.

`--debug`

Enables debug logging.

`--end-time` ***`END_TIME`***

Specifies the end of the time range, expressed in ISO 8601 format (`YYYY-MM-DDThh:mm:ssZ`, for example `2021-01-01T20:00:00Z`). Events with a timestamp equal to or later than this time are not included.

`--limit` ***`LIMIT`***

Specifies the maximum number of log events returned. If a value is not specified, the maximum is as many log events as can fit in a response size of 1 MB, up to 10,000 log events.

`--next-token` ***`NEXT_TOKEN`***

The token for the next set of results.

`--query` ***`QUERY`***

Specifies the JMESPath query to perform on the output.

`--region, -r` ***`REGION`***

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

`--start-from-head` ***`START_FROM_HEAD`***

If the value is `true`, the earliest log events are returned first. If the value is `false`, the most recent log events are returned first. (Defaults to `false`.)

`--start-time` *`START_TIME`*

Specifies the start of the time range, expressed in ISO 8601 format (`YYYY-MM-DDThh:mm:ssZ`, for example `2021-01-01T20:00:00Z`). Events with a timestamp equal to this time or later than this time are included.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster get-cluster-log-events \
    -c cluster-v3 \
    -r us-east-1 \
    --log-stream-name ip-198-51-100-44.i-1234567890abcdef0.clustermgtd \
    --limit 3
{
  "nextToken": "f/36966906399261933213029082268132291405859205452101451780/s",
  "prevToken": "b/36966906399239632467830551644990755687586557090595471362/s",
  "events": [
    {
      "message": "2022-07-12 19:16:53,379 - [slurm_plugin.clustermgtd:_maintain_nodes]
 - INFO - Performing node maintenance actions",
      "timestamp": "2022-07-12T19:16:53.379Z"
    },
    {
      "message": "2022-07-12 19:16:53,380 - [slurm_plugin.clustermgtd:_maintain_nodes]
 - INFO - Following nodes are currently in replacement: (x0) []",
      "timestamp": "2022-07-12T19:16:53.380Z"
    },
    {
      "message": "2022-07-12 19:16:53,380 -
 [slurm_plugin.clustermgtd:_terminate_orphaned_instances] - INFO - Checking for
 orphaned instance",
      "timestamp": "2022-07-12T19:16:53.380Z"
    }
  ]
}
```

`pcluster get-cluster-stack-events`

Retrieve the events associated with the stack for the specified cluster.

Note

Starting in version 3.6.0, AWS ParallelCluster uses nested stacks to create the resources associated with queues and compute resources. The `GetClusterStackEvents` API and the `pcluster get-cluster-stack-events` command only return the cluster main stack events. You can view the cluster stack events, including those related to queues and compute resources, in the CloudFormation console.

```
pcluster get-cluster-stack-events [-h]
                 --cluster-name CLUSTER_NAME
                [--debug]
                [--next-token NEXT_TOKEN]
                [--query QUERY]
                [--region REGION]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster get-cluster-stack-events`.

`--cluster-name, -n CLUSTER_NAME`

Specifies the name of the cluster.

`--debug`

Enables debug logging.

`--next-token NEXT_TOKEN`

The token for the next set of results.

`--query QUERY`

Specifies the JMESPath query to perform on the output.

`--region, -r REGION`

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster get-cluster-stack-events \
    -n cluster-v3 \
    -r us-east-1 \
    --query "events[0]"
{
  "eventId": "1234abcd-56ef-78gh-90ij-abcd1234efgh",
  "physicalResourceId": "arn:aws:cloudformation:us-east-1:123456789012:stack/cluster-
v3/1234abcd-56ef-78gh-90ij-abcd1234efgh",
  "resourceStatus": "CREATE_COMPLETE",
  "stackId": "arn:aws:cloudformation:us-east-1:123456789012:stack/cluster-
v3/1234abcd-56ef-78gh-90ij-abcd1234efgh",
  "stackName": "cluster-v3",
  "logicalResourceId": "cluster-v3",
  "resourceType": "AWS::CloudFormation::Stack",
  "timestamp": "2022-07-12T18:29:12.140Z"
}
```

pcluster get-image-log-events

Retrieve the log events associated with an image build.

```
pcluster get-image-log-events [-h]
                 --image-id IMAGE_ID
                 --log-stream-name LOG_STREAM_NAME
                [--debug]
                [--end-time END_TIME]
                [--limit LIMIT]
                [--next-token NEXT_TOKEN]
                [--query QUERY]
                [--region REGION]
                [--start-from-head START_FROM_HEAD]
                [--start-time START_TIME]
```

Named arguments

-h, --help

Shows the help text for `pcluster get-image-log-events`.

--image-id, -i *IMAGE_ID*

Specifies the Id of the image.

--log-stream-name *LOG_STREAM_NAME*

Specifies the name of the log stream. You can use the `list-image-log-streams` command to retrieve a log stream associated with an event or events.

--debug

Enables debug logging.

--end-time *END_TIME*

Specifies the end of the time range, expressed in ISO 8601 format (`YYYY-MM-DDThh:mm:ssZ`, for example `2021-01-01T20:00:00Z`). Events with a timestamp equal to or later than this time are not included.

--limit *LIMIT*

Specifies the maximum number of log events returned. If a value is not specified, the maximum is as many log events as can fit in a response size of 1 MB, up to 10,000 log events.

--next-token *NEXT_TOKEN*

The token for the next set of results.

--query *QUERY*

Specifies the JMESPath query to perform on the output.

--region, -r *REGION*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

--start-from-head *START_FROM_HEAD*

If the value is `true`, the earliest log events are returned first. If the value is `false`, the most recent log events are returned first. (Defaults to `false`.)

--start-time *START_TIME*

Specifies the start of the time range, expressed in ISO 8601 format (`YYYY-MM-DDThh:mm:ssZ`, for example `2021-01-01T20:00:00Z`). Events with a timestamp equal to or later than this time are included.

Example using AWS ParallelCluster version 3.1.2:

```
$ pcluster get-image-log-events --image-id custom-alinux2-image --region us-east-1 --
log-stream-name 3.1.2/1 --limit 3
   {
  "nextToken": "f/36778317771100849897800729464621464113270312017760944178/s",
  "prevToken": "b/36778317766952911290874033560295820514557716777648586800/s",
  "events": [
    {
      "message": "ExecuteBash: FINISHED EXECUTION",
      "timestamp": "2022-04-05T22:13:26.633Z"
    },
    {
      "message": "Document arn:aws:imagebuilder:us-east-1:123456789012:component/
parallelclusterimage-test-1234abcd-56ef-78gh-90ij-abcd1234efgh/3.1.2/1",
      "timestamp": "2022-04-05T22:13:26.741Z"
    },
    {
      "message": "TOE has completed execution successfully",
      "timestamp": "2022-04-05T22:13:26.819Z"
    }
  ]
}
```

pcluster get-image-stack-events

Retrieve the events associated with the stack for the specified image build.

```
pcluster get-image-stack-events [-h]
                 --image-id IMAGE_ID
                [--debug]
                [--next-token NEXT_TOKEN]
                [--query QUERY]
                [--region REGION]
```

Named arguments

-h, --help

Shows the help text for `pcluster get-image-stack-events`.

--image-id, -i *IMAGE_ID*

Specifies the ID of the image.

--debug

Enables debug logging.

--next-token *NEXT_TOKEN*

The token for the next set of results.

--query *QUERY*

Specifies the JMESPath query to perform on the output.

--region, -r *REGION*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.2:

```
$ pcluster get-image-stack-events --image-id custom-alinux2-image --region us-east-1 --
query "events[0]"
   {
  "eventId": "ParallelClusterImage-CREATE_IN_PROGRESS-2022-04-05T21:39:24.725Z",
  "physicalResourceId": "arn:aws:imagebuilder:us-east-1:123456789012:image/
parallelclusterimage-custom-alinux2-image/3.1.2/1",
  "resourceStatus": "CREATE_IN_PROGRESS",
  "resourceStatusReason": "Resource creation Initiated",
  "resourceProperties": "{\"InfrastructureConfigurationArn\":
\"arn:aws:imagebuilder:us-east-1:123456789012:infrastructure-configuration/
parallelclusterimage-1234abcd-56ef-78gh-90ij-abcd1234efgh\",\"ImageRecipeArn
\":\"arn:aws:imagebuilder:us-east-1:123456789012:image-recipe/
parallelclusterimage-custom-alinux2-image/3.1.2\",\"DistributionConfigurationArn
\":\"arn:aws:imagebuilder:us-east-1:123456789012:distribution-
configuration/parallelclusterimage-1234abcd-56ef-78gh-90ij-abcd1234efgh\",
\"EnhancedImageMetadataEnabled\":\"false\",\"Tags\":{\"parallelcluster:image_name\":
\"custom-alinux2-image\",\"parallelcluster:image_id\":\"custom-alinux2-image\"}}",
  "stackId": "arn:aws:cloudformation:us-east-1:123456789012:stack/custom-alinux2-
image/1234abcd-56ef-78gh-90ij-abcd1234efgh",
  "stackName": "custom-alinux2-image",
  "logicalResourceId": "ParallelClusterImage",
```

```
  "resourceType": "AWS::ImageBuilder::Image",
  "timestamp": "2022-04-05T21:39:24.725Z"
}
```

pcluster list-clusters

Retrieve the list of existing clusters.

```
pcluster list-clusters [-h]
                [--cluster-status {CREATE_IN_PROGRESS,CREATE_FAILED,CREATE_COMPLETE,
                           DELETE_IN_PROGRESS,DELETE_FAILED,UPDATE_IN_PROGRESS,
                           UPDATE_COMPLETE,UPDATE_FAILED}
                           [{CREATE_IN_PROGRESS,CREATE_FAILED,CREATE_COMPLETE,
                             DELETE_IN_PROGRESS,DELETE_FAILED,UPDATE_IN_PROGRESS,
                             UPDATE_COMPLETE,UPDATE_FAILED} ...]]
                [--debug]
                [--next-token NEXT_TOKEN]
                [--query QUERY]
                [--region REGION]
```

Named arguments

-h, --help

Shows the help text for `pcluster list-clusters`.

--cluster-status {CREATE_IN_PROGRESS, CREATE_FAILED, CREATE_COMPLETE, DELETE_IN_PROGRESS, DELETE_FAILED, UPDATE_IN_PROGRESS, UPDATE_COMPLETE, UPDATE_FAILED} [{CREATE_IN_PROGRESS, CREATE_FAILED, CREATE_COMPLETE, DELETE_IN_PROGRESS, DELETE_FAILED, UPDATE_IN_PROGRESS, UPDATE_COMPLETE, UPDATE_FAILED} ...]

Specifies the list of cluster statuses to filter for. (Defaults to `all`.)

--debug

Enables debug logging.

--next-token *NEXT_TOKEN*

The token for the next set of results.

--query *QUERY*

Specifies the JMESPath query to perform on the output.

`--region, -r REGION`

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster list-clusters
{
  "clusters": [
    {
      "clusterName": "cluster-v3",
      "cloudformationStackStatus": "CREATE_COMPLETE",
      "cloudformationStackArn": "arn:aws:cloudformation:us-east-1:123456789012:stack/
cluster-v3/1234abcd-56ef-78gh-90ij-abcd1234efgh",
      "region": "us-east-1",
      "version": "3.1.4",
      "clusterStatus": "CREATE_COMPLETE"
    }
  ]
}
```

pcluster list-cluster-log-streams

Retrieve the list of log streams associated with a cluster.

```
pcluster list-cluster-log-streams [-h]
                 --cluster-name CLUSTER_NAME
                [--filters FILTERS [FILTERS ...]]
                [--next-token NEXT_TOKEN] [--debug]
                [--query QUERY]
                [--region REGION]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster list-cluster-log-streams`.

--cluster-name, -n *CLUSTER_NAME*

Specifies the name of the cluster.

--debug

Enables debug logging.

--filters *FILTERS* [*FILTERS* ...]

Specifies filters for the log streams. Format: `Name=a,Values=1 Name=b,Values=2,3`. Supported filters are:

`private-dns-name`

Specifies the short form of the private DNS name of the instance (e.g. `ip-10-0-0-101`).

`node-type`

Specifies the node type, the only accepted value for this filter is `HeadNode`.

--next-token *NEXT_TOKEN*

The token for the next set of results.

--query *QUERY*

Specifies the JMESPath query to perform on the output.

--region, -r *REGION*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster list-cluster-log-streams \
    -n cluster-v3 \
    -r us-east-1 \
    --query 'logStreams[*].logStreamName'
[
  "ip-172-31-58-205.i-1234567890abcdef0.cfn-init",
  "ip-172-31-58-205.i-1234567890abcdef0.chef-client",
  "ip-172-31-58-205.i-1234567890abcdef0.cloud-init",
```

```
  "ip-172-31-58-205.i-1234567890abcdef0.clustermgtd",
  "ip-172-31-58-205.i-1234567890abcdef0.slurmctld",
  "ip-172-31-58-205.i-1234567890abcdef0.supervisord",
  "ip-172-31-58-205.i-1234567890abcdef0.system-messages"
]
```

pcluster list-images

Retrieve the list of existing custom images.

```
pcluster list-images [-h]
                 --image-status {AVAILABLE,PENDING,FAILED}
                [--debug]
                [--next-token NEXT_TOKEN]
                [--query QUERY]
                [--region REGION]
```

Named arguments

-h, --help

Shows the help text for `pcluster list-images`.

--image-status {AVAILABLE,PENDING,FAILED}

Filter returned images by the status provided.

--debug

Enables debug logging.

--next-token *NEXT_TOKEN*

The token for the next set of results.

--query *QUERY*

Specifies the JMESPath query to perform on the output.

--region, -r *REGION*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.2:

```
$ pcluster list-images --image-status AVAILABLE
{
  "images": [
    {
      "imageId": "custom-alinux2-image",
      "imageBuildStatus": "BUILD_COMPLETE",
      "ec2AmiInfo": {
        "amiId": "ami-1234abcd5678efgh"
      },
      "region": "us-east-1",
      "version": "3.1.2"
    }
  ]
}
```

pcluster list-image-log-streams

Retrieve the list of log streams associated with an image.

```
pcluster list-image-log-streams [-h]
                 --image-id IMAGE_ID
                [--next-token NEXT_TOKEN] [--debug]
                [--query QUERY]
                [--region REGION]
```

Named arguments

-h, --help

Shows the help text for `pcluster list-image-log-streams`.

--image-id, -i *IMAGE_ID*

Specifies the ID of the image.

--debug

Enables debug logging.

--next-token *NEXT_TOKEN*

The token for the next set of results.

--query *QUERY*

Specifies the JMESPath query to perform on the output.

--region, -r *REGION*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.2:

```
$ pcluster list-image-log-streams --image-id custom-alinux2-image --region us-east-1 --
query 'logStreams[*].logStreamName'
[
  "3.0.0/1",
  "3.1.2/1"
]
```

pcluster list-official-images

Describe official AWS ParallelCluster AMIs.

```
pcluster list-official-images [-h]
                [--architecture ARCHITECTURE]
                [--debug]
                [--os OS]
                [--query QUERY]
                [--region REGION]
```

Named arguments

-h, --help

Shows the help text for `pcluster list-official-images`.

--architecture *ARCHITECTURE*

Specifies the architecture to use to filter the results. If this parameter is not specified, all architectures are returned.

--debug

Enables debug logging.

--os *OS*

Specifies the operating system to use to filter the results. If this parameter is not specified, all operating systems are returned.

--query *QUERY*

Specifies the JMESPath query to perform on the output.

--region, -r *REGION*

Specifies the AWS Region to use. The AWS Region must be specified, using the Region setting in the image configuration file, the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.2:

```
$ pcluster list-official-images
{
  "images": [
    {
      "amiId": "ami-015cfeb4e0d6306b2",
      "os": "ubuntu2004",
      "name": "aws-parallelcluster-3.1.2-ubuntu-2004-lts-hvm-x86_64-202202261505
 2022-02-26T15-08-34.759Z",
      "version": "3.1.2",
      "architecture": "x86_64"
    },
    {
      "amiId": "ami-036f23237ce49d25b",
      "os": "ubuntu2204",
      "name": "aws-parallelcluster-3.1.2-ubuntu-1804-lts-hvm-x86_64-202202261505
 2022-02-26T15-08-17.558Z",
      "version": "3.1.2",
      "architecture": "x86_64"
    },
    {
      "amiId": "ami-09e5327e694d89ef4",
      "os": "ubuntu2004",
      "name": "aws-parallelcluster-3.1.2-ubuntu-2004-lts-hvm-arm64-202202261505
 2022-02-26T15-08-45.736Z",
      "version": "3.1.2",
      "architecture": "arm64"
    },
```

```
    {
      "amiId": "ami-0b9b0874c35f626ae",
      "os": "alinux2",
      "name": "aws-parallelcluster-3.1.2-amzn2-hvm-x86_64-202202261505
 2022-02-26T15-08-31.311Z",
      "version": "3.1.2",
      "architecture": "x86_64"
    },
    {
      "amiId": "ami-0d0de4f95f56374bc",
      "os": "alinux2",
      "name": "aws-parallelcluster-3.1.2-amzn2-hvm-arm64-202202261505
 2022-02-26T15-08-46.088Z",
      "version": "3.1.2",
      "architecture": "arm64"
    },
    {
      "amiId": "ami-0ebf7bc54b8740dc6",
      "os": "ubuntu2204",
      "name": "aws-parallelcluster-3.1.2-ubuntu-1804-lts-hvm-arm64-202202261505
 2022-02-26T15-08-45.293Z",
      "version": "3.1.2",
      "architecture": "arm64"
    }
  ]
}
```


pcluster ssh

Runs a `ssh` command with the cluster user name and IP address pre-populated. Arbitrary arguments are appended to the end of the `ssh` command line.

```
pcluster ssh [-h]
                 --cluster-name CLUSTER_NAME
               [--debug]
               [--dryrun DRYRUN]
               [--region REGION]
```

Named arguments

-h, --help

Shows the help text for `pcluster ssh`.

`--cluster-name, -n` *`CLUSTER_NAME`*

Specifies the name of the cluster to connect to.

`--debug`

Enables debug logging.

`--dryrun` *`DRYRUN`*

When `true`, prints the command line that would be run and exits. (Defaults to `false`.)

`--region, -r` *`REGION`*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example:

```
$ pcluster ssh --cluster-name mycluster -i ~/.ssh/id_rsa
```

Runs an `ssh` command with the user name and IP address of the cluster pre-populated:

```
ssh ec2-user@1.1.1.1 -i ~/.ssh/id_rsa
```

`pcluster update-cluster`

Updates an existing cluster to match the settings of a specified configuration file.

```
pcluster update-cluster [-h]
                 --cluster-configuration CLUSTER_CONFIGURATION
                 --cluster-name CLUSTER_NAME
                [--debug]
                [--dryrun DRYRUN]
                [--force-update FORCE_UPDATE]
                [--query QUERY]
                [--region REGION]
                [--suppress-validators SUPPRESS_VALIDATORS [SUPPRESS_VALIDATORS ...]]
                [--validation-failure-level {INFO,WARNING,ERROR}]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster update-cluster`.

`--cluster-configuration, -c` *`CLUSTER_CONFIGURATION`*

Specifies the YAML cluster configuration file.

`--cluster-name, -n` *`CLUSTER_NAME`*

Specifies the name of the cluster.

`--debug`

Enables debug logging.

`--dryrun` *`DRYRUN`*

When `true`, performs the validation without updating the cluster and creating any resources. It can be used to validate the image configuration and update requirements. (Defaults to `false`.)

`--force-update` *`FORCE_UPDATE`*

When `true`, forces the update by ignoring the update validation errors. (Defaults to `false`.)

`--query` *`QUERY`*

Specifies the JMESPath query to perform on the output.

`--region, -r` *`REGION`*

Specifies the AWS Region to use. The AWS Region must be specified, using the `Region` setting in the cluster configuration file, the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

`--suppress-validators` *`SUPPRESS_VALIDATORS`* `[`*`SUPPRESS_VALIDATORS ...`*`]`

Identifies one or more config validators to suppress.

Format: (`ALL|type:[A-Za-z0-9]+`)

`--validation-failure-level` *`{INFO,WARNING,ERROR}`*

Specifies the level of validation failures reported for update.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster update-cluster -c cluster-config.yaml -n cluster-v3 -r us-east-1
{
  "cluster": {
    "clusterName": "cluster-v3",
    "cloudformationStackStatus": "UPDATE_IN_PROGRESS",
    "cloudformationStackArn": "arn:aws:cloudformation:us-east-1:123456789012:stack/
cluster-v3/1234abcd-56ef-78gh-90ij-abcd1234efgh",
    "region": "us-east-1",
    "version": "3.1.4",
    "clusterStatus": "UPDATE_IN_PROGRESS"
  },
  "changeSet": [
    {
      "parameter": "HeadNode.Iam.S3Access",
      "requestedValue": {
        "BucketName": "amzn-s3-demo-bucket1",
        "KeyName": "output",
        "EnableWriteAccess": false
      }
    },
    {
      "parameter": "HeadNode.Iam.S3Access",
      "currentValue": {
        "BucketName": "amzn-s3-demo-bucket2",
        "KeyName": "logs",
        "EnableWriteAccess": true
      }
    }
  ]
}
```

pcluster update-compute-fleet

Updates the status of the cluster compute fleet.

```
pcluster update-compute-fleet [-h]
                 --cluster-name CLUSTER_NAME
                 --status {START_REQUESTED,STOP_REQUESTED,ENABLED,DISABLED}

                [--debug]
                [--query QUERY]
                [--region REGION]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster update-compute-fleet`.

`--cluster-name, -n` *`CLUSTER_NAME`*

Specifies the name of the cluster.

`--status {START_REQUESTED,STOP_REQUESTED,ENABLED,DISABLED}`

Specifies the status applied to the cluster compute fleet. The statuses `START_REQUESTED` and `STOP_REQUESTED` correspond to the Slurm scheduler while the statuses `ENABLED` and `DISABLED` correspond to the AWS Batch scheduler.

`--debug`

Enables debug logging.

`--query` *`QUERY`*

Specifies the JMESPath query to perform on the output.

`--region, -r` *`REGION`*

Specifies the AWS Region to use. The AWS Region must be specified, using the `AWS_DEFAULT_REGION` environment variable, the `region` setting in the `[default]` section of the `~/.aws/config` file, or the `--region` parameter.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster update-compute-fleet -n cluster-v3 --status STOP_REQUESTED
{
  "status": "STOP_REQUESTED",
  "lastStatusUpdatedTime": "2022-07-12T20:19:47.653Z"
}
```

`pcluster version`

Displays the version of AWS ParallelCluster.

```
pcluster version [-h] [--debug]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster version`.

`--debug`

Enables debug logging.

Example using AWS ParallelCluster version 3.1.4:

```
$ pcluster version
{
  "version": "3.1.4"
}
```

pcluster3-config-converter

Reads a AWS ParallelCluster version 2 configuration file and writes a AWS ParallelCluster version 3 configuration file.

```
pcluster3-config-converter [-h]
                [-t CLUSTER_TEMPLATE]
                [-c CONFIG_FILE]
                [--force-convert]
                [-o OUTPUT_FILE]
```

Named arguments

`-h, --help`

Shows the help text for `pcluster3-config-converter`.

`-t CLUSTER_TEMPLATE, --cluster-template CLUSTER_TEMPLATE`

Specifies the `[cluster]` section of the configuration file to convert. If not specified the script will look for the cluster-template parameter in the `[global]` section or will search for `[cluster default]`.

`-c CONFIG_FILE, --config-file CONFIG_FILE`

Specifies the AWS ParallelCluster version 2 configuration file to be read.

`--force-convert`

Enables a conversion even if one or more settings is not supported and not recommended.

`-o` *`OUTPUT_FILE`*`, --output-file` *`OUTPUT_FILE`*

Specifies the AWS ParallelCluster version 3 configuration file to be written. If this parameter is not specified, the configuration is written to stdout.

> **Note**
>
> The `pcluster3-config-converter` command was added in AWS ParallelCluster version 3.0.1.

Configuration files

AWS ParallelCluster uses YAML 1.1 files for configuration parameters.

Topics

Cluster configuration file

AWS ParallelCluster version 3 uses separate configuration files to control the definition of cluster infrastructure and the definition of custom AMIs. All configuration files use YAML 1.1 files. Detailed information for each of these configuration files is linked below. For some example configurations, see https://github.com/aws/aws-parallelcluster/tree/release-3.0/cli/tests/pcluster/example_configs.

These objects are used for the AWS ParallelCluster version 3 cluster configuration.

Topics

- Scheduling section
- SharedStorage section
- Iam section
- LoginNodes section
- Monitoring section
- Tags section
- AdditionalPackages section
- DirectoryService section
- DeploymentSettings section

Cluster configuration file properties

`Region` (**Optional**, `String`)

Specifies the AWS Region for the cluster. For example, `us-east-2`.

Update policy: If this setting is changed, the update is not allowed.

`CustomS3Bucket` (**Optional**, `String`)

Specifies the name of an Amazon S3 bucket that is created in your AWS account to store resources that are used are used by your clusters, such as the cluster configuration file, and to export logs. AWS ParallelCluster maintains one Amazon S3 bucket in each AWS Region that you create clusters in. By default, these Amazon S3 buckets are named `parallelcluster-hash-v1-DO-NOT-DELETE`.

Update policy: If this setting is changed, the update is not allowed. If you force the update, the new value will be ignored and the old value will be used.

`AdditionalResources` (**Optional**, `String`)

Defines an additional AWS CloudFormation template to launch along with the cluster. This additional template is used for creating resources that are outside of the cluster but are part of the cluster's lifecycle.

The value must be an HTTPS URL to a public template, with all parameters provided.

There is no default value.

Update policy: This setting can be changed during an update.

Imds section

(Optional) Specifies the global instance metadata service (IMDS) configuration.

```
Imds:
  ImdsSupport: string
```

Imds properties

`ImdsSupport` (**Optional**, `String`)

Specifies which IMDS versions are supported in the cluster nodes. Supported values are `v1.0` and `v2.0`. The default value is `v2.0`.

If `ImdsSupport` is set to `v1.0`, both IMDSv1 and IMDSv2 are supported.

If `ImdsSupport` is set to `v2.0`, only IMDSv2 is supported.

For more information, see Use IMDSv2 in the *Amazon EC2 User Guide for Linux instances*.

Update policy: If this setting is changed, the update is not allowed.

> **Note**
>
> Starting with AWS ParallelCluster 3.7.0, the `ImdsSupport` default value is `v2.0`. We recommend that you set `ImdsSupport` to `v2.0` and replace IMDSv1 with IMDSv2 in your custom actions calls.
> Support for `Imds` / `ImdsSupport` is added with AWS ParallelCluster version 3.3.0.

Image section

> **Note**
>
> Unsupported versions of the official AMIs distributed by AWS ParallelCluster will be made unavailable after 18 months of inactivity. These old images contain outdated software and cannot receive support in case of issues. We strongly suggest to move to the latest supported version.

(Required) Defines the operating system for the cluster.

```
Image:
  Os: string
  CustomAmi: string
```

`Image` properties

`Os` (**Required**, `String`)

Specifies the operating system to use for the cluster. The supported values are `alinux2`, `alinux2023`, `ubuntu2404`, `ubuntu2204`, `ubuntu2004`, `rhel8`, `rocky8`, `rhel9`, `rocky9`.

> **Note**
>
> RedHat Enterprise Linux 8.7 (`rhel8`) is added starting in AWS ParallelCluster version 3.6.0.
> If you configure your cluster to use `rhel`, the on-demand cost for any instance type is higher than when you configure your cluster to use other supported operation systems. For more information about pricing, see On-Demand Pricing and How is Red Hat Enterprise Linux on Amazon EC2 offered and priced?.
> RedHat Enterprise Linux 9 (rhel9) is added starting in AWS ParallelCluster version 3.9.0.

All AWS commercial Regions support all of the following operating systems.

Partition (AWS Regions)	alinux2	ubuntu2204 and ubuntu2004	ubuntu2404	rhel8	rhel9	alinux2023
Commercial (All AWS Regions not specifically mentioned)	True	True	True	True	True	True
AWS GovCloud (US-East) (us-gov-east-1)	True	True	True	True	True	True

Partition (AWS Regions)	alinux2	ubuntu2204 and ubuntu2004	ubuntu2404	rhel8	rhel9	alinux2023
AWS GovCloud (US-West) (us-gov-west-1)	True	True	True	True	True	True
China (Beijing) (cn-north-1)	True	True	True	True	True	True
China (Ningxia) (cn-northwest-1)	True	True	True	True	True	True

Update policy: If this setting is changed, the update is not allowed.

> **Note**
>
> AWS ParallelCluster 3.8.0 supports Rocky Linux 8, but pre-built Rocky Linux 8 AMIs (for x86 and ARM architectures) are not available. AWS ParallelCluster 3.8.0 supports creating clusters with Rocky Linux 8 using custom AMIs. For more information refer to Operating system considerations. AWS ParallelCluster 3.9.0 supports Rocky Linux 9, but pre-built Rocky Linux 9 AMIs (for x86 and ARM architectures) are not available. AWS ParallelCluster 3.9.0 supports creating clusters with Rocky Linux 9 using custom AMIs. For more information refer to Operating System Considerations.

`CustomAmi` (**Optional**, `String`)

Specifies the ID of a custom AMI to use for the head and compute nodes instead of the default AMI. For more information, see AWS ParallelCluster AMI customization.

If the custom AMI requires additional permissions for its launch, these permissions must be added to both the user and head node policies.

For example, if a custom AMI has an encrypted snapshot associated with it, the following additional policies are required in both the user and head node policies:

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:DescribeKey",
                "kms:ReEncrypt*",
                "kms:CreateGrant",
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:us-east-1:111122223333:key/<AWS_KMS_KEY_ID>"
            ]
        }
    ]
}
```

To build a RedHat Enterprise Linux custom AMI, you must configure the OS for installing the packages that are provided by the RHUI (AWS) repositories: `rhel-<version>-baseos-rhui-rpms`, `rhel-<version>-appstream-rhui-rpms`, and `codeready-builder-for-rhel-<version>-rhui-rpms`. Moreover, the repositories on the custom AMI must contain `kernel-devel` packages on the same version as the running kernel version. kernel.

Known limitations:

- Only RHEL 8.2 and later versions support FSx for Lustre.
- RHEL 8.7 kernel version 4.18.0-425.3.1.el8 doesn't support FSx for Lustre.
- Only RHEL 8.4 and later versions support EFA.
- AL23 doesn't support NICE DCV, as it doesn't include a graphical desktop environment, which is required to run NICE DCV. For more information, see the official NICE DCV documentation.

To troubleshoot custom AMI validation warnings, see Troubleshooting custom AMI issues.

Update policy: If this setting is changed, the update is not allowed.

HeadNode section

(Required) Specifies the configuration for the head node.

```
HeadNode:
  InstanceType: string
  Networking:
    SubnetId: string
    ElasticIp: string/boolean
    SecurityGroups:
      - string
    AdditionalSecurityGroups:
      - string
    Proxy:
      HttpProxyAddress: string
  DisableSimultaneousMultithreading: boolean
  Ssh:
    KeyName: string
    AllowedIps: string
  LocalStorage:
    RootVolume:
      Size: integer
      Encrypted: boolean
      VolumeType: string
      Iops: integer
      Throughput: integer
      DeleteOnTermination: boolean
    EphemeralVolume:
      MountDir: string
  SharedStorageType: string
  Dcv:
    Enabled: boolean
    Port: integer
    AllowedIps: string
  CustomActions:
    OnNodeStart:
      Sequence:
        - Script: string
          Args:
            - string
      Script: string
      Args:
        - string
    OnNodeConfigured:
      Sequence:
        - Script: string
          Args:
```

```
            - string
      Script: string
      Args:
        - string
    OnNodeUpdated:
      Sequence:
        - Script: string
          Args:
            - string
      Script: string
      Args:
        - string
  Iam:
    InstanceRole: string
    InstanceProfile: string
    S3Access:
      - BucketName: string
        EnableWriteAccess: boolean
        KeyName: string
    AdditionalIamPolicies:
      - Policy: string
  Imds:
    Secured: boolean
  Image:
    CustomAmi: string
```

HeadNode properties

`InstanceType` (**Required**, `String`)

Specifies the instance type for the head node.

Specifies the Amazon EC2 instance type that's used for the head node. The architecture of the instance type must be the same as the architecture used for the AWS Batch `InstanceType` or Slurm `InstanceType` setting.

> **Note**
>
> AWS ParallelCluster doesn't support the following instance types for the `HeadNode` setting.
>
> - hpc6id

If you define a p4d instance type or another instance type that has multiple network interfaces or a network interface card, you must set `ElasticIp` to `true` to provide public access. AWS public IPs can only be assigned to instances launched with a single network interface. For this case, we recommend that you use a NAT gateway to provide public access to the cluster compute nodes. For more information, see Assign a public IPv4 address during instance launch in the *Amazon EC2 User Guide for Linux Instances*.

Update policy: If this setting is changed, the update is not allowed.

`DisableSimultaneousMultithreading` (**Optional**, `Boolean`)

If `true`, disables hyper-threading on the head node. The default value is `false`.

Not all instance types can disable hyper-threading. For a list of instance types that support disabling hyperthreading, see CPU cores and threads for each CPU core per instance type in the *Amazon EC2 User Guide*.

Update policy: If this setting is changed, the update is not allowed.

`SharedStorageType` (**Optional**, `String`)

Specifies the type of storage used for internally shared data. Internally shared data includes data that AWS ParallelCluster uses to manage the cluster and the default shared `/home` if not specified in the `SharedStorage` section as a Mount directory to mount a shared filesystem volume. For more details on internal shared data refer AWS ParallelCluster internal directories.

If `Ebs`, which is the default storage type, the head node will export portions of its root volume as shared directories for compute nodes and login nodes using NFS.

If `Efs`, ParallelCluster will create an EFS filesystem to use for shared internal data and `/home`.

Update policy: If this setting is changed, the update is not allowed.

> **Note**
>
> When the cluster scales out, the EBS storage type may present performance bottlenecks as the head node shares data from the root volume with the compute nodes using NFS exports. Using EFS, you can avoid NFS exports as your cluster scales out and avoid performance bottlenecks associated with them. It is recommended to choose EBS for max read/write potential for small files and installation process. Choose EFS for scale.

`Networking`

(Required) Defines the networking configuration for the head node.

```
Networking:
  SubnetId: string
  ElasticIp: string/boolean
  SecurityGroups:
    - string
  AdditionalSecurityGroups:
    - string
  Proxy:
    HttpProxyAddress: string
```

Update policy: If this setting is changed, the update is not allowed.

`Networking` properties

`SubnetId` (**Required**, `String`)

Specifies the ID of an existing subnet in which to provision the head node.

Update policy: If this setting is changed, the update is not allowed.

`ElasticIp` (**Optional**, `String`)

Creates or assigns an Elastic IP address to the head node. Supported values are `true`, `false`, or the ID of an existing Elastic IP address. The default is `false`.

Update policy: If this setting is changed, the update is not allowed.

`SecurityGroups` (**Optional**, `[String]`)

List of Amazon VPC security group ids to use for the head node. These replace the security groups that AWS ParallelCluster creates if this property is not included.

Verify that the security groups are configured correctly for your SharedStorage systems.

Update policy: This setting can be changed during an update.

`AdditionalSecurityGroups` (**Optional**, `[String]`)

List of additional Amazon VPC security group ids to use for the head node.

Update policy: This setting can be changed during an update.

`Proxy` (**Optional**)

Specifies the proxy settings for the head node.

```
Proxy:
                              HttpProxyAddress:
                              string
```

`HttpProxyAddress` (**Optional**, `String`)

Defines an HTTP or HTTPS proxy server, typically `https://x.x.x.x:8080`.

There is no default value.

Update policy: If this setting is changed, the update is not allowed.

Ssh

(Optional) Defines the configuration for SSH access to the head node.

```
Ssh:
      KeyName: string
      AllowedIps: string
```

Update policy: This setting can be changed during an update.

Ssh properties

`KeyName` (**Optional**, `String`)

Names an existing Amazon EC2 key pair to enable SSH access to the head node.

Update policy: If this setting is changed, the update is not allowed.

`AllowedIps` (**Optional**, `String`)

Specifies the CIDR-formatted IP range or a prefix list id for SSH connections to the head node. The default is `0.0.0.0/0`.

Update policy: This setting can be changed during an update.

`LocalStorage`

(Optional) Defines the local storage configuration for the head node.

```
LocalStorage:
  RootVolume:
    Size: integer
    Encrypted: boolean
    VolumeType: string
    Iops: integer
    Throughput: integer
    DeleteOnTermination: boolean
  EphemeralVolume:
    MountDir: string
```

Update policy: This setting can be changed during an update.

`LocalStorage` properties

`RootVolume` (**Required**)

Specifies the root volume storage for the head node.

```
RootVolume:
  Size: integer
  Encrypted: boolean
  VolumeType: string
  Iops: integer
  Throughput: integer
  DeleteOnTermination: boolean
```

Update policy: This setting can be changed during an update.

`Size` (**Optional**, `Integer`)

Specifies the head node root volume size in gibibytes (GiB). The default size comes from the AMI. Using a different size requires that the AMI supports `growroot`.

Update policy: If this setting is changed, the update is not allowed.

`Encrypted` (**Optional**, `Boolean`)

Specifies if the root volume is encrypted. The default value is `true`.

Update policy: If this setting is changed, the update is not allowed.

`VolumeType` (**Optional**, `String`)

Specifies the Amazon EBS volume type. Supported values are `gp2`, `gp3`, `io1`, `io2`, `sc1`, `st1`, and `standard`. The default value is `gp3`.

For more information, see Amazon EBS volume types in the *Amazon EC2 User Guide*.

Update policy: If this setting is changed, the update is not allowed.

`Iops` (**Optional**, `Integer`)

Defines the number of IOPS for `io1`, `io2`, and `gp3` type volumes.

The default value, supported values, and `volume_iops` to `volume_size` ratio varies by `VolumeType` and `Size`.

Update policy: If this setting is changed, the update is not allowed.

`VolumeType = io1`

Default `Iops` = 100

Supported values `Iops` = 100–64000 †

Maximum `Iops` to `Size` ratio = 50 IOPS per GiB. 5000 IOPS requires a `Size` of at least 100 GiB.

`VolumeType = io2`

Default `Iops` = 100

Supported values `Iops` = 100–64000 (256000 for `io2` Block Express volumes) †

Maximum `Iops` to `Size` ratio = 500 IOPS per GiB. 5000 IOPS requires a `Size` of at least 10 GiB.

`VolumeType = gp3`

Default `Iops` = 3000

Supported values `Iops` = 3000–16000

Maximum `Iops` to `Size` ratio = 500 IOPS per GiB. 5000 IOPS requires a `Size` of at least 10 GiB.

† Maximum IOPS is guaranteed only on Instances built on the Nitro System provisioned with more than 32,000 IOPS. Other instances guarantee up to 32,000 IOPS. Older `io1` volumes might not reach full performance unless you modify the volume. `io2` Block Express volumes support `Iops` values up to 256000 on `R5b` instance types. For more information, see `io2` Block Express volumes in the *Amazon EC2 User Guide*.

Update policy: This setting can be changed during an update.

`Throughput` (**Optional**, `Integer`)

Defines the throughput for `gp3` volume types, in MiB/s. This setting is valid only when `VolumeType` is `gp3`. The default value is `125`. Supported values: 125–1000 MiB/s

The ratio of `Throughput` to `Iops` can be no more than 0.25. The maximum throughput of 1000 MiB/s requires that the `Iops` setting is at least 4000.

Update policy: If this setting is changed, the update is not allowed.

`DeleteOnTermination` (**Optional**, `Boolean`)

Specifies whether the root volume should be deleted when the head node is terminated. The default value is `true`.

Update policy: If this setting is changed, the update is not allowed.

`EphemeralVolume` (**Optional**)

Specifies details for any instance store volume. For more information, see Instance store volumes in the *Amazon EC2 User Guide*.

```
EphemeralVolume:
  MountDir: string
```

Update policy: If this setting is changed, the update is not allowed.

`MountDir` (**Optional**, `String`)

Specifies the mount directory for the instance store volume. The default is `/scratch`.

Update policy: If this setting is changed, the update is not allowed.

`Dcv`

(Optional) Defines configuration settings for the Amazon DCV server that runs on the head node.

For more information, see Connect to the head and login nodes through Amazon DCV.

```
Dcv:
  Enabled: boolean
  Port: integer
  AllowedIps: string
```

Important

By default, the Amazon DCV port setup by AWS ParallelCluster is open to all IPv4 addresses. However, you can connect to a Amazon DCV port only if you have the URL for the Amazon DCV session and connect to the Amazon DCV session within 30 seconds of when the URL is returned from `pcluster dcv-connect`. Use the `AllowedIps` setting to further restrict access to the Amazon DCV port with a CIDR-formatted IP range, and use the `Port` setting to set a nonstandard port.

Update policy: If this setting is changed, the update is not allowed.

`Dcv` properties

`Enabled` (**Required**, `Boolean`)

Specifies whether Amazon DCV is enabled on the head node. The default value is `false`.

Update policy: If this setting is changed, the update is not allowed.

Note

Amazon DCV automatically generates a self-signed certificate that's used to secure traffic between the Amazon DCV client and the Amazon DCV server that runs on the head node. To configure your own certificate, see Amazon DCV HTTPS certificate.

`Port` (**Optional**, `Integer`)

Specifies the port for Amazon DCV. The default value is 8443.

Update policy: If this setting is changed, the update is not allowed.

`AllowedIps` (**Optional, Recommended**, `String`)

Specifies the CIDR-formatted IP range for connections to Amazon DCV. This setting is used only when AWS ParallelCluster creates the security group. The default value is `0.0.0.0/0`, which allows access from any internet address.

Update policy: This setting can be changed during an update.

`CustomActions`

(Optional) Specifies custom scripts to run on the head node.

```
CustomActions:
  OnNodeStart:
    Sequence:
      - Script: string
        Args:
          - string
    Script: string
    Args:
      - string
  OnNodeConfigured:
    Sequence:
      - Script: string
        Args:
          - string
    Script: string
    Args:
      - string
  OnNodeUpdated:
    Sequence:
      - Script: string
        Args:
          - string
    Script: string
    Args:
      - string
```

`CustomActions` properties

`OnNodeStart` (**Optional**)

Specifies single script or a sequence of scripts to run on the head node before any node deployment bootstrap action is started. For more information, see Custom bootstrap actions.

`Sequence` (**Optional**)

List of scripts to run. AWS ParallelCluster runs the scripts in the same order as they are listed in the configuration file, starting with the first.

`Script` (**Required**, `String`)

Specifies the file to use. The file path can start with `https://` or `s3://`.

`Args` (**Optional**, `[String]`)

List of arguments to pass to the script.

`Script` (**Required**, `String`)

Specifies the file to use for a single script. The file path can start with `https://` or `s3://`.

`Args` (**Optional**, `[String]`)

List of arguments to pass to the single script.

Update policy: If this setting is changed, the update is not allowed.

`OnNodeConfigured` (**Optional**)

Specifies a single script or a sequence of scripts to run on the head node after the node bootstrap actions are complete. For more information, see Custom bootstrap actions.

`Sequence` (**Optional**)

Specifies the list of scripts to run.

`Script` (**Required**, `String`)

Specifies the file to use. The file path can start with `https://` or `s3://`.

`Args` (**Optional**, `[String]`)

List of arguments to pass to the script.

`Script` (**Required**, `String`)

Specifies the file to use for a single script. The file path can start with `https://` or `s3://`.

`Args` (**Optional**, `[String]`)

List of arguments to pass to the single script.

Update policy: If this setting is changed, the update is not allowed.

`OnNodeUpdated` (**Optional**)

Specifies a single script or a sequence of scripts to run on the head node after node update actions are complete. For more information, see Custom bootstrap actions.

`Sequence` (**Optional**)

Specifies the list of scripts to run.

`Script` (**Required**, `String`)

Specifies the file to use. The file path can start with `https://` or `s3://`.

`Args` (**Optional**, `[String]`)

List of arguments to pass to the script.

`Script` (**Required**, `String`)

Specifies the file to use for the single script. The file path can start with `https://` or `s3://`.

`Args` (**Optional**, `[String]`)

List of arguments to pass to the single script.

Update policy: This setting can be changed during an update.

> **Note**
>
> `OnNodeUpdated` is added starting with AWS ParallelCluster 3.4.0.
> `Sequence` is added starting with AWS ParallelCluster version 3.6.0. When you specify `Sequence`, you can list multiple scripts for a custom action. AWS ParallelCluster continues to support configuring a custom action with a single script, without including `Sequence`.
> AWS ParallelCluster doesn't support including both a single script and `Sequence` for the same custom action.

`Iam`

(Optional) Specifies either an instance role or an instance profile to use on the head node to override the default instance role or instance profile for the cluster.

```
Iam:
  InstanceRole: string
  InstanceProfile: string
  S3Access:
    - BucketName: string
      EnableWriteAccess: boolean
      KeyName: string
  AdditionalIamPolicies:
    - Policy: string
```

Update policy: This setting can be changed during an update.

`Iam` properties

`InstanceProfile` (**Optional**, `String`)

Specifies an instance profile to override the default head node instance profile. You can't specify both `InstanceProfile` and `InstanceRole`. The format is `arn:`*`Partition`*`:iam::`*`Account`*`:instance-profile/`*`InstanceProfileName`*.

If this is specified, the `S3Access` and `AdditionalIamPolicies` settings can't be specified.

We recommend that you specify one or both of the `S3Access` and `AdditionalIamPolicies` settings because features added to AWS ParallelCluster often require new permissions.

Update policy: If this setting is changed, the update is not allowed.

`InstanceRole` (**Optional**, `String`)

Specifies an instance role to override the default head node instance role. You can't specify both `InstanceProfile` and `InstanceRole`. The format is `arn:`*`Partition`*`:iam::`*`Account`*`:role/`*`RoleName`*.

If this is specified, the `S3Access` and `AdditionalIamPolicies` settings can't be specified.

We recommend that you specify one or both of the `S3Access` and `AdditionalIamPolicies` settings because features added to AWS ParallelCluster often require new permissions.

Update policy: This setting can be changed during an update.

S3Access

S3Access (**Optional**)

Specifies a bucket. This is used to generate policies to grant the specified access to the bucket.

If this is specified, the InstanceProfile and InstanceRole settings can't be specified.

We recommend that you specify one or both of the S3Access and AdditionalIamPolicies settings because features added to AWS ParallelCluster often require new permissions.

```
S3Access:
  - BucketName: string
    EnableWriteAccess: boolean
    KeyName: string
```

Update policy: This setting can be changed during an update.

BucketName (**Required**, String)

The name of the bucket.

Update policy: This setting can be changed during an update.

KeyName (**Optional**, String)

The key for the bucket. The default value is "*".

Update policy: This setting can be changed during an update.

EnableWriteAccess (**Optional**, Boolean)

Indicates whether write access is enabled for the bucket. The default value is false.

Update policy: This setting can be changed during an update.

AdditionalIamPolicies

AdditionalIamPolicies (**Optional**)

Specifies a list of Amazon Resource Names (ARNs) of IAM policies for Amazon EC2. This list is attached to the root role used for the head node in addition to the permissions required by AWS ParallelCluster.

An IAM policy name and its ARN are different. Names can't be used.

If this is specified, the `InstanceProfile` and `InstanceRole` settings can't be specified.

We recommend that you use `AdditionalIamPolicies` because `AdditionalIamPolicies` are added to the permissions that AWS ParallelCluster requires, and the `InstanceRole` must include all permissions required. The permissions required often change from release to release as features are added.

There is no default value.

```
AdditionalIamPolicies:
  - Policy: string
```

Update policy: This setting can be changed during an update.

`Policy` (**Optional**, `[String]`)

List of IAM policies.

Update policy: This setting can be changed during an update.

`Imds`

(Optional) Specifies the properties for instance metadata service (IMDS). For more information, see How instance metadata service version 2 works in the *Amazon EC2 User Guide*.

```
Imds:
    Secured: boolean
```

Update policy: If this setting is changed, the update is not allowed.

`Imds` properties

`Secured` (**Optional**, `Boolean`)

If `true`, restricts access to the head node's IMDS (and the instance profile credentials) to a subset of superusers.

If `false`, every user in the head node has access to the head node's IMDS.

The following users are permitted access to the head node's IMDS:

- root user
- cluster administrative user (`pc-cluster-admin` by default)
- operating system specific default user (`ec2-user` on Amazon Linux 2 and RedHat, and `ubuntu` on Ubuntu 18.04.

The default is `true`.

The `default` users are responsible for ensuring a cluster has the permissions it needs to interact with AWS resources. If you disable `default` user IMDS access, AWS ParallelCluster can't manage the compute nodes and stops working. Don't disable `default` user IMDS access.

When a user is granted access to the head node's IMDS, they can use the permissions included in the head node's instance profile. For example, they can use these permissions to launch Amazon EC2 instances or to read the password for an AD domain that the cluster is configured to use for authentication.

To restrict IMDS access, AWS ParallelCluster manages a chain of `iptables`.

Cluster users with `sudo` access can selectively enable or disable access to the head node's IMDS for other individual users, including `default` users, by running the command:

```
$ sudo /opt/parallelcluster/scripts/imds/imds-access.sh --allow <USERNAME>
```

You can disable user IMDS access with the `--deny` option for this command.

If you unknowingly disable `default` user IMDS access, you can restore the permission by using the `--allow` option.

> **Note**
>
> Any customization of `iptables` or `ip6tables` rules can interfere with the mechanism used to restrict IMDS access on the head node.

Update policy: If this setting is changed, the update is not allowed.

`Image`

(Optional) Defines a custom image for the head node.

```
Image:
     CustomAmi: string
```

Update policy: If this setting is changed, the update is not allowed.

`Image` properties

`CustomAmi` (**Optional**, `String`)

Specifies the ID of a custom AMI to use for the head node instead of the default AMI. For more information, see AWS ParallelCluster AMI customization.

If the custom AMI requires additional permissions for its launch, these permissions must be added to both the user and head node policies.

For example, if a custom AMI has an encrypted snapshot associated with it, the following additional policies are required in both the user and head node policies:

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:DescribeKey",
                "kms:ReEncrypt*",
                "kms:CreateGrant",
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:us-east-1:111122223333:key/<AWS_KMS_KEY_ID>"
            ]
        }
    ]
}
```

To troubleshoot custom AMI validation warnings, see Troubleshooting custom AMI issues.

Update policy: If this setting is changed, the update is not allowed.

`Scheduling` section

(Required) Defines the job scheduler that's used in the cluster and the compute instances that the job scheduler manages. You can either use the Slurm or AWS Batch scheduler. Each supports a different set of settings and properties.

Topics


```
Scheduling:
  Scheduler: slurm
  ScalingStrategy: string
  SlurmSettings:
    MungeKeySecretArn: string
    ScaledownIdletime: integer
    QueueUpdateStrategy: string
    EnableMemoryBasedScheduling: boolean
    CustomSlurmSettings: [dict]
    CustomSlurmSettingsIncludeFile: string
    Database:
      Uri: string
      UserName: string
      PasswordSecretArn: string
      DatabaseName: string
    ExternalSlurmdbd: boolean
      Host: string
      Port: integer
    Dns:
      DisableManagedDns: boolean
      HostedZoneId: string
      UseEc2Hostnames: boolean
  SlurmQueues:
    - Name: string
      ComputeSettings:
        LocalStorage:
          RootVolume:
            Size: integer
```

```
          Encrypted: boolean
          VolumeType: string
          Iops: integer
          Throughput: integer
        EphemeralVolume:
          MountDir: string
    CapacityReservationTarget:
      CapacityReservationId: string
      CapacityReservationResourceGroupArn: string
    CapacityType: string
    AllocationStrategy: string
    JobExclusiveAllocation: boolean
    CustomSlurmSettings: dict
    Tags:
      - Key: string
        Value: string
    HealthChecks:
      Gpu:
        Enabled: boolean
    Networking:
      SubnetIds:
        - string
      AssignPublicIp: boolean
      SecurityGroups:
        - string
      AdditionalSecurityGroups:
        - string
      PlacementGroup:
        Enabled: boolean
        Id: string
        Name: string
      Proxy:
        HttpProxyAddress: string
    ComputeResources:
      - Name: string
        InstanceType: string
        Instances:
          - InstanceType: string
        MinCount: integer
        MaxCount: integer
        DynamicNodePriority: integer
        StaticNodePriority: integer
        SpotPrice: float
        DisableSimultaneousMultithreading: boolean
```

```
          SchedulableMemory: integer
          HealthChecks:
            Gpu:
              Enabled: boolean
          Efa:
            Enabled: boolean
            GdrSupport: boolean
          CapacityReservationTarget:
            CapacityReservationId: string
            CapacityReservationResourceGroupArn: string
          Networking:
            PlacementGroup:
              Enabled: boolean
              Name: string
          CustomSlurmSettings: dict
          Tags:
            - Key: string
              Value: string
      CustomActions:
        OnNodeStart:
          Sequence:
            - Script: string
              Args:
                - string
          Script: string
          Args:
            - string
        OnNodeConfigured:
          Sequence:
            - Script: string
              Args:
                - string
          Script: string
          Args:
            - string
      Iam:
        InstanceProfile: string
        InstanceRole: string
        S3Access:
          - BucketName: string
            EnableWriteAccess: boolean
            KeyName: string
        AdditionalIamPolicies:
          - Policy: string
```

```
      Image:
        CustomAmi: string
```

```
Scheduling:
  Scheduler: awsbatch
  AwsBatchQueues:
    - Name: string
      CapacityType: string
      Networking:
        SubnetIds:
          - string
        AssignPublicIp: boolean
        SecurityGroups:
          - string
        AdditionalSecurityGroups:
          - string
      ComputeResources:  # this maps to a Batch compute environment (initially we
 support only 1)
        - Name: string
          InstanceTypes:
            - string
          MinvCpus: integer
          DesiredvCpus: integer
          MaxvCpus: integer
          SpotBidPercentage: float
```

`Scheduling` properties

`Scheduler` (Required, `String`)

Specifies the type of scheduler that's used. Supported values are `slurm` and `awsbatch`.

Update policy: If this setting is changed, the update is not allowed.

> **Note**
>
> `awsbatch` only supports the `alinux2` operating system and `x86_64` platform.

`ScalingStrategy` (Optional, `String`)

Allows you to choose how dynamic Slurm nodes scale up. Supported values are `all-or-nothing`, `greedy-all-or-nothing` and `best-effort` The default value is `all-or-nothing`.

Update policy: This setting can be changed during an update.

> **Note**
>
> The scaling strategy applies only to nodes to be resumed by Slurm, not to nodes that are eventually already running.

- `all-or-nothing`This strategy strictly follows an all-or-nothing-approach, aimed at avoiding idle instances at the end of the scaling process. It operates on an all-or-nothing basis, meaning it either scales up completely or not at all. Be aware that there may be additional costs due to temporarily launched instances, when jobs require over 500 nodes or span multiple compute resources. This strategy has the lowest throughput among the three possible Scaling Strategies. The scaling time depends on the number of jobs submitted per Slurm resume program execution. Also, you can't scale far beyond the default RunInstances resource account limit per execution, which is 1000 instances by defaults. More details can be found at the Amazon EC2 API throttling documentation
- `greedy-all-or-nothing` Similar to the all-or-nothing strategy, it aims to avoid idle instances post-scaling. This strategy allows for temporary over-scaling during the scaling process in order to achieve higher throughput than the all-or-nothing approach but also comes with the same scaling limit of 1000 instances as per the RunInstances resource account limit.
- `best-effort` This strategy prioritizes high throughput, even if it means that some instances might be idle at the end of the scaling process. It attempts to allocate as many nodes as requested by the jobs, but there's a possibility of not fulfilling the entire request. Unlike the other strategies, the best-effort approach can accumulate more instances than the standard RunInstances limit, at the cost of having idle resources along the multiple scaling process executions.

Each strategy is designed to cater to different scaling needs, allowing you to select one that meets your specific requirements and constraints.

AwsBatchQueues

(Optional) The AWS Batch queue settings. Only one queue is supported. If Scheduler is set to awsbatch, this section is required. For more information about the awsbatch scheduler, see networking setup and Using AWS Batch (awsbatch) scheduler with AWS ParallelCluster.

```
AwsBatchQueues:
  - Name: string
    CapacityType: string
    Networking:
      SubnetIds:
        - string
      AssignPublicIp: boolean
      SecurityGroups:
        - string
      AdditionalSecurityGroups:
        - string
    ComputeResources:  # this maps to a Batch compute environment (initially we support
 only 1)
      - Name: string
        InstanceTypes:
          - string
        MinvCpus: integer
        DesiredvCpus: integer
        MaxvCpus: integer
        SpotBidPercentage: float
```

Update policy: This setting can be changed during an update.

AwsBatchQueues properties

Name (Required, String)

The name of the AWS Batch queue.

Update policy: If this setting is changed, the update is not allowed.

CapacityType (Optional, String)

The type of the compute resources that the AWS Batch queue uses. Supported values are ONDEMAND , SPOT or CAPACITY_BLOCK. The default value is ONDEMAND.

> **Note**
>
> If you set `CapacityType` to `SPOT`, your account must contain an `AWSServiceRoleForEC2Spot` service-linked role. You can create this role using the following AWS CLI command.
>
> ```
> $ aws iam create-service-linked-role --aws-service-name spot.amazonaws.com
> ```
>
> For more information, see Service-linked role for Spot Instance requests in the *Amazon Amazon EC2 User Guide for Linux Instances*.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`Networking`

(Required) Defines the networking configuration for the AWS Batch queue.

```
Networking:
  SubnetIds:
    - string
  AssignPublicIp: boolean
  SecurityGroups:
    - string
  AdditionalSecurityGroups:
    - string
```

`Networking` properties

`SubnetIds` (Required, `[String]`)

Specifies the ID of an existing subnet to provision the AWS Batch queue in. Currently only one subnet is supported.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`AssignPublicIp` (Optional, `String`)

Creates or assigns a public IP address to the nodes in the AWS Batch queue. Supported values are `true` and `false`. The default depends on the subnet that you specified.

Update policy: If this setting is changed, the update is not allowed.

`SecurityGroups` (Optional, `[String]`)

List of security groups that the AWS Batch queue uses. If you don't specify security groups, AWS ParallelCluster creates new security groups.

Update policy: This setting can be changed during an update.

`AdditionalSecurityGroups` (Optional, `[String]`)

List of security groups that the AWS Batch queue uses.

Update policy: This setting can be changed during an update.

`ComputeResources`

(Required) Defines the ComputeResources configuration for the AWS Batch queue.

```
ComputeResources:  # this maps to a Batch compute environment (initially we support
 only 1)
  - Name: string
    InstanceTypes:
      - string
    MinvCpus: integer
    DesiredvCpus: integer
    MaxvCpus: integer
    SpotBidPercentage: float
```

`ComputeResources` properties

`Name` (Required, `String`)

The name of the AWS Batch queue compute environment.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`InstanceTypes` (Required, `[String]`)

The AWS Batch compute environment array of instance types. All of the instance types must use the `x86_64` architecture.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`MinvCpus` (Optional, `Integer`)

The minimum number of VCPUs that an AWS Batch compute environment can use.

Update policy: This setting can be changed during an update.

`DesiredVcpus` (Optional, `Integer`)

The desired number of VCPUs in the AWS Batch compute environment. AWS Batch adjusts this value between `MinvCpus` and `MaxvCpus` based on the demand in the job queue.

Update policy: This setting is not analyzed during an update.

`MaxvCpus` (Optional, `Integer`)

The maximum number of VCPUs for the AWS Batch compute environment. You can't set this to a value that's lower than `DesiredVcpus`.

Update policy: This setting can't be decreased during an update.

`SpotBidPercentage` (Optional, `Float`)

The maximum percentage of the On-Demand price for the instance type that an Amazon EC2 Spot Instance price can reach before instances are launched. The default value is `100` (100%). The supported range is `1-100`.

Update policy: This setting can be changed during an update.

`SlurmQueues`

(Optional) Settings for the Slurm queue. If `Scheduler` is set to `slurm`, this section is required.

```
SlurmQueues:
  - Name: string
    ComputeSettings:
      LocalStorage:
        RootVolume:
```

```
          Size: integer
          Encrypted: boolean
          VolumeType: string
          Iops: integer
          Throughput: integer
        EphemeralVolume:
          MountDir: string
    CapacityReservationTarget:
      CapacityReservationId: string
      CapacityReservationResourceGroupArn: string
    CapacityType: string
    AllocationStrategy: string
    JobExclusiveAllocation: boolean
    CustomSlurmSettings: dict
    Tags:
      - Key: string
        Value: string
    HealthChecks:
      Gpu:
        Enabled: boolean
    Networking:
      SubnetIds:
        - string
      AssignPublicIp: boolean
      SecurityGroups:
        - string
      AdditionalSecurityGroups:
        - string
      PlacementGroup:
        Enabled: boolean
        Id: string
        Name: string
      Proxy:
        HttpProxyAddress: string
    ComputeResources:
      - Name: string
        InstanceType: string
        Instances:
          - InstanceType: string
        MinCount: integer
        MaxCount: integer
        DynamicNodePriority: integer
        StaticNodePriority: integer
        SpotPrice: float
```

```
        DisableSimultaneousMultithreading: boolean
        SchedulableMemory: integer
        HealthChecks:
          Gpu:
            Enabled: boolean
        Efa:
          Enabled: boolean
          GdrSupport: boolean
        CapacityReservationTarget:
          CapacityReservationId: string
          CapacityReservationResourceGroupArn: string
        Networking:
          PlacementGroup:
            Enabled: boolean
            Name: string
        CustomSlurmSettings: dict
        Tags:
          - Key: string
            Value: string
    CustomActions:
      OnNodeStart:
        Sequence:
          - Script: string
            Args:
              - string
        Script: string
        Args:
          - string
      OnNodeConfigured:
        Sequence:
          - Script: string
            Args:
              - string
        Script: string
        Args:
          - string
    Iam:
      InstanceProfile: string
      InstanceRole: string
      S3Access:
        - BucketName: string
          EnableWriteAccess: boolean
          KeyName: string
      AdditionalIamPolicies:
```

```
        - Policy: string
    Image:
      CustomAmi: string
```

Update policy: For this list values setting, a new value can be added during an update or the compute fleet must be stopped when removing an existing value.

`SlurmQueues` properties

`Name` (Required, `String`)

The name of the Slurm queue.

> **Note**
>
> Cluster size may change during an update. For more information, see Cluster capacity size and update

Update policy: If this setting is changed, the update is not allowed.

`CapacityReservationTarget`

> **Note**
>
> `CapacityReservationTarget` is added with AWS ParallelCluster version 3.3.0.

```
CapacityReservationTarget:
   CapacityReservationId: string
   CapacityReservationResourceGroupArn: string
```

Specifies the On-Demand capacity reservation for the queue's compute resources.

`CapacityReservationId` (Optional, `String`)

The ID of the existing capacity reservation to target for the queue's compute resources. The ID can refer to an ODCR or a Capacity Block for ML.

The reservation must use the same platform that the instance uses. For example, if your instances run on `rhel8`, your capacity reservation must run on the Red Hat Enterprise Linux

platform. For more information, see Supported platforms in the *Amazon EC2 User Guide for Linux Instances*.

> **Note**
>
> If you include `Instances` in the cluster configuration, you must exclude this queue level `CapacityReservationId` setting from the configuration.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`CapacityReservationResourceGroupArn` (Optional, `String`)

The Amazon Resource Name (ARN) of the resource group that serves as the service-linked group of capacity reservations for the queue's compute resources. AWS ParallelCluster identifies and uses the most appropriate capacity reservation from the resource group based on the following conditions:

- If `PlacementGroup` is enabled in `SlurmQueues` / `Networking` or `SlurmQueues` / `ComputeResources` / `Networking`, AWS ParallelCluster selects a resource group that targets the instance type and `PlacementGroup` for a compute resource, if the compute resource exists.

  The `PlacementGroup` must target one of the instance types that's defined in `ComputeResources`.
- If `PlacementGroup` isn't enabled in `SlurmQueues` / `Networking` or `SlurmQueues` / `ComputeResources` / `Networking`, AWS ParallelCluster selects a resource group that targets only the instance type of a compute resource, if the compute resource exists.

The resource group must have at least one ODCR for each instance type reserved in an Availability Zone across all of the queue's compute resources and Availability Zones. For more information, see Launch instances with On-Demand Capacity Reservations (ODCR).

For more information on multiple subnet configuration requirements, see `Networking` / `SubnetIds`.

> **Note**
>
> Multiple Availability Zones is added in AWS ParallelCluster version 3.4.0.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`CapacityType` (Optional, `String`)

The type of the compute resources that the Slurm queue uses. Supported values are `ONDEMAND` , `SPOT` or `CAPACITY_BLOCK`. The default value is `ONDEMAND`.

> **Note**
>
> If you set the `CapacityType` to `SPOT`, your account must have an `AWSServiceRoleForEC2Spot` service-linked role. You can use the following AWS CLI command to create this role.
>
> ```
> $ aws iam create-service-linked-role --aws-service-name spot.amazonaws.com
> ```
>
> For more information, see Service-linked role for Spot Instance requests in the *Amazon Amazon EC2 User Guide for Linux Instances*.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`AllocationStrategy` (Optional, `String`)

Specify the allocation strategy for all the compute resources defined in `Instances`.

Valid values: `lowest-price` | `capacity-optimized` | `price-capacity-optimized`

Default: `lowest-price`

`lowest-price`

- If you use `CapacityType = ONDEMAND`, Amazon EC2 Fleet uses price to determine the order and launches the lowest price instances first.
- If you use `CapacityType = SPOT`, Amazon EC2 Fleet launches instances from the lowest price Spot Instance pool that has available capacity. If a pool runs out of capacity before it fulfills your required capacity, Amazon EC2 Fleet fulfills your request by launching instances for you. In particular, Amazon EC2 Fleet launches instances from the lowest price Spot Instance pool that has available capacity. Amazon EC2 Fleet might launch Spot Instances from several different pools.

- If you set `CapacityType = CAPACITY_BLOCK`, there are no allocation strategies, thus `AllocationStrategy` parameter cannot be configured.

`capacity-optimized`

- If you set `CapacityType = ONDEMAND`, `capacity-optimized` isn't available.
- If you set `CapacityType = SPOT`, Amazon EC2 Fleet launches instances from Spot Instance pools with optimal capacity for the number of instances to be launched.

`price-capacity-optimized`

- If you set `CapacityType = ONDEMAND`, `capacity-optimized` isn't available.
- If you set `CapacityType = SPOT`, Amazon EC2 Fleet identifies the pools with the highest capacity availability for the number of instances that are launching. This means that we will request Spot Instances from the pools that we believe have the lowest chance of interruption in the near term. Amazon EC2 Fleet then requests Spot Instances from the lowest priced of these pools.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

> **Note**
>
> `AllocationStrategy` is supported starting in AWS ParallelCluster version 3.3.0.

`JobExclusiveAllocation` (Optional, `String`)

If set to `true`, the Slurm partition `OverSubscribe` flag is set to `EXCLUSIVE`. When `OverSubscribe=EXCLUSIVE`, jobs in the partition have exclusive access to all allocated nodes. For more information, see EXCLUSIVE in the Slurm documentation.

Valid values: `true` | `false`

Default: `false`

Update policy: This setting can be changed during an update.

> **Note**
>
> `JobExclusiveAllocation` is supported starting in AWS ParallelCluster version 3.7.0.

`CustomSlurmSettings` (Optional, `Dict`)

Defines the custom Slurm partition (queue) configuration settings.

Specifies a dictionary of custom Slurm configuration parameter key-value pairs that apply to queues (partitions).

Each separate key-value pair, such as `Param1: Value1`, is added separately to the end of the Slurm partition configuration line in the format `Param1=Value1`.

You can only specify Slurm configuration parameters that aren't deny-listed in `CustomSlurmSettings`. For information about deny-listed Slurm configuration parameters, see Deny-listed Slurm configuration parameters for `CustomSlurmSettings`.

AWS ParallelCluster only checks whether a parameter is in a deny list. AWS ParallelCluster doesn't validate your custom Slurm configuration parameter syntax or semantics. It is your responsibility to validate your custom Slurm configuration parameters. Invalid custom Slurm configuration parameters can cause Slurm daemon failures that can lead to cluster create and update failures.

For more information about how to specify custom Slurm configuration parameters with AWS ParallelCluster, see Slurm configuration customization.

For more information about Slurm configuration parameters, see slurm.conf in the Slurm documentation.

Update policy: This setting can be changed during an update.

> **Note**
>
> `CustomSlurmSettings` is supported starting with AWS ParallelCluster version 3.6.0.

`Tags` (Optional, [String])

A list of tag key-value pairs. `ComputeResource` tags override duplicate tags specified in the `Tags` section or in `SlurmQueues` / `Tags`.

`Key` (Optional, `String`)

The tag key.

`Value` (Optional, `String`)

The tag value.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`HealthChecks` (Optional)

Specify compute node health checks on all compute resources in the queue.

`Gpu` (Optional)

Specify GPU health checks on all compute resources in a queue.

> **Note**
>
> AWS ParallelCluster doesn't support `HealthChecks` / `Gpu` in nodes that use `alinux2` ARM operating systems. These platforms don't support the NVIDIA Data Center GPU Manager (DCGM).

`Enabled` (Optional, `Boolean`)

Whether AWS ParallelCluster performs GPU health checks on compute nodes. The default is `false`.

`Gpu` health check behavior

- If `Gpu` / `Enabled` is set to `true`, AWS ParallelCluster performs GPU health checks on compute resources in the queue.
- The `Gpu` health check performs GPU health checks on compute resources to prevent the submission of jobs on nodes with a degraded GPU.
- If a compute node fails a `Gpu` health check, the compute node state changes to `DRAIN`. New jobs don't start on this node. Existing jobs run to completion. After all running jobs complete, the compute node terminates if it's a dynamic node, and it's replaced if it's a static node.
- The duration of the `Gpu` health check depends on the selected instance type, the number of GPUs in the instance, and the number of `Gpu` health check targets (equivalent to the number of job GPU targets). For an instance with 8 GPUs, the typical duration is less than 3 minutes.

- If the Gpu health check runs on an instance that's not supported, it exits and the job runs on the compute node. For example, if an instance doesn't have a GPU, or, if an instance has a GPU, but it isn't an NVIDIA GPU, the health check exits and the job runs on the compute node. Only NVIDIA GPUs are supported.
- The Gpu health check uses the `dcgmi` tool to perform health checks on a node and takes the following steps:

  When the Gpu health check begins in a node:

  1. It detects whether the `nvidia-dcgm` and `nvidia-fabricmanager` services are running.
  2. If these services aren't running, the Gpu health check starts them.
  3. It detects whether the persistence mode is enabled.
  4. If the persistence mode isn't enabled, the Gpu health check enables it.

  At the end of the health check, the Gpu health check restores these services and resources to their initial state.
- If the job is assigned to a specific set of node GPUs, the Gpu health check runs only on that specific set. Otherwise, the Gpu health check runs on all GPUs in the node.
- If a compute node receives 2 or more Gpu health check requests at the same time, only the first health check runs and the others are skipped. This is also the case for health checks that target node GPUs. You can check the log files for additional information regarding this situation.
- The health check log for a specific compute node is available in the `/var/log/parallelcluster/slurm_health_check.log` file. The file is available in Amazon CloudWatch, in the cluster CloudWatch log group, where you can find:
  - Details on the action run by the Gpu health check, including enabling and disabling services and persistence mode.
  - The GPU identifier, serial ID, and the UUID.
  - The health check output.

Update policy: This setting can be changed during an update.

> **Note**
>
> `HealthChecks` is supported starting in AWS ParallelCluster version 3.6.0.

`Networking`

(Required) Defines the networking configuration for the Slurm queue.

```
Networking:
  SubnetIds:
    - string
  AssignPublicIp: boolean
  SecurityGroups:
    - string
  AdditionalSecurityGroups:
    - string
  PlacementGroup:
    Enabled: boolean
    Id: string
    Name: string
  Proxy:
    HttpProxyAddress: string
```

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Networking` properties

`SubnetIds` (Required, `[String]`)

The IDs of existing subnets that you provision the Slurm queue in.

If you configure instance types in `SlurmQueues` / `ComputeResources` / `InstanceType`, you can only define one subnet.

If you configure instance types in `SlurmQueues` / `ComputeResources` / `Instances`, you can define a single subnet or multiple subnets.

If you use multiple subnets, all subnets defined for a queue must be in the same VPC, with each subnet in a separate Availability Zone (AZ).

For example, suppose you define subnet-1 and subnet-2 for your queue.

`subnet-1` and `subnet-2` can't both be in AZ-1.

`subnet-1` can be in AZ-1 and `subnet-2` can be in AZ-2.

If you configure only one instance type and want to use multiple subnets, define your instance type in `Instances` rather than `InstanceType`.

For example, define `ComputeResources` / `Instances` / `InstanceType=instance.type` instead of `ComputeResources` / `InstanceType=instance.type`.

> **Note**
>
> Elastic Fabric Adapter (EFA) isn't supported over different availability zones.

The use of multiple Availability Zones might cause increases in storage networking latency and added inter-AZ data transfer costs. For example, this could occur when an instance accesses file storage that's located in a different AZ. For more information, see Data Transfer within the same AWS Region.

Cluster updates to change from the use of a single subnet to multiple subnets:

- Suppose the subnet definition of a cluster is defined with a single subnet and an AWS ParallelCluster managed FSx for Lustre file system. Then, you can't update this cluster with an updated subnet ID definition directly. To make the cluster update, you must first change the managed file system to an external file system. For more information, see Convert AWS ParallelCluster managed storage to external storage.
- Suppose the subnet definition of a cluster is defined with a single subnet and an external Amazon EFS file system if EFS mount targets don't exist for all of the AZs for the multiple subnets defined to be added. Then, you can't update this cluster with an updated subnet ID definition directly. To make the cluster update or to create a cluster, you must first create all of the mount targets for all of the AZs for the defined multiple subnets.

Availability Zones and cluster capacity reservations defined in CapacityReservationResourceGroupArn:

- You can't create a cluster if there is no overlap between the set of instance types and availability zones covered by the defined capacity reservation resource group and the set of instance types and availability zones defined for the queue.
- You can create a cluster if there is a partial overlap between the set of instance types and availability zones covered by the defined capacity reservation resource group and the set of instance types and availability zones defined for the queue. AWS ParallelCluster sends a warning message about the partial overlap for this case.

- For more information, see Launch instances with On-Demand Capacity Reservations (ODCR).

> **Note**
>
> Multiple Availability Zones is added in AWS ParallelCluster version 3.4.0.

> **Warning**
>
> This warning applies to all 3.x.y AWS ParallelCluster versions prior to version 3.3.1. AWS ParallelCluster version 3.3.1 isn't impacted if this parameter is changed.
> For AWS ParallelCluster 3 versions prior to version 3.3.1:
> If you change this parameter and update a cluster this creates a new managed FSx for Lustre file system and deletes the existing managed FSx for Lustre file system without preserving the existing data. This results in data loss. Before you proceed, make sure you back up the data from the existing FSx for Lustre file system if you want to preserve data. For more information, see Working with backups in the *FSx for Lustre User Guide*.

If a new subnet value is added, **Update policy: This setting can be changed during an update.**

If a subnet value is removed, Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`AssignPublicIp` (Optional, `String`)

Creates or assigns a public IP address to the nodes in the Slurm queue. Supported values are `true` and `false`. The subnet that you specify determines the default value. A subnet with public IPs default to assigning public IP addresses.

If you define a p4d or hpc6id instance type, or another instance type that has multiple network interfaces or a network interface card, you must set `HeadNode` / `Networking` / `ElasticIp` to `true` to provide public access. AWS public IPs can only be assigned to instances launched with a single network interface. For this case, we recommend that you use a NAT gateway to provide public access to the cluster compute nodes. In this case, set `AssignPublicIp` to `false`. For more information on IP addresses, see Assign a public IPv4 address during instance launch in the *Amazon EC2 User Guide for Linux Instances*.

Update policy: If this setting is changed, the update is not allowed.

`SecurityGroups` (Optional, `[String]`)

A list of security groups to use for the Slurm queue. If no security groups are specified, AWS ParallelCluster creates security groups for you.

Verify that the security groups are configured correctly for your SharedStorage systems.

> **Warning**
>
> This warning applies to all 3.*x*.*y* AWS ParallelCluster versions prior to version 3.3.0. AWS ParallelCluster version 3.3.0 isn't impacted if this parameter is changed.
> For AWS ParallelCluster 3 versions prior to version 3.3.0:
> If you change this parameter and update a cluster this creates a new managed FSx for Lustre file system and deletes the existing managed FSx for Lustre file system without preserving the existing data. This results in data loss. Make sure to back up the data from the existing FSx for Lustre file system if you want to preserve data. For more information, see Working with backups in the *FSx for Lustre User Guide*.

> **Warning**
>
> If you enable Efa for your compute instances, make sure that your EFA-enabled instances are members of a security group that allows all inbound and outbound traffic to itself.

Update policy: This setting can be changed during an update.

`AdditionalSecurityGroups` (Optional, `[String]`)

A list of additional security groups to use for the Slurm queue.

Update policy: This setting can be changed during an update.

`PlacementGroup` (Optional)

Specifies the placement group settings for the Slurm queue.

```
PlacementGroup:
  Enabled: boolean
  Id: string
```

```
  Name: string
```

Update policy: All compute nodes must be stopped for a managed placement group deletion. The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Enabled` (Optional, `Boolean`)

Indicates whether a placement group is used for the Slurm queue. The default is `false`.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Id` (Optional, `String`)

The placement group ID for an existing cluster placement group that the Slurm queue uses. Make sure to provide the placement group *ID* and *not the name*.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Name` (Optional, `String`)

The placement group name for an existing cluster placement group that the Slurm queue uses. Make sure to provide the placement group *name* and *not the ID*.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

Note

- If `PlacementGroup` / `Enabled` is set to `true`, without a `Name` or `Id` defined, each compute resource is assigned its own managed placement group, unless `ComputeResources` / `Networking` / `PlacementGroup` is defined to override this setting.
- Starting with AWS ParallelCluster version 3.3.0, `SlurmQueues` / `Networking` / `PlacementGroup` / `Name` was added as a preferred alternative to `SlurmQueues` / `Networking` / `PlacementGroup` / `Id`.

  `PlacementGroup` / `Id` and `PlacementGroup` / `Name` are equivalent. You can use either one.

If you include both `PlacementGroup` / `Id` and `PlacementGroup` / `Name`, AWS ParallelCluster fails. You can only choose one or the other.

You don't need to update your cluster to use `PlacementGroup` / `Name`.

`Proxy` (Optional)

Specifies the proxy settings for the Slurm queue.

```
Proxy:
  HttpProxyAddress: string
```

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`HttpProxyAddress` (Optional, `String`)

Defines an HTTP or HTTPS proxy server for the Slurm queue. Typically, it's `https://x.x.x.x:8080`.

There's no default value.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Image`

(Optional) Specifies the image to use for the Slurm queue. To use the same AMI for all nodes, use the `CustomAmi` setting in the `Image` section.

```
Image:
  CustomAmi: string
```

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Image` Properties

`CustomAmi` (Optional, `String`)

The AMI to use for the Slurm queue instead of the default AMIs. You can use the pcluster CLI command to view a list of the default AMIs.

> **Note**
>
> The AMI must be based on the same operating system that's used by the head node.

```
pcluster list-official-images
```

If the custom AMI requires additional permissions for its launch, you must add these permissions to the head node policy.

For example, if a custom AMI has an encrypted snapshot associated with it, the following additional policies are required in the head node policies.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:DescribeKey",
                "kms:ReEncrypt*",
                "kms:CreateGrant",
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:us-east-1:111122223333:key/<AWS_KMS_KEY_ID>"
            ]
        }
    ]
}
```

To troubleshoot custom AMI validation warnings, see Troubleshooting custom AMI issues.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`ComputeResources`

(Required) Defines the `ComputeResources` configuration for the Slurm queue.

> **Note**
>
> - Cluster size may change during an update. For more information, see Cluster capacity size and update.
> - New compute resources can be added to the cluster only if they are deployed in subnets that belong to CIDR blocks that exist when the cluster is created.

```
ComputeResources:
  - Name: string
    InstanceType: string
    Instances:
      - InstanceType: string
    MinCount: integer
    MaxCount: integer
    DynamicNodePriority: integer
    StaticNodePriority: integer
    SpotPrice: float
    DisableSimultaneousMultithreading: boolean
    SchedulableMemory: integer
    HealthChecks:
      Gpu:
        Enabled: boolean
    Efa:
      Enabled: boolean
      GdrSupport: boolean
    CapacityReservationTarget:
      CapacityReservationId: string
      CapacityReservationResourceGroupArn: string
    Networking:
      PlacementGroup:
        Enabled: boolean
        Name: string
```

```
    CustomSlurmSettings: dict
    Tags:
      - Key: string
        Value: string
```

Update policy: For this list values setting, a new value can be added during an update or the compute fleet must be stopped when removing an existing value.

`ComputeResources` properties

`Name` (Required, `String`)

The name of the Slurm queue compute environment. The name can have up to 25 characters.

Update policy: If this setting is changed, the update is not allowed.

`InstanceType` (Required, `String`)

The instance type that's used in this Slurm compute resource. All of the instance types in a cluster must use the same processor architecture. Instances can use either the `x86_64` or `arm64` architecture.

The cluster configuration must define either `InstanceType` or `Instances`. If both are defined, AWS ParallelCluster fails.

When you define `InstanceType`, you can't define multiple subnets. If you configure only one instance type and want to use multiple subnets, define your instance type in `Instances` rather than in `InstanceType`. For more information, see `Networking` / `SubnetIds`.

If you define a p4d or hpc6id instance type, or another instance type that has multiple network interfaces or a network interface card, you must launch the compute instances in private subnet as described in AWS ParallelCluster using two subnets. AWS public IPs can only be assigned to instances that are launched with a single network interface. For more information, see Assign a public IPv4 address during instance launch in the *Amazon EC2 User Guide for Linux Instances*.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`Instances` (Required)

Specifies the list of instance types for a compute resource. To specify the allocation strategy for the list of instance types, see `AllocationStrategy`.

The cluster configuration must define either `InstanceType` or `Instances`. If both are defined, AWS ParallelCluster fails.

For more information, see Multiple instance type allocation with Slurm.

```
Instances:
   - InstanceType: string
```

> **Note**
>
> Starting with AWS ParallelCluster version 3.7.0, `EnableMemoryBasedScheduling` can be enabled if you configure multiple instance types in `Instances`.
> For AWS ParallelCluster versions 3.2.0 to 3.6.*x*, `EnableMemoryBasedScheduling` can't be enabled if you configure multiple instance types in `Instances`.

Update policy: For this list values setting, a new value can be added during an update or the compute fleet must be stopped when removing an existing value.

`InstanceType` (Required, `String`)

The instance type to use in this Slurm compute resource. All of the instance types in a cluster must use the same processor architecture, either `x86_64` or `arm64`.

The instance types listed in `Instances` must have:

- The same number of vCPUs, or, if `DisableSimultaneousMultithreading` is set to `true`, the same number of cores.
- The same number of accelerators of the same manufacturers.
- EFA supported, if `Efa` / `Enabled` set to `true`.

The instance types that are listed in `Instances` can have:

- Different amount of memory.

  In this case, the minimum memory is to be set as a consumable Slurm resource.

> **Note**
>
> Starting with AWS ParallelCluster version 3.7.0, `EnableMemoryBasedScheduling` can be enabled if you configure multiple instance types in Instances.
> For AWS ParallelCluster versions 3.2.0 to 3.6.*x*, `EnableMemoryBasedScheduling` can't be enabled if you configure multiple instance types in Instances.

- Different network cards.

  In this case, the number of network interfaces configured for the compute resource is defined by the instance type with the smallest number of network cards.
- Different network bandwidth.
- Different instance store size.

If you define a p4d or hpc6id instance type, or another instance type that has multiple network interfaces or a network interface card, you must launch the compute instances in private subnet as described in AWS ParallelCluster using two subnets. AWS public IPs can only be assigned to instances launched with a single network interface. For more information, see Assign a public IPv4 address during instance launch in the *Amazon EC2 User Guide for Linux Instances*.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

> **Note**
>
> `Instances` is supported starting with AWS ParallelCluster version 3.3.0.

`MinCount` (Optional, `Integer`)

The minimum number of instances that the Slurm compute resource uses. The default is 0.

> **Note**
>
> Cluster size may change during an update. For more information, see Cluster capacity size and update

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`MaxCount` (Optional, `Integer`)

The maximum number of instances that the Slurm compute resource uses. The default is 10.

When you use `CapacityType = CAPACITY_BLOCK`, `MaxCount` must be equal to `MinCount` and greater than 0, because all the instances part of the Capacity Block reservation are managed as static nodes.

At cluster creation time, the head node waits for all the static nodes to be ready before signaling the success of cluster creation. However, when you use `CapacityType = CAPACITY_BLOCK`, the nodes part of the compute resources associated to Capacity Blocks won't be considered for this check. The cluster will be created even if not all the configured Capacity Blocks are active.

> **Note**
>
> Cluster size may change during an update. For more information, see Cluster capacity size and update

`DynamicNodePriority` (Optional, `Integer`)

The priority of dynamic nodes in a queue compute resource. The priority maps to the Slurm node `Weight` configuration parameter for the compute resource dynamic nodes. The default value is `1000`.

Slurm prioritizes nodes with the lowest `Weight` values first.

> **Warning**
>
> The use of many different `Weight` values in a Slurm partition (queue) might slow down the rate of job scheduling in the queue.

In AWS ParallelCluster versions earlier than version 3.7.0, both static and dynamic nodes were assigned the same default weight of 1. In this case, Slurm might prioritize idle dynamic nodes over idle static nodes due to the naming schema for static and dynamic nodes. When all else is equal, Slurm schedules nodes alphabetically by name.

Note

`DynamicNodePriority` is added in AWS ParallelCluster version 3.7.0.

Update policy: This setting can be changed during an update.

`StaticNodePriority` (Optional, `Integer`)

The priority of static nodes in a queue compute resource. The priority maps to the Slurm node `Weight` configuration parameter for the compute resource static nodes. The default value is 1.

Slurm prioritizes nodes with the lowest `Weight` values first.

Warning

The use of many different `Weight` values in a Slurm partition (queue) might slow down the rate of job scheduling in the queue.

Note

`StaticNodePriority` is added in AWS ParallelCluster version 3.7.0.

Update policy: This setting can be changed during an update.

`SpotPrice` (Optional, `Float`)

The maximum price that paid for an Amazon EC2 Spot Instance before any instances are launched. The default value is the On-Demand price.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`DisableSimultaneousMultithreading` (Optional, `Boolean`)

If `true`, multithreading on the nodes in the Slurm queue is disabled. The default value is `false`.

Not all instance types can disable multithreading. For a list of instance types that support disabling multithreading, see CPU cores and threads for each CPU core per instance type in the *Amazon EC2 User Guide*.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`SchedulableMemory` (Optional, `Integer`)

The amount of memory in MiB that's configured in the Slurm parameter `RealMemory` for the compute nodes of a compute resource. This value is the upper limit for the node memory available to jobs when `SlurmSettings` / `EnableMemoryBasedScheduling` is enabled. The default value is 95 percent of the memory that's listed in Amazon EC2 Instance Types and returned by the Amazon EC2 API DescribeInstanceTypes. Make sure to convert values that are given in GiB to MiB.

Supported values: `1-EC2Memory`

`EC2Memory` is the memory (in MiB) that's listed in Amazon EC2 Instance Types and returned by the Amazon EC2 API DescribeInstanceTypes. Make sure to convert values that are given in GiB to MiB.

This option is most relevant when `SlurmSettings` / `EnableMemoryBasedScheduling` is enabled. For more information, see Slurm memory-based scheduling.

> **Note**
>
> `SchedulableMemory` is supported starting with AWS ParallelCluster version 3.2.0. Starting with version 3.2.0, by default, AWS ParallelCluster configures `RealMemory` for Slurm compute nodes to 95 percent of the memory that's returned by the Amazon EC2 API `DescribeInstanceTypes`. This configuration is independent of the value of `EnableMemoryBasedScheduling`.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`HealthChecks` (Optional)

Specify health checks on a compute resource.

`Gpu` (Optional)

Specify GPU health checks on a compute resource.

`Enabled` (Optional, `Boolean`)

Whether AWS ParallelCluster performs GPU health checks on compute a resource in a queue. The default is `false`.

> **Note**
>
> AWS ParallelCluster doesn't support `HealthChecks` / `Gpu` in nodes that use `alinux2` ARM operating systems. These platforms don't support the NVIDIA Data Center GPU Manager (DCGM).

Gpu health check behavior

- If `Gpu` / `Enabled` is set to `true`, AWS ParallelCluster performs health GPU health checks on a compute resource.
- The `Gpu` health check performs health checks on a compute resource to prevent the submission of jobs on nodes with a degraded GPU.
- If a compute node fails a `Gpu` health check, the compute node state changes to `DRAIN`. New jobs don't start on this node. Existing jobs run to completion. After all running jobs complete, the compute node terminates if it's a dynamic node, and it's replaced if it's a static node.
- The duration of the `Gpu` health check depends on the selected instance type, the number of GPUs in the instance, and the number of `Gpu` health check targets (equivalent to the number of job GPU targets). For an instance with 8 GPUs, the typical duration is less than 3 minutes.
- If the `Gpu` health check runs on an instance that's not supported, it exits and the job runs on the compute node. For example, if an instance doesn't have a GPU, or, if an instance has a GPU, but it isn't an NVIDIA GPU, the health check exits and the job runs on the compute node. Only NVIDIA GPUs are supported.

- The Gpu health check uses the `dcgmi` tool to perform health checks on a node and takes the following steps:

  When the Gpu health check begins in a node:

  1. It detects whether the `nvidia-dcgm` and `nvidia-fabricmanager` services are running.
  2. If these services aren't running, the Gpu health check starts them.
  3. It detects whether the persistence mode is enabled.
  4. If the persistence mode isn't enabled, the Gpu health check enables it.

  At the end of the health check, the Gpu health check restores these services and resources to their initial state.
- If the job is assigned to a specific set of node GPUs, the Gpu health check runs only on that specific set. Otherwise, the Gpu health check runs on all GPUs in the node.
- If a compute node receives 2 or more Gpu health check requests at the same time, only the first health check runs and the others are skipped. This is also the case for health checks targeting node GPUs. You can check the log files for additional information regarding this situation.
- The health check log for a specific compute node is available in the `/var/log/parallelcluster/slurm_health_check.log` file. This file is available in Amazon CloudWatch, in the cluster CloudWatch log group, where you can find:
  - Details on the action run by the Gpu health check, including enabling and disabling services and persistence mode.
  - The GPU identifier, serial ID, and the UUID.
  - The health check output.

Update policy: This setting can be changed during an update.

> **Note**
>
> `HealthChecks` is supported starting in AWS ParallelCluster version 3.6.0.

`Efa` (Optional)

Specifies the Elastic Fabric Adapter (EFA) settings for the nodes in the Slurm queue.

```
Efa:
  Enabled: boolean
  GdrSupport: boolean
```

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Enabled` (Optional, `Boolean`)

Specifies that Elastic Fabric Adapter (EFA) is enabled. To view the list of Amazon EC2 instances that support EFA, see Supported instance types in the *Amazon EC2 User Guide for Linux Instances*. For more information, see Elastic Fabric Adapter. We recommend that you use a cluster `SlurmQueues` / `Networking` / `PlacementGroup` to minimize latencies between instances.

The default value is `false`.

> **Note**
>
> Elastic Fabric Adapter (EFA) isn't supported over different availability zones. For more information, see SubnetIds.

> **Warning**
>
> If you're defining a custom security group in SecurityGroups, make sure that your EFA-enabled instances are members of a security group that allows all inbound and outbound traffic to itself.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`GdrSupport` (Optional, `Boolean`)

(Optional) Starting with AWS ParallelCluster version 3.0.2, this setting has no effect. Elastic Fabric Adapter (EFA) support for GPUDirect RDMA (remote direct memory access) is always enabled if it's supported by the instance type for the Slurm compute resource and the operating system.

> **Note**
>
> AWS ParallelCluster version 3.0.0 through 3.0.1: Support for GPUDirect RDMA is enabled for Slurm compute resources. Support for GPUDirect RDMA is supported by specific instance types (`p4d.24xlarge`) on specific operating systems (`Os` is `alinux2`, `ubuntu1804`, or `ubuntu2004`). The default value is false.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`CapacityReservationTarget`

```
CapacityReservationTarget:
   CapacityReservationId: string
   CapacityReservationResourceGroupArn: string
```

Specifies the on-demand capacity reservation to use for the compute resource.

`CapacityReservationId` (Optional, `String`)

The ID of the existing capacity reservation to target for the queue's compute resources. The id can refer to an ODCR or a Capacity Block for ML.

When this parameter is specified at compute resource level, InstanceType is optional, it will be automatically retrieved from the reservation.

`CapacityReservationResourceGroupArn` (Optional, `String`)

Indicates the Amazon Resource Name (ARN) of the resource group that serves as the service linked group of capacity reservations for the compute resource. AWS ParallelCluster identifies and uses the most appropriate capacity reservation from the group. The resource group must have at least one ODCR for each instance type that's listed for the compute resource. For more information, see Launch instances with On-Demand Capacity Reservations (ODCR).

- If `PlacementGroup` is enabled in `SlurmQueues` / `Networking` or `SlurmQueues` / `ComputeResources` / `Networking`, AWS ParallelCluster selects a resource group that targets the instance type and `PlacementGroup` for a compute resource if it exists.

  The `PlacementGroup` must target one of the instances types defined in `ComputeResources`.

- If `PlacementGroup` isn't enabled in `SlurmQueues` / `Networking` or `SlurmQueues` / `ComputeResources` / `Networking`, AWS ParallelCluster selects a resource group that targets only the instance type of a compute resource, if it exists.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

> **Note**
>
> `CapacityReservationTarget` is added with AWS ParallelCluster version 3.3.0.

`Networking`

```
Networking:
  PlacementGroup:
    Enabled: boolean
    Name: string
```

Update policy: All compute nodes must be stopped for a managed placement group deletion. The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`PlacementGroup` (Optional)

Specifies the placement group settings for the compute resource.

`Enabled` (Optional, `Boolean`)

Indicates whether a placement group is used for the compute resource.

- If set to `true`, without a `Name` defined, that compute resource is assigned its own managed placement group, regardless of the `SlurmQueues` / `Networking` / `PlacementGroup` setting.
- If set to `true`, with a `Name` defined, that compute resource is assigned the named placement group, regardless of `SlurmQueues` / `Networking` / `PlacementGroup` settings.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

Name (Optional, `String`)

The placement group name for an existing cluster placement group that's used for the compute resource.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

Note

- If both `PlacementGroup` / `Enabled` and `Name` aren't set, their respective values default to the `SlurmQueues` / `Networking` / `PlacementGroup` settings.
- `ComputeResources` / `Networking` / `PlacementGroup` is added with AWS ParallelCluster version 3.3.0.

`CustomSlurmSettings` (Optional, `Dict`)

(Optional) Defines the custom Slurm node (compute resource) configuration settings.

Specifies a dictionary of custom Slurm configuration parameter key-value pairs that apply to Slurm nodes (compute resources).

Each separate key-value pair, such as `Param1: Value1`, is added separately to the end of the Slurm node configuration line in the format `Param1=Value1`.

You can only specify Slurm configuration parameters that aren't deny-listed in `CustomSlurmSettings`. For information about deny-listed Slurm configuration parameters, see Deny-listed Slurm configuration parameters for `CustomSlurmSettings`.

AWS ParallelCluster only checks whether a parameter is in a deny list. AWS ParallelCluster doesn't validate your custom Slurm configuration parameter syntax or semantics. It is your responsibility to validate your custom Slurm configuration parameters. Invalid custom Slurm configuration parameters can cause Slurm daemon failures that can lead to cluster create and update failures.

For more information about how to specify custom Slurm configuration parameters with AWS ParallelCluster, see Slurm configuration customization.

For more information about Slurm configuration parameters, see slurm.conf in the Slurm documentation.

Update policy: This setting can be changed during an update.

> **Note**
>
> `CustomSlurmSettings` is supported starting with AWS ParallelCluster version 3.6.0.

`Tags` (Optional, [String])

A list of tag key-value pairs. `ComputeResource` tags override duplicate tags specified in the `Tags` section or `SlurmQueues` / `Tags`.

`Key` (Optional, `String`)

The tag key.

`Value` (Optional, `String`)

The tag value.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`ComputeSettings`

(Required) Defines the `ComputeSettings` configuration for the Slurm queue.

`ComputeSettings` properties

Specifies the properties of `ComputeSettings` of the nodes in the Slurm queue.

```
ComputeSettings:
  LocalStorage:
    RootVolume:
      Size: integer
      Encrypted: boolean
      VolumeType: string
      Iops: integer
      Throughput: integer
     EphemeralVolume:
      MountDir: string
```

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`LocalStorage` (Optional)

Specifies the properties of `LocalStorage` of the nodes in the Slurm queue.

```
LocalStorage:
  RootVolume:
    Size: integer
    Encrypted: boolean
    VolumeType: string
    Iops: integer
    Throughput: integer
  EphemeralVolume:
    MountDir: string
```

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`RootVolume` (Optional)

Specifies the details of the root volume of the nodes in the Slurm queue.

```
RootVolume:
  Size: integer
  Encrypted: boolean
  VolumeType: string
  Iops: integer
  Throughput: integer
```

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Size` (Optional, `Integer`)

Specifies the root volume size in gibibytes (GiB) for the nodes in the Slurm queue. The default size comes from the AMI. Using a different size requires that the AMI supports `growroot`.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Encrypted` (Optional, `Boolean`)

If `true`, the root volume of the nodes in the Slurm queue are encrypted. The default value is `false`.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`VolumeType` (Optional, `String`)

Specifies the Amazon EBS volume type of the nodes in the Slurm queue. Supported values are gp2, gp3, `io1`, `io2`, `sc1`, `st1`, and `standard`. The default value is gp3.

For more information, see Amazon EBS volume types in the *Amazon EC2 User Guide*.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Iops` (Optional, `Boolean`)

Defines the number of IOPS for `io1`, `io2`, and gp3 type volumes.

The default value, supported values, and `volume_iops` to `volume_size` ratio varies by `VolumeType` and `Size`.

`VolumeType` = `io1`

Default `Iops` = 100

Supported values `Iops` = 100–64000 †

Maximum `volume_iops` to `volume_size` ratio = 50 IOPS per GiB. 5000 IOPS requires a `volume_size` of at least 100 GiB.

`VolumeType` = `io2`

Default `Iops` = 100

Supported values `Iops` = 100–64000 (256000 for `io2` Block Express volumes) †

Maximum `Iops` to `Size` ratio = 500 IOPS per GiB. 5000 IOPS requires a `Size` of at least 10 GiB.

`VolumeType = gp3`

Default `Iops` = 3000

Supported values `Iops` = 3000–16000 †

Maximum `Iops` to `Size` ratio = 500 IOPS per GiB for volumes with IOPS greater than 3000.

† Maximum IOPS is guaranteed only on Instances built on the Nitro System that are also provisioned with more than 32,000 IOPS. Other instances can have up to 32,000 IOPS. Earlier `io1` volumes might not reach full performance unless you modify the volume. `io2` Block Express volumes support `volume_iops` values up to 256000 on `R5b` instance types. For more information, see `io2` Block Express volumes in the *Amazon EC2 User Guide*.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Throughput` (Optional, `Integer`)

Defines the throughput for gp3 volume types, in MiB/s. This setting is valid only when `VolumeType` is gp3. The default value is `125`. Supported values: 125–1000 MiB/s

The ratio of `Throughput` to `Iops` can be no more than 0.25. The maximum throughput of 1000 MiB/s requires that the `Iops` setting is at least 4000.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`EphemeralVolume` (Optional, `Boolean`)

Specifies the settings for the ephemeral volume. The ephemeral volume is created by combining all instance store volumes into a single logical volume formatted with the `ext4` file system. The default is `/scratch`. If the instance type doesn't have any instance store volumes, no ephemeral volume is created. For more information, see Instance store volumes in the *Amazon EC2 User Guide*.

```
EphemeralVolume:
  MountDir: string
```

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`MountDir` (Optional, `String`)

The mount directory for the ephemeral volume for each node in the Slurm queue.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`CustomActions`

(Optional) Specifies custom scripts to run on the nodes in the Slurm queue.

```
CustomActions:
  OnNodeStart:
    Sequence:
      - Script: string
        Args:
          - string
    Script: string
    Args:
      - string
  OnNodeConfigured:
    Sequence:
      - Script: string
        Args:
          - string
    Script: string
    Args:
      - string
```

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`CustomActions` Properties

`OnNodeStart` (Optional, `String`)

Specifies a sequence of scripts or single script to run on the nodes in the Slurm queue before any node deployment bootstrap action is started. AWS ParallelCluster doesn't support including

both a single script and `Sequence` for the same custom action. For more information, see Custom bootstrap actions.

`Sequence` (Optional)

List of scripts to run.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Script` (Required, `String`)

The file to use. The file path can start with `https://` or `s3://`.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Args` (Optional, `[String]`)

The list of arguments to pass to the script.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Script` (Required, `String`)

The file to use for a single script. The file path can start with `https://` or `s3://`.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Args` (Optional, `[String]`)

The list of arguments to pass to the single script.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`OnNodeConfigured` (Optional, `String`)

Specifies a sequence of scripts or a single script to run on the nodes in the Slurm queue after all of the node bootstrap actions are complete. AWS ParallelCluster doesn't support including both a single script and `Sequence` for the same custom action. For more information, see Custom bootstrap actions.

Sequence (Optional)

List of scripts to run.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Script` (Required, `String`)

The file to use. The file path can start with `https://` or `s3://`.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Args` (Optional, `[String]`)

The list of arguments to pass to the script.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Script` (Required, `String`)

The file to use for a single script. The file path can start with `https://`or `s3://`.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

`Args` (Optional, `[String]`)

A list of arguments to pass to the single script.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

Note

`Sequence` is added starting with AWS ParallelCluster version 3.6.0. When you specify `Sequence`, you can list multiple scripts for a custom action. AWS ParallelCluster continues to support configuring a custom action with a single script, without including `Sequence`.

AWS ParallelCluster doesn't support including both a single script and Sequence for the same custom action.

Iam

(Optional) Defines optional IAM settings for the Slurm queue.

```
Iam:
  S3Access:
    - BucketName: string
      EnableWriteAccess: boolean
      KeyName: string
  AdditionalIamPolicies:
    - Policy: string
  InstanceProfile: string
  InstanceRole: string
```

Update policy: This setting can be changed during an update.

Iam Properties

InstanceProfile (Optional, String)

Specifies an instance profile to override the default instance role or instance profile for the Slurm queue. You cannot specify both `InstanceProfile` and `InstanceRole`. The format is `arn:${Partition}:iam::${Account}:instance-profile/${InstanceProfileName}`.

If this is specified, the `S3Access` and `AdditionalIamPolicies` settings can't be specified.

We recommend that you specify one or both of the `S3Access` and `AdditionalIamPolicies` settings because features added to AWS ParallelCluster often require new permissions.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

InstanceRole (Optional, String)

Specifies an instance role to override the default instance role or instance profile for the Slurm queue. You cannot specify both `InstanceProfile` and `InstanceRole`. The format is `arn:${Partition}:iam::${Account}:role/${RoleName}`.

If this is specified, the `S3Access` and `AdditionalIamPolicies` settings can't be specified.

We recommend that you specify one or both of the `S3Access` and `AdditionalIamPolicies` settings because features added to AWS ParallelCluster often require new permissions.

Update policy: This setting can be changed during an update.

`S3Access` (Optional)

Specifies a bucket for the Slurm queue. This is used to generate policies to grant the specified access to the bucket in the Slurm queue.

If this is specified, the `InstanceProfile` and `InstanceRole` settings can't be specified.

We recommend that you specify one or both of the `S3Access` and `AdditionalIamPolicies` settings because features added to AWS ParallelCluster often require new permissions.

```
S3Access:
  - BucketName: string
    EnableWriteAccess: boolean
    KeyName: string
```

Update policy: This setting can be changed during an update.

`BucketName` (Required, `String`)

The name of the bucket.

Update policy: This setting can be changed during an update.

`KeyName` (Optional, `String`)

The key for the bucket. The default value is `*`.

Update policy: This setting can be changed during an update.

`EnableWriteAccess` (Optional, `Boolean`)

Indicates whether write access is enabled for the bucket.

Update policy: This setting can be changed during an update.

`AdditionalIamPolicies` (Optional)

Specifies a list of Amazon Resource Names (ARNs) of IAM policies for Amazon EC2 . This list is attached to the root role used for the Slurm queue in addition to the permissions that are required by AWS ParallelCluster.

An IAM policy name and its ARN are different. Names can't be used.

If this is specified, the `InstanceProfile` and `InstanceRole` settings can't be specified.

We recommend that you use `AdditionalIamPolicies` because `AdditionalIamPolicies` are added to the permissions that AWS ParallelCluster requires, and the `InstanceRole` must include all permissions required. The permissions required often change from release to release as features are added.

There's no default value.

```
AdditionalIamPolicies:
  - Policy: string
```

Update policy: This setting can be changed during an update.

`Policy` (Required, `[String]`)

List of IAM policies.

Update policy: This setting can be changed during an update.

`SlurmSettings`

(Optional) Defines the settings for Slurm that apply to the entire cluster.

```
SlurmSettings:
  ScaledownIdletime: integer
  QueueUpdateStrategy: string
  EnableMemoryBasedScheduling: boolean
  CustomSlurmSettings: [dict]
  CustomSlurmSettingsIncludeFile: string
  Database:
    Uri: string
    UserName: string
    PasswordSecretArn: string
  ExternalSlurmdbd:
    Host: string
    Port: integer
  Dns:
    DisableManagedDns: boolean
    HostedZoneId: string
    UseEc2Hostnames: boolean
```

`SlurmSettings` Properties

`ScaledownIdletime` (Optional, `Integer`)

Defines the amount of time (in minutes) that there's no job and the Slurm node terminates.

The default value is `10`.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`MungeKeySecretArn` (Optional, `String`)

The Amazon Resource Name (ARN) of the plaintext AWS Secrets Manager secret that contains the base64-encoded munge key to be used in the Slurm cluster. This munge key will be used to authenticate RPC calls between Slurm client commands and Slurm daemons acting as remote servers. If MungeKeySecretArn is not provided, AWS ParallelCluster will generate a random munge key for the cluster.

> **Note**
>
> `MungeKeySecretArn` is supported starting with AWS ParallelCluster version 3.8.0.

> **Warning**
>
> If the MungeKeySecretArn is newly added to an existing cluster, ParallelCluster will not restore the previous munge Key in the event of a Rollback or when later removing the MungeKeySecretArn. Instead, a new random munge key will be generated.

If the AWS ParallelCluster user has the permission to DescribeSecret on that specific secret resource, MungeKeySecretArn is validated. MungeKeySecretArn is valid if:

- The specified secret exists, and
- The secret is plaintext and contains a valid base64-encoded string, and
- The decoded binary munge key has a size between 256 and 8192 bits.

If the pcluster user IAM policy doesn't include DescribeSecret, MungeKeySecretArn is not validated and a warning message is displayed. For more information, see Base AWS ParallelCluster `pcluster` user policy.

When you update MungeKeySecretArn, the compute fleet and all login nodes must be stopped.

If the secret value in the secret ARN is modified while the ARN remains the same, the cluster won't automatically be updated with the new munge key. In order to use the secret ARN's new munge key, you must stop the compute fleet and login nodes then run the following command from the head node.

```
sudo /opt/parallelcluster/scripts/slurm/update_munge_key.sh
```

After you run the command, you can resume both the compute fleet and login nodes: the newly provisioned compute and login nodes will automatically start using the new munge key.

To generate a base64-encoded custom munge key, you can use the mungekey utility distributed with the munge software and then encode it using the base64 utility generally available in your OS. Alternatively, you either use bash (please set the bs parameter between 32 and 1024)

```
dd if=/dev/random bs=128 count=1 2>/dev/null | base64 -w 0
```

or Python as follows:

```
import random
import os
import base64

# key length in bytes
key_length=128

base64.b64encode(os.urandom(key_length)).decode("utf-8")
```

Update Policy: The compute fleet and login nodes must be stopped for this setting to be changed for an update.

`QueueUpdateStrategy` (Optional, `String`)

Specifies the replacement strategy for the `SlurmQueues` section parameters that have the following update policy:

Update policy: The compute fleet must be stopped or `QueueUpdateStrategy` must be set for this setting to be changed for an update.

The `QueueUpdateStrategy` value is used only when a cluster update process starts.

Valid values: `COMPUTE_FLEET_STOP` | `DRAIN` | `TERMINATE`

Default value: `COMPUTE_FLEET_STOP`

`DRAIN`

Nodes in queues with changed parameter values are set to `DRAINING`. Nodes in this state don't accept new jobs and running jobs continue to completion.

After a node becomes `idle` (`DRAINED`), a node is replaced if the node is static, and the node is terminated if the node is dynamic. Other nodes in other queues without changed parameter values aren't impacted.

The time this strategy needs to replace all of the queue nodes with changed parameter values depends on the running workload.

`COMPUTE_FLEET_STOP`

The default value of the `QueueUpdateStrategy` parameter. With this setting, updating parameters under the `SlurmQueues` section requires you to stop the compute fleet before you perform a cluster update:

```
$ pcluster update-compute-fleet --status STOP_REQUESTED
```

`TERMINATE`

In queues with changed parameter values, running jobs are terminated and the nodes are powered down immediately.

Static nodes are replaced and dynamic nodes are terminated.

Other nodes in other queues without changed parameter values aren't impacted.

Update policy: This setting is not analyzed during an update.

> **Note**
>
> `QueueUpdateStrategy` is supported starting with AWS ParallelCluster version 3.2.0.

`EnableMemoryBasedScheduling` (Optional, `Boolean`)

If `true`, memory-based scheduling is enabled in Slurm. For more information, see `SlurmQueues` / `ComputeResources` / `SchedulableMemory`.

The default value is `false`.

Warning

Enabling memory-based scheduling impacts the way that the Slurm scheduler handles jobs and node allocation.
For more information, see Slurm memory-based scheduling.

Note

`EnableMemoryBasedScheduling` is supported starting with AWS ParallelCluster version 3.2.0.

Note

Starting with AWS ParallelCluster version 3.7.0, `EnableMemoryBasedScheduling` can be enabled if you configure multiple instance types in Instances.
For AWS ParallelCluster versions 3.2.0 to 3.6.*x*, `EnableMemoryBasedScheduling` can't be enabled if you configure multiple instance types in Instances.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`CustomSlurmSettings` (Optional, `[Dict]`)

Defines the custom Slurm settings that apply to the entire cluster.

Specifies a list of Slurm configuration dictionaries of key-value pairs to be appended to the end of the `slurm.conf` file that AWS ParallelCluster generates.

Each dictionary in the list appears as a separate line added to the Slurm configuration file. You can specify either simple or complex parameters.

Simple parameters consist of a single key pair, as shown in the following examples:

```
- Param1: 100
- Param2: "SubParam1,SubParam2=SubValue2"
```

Example rendered in Slurm configuration:

```
Param1=100
Param2=SubParam1,SubParam2=SubValue2
```

Complex Slurm configuration parameters consist of multiple space-separated key-value, pairs as shown in the next examples:

```
- NodeName: test-nodes[1-10]
  CPUs: 4
  RealMemory: 4196
  ... # other node settings
- NodeSet: test-nodeset
  Nodes: test-nodes[1-10]
  ... # other nodeset settings
- PartitionName: test-partition
  Nodes: test-nodeset
  ... # other partition settings
```

Example, rendered in Slurm configuration:

```
NodeName=test-nodes[1-10] CPUs=4 RealMemory=4196 ... # other node settings
NodeSet=test-nodeset Nodes=test-nodes[1-10] ... # other nodeset settings
PartitionName=test-partition Nodes=test-nodeset ... # other partition settings
```

Note

Custom Slurm nodes must not contain the `-st-` or `-dy-` patterns in their names. These patterns are reserved for nodes managed by AWS ParallelCluster.

If you specify custom Slurm configuration parameters in `CustomSlurmSettings`, you must not specify custom Slurm configuration parameters for `CustomSlurmSettingsIncludeFile`.

You can only specify Slurm configuration parameters that aren't deny-listed in `CustomSlurmSettings`. For information about deny-listed Slurm configuration parameters, see Deny-listed Slurm configuration parameters for `CustomSlurmSettings`.

AWS ParallelCluster only checks whether a parameter is in a deny list. AWS ParallelCluster doesn't validate your custom Slurm configuration parameter syntax or semantics. It is your responsibility to validate your custom Slurm configuration parameters. Invalid custom Slurm

configuration parameters can cause Slurm daemon failures that can lead to cluster create and update failures.

For more information about how to specify custom Slurm configuration parameters with AWS ParallelCluster, see Slurm configuration customization.

For more information about Slurm configuration parameters, see slurm.conf in the Slurm documentation.

Update policy: This setting can be changed during an update.

> **Note**
>
> `CustomSlurmSettings` is supported starting with AWS ParallelCluster version 3.6.0.

`CustomSlurmSettingsIncludeFile` (Optional, `String`)

Defines the custom Slurm settings that apply to the entire cluster.

Specifies the custom Slurm file consisting of custom Slurm configuration parameters to be appended at the end of the `slurm.conf` file that AWS ParallelCluster generates.

You must include the path to the file. The path can start with `https://` or `s3://`.

If you specify custom Slurm configuration parameters for `CustomSlurmSettingsIncludeFile`, you must not specify custom Slurm configuration parameters for `CustomSlurmSettings`.

> **Note**
>
> Custom Slurm nodes must not contain the `-st-` or `-dy-` patterns in their names. These patterns are reserved for nodes managed by AWS ParallelCluster.

You can only specify Slurm configuration parameters that aren't deny-listed in `CustomSlurmSettingsIncludeFile`. For information about deny-listed Slurm configuration parameters, see Deny-listed Slurm configuration parameters for `CustomSlurmSettings`.

AWS ParallelCluster only checks whether a parameter is in a deny list. AWS ParallelCluster doesn't validate your custom Slurm configuration parameter syntax or semantics. It is your

responsibility to validate your custom Slurm configuration parameters. Invalid custom Slurm configuration parameters can cause Slurm daemon failures that can lead to cluster create and update failures.

For more information about how to specify custom Slurm configuration parameters with AWS ParallelCluster, see Slurm configuration customization.

For more information about Slurm configuration parameters, see slurm.conf in the Slurm documentation.

Update policy: This setting can be changed during an update.

> **Note**
>
> `CustomSlurmSettings` is supported starting with AWS ParallelCluster version 3.6.0.

`Database`

(Optional) Defines the settings to enable Slurm Accounting on the cluster. For more information, see Slurm accounting with AWS ParallelCluster.

```
Database:
   Uri: string
   UserName: string
   PasswordSecretArn: string
```

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`Database` properties

`Uri` (Required, `String`)

The address to the database server that's used as the backend for Slurm accounting. This URI must be formatted as `host:port` and must not contain a scheme, such as `mysql://`. The host can either be an IP address or a DNS name that's resolvable by the head node. If a port isn't provided, AWS ParallelCluster uses the MySQL default port 3306.

AWS ParallelCluster bootstraps the Slurm accounting database to the cluster and must access the database.

The database must be reachable before the following occurs:

- A cluster is created.
- Slurm accounting is enabled with a cluster update.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`UserName` (Required, `String`)

The identity that Slurm uses to connect to the database, write accounting logs, and perform queries. The user must have both read and write permissions on the database.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`PasswordSecretArn` (Required, `String`)

The Amazon Resource Name (ARN) of the AWS Secrets Manager secret that contains the `UserName` plaintext password. This password is used together with `UserName` and Slurm accounting to authenticate on the database server.

> **Note**
>
> - When you create a secret using the AWS Secrets Manager console be sure to select "Other type of secret", select plaintext, and only include the password text in the secret.
> - You cannot use the '#' character in the Database password as Slurm does not support it in slurmdbd.conf.
> - For more information on how to use AWS Secrets Manager to create a secret refer to Create an AWS Secrets Manager Secret.

If the user has the permission to DescribeSecret, `PasswordSecretArn` is validated. `PasswordSecretArn` is valid if the specified secret exists. If the user IAM policy doesn't include `DescribeSecret`, `PasswordSecretArn` isn't validated and a warning message is displayed. For more information, see Base AWS ParallelCluster `pcluster` user policy.

When you update `PasswordSecretArn`, the compute fleet must be stopped. If the secret value changes, and the secret ARN doesn't change, the cluster isn't automatically updated with the new database password. To update the cluster for the new secret value, you must run the following command from within the head node after the compute fleet is stopped.

```
$ sudo /opt/parallelcluster/scripts/slurm/update_slurm_database_password.sh
```

> **Warning**
>
> We recommend that you only change the database password when the compute fleet is stopped to avoid loss of accounting data.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`DatabaseName` (Optional, `String`)

Name of the database on the database server (defined by the Uri parameter) to be used for Slurm Accounting.

The name of the database may contain lowercase letters, numbers and underscores. The name may not be longer than 64 characters.

This parameter maps to the `StorageLoc` parameter of slurmdbd.conf.

If `DatabaseName` is not provided, ParallelCluster will use the name of the cluster to define a value for `StorageLoc`.

Updating the `DatabaseName` is allowed, with the following considerations:

- If a database with a name DatabaseName does not yet exist on the database server, slurmdbd will create it. It will be your responsibility to reconfigure the new database as needed (e.g. adding the accounting entities — clusters, accounts, users, associations, QOSs, etc.).
- If a database with a name DatabaseName already exists on the database server, slurmdbd will use it for the Slurm Accounting functionality.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

> **Note**
>
> `Database` is added starting with release 3.3.0.

ExternalSlurmdbd

(Optional) Defines the settings to enable Slurm Accounting with an external slurmdbd server. For more information, see Slurm accounting with AWS ParallelCluster.

```
ExternalSlurmdbd:
  Host: string
  Port: integer
```

`ExternalSlurmdbd` properties

`Host` (Required, `String`)

The address to the external slurmdbd server for Slurm accounting. The host can either be an IP address or a DNS name that's resolvable by the head node.

Update policy: This setting can be changed during an update.

`Port` (Optional, `Integer`)

The port the slurmdbd service listens to. The default value is 6819.

Update policy: This setting can be changed during an update.

`Dns`

(Optional) Defines the settings for Slurm that apply to the entire cluster.

```
Dns:
  DisableManagedDns: boolean
  HostedZoneId: string
  UseEc2Hostnames: boolean
```

`Dns` properties

`DisableManagedDns` (Optional, `Boolean`)

If `true`, the DNS entries for the cluster aren't created and Slurm node names aren't resolvable.

By default, AWS ParallelCluster creates a Route 53 hosted zone where nodes are registered when launched. The default value is `false`. If `DisableManagedDns` is set to `true`, the hosted zone isn't created by AWS ParallelCluster.

To learn how to use this setting to deploy clusters in subnets with no internet access, see AWS ParallelCluster in a single subnet with no internet access.

> **Warning**
>
> A name resolution system is required for the cluster to operate properly. If `DisableManagedDns` is set to `true`, you must provide a name resolution system. To use the Amazon EC2 default DNS, set `UseEc2Hostnames` to `true`. Alternatively, configure your own DNS resolver and make sure that node names are registered when instances are launched. For example, you can do this by configuring `CustomActions` / `OnNodeStart`.

Update policy: If this setting is changed, the update is not allowed.

`HostedZoneId` (Optional, `String`)

Defines a custom Route 53 hosted zone ID to use for DNS name resolution for the cluster. When provided, AWS ParallelCluster registers cluster nodes in the specified hosted zone and doesn't create a managed hosted zone.

Update policy: If this setting is changed, the update is not allowed.

`UseEc2Hostnames` (Optional, `Boolean`)

If `true`, cluster compute nodes are configured with the default EC2 hostname. The Slurm `NodeHostName` is also updated with this information. The default is `false`.

To learn how to use this setting to deploy clusters in subnets with no internet access, see AWS ParallelCluster in a single subnet with no internet access.

> **Note**
>
> **This note isn't relevant starting with AWS ParallelCluster version 3.3.0.**
> For AWS ParallelCluster supported versions prior to 3.3.0:
> When `UseEc2Hostnames` is set to `true`, the Slurm configuration file is set with the AWS ParallelCluster `prolog` and `epilog` scripts:
>
> - `prolog` runs to add nodes info to `/etc/hosts` on compute nodes when each job is allocated.
> - `epilog` runs to clean contents written by `prolog`.

To add custom `prolog` or `epilog` scripts, add them to the `/opt/slurm/etc/pcluster/prolog.d/` or `/opt/slurm/etc/pcluster/epilog.d/` folders respectively.

Update policy: If this setting is changed, the update is not allowed.

`SharedStorage` section

(Optional) The shared storage settings for the cluster.

AWS ParallelCluster supports either using Amazon EBS, FSx for ONTAP, and FSx for OpenZFS shared storage volumes, Amazon EFS and FSx for Lustre shared storage file systems, or File Caches.

In the `SharedStorage` section, you can define either external or managed storage:

- **External storage** refers to an existing volume or file system that you manage. AWS ParallelCluster doesn't create or delete it.
- **AWS ParallelCluster managed storage** refers to a volume or file system that AWS ParallelCluster created and can delete.

For shared storage quotas and more information about how to configure your shared storage, see Shared storage in *Using AWS ParallelCluster*.

Note

If AWS Batch is used as a scheduler, FSx for Lustre is only available on the cluster head node.

```
SharedStorage:
  - MountDir: string
    Name: string
    StorageType: Ebs
    EbsSettings:
      VolumeType: string
      Iops: integer
      Size: integer
      Encrypted: boolean
```

```
      KmsKeyId: string
      SnapshotId: string
      Throughput: integer
      VolumeId: string
      DeletionPolicy: string
      Raid:
        Type: string
        NumberOfVolumes: integer
  - MountDir: string
    Name: string
    StorageType: Efs
    EfsSettings:
      Encrypted: boolean
      KmsKeyId: string
      EncryptionInTransit: boolean
      IamAuthorization: boolean
      PerformanceMode: string
      ThroughputMode: string
      ProvisionedThroughput: integer
      FileSystemId: string
      DeletionPolicy: string
      AccessPointId: string
  - MountDir: string
    Name: string
    StorageType: FsxLustre
    FsxLustreSettings:
      StorageCapacity: integer
      DeploymentType: string
      ImportedFileChunkSize: integer
      DataCompressionType: string
      ExportPath: string
      ImportPath: string
      WeeklyMaintenanceStartTime: string
      AutomaticBackupRetentionDays: integer
      CopyTagsToBackups: boolean
      DailyAutomaticBackupStartTime: string
      PerUnitStorageThroughput: integer
      BackupId: string
      KmsKeyId: string
      FileSystemId: string
      AutoImportPolicy: string
      DriveCacheType: string
      StorageType: string
      DeletionPolicy: string
```

```
      DataRepositoryAssociations:
      - Name: string
        BatchImportMetaDataOnCreate: boolean
        DataRepositoryPath: string
        FileSystemPath: string
        ImportedFileChunkSize: integer
        AutoExportPolicy: string
        AutoImportPolicy: string
  - MountDir: string
    Name: string
    StorageType: FsxOntap
    FsxOntapSettings:
      VolumeId: string
  - MountDir: string
    Name: string
    StorageType: FsxOpenZfs
    FsxOpenZfsSettings:
      VolumeId: string
  - MountDir: string
    Name: string
    StorageType: FileCache
    FileCacheSettings:
      FileCacheId: string
```

`SharedStorage` update policies

- For managed/external EBS, managed EFS and managed FSx Lustre, the update policy is Update policy: For this list values setting, the compute fleet must be stopped or `QueueUpdateStrategy` must be set to add a new value; the compute fleet must be stopped when removing an existing value.
- For external EFS, FSx Lustre, FSx ONTAP, FSx OpenZfs and File Cache, the update policy is, **Update policy: This setting can be changed during an update.**

`SharedStorage` properties

`MountDir` (**Required**, `String`)

The path where the shared storage is mounted.

Update policy: If this setting is changed, the update is not allowed.

Name (**Required**, String)

The name of the shared storage. You use this name when you update the settings.

> **Warning**
>
> If you specify AWS ParallelCluster managed shared storage, and you change the value for Name, the existing managed shared storage and data is deleted and new managed shared storage is created. Changing the value for Name with a cluster update is equivalent to replacing the existing managed shared storage with a new one. Make sure you back up your data before you change Name if you need to retain the data from the existing shared storage.

Update policy: For this list values setting, the compute fleet must be stopped or QueueUpdateStrategy must be set to add a new value; the compute fleet must be stopped when removing an existing value.

StorageType (**Required**, String)

The type of the shared storage. Supported values are Ebs, Efs, FsxLustre, FsxOntap, and FsxOpenZfs.

For more information, see FsxLustreSettings, FsxOntapSettings, and FsxOpenZfsSettings.

> **Note**
>
> If you use AWS Batch as a scheduler, FSx for Lustre is only available on the cluster head node.

Update policy: If this setting is changed, the update is not allowed.

EbsSettings

(Optional) The settings for an Amazon EBS volume.

```
EbsSettings:
```

```
  VolumeType: string
  Iops: integer
  Size: integer
  Encrypted: boolean
  KmsKeyId: string
  SnapshotId: string
  VolumeId: string
  Throughput: integer
  DeletionPolicy: string
  Raid:
    Type: string
    NumberOfVolumes: integer
```

Update policy: If this setting is changed, the update is not allowed.

`EbsSettings` properties

When the DeletionPolicy is set to `Delete`, a managed volume, with its data, is deleted if the cluster is deleted or if the volume is removed with a cluster update.

For more information, see Shared storage in *Using AWS ParallelCluster*.

`VolumeType` (**Optional**, `String`)

Specifies the Amazon EBS volume type. Supported values are `gp2`, `gp3`, `io1`, `io2`, `sc1`, `st1`, and `standard`. The default value is `gp3`.

For more information, see Amazon EBS volume types in the *Amazon EC2 User Guide*.

Update policy: If this setting is changed, the update is not allowed.

`Iops` (**Optional**, `Integer`)

Defines the number of IOPS for `io1`, `io2`, and `gp3` type volumes.

The default value, supported values, and `volume_iops` to `volume_size` ratio varies by `VolumeType` and `Size`.

`VolumeType` = `io1`

Default `Iops` = 100

Supported values `Iops` = 100–64000 †

Maximum `volume_iops` to `volume_size` ratio = 50 IOPS for each GiB. 5000 IOPS requires a `volume_size` of at least 100 GiB.

`VolumeType` = `io2`

Default `Iops` = 100

Supported values `Iops` = 100–64000 (256000 for `io2` Block Express volumes) †

Maximum `Iops` to `Size` ratio = 500 IOPS for each GiB. 5000 IOPS requires a `Size` of at least 10 GiB.

`VolumeType` = `gp3`

Default `Iops` = 3000

Supported values `Iops` = 3000–16000

Maximum `Iops` to `Size` ratio = 500 IOPS for each GiB. 5000 IOPS requires a `Size` of at least 10 GiB.

† Maximum IOPS is guaranteed only on Instances built on the Nitro System provisioned with more than 32,000 IOPS. Other instances guarantee up to 32,000 IOPS. Unless you modify the volume, earlier `io1` volumes might not reach full performance. `io2` Block Express volumes support `volume_iops` values up to 256000 on R5b instance types. For more information, see `io2` Block Express volumes in the *Amazon EC2 User Guide*.

Update policy: This setting can be changed during an update.

`Size` (**Optional**, `Integer`)

Specifies the volume size in gibibytes (GiB). The default value is 35.

Update policy: If this setting is changed, the update is not allowed.

`Encrypted` (**Optional**, `Boolean`)

Specifies if the volume is encrypted. The default value is `true`.

Update policy: If this setting is changed, the update is not allowed.

`KmsKeyId` (**Optional**, `String`)

Specifies a custom AWS KMS key to use for encryption. This setting requires that the `Encrypted` setting is set to `true`.

Update policy: If this setting is changed, the update is not allowed.

`SnapshotId` (**Optional**, `String`)

Specifies the Amazon EBS snapshot ID if you use a snapshot as the source for the volume.

Update policy: If this setting is changed, the update is not allowed.

`VolumeId` (**Optional**, `String`)

Specifies the Amazon EBS volume ID. When this is specified for an `EbsSettings` instance, only the `MountDir` parameter can also be specified.

The volume must be created in the same Availability Zone as the `HeadNode`.

> **Note**
>
> Multiple Availability Zones is added in AWS ParallelCluster version 3.4.0.

Update policy: If this setting is changed, the update is not allowed.

`Throughput` (**Optional**, `Integer`)

The throughput, in MiB/s to provision for a volume, with a maximum of 1,000 MiB/s.

This setting is valid only when `VolumeType` is gp3. The supported range is 125 to 1000, with a default value of 125.

Update policy: This setting can be changed during an update.

`DeletionPolicy` (**Optional**, `String`)

Specifies whether the volume should be retained, deleted, or snapshotted when the cluster is deleted or the volume is removed. The supported values are `Delete`, `Retain`, and `Snapshot`. The default value is `Delete`.

When the DeletionPolicy set to `Delete`, a managed volume, with its data, is deleted if the cluster is deleted or if the volume is removed with a cluster update.

For more information, see Shared storage.

Update policy: This setting can be changed during an update.

> **Note**
>
> `DeletionPolicy` is supported starting with AWS ParallelCluster version 3.2.0.

`Raid`

(Optional) Defines the configuration of a RAID volume.

```
Raid:
  Type: string
  NumberOfVolumes: integer
```

Update policy: If this setting is changed, the update is not allowed.

`Raid` properties

`Type` (**Required**, `String`)

Defines the type of RAID array. Supported values are "0" (striped) and "1" (mirrored).

Update policy: If this setting is changed, the update is not allowed.

`NumberOfVolumes` (**Optional**, `Integer`)

Defines the number of Amazon EBS volumes to use to create the RAID array. The supported range of values is 2-5. The default value (when the `Raid` setting is defined) is 2.

Update policy: If this setting is changed, the update is not allowed.

`EfsSettings`

(Optional) The settings for an Amazon EFS file system.

```
EfsSettings:
  Encrypted: boolean
  KmsKeyId: string
  EncryptionInTransit: boolean
  IamAuthorization: boolean
  PerformanceMode: string
  ThroughputMode: string
```

```
  ProvisionedThroughput: integer
  FileSystemId: string
  DeletionPolicy: string
  AccessPointId: string
```

Update policy: If this setting is changed, the update is not allowed.

`EfsSettings` properties

When the DeletionPolicy set to `Delete`, a managed file system, with its data, is deleted if the cluster is deleted, or if the file system is removed with a cluster update.

For more information, see Shared storage in *Using AWS ParallelCluster*.

`Encrypted` (**Optional**, `Boolean`)

Specifies if the Amazon EFS file system is encrypted. The default value is `false`.

Update policy: If this setting is changed, the update is not allowed.

`KmsKeyId` (**Optional**, `String`)

Specifies a custom AWS KMS key to use for encryption. This setting requires that the `Encrypted` setting is set to `true`.

Update policy: If this setting is changed, the update is not allowed.

`EncryptionInTransit` (**Optional**, `Boolean`)

If set to `true`, Amazon EFS file systems are mounted using Transport Layer Security (TLS). By default, this is set to `false`.

> **Note**
>
> If AWS Batch is used as scheduler, `EncryptionInTransit` isn't supported.

> **Note**
>
> `EncryptionInTransit` is added starting with AWS ParallelCluster version 3.4.0.

Update policy: If this setting is changed, the update is not allowed.

`IamAuthorization` (**Optional**, `Boolean`)

`IamAuthorization` is added starting with AWS ParallelCluster version 3.4.0.

If set to `true`, Amazon EFS is authenticated by using the system's IAM identity. By default, this is set to `false`.

> **Note**
>
> If `IamAuthorization` is set to `true`, `EncryptionInTransit` must also be set to `true`.

> **Note**
>
> If AWS Batch is used as scheduler, `IamAuthorization` isn't supported.

Update policy: If this setting is changed, the update is not allowed.

`PerformanceMode` (**Optional**, `String`)

Specifies the performance mode of the Amazon EFS file system. Supported values are `generalPurpose` and `maxIO`. The default value is `generalPurpose`. For more information, see Performance modes in the *Amazon Elastic File System User Guide*.

We recommend the `generalPurpose` performance mode for most file systems.

File systems that use the `maxIO` performance mode can scale to higher levels of aggregate throughput and operations per second. However, there's a trade-off of slightly higher latencies for most file operations.

Update policy: If this setting is changed, the update is not allowed.

`ThroughputMode` (**Optional**, `String`)

Specifies the throughput mode of the Amazon EFS file system. Supported values are `bursting` and `provisioned`. The default value is `bursting`. When `provisioned` is used, `ProvisionedThroughput` must be specified.

Update policy: This setting can be changed during an update.

`ProvisionedThroughput` (**Required** when `ThroughputMode` is `provisioned`, `Integer`)

Defines the provisioned throughput (in MiB/s) of the Amazon EFS file system, measured in MiB/s. This corresponds to the ProvisionedThroughputInMibps parameter in the *Amazon EFS API Reference*.

If you use this parameter, you must set `ThroughputMode` to `provisioned`.

The supported range is `1-1024`. To request a limit increase, contact Support.

Update policy: This setting can be changed during an update.

`FileSystemId` (**Optional**, `String`)

Defines the Amazon EFS file system ID for an existing file system.

If the cluster is configured to span multiple Availability Zones, you must define a file system mount target in each Availability Zone that's used by the cluster.

When this is specified, only `MountDir` can be specified. No other `EfsSettings` can be specified.

If you set this option, the following must be true for the file systems that you define:

- The file systems have an existing mount target in each of the cluster's Availability Zones, with inbound and outbound NFS traffic allowed from the `HeadNode` and `ComputeNodes`. Multiple availability zones are configured in Scheduling / SlurmQueues / Networking / SubnetIds.

  To make sure traffic is allowed between the cluster and file system, you can do one of the following:

  - Configure the security groups of the mount target to allow the traffic to and from the CIDR or prefix list of cluster subnets.

    > **Note**
    >
    > AWS ParallelCluster validates that ports are open and that the CIDR or prefix list is configured. AWS ParallelCluster doesn't validate the content of CIDR block or prefix list.

- Set custom security groups for cluster nodes by using `SlurmQueues` / `Networking` / `SecurityGroups` and `HeadNode` / `Networking` / `SecurityGroups`. The custom security groups must be configured to allow traffic between the cluster and the file system.

  Note

  If all cluster nodes use custom security groups, AWS ParallelCluster only validates that the ports are open. AWS ParallelCluster doesn't validate that the source and destination are properly configured.

Warning

EFS OneZone is only supported if all compute nodes and the head node are in the same Availability Zone. EFS OneZone can have only one mount target.

Note

Multiple Availability Zones is added in AWS ParallelCluster version 3.4.0.

Update policy: If this setting is changed, the update is not allowed.

`DeletionPolicy` (**Optional**, `String`)

Specifies whether the file system should be retained or deleted when the file system is removed from the cluster or the cluster is deleted. The supported values are `Delete` and `Retain`. The default value is `Delete`.

When the DeletionPolicy is set to `Delete`, a managed file system, with its data, is deleted if the cluster is deleted, or if the file system is removed with a cluster update.

For more information, see Shared storage.

Update policy: This setting can be changed during an update.

Note

`DeletionPolicy` is supported starting with AWS ParallelCluster version 3.3.0.

`AccessPointId` (**Optional**, `String`)

If this option is specified, the filesystem entry point defined by the `access point ID` will be mounted rather than the filesystem root.

For more information, see Shared storage.

Update policy: If this setting is changed, the update is not allowed.

`FsxLustreSettings`

> **Note**
>
> You must define `FsxLustreSettings` if `FsxLustre` is specified for `StorageType`.

(Optional) The settings for an FSx for Lustre file system.

```
FsxLustreSettings:
  StorageCapacity: integer
  DeploymentType: string
  ImportedFileChunkSize: integer
  DataCompressionType: string
  ExportPath: string
  ImportPath: string
  WeeklyMaintenanceStartTime: string
  AutomaticBackupRetentionDays: integer
  CopyTagsToBackups: boolean
  DailyAutomaticBackupStartTime: string
  PerUnitStorageThroughput: integer
  BackupId: string # BackupId cannot coexist with some of the fields
  KmsKeyId: string
  FileSystemId: string # FileSystemId cannot coexist with other fields
  AutoImportPolicy: string
  DriveCacheType: string
  StorageType: string
  DeletionPolicy: string
```

Update policy: If this setting is changed, the update is not allowed.

Note

If AWS Batch is used as a scheduler, FSx for Lustre is only available on the cluster head node.

`FsxLustreSettings` properties

When the DeletionPolicy is set to `Delete`, a managed file system, with its data, is deleted if the cluster is deleted, or if the file system is removed with a cluster update.

For more information, see Shared storage.

`StorageCapacity` (**Required**, `Integer`)

Sets the storage capacity of the FSx for Lustre file system, in GiB. `StorageCapacity` is required if you're creating a new file system. Do not include `StorageCapacity` if `BackupId` or `FileSystemId` is specified.

- For `SCRATCH_2`, `PERSISTENT_1`, and `PERSISTENT_2` deployment types, valid values are 1200 GiB, 2400 GiB, and increments of 2400 GiB.
- For `SCRATCH_1` deployment type, valid values are 1200 GiB, 2400 GiB, and increments of 3600 GiB.

Update policy: If this setting is changed, the update is not allowed.

`DeploymentType` (**Optional**, `String`)

Specifies the deployment type of the FSx for Lustre file system. Supported values are `SCRATCH_1`, `SCRATCH_2`, `PERSISTENT_1`, and `PERSISTENT_2`. The default value is `SCRATCH_2`.

Choose `SCRATCH_1` and `SCRATCH_2` deployment types when you need temporary storage and shorter term processing of data. The `SCRATCH_2` deployment type provides in transit encryption of data and higher burst throughput capacity than `SCRATCH_1`.

Choose `PERSISTENT_1` deployment type for longer term storage and for throughput focused workloads that aren't latency sensitive. `PERSISTENT_1` supports encryption of data in transit. It's available in all AWS Regions where FSx for Lustre is available.

Choose `PERSISTENT_2` deployment type for longer term storage and for latency sensitive workloads that require the highest levels of IOPS and throughput. `PERSISTENT_2` supports

SSD storage and offers higher `PerUnitStorageThroughput` (up to 1000 MB/s/TiB). `PERSISTENT_2` is available in a limited number of AWS Regions. For more information about deployment types and the list of AWS Regions where `PERSISTENT_2` is available, see File system deployment options for FSx for Lustre in the *Amazon FSx for Lustre User Guide*.

Encryption of data in transit is automatically enabled when you access `SCRATCH_2`, `PERSISTENT_1`, or `PERSISTENT_2` deployment type file systems from Amazon EC2 instances that support this feature.

Encryption of data in transit for `SCRATCH_2`, `PERSISTENT_1`, and `PERSISTENT_2` deployment types is supported when accessed from supported instance types in supported AWS Regions. For more information, see Encrypting data in transit in the *Amazon FSx for Lustre User Guide*.

> **Note**
>
> Support for the `PERSISTENT_2` deployment type was added with AWS ParallelCluster version 3.2.0.

Update policy: If this setting is changed, the update is not allowed.

`ImportedFileChunkSize` (**Optional**, `Integer`)

For files that are imported from a data repository, this value determines the stripe count and maximum amount of data for each file (in MiB) that's stored on a single physical disk. The maximum number of disks that a single file can be striped across is limited by the total number of disks that make up the file system.

The default chunk size is 1,024 MiB (1 GiB) and can go as high as 512,000 MiB (500 GiB). Amazon S3 objects have a maximum size of 5 TB.

> **Note**
>
> This parameter isn't supported for file systems that use the `PERSISTENT_2` deployment type. For instructions on how to configure data repositories associations, see Linking your file system to an S3 bucket in the *Amazon FSx for Lustre User Guide*.

Update policy: If this setting is changed, the update is not allowed.

`DataCompressionType` (**Optional**, `String`)

Sets the data compression configuration for the FSx for Lustre file system. The supported value is `LZ4`. `LZ4` indicates that data compression is turned on with the LZ4 algorithm. When `DataCompressionType` isn't specified, data compression is turned off when the file system is created.

For more information, see Lustre data compression.

Update policy: This setting can be changed during an update.

`ExportPath` (**Optional**, `String`)

The path in Amazon S3 where the root of your FSx for Lustre file system is exported. This setting is only supported when the `ImportPath` parameter is specified. The path must use the same Amazon S3 bucket as specified in `ImportPath`. You can provide an optional prefix to which new and changed data is to be exported from your FSx for Lustre file system. If an `ExportPath` value is not provided, FSx for Lustre sets a default export path, `s3://amzn-s3-demo-bucket/FSxLustre[creation-timestamp]`. The timestamp is in UTC format, for example `s3://amzn-s3-demo-bucket/FSxLustre20181105T222312Z`.

The Amazon S3 export bucket must be the same as the import bucket specified by `ImportPath`. If you only specify a bucket name, such as `s3://amzn-s3-demo-bucket`, you get a 1:1 mapping of file system objects to Amazon S3 bucket objects. This mapping means that the input data in Amazon S3 is overwritten on export. If you provide a custom prefix in the export path, such as `s3://amzn-s3-demo-bucket/[custom-optional-prefix]`, FSx for Lustre exports the contents of your file system to that export prefix in the Amazon S3 bucket.

> **Note**
>
> This parameter isn't supported for file systems that use the `PERSISTENT_2` deployment type. Configure data repositories associations as described in Linking your file system to an S3 bucket in the *Amazon FSx for Lustre User Guide*.

Update policy: If this setting is changed, the update is not allowed.

`ImportPath` (**Optional**, `String`)

The path to the Amazon S3 bucket (including the optional prefix) that you're using as the data repository for your FSx for Lustre file system. The root of your FSx for Lustre file system will be

mapped to the root of the Amazon S3 bucket you select. An example is `s3://amzn-s3-demo-bucket/optional-prefix`. If you specify a prefix after the Amazon S3 bucket name, only object keys with that prefix are loaded into the file system.

> **Note**
>
> This parameter isn't supported for file systems that use the `PERSISTENT_2` deployment type. Configure data repositories associations as described in Linking your file system to an S3 bucket in the *Amazon FSx for Lustre User Guide*.

Update policy: If this setting is changed, the update is not allowed.

`WeeklyMaintenanceStartTime` (**Optional**, `String`)

The preferred start time to perform weekly maintenance. It's in the `"d:HH:MM"` format in the UTC+0 time zone. For this format, d is the weekday number from 1 through 7, beginning with Monday and ending with Sunday. Quotation marks are required for this field.

Update policy: This setting can be changed during an update.

`AutomaticBackupRetentionDays` (**Optional**, `Integer`)

The number of days to retain automatic backups. Setting this to 0 disables automatic backups. The supported range is 0-90. The default is 0. This setting is only valid for use with `PERSISTENT_1` and `PERSISTENT_2` deployment types. For more information, see Working with backups in the *Amazon FSx for Lustre User Guide*.

Update policy: This setting can be changed during an update.

`CopyTagsToBackups` (**Optional**, `Boolean`)

If `true`, copy the tags for the FSx for Lustre file system to backups. This value defaults to `false`. If it's set to `true`, all tags for the file system are copied to all automatic and user-initiated backups where the user doesn't specify tags. If this value is `true`, and you specify one or more tags, only the specified tags are copied to backups. If you specify one or more tags when you create a user-initiated backup, no tags are copied from the file system, regardless of this value. This setting is only valid for use with `PERSISTENT_1` and `PERSISTENT_2` deployment types.

Update policy: If this setting is changed, the update is not allowed.

`DailyAutomaticBackupStartTime` (**Optional**, `String`)

A recurring daily time, in the `HH:MM` format. `HH` is the zero-padded hour of the day (00-23). `MM` is the zero-padded minute of the hour (00-59). For example, `05:00` specifies 5 A.M. daily. This setting is only valid for use with `PERSISTENT_1` and `PERSISTENT_2` deployment types.

Update policy: This setting can be changed during an update.

`PerUnitStorageThroughput` (**Required for `PERSISTENT_1` and `PERSISTENT_2` deployment types**, `Integer`)

Describes the amount of read and write throughput for each 1 tebibyte of storage, in MB/s/TiB. File system throughput capacity is calculated by multiplying file system storage capacity (TiB) by the `PerUnitStorageThroughput` (MB/s/TiB). For a 2.4 TiB file system, provisioning 50 MB/s/TiB of `PerUnitStorageThroughput` yields 120 MB/s of file system throughput. You pay for the amount of throughput that you provision. This corresponds to the PerUnitStorageThroughput property.

Valid values:

PERSISTENT_1 SSD storage: 50, 100, 200 MB/s/TiB.

PERSISTENT_1 HDD storage: 12, 40 MB/s/TiB.

PERSISTENT_2 SSD storage: 125, 250, 500, 1000 MB/s/TiB.

Update policy: If this setting is changed, the update is not allowed.

`BackupId` (**Optional**, `String`)

Specifies the ID of the backup to use to restore the FSx for Lustre file system from an existing backup. When the `BackupId` setting is specified, the `AutoImportPolicy`, `DeploymentType`, `ExportPath`, `KmsKeyId`, `ImportPath`, `ImportedFileChunkSize`, `StorageCapacity`, and `PerUnitStorageThroughput` settings must not be specified. These settings are read from the backup. Additionally, the `AutoImportPolicy`, `ExportPath`, `ImportPath`, and `ImportedFileChunkSize` settings must not be specified. This corresponds to the BackupId property.

Update policy: If this setting is changed, the update is not allowed.

`KmsKeyId` (**Optional**, `String`)

The ID of the AWS Key Management Service (AWS KMS) key ID that's used to encrypt the FSx for Lustre file system's data for persistent FSx for Lustre file systems at rest. If not specified, the FSx

for Lustre managed key is used. The SCRATCH_1 and SCRATCH_2 FSx for Lustre file systems are always encrypted at rest using FSx for Lustre managed keys. For more information, see Encrypt in the *AWS Key Management Service API Reference*.

Update policy: If this setting is changed, the update is not allowed.

`FileSystemId` (**Optional**, `String`)

Specifies the ID of an existing FSx for Lustre file system.

If this option is specified, only the `MountDir` and `FileSystemId` settings in the `FsxLustreSettings` are used. All other settings in the `FsxLustreSettings` are ignored.

> **Note**
>
> If AWS Batch scheduler is used, FSx for Lustre is only available on the head node.

> **Note**
>
> The file system must be associated to a security group that allows inbound and outbound TCP traffic through ports 988, 1021, 1022, and 1023.

Make sure that traffic is allowed between the cluster and file system by doing one of the following:

- Configure the security groups of the file system to allow the traffic to and from the CIDR or prefix list of cluster subnets.

  > **Note**
  >
  > AWS ParallelCluster validates that ports are open and that the CIDR or prefix list is configured. AWS ParallelCluster doesn't validate the content of CIDR block or prefix list.

- Set custom security groups for cluster nodes by using `SlurmQueues` / `Networking` / `SecurityGroups` and `HeadNode` / `Networking` / `SecurityGroups`. The custom security groups must be configured to allow traffic between the cluster and the file system.

> **Note**
>
> If all cluster nodes use custom security groups, AWS ParallelCluster only validates that the ports are open. AWS ParallelCluster doesn't validate that the source and destination are properly configured.

Update policy: If this setting is changed, the update is not allowed.

`AutoImportPolicy` (**Optional**, `String`)

When you create your FSx for Lustre file system, your existing Amazon S3 objects appear as file and directory listings. Use this property to choose how FSx for Lustre keeps your file and directory listings up to date as you add or modify objects in your linked Amazon S3 bucket. `AutoImportPolicy` can have the following values:

- `NEW` - Automatic import is on. FSx for Lustre automatically imports directory listings of any new objects added to the linked Amazon S3 bucket that do not currently exist in the FSx for Lustre file system.
- `NEW_CHANGED` - Automatic import is on. FSx for Lustre automatically imports file and directory listings of any new objects added to the Amazon S3 bucket and any existing objects that are changed in the Amazon S3 bucket after you choose this option.
- `NEW_CHANGED_DELETED` - Automatic import is on. FSx for Lustre automatically imports file and directory listings of any new objects added to the Amazon S3 bucket, any existing objects that are changed in the Amazon S3 bucket, and any objects that were deleted in the Amazon S3 bucket after you choose this option.

> **Note**
>
> Support for `NEW_CHANGED_DELETED` was added in AWS ParallelCluster version 3.1.1.

If `AutoImportPolicy` isn't specified, automatic import is off. FSx for Lustre only updates file and directory listings from the linked Amazon S3 bucket when the file system is created. FSx for Lustre doesn't update file and directory listings for any new or changed objects after choosing this option.

For more information, see Automatically import updates from your S3 bucket in the *Amazon FSx for Lustre User Guide*.

> **Note**
>
> This parameter isn't supported for file systems using the `PERSISTENT_2` deployment type. For instructions on how to configure data repositories associations, see Linking your file system to an S3 bucket in the *Amazon FSx for Lustre User Guide*.

Update policy: If this setting is changed, the update is not allowed.

`DriveCacheType` (**Optional**, `String`)

Specifies that the file system has an SSD drive cache. This can only be set if the `StorageType` setting is set to `HDD`, and the `DeploymentType` setting is set to `PERSISTENT_1`. This corresponds to the DriveCacheType property. For more information, see FSx for Lustre deployment options in the *Amazon FSx for Lustre User Guide*.

The only valid value is `READ`. To disable the SSD drive cache, don't specify the `DriveCacheType` setting.

Update policy: If this setting is changed, the update is not allowed.

`StorageType` (**Optional**, `String`)

Sets the storage type for the FSx for Lustre file system that you're creating. Valid values are `SSD` and `HDD`.

- Set to `SSD` to use solid state drive storage.
- Set to `HDD` to use hard disk drive storage. `HDD` is supported on `PERSISTENT` deployment types.

The default value is `SSD`. For more information, see Storage Type Options in the *Amazon FSx for Windows User Guide* and Multiple Storage Options in the *Amazon FSx for Lustre User Guide*.

Update policy: If this setting is changed, the update is not allowed.

`DeletionPolicy` (**Optional**, `String`)

Specifies whether the file system should be retained or deleted when the file system is removed from the cluster or the cluster is deleted. The supported values are `Delete` and `Retain`. The default value is `Delete`.

When the DeletionPolicy is set to `Delete`, a managed file system, with its data, is deleted if the cluster is deleted, or if the file system is removed with a cluster update.

For more information, see Shared storage.

Update policy: This setting can be changed during an update.

> **Note**
>
> `DeletionPolicy` is supported starting with AWS ParallelCluster version 3.3.0.

`DataRepositoryAssociations` (**Optional**, `String`)

List of DRAs (up to 8 per file system)

Each data repository association must have a unique Amazon FSx file system directory and a unique S3 bucket or prefix associated with it.

You can not use ExportPath and ImportPath in the FsxLustreSettings at the same time as using DRAs.

Update policy: This setting can be changed during an update.

`Name` (**Required**, `String`)

The name of the DRA. You use this name when you update the settings.

Update policy: If this setting is changed, the update is not allowed.

`BatchImportMetaDataOnCreate` (**Optional**, `Boolean`)

A boolean flag indicating whether an import data repository task to import metadata should run after the data repository association is created. The task runs if this flag is set to `true`.

Default value: `false`

Update policy: If this setting is changed, the update is not allowed.

`DataRepositoryPath` (**Required**, `String`)

The path to the Amazon S3 data repository that will be linked to the file system. The path can be an S3 bucket or prefix in the format `s3://`*`amzn-s3-demo-bucket`*`/myPrefix/`. This path specifies where in the S3 data repository files will be imported from or exported to.

Cannot overlap with other DRAs

Pattern: `^[^\u0000\u0085\u2028\u2029\r\n]{3,4357}$`

Minimum: `3`

Maximum: `4357`

Update policy: If this setting is changed, the update is not allowed.

`FileSystemPath` (**Required**, `String`)

A path on the Amazon FSx for Lustre file system that points to a high-level directory (such as `/ns1/`) or subdirectory (such as `/ns1/subdir/`) that will be mapped 1-1 with `DataRepositoryPath`. The leading forward slash in the name is required. Two data repository associations cannot have overlapping file system paths. For example, if a data repository is associated with file system path `/ns1/`, then you cannot link another data repository with file system path `/ns1/ns2`.

This path specifies where in your file system files will be exported from or imported to. This file system directory can be linked to only one Amazon S3 bucket, and no other S3 bucket can be linked to the directory.

Cannot overlap with other DRAs

> **Note**
>
> If you specify only a forward slash (`/`) as the file system path, you can link only one data repository to the file system. You can only specify "`/`" as the file system path for the first data repository associated with a file system.

Pattern: `^[^\u0000\u0085\u2028\u2029\r\n]{1,4096}$`

Minimum: `1`

Maximum: `4096`

Update policy: If this setting is changed, the update is not allowed.

`ImportedFileChunkSize` (**Optional**, `Integer`)

For files imported from a data repository, this value determines the stripe count and maximum amount of data per file (in MiB) stored on a single physical disk. The maximum number of disks

that a single file can be striped across is limited by the total number of disks that make up the file system or cache.

The default chunk size is 1,024 MiB (1 GiB) and can go as high as 512,000 MiB (500 GiB). Amazon S3 objects have a maximum size of 5 TB.

Minimum: 1

Maximum: 4096

Update policy: This setting can be changed during an update.

`AutoExportPolicy` (**Optional**, `Array of strings`)

The list can contain one or more of the following values:

- `NEW` - New files and directories are automatically exported to the data repository as they are added to the file system.
- `CHANGED` - Changes to files and directories on the file system are automatically exported to the data repository.
- `DELETED` - Files and directories are automatically deleted on the data repository when they are deleted on the file system.

You can define any combination of event types for your `AutoExportPolicy`.

Maximum: 3

Update policy: This setting can be changed during an update.

`AutoImportPolicy` (**Optional**, `Array of strings`)

The list can contain one or more of the following values:

- `NEW` - Amazon FSx automatically imports metadata of files added to the linked S3 bucket that do not currently exist in the FSx file system.
- `CHANGED` - Amazon FSx automatically updates file metadata and invalidates existing file content on the file system as files change in the data repository.
- `DELETED` - Amazon FSx automatically deletes files on the file system as corresponding files are deleted in the data repository.

You can define any combination of event types for your `AutoImportPolicy`.

Maximum: 3

Update policy: This setting can be changed during an update.

`FsxOntapSettings`

> **Note**
>
> You must define `FsxOntapSettings` if `FsxOntap` is specified for `StorageType`.

(Optional) The settings for an FSx for ONTAP file system.

```
FsxOntapSettings:
  VolumeId: string
```

`FsxOntapSettings` properties

`VolumeId` (**Required**, `String`)

Specifies the volume ID of the existing FSx for ONTAP system.

> **Note**
>
> - If an AWS Batch scheduler is used, FSx for ONTAP is only available on the head node.
> - If the FSx for ONTAP deployment type is `Multi-AZ`, make sure that the head node subnet's route table is properly configured.
> - Support for FSx for ONTAP was added in AWS ParallelCluster version 3.2.0.
> - The file system must be associated to a security group that allows inbound and outbound TCP and UDP traffic through ports 111, 635, 2049, and 4046.

Make sure traffic is allowed between the cluster and file system by doing one of the following actions:

- Configure the security groups of the file system to allow the traffic to and from the CIDR or prefix list of cluster subnets.

> **Note**
>
> AWS ParallelCluster validates that ports are open and that the CIDR or prefix list is configured. AWS ParallelCluster doesn't validate the content of CIDR block or prefix list.

- Set custom security groups for cluster nodes by using `SlurmQueues` / `Networking` / `SecurityGroups` and `HeadNode` / `Networking` / `SecurityGroups`. The custom security groups must be configured to allow traffic between the cluster and the file system.

> **Note**
>
> If all cluster nodes use custom security groups, AWS ParallelCluster only validates that the ports are open. AWS ParallelCluster doesn't validate that the source and destination are properly configured.

Update policy: If this setting is changed, the update is not allowed.

`FsxOpenZfsSettings`

> **Note**
>
> You must define `FsxOpenZfsSettings` if `FsxOpenZfs` is specified for `StorageType`.

(Optional) The settings for a FSx for OpenZFS file system.

```
FsxOpenZfsSettings:
  VolumeId: string
```

Update policy: If this setting is changed, the update is not allowed.

`FsxOpenZfsSettings` properties

`VolumeId` (**Required**, `String`)

Specifies the volume ID of the existing FSx for OpenZFS system.

Note

- If an AWS Batch scheduler is used, FSx for OpenZFS is only available on the head node.
- Support for FSx for OpenZFS was added in AWS ParallelCluster version 3.2.0.
- The file system must be associated to a security group that allows inbound and outbound TCP and UDP traffic through ports 111, 2049, 20001, 20002, and 20003.

Make sure that traffic is allowed between the cluster and file system by doing one of the following:

- Configure the security groups of the file system to allow the traffic to and from the CIDR or prefix list of cluster subnets.

  Note

  AWS ParallelCluster validates that ports are open and that the CIDR or prefix list is configured. AWS ParallelCluster doesn't validate the content of CIDR block or prefix list.

- Set custom security groups for cluster nodes by using `SlurmQueues` / `Networking` / `SecurityGroups` and `HeadNode` / `Networking` / `SecurityGroups`. The custom security groups must be configured to allow traffic between the cluster and the file system.

  Note

  If all cluster nodes use custom security groups, AWS ParallelCluster only validates that the ports are open. AWS ParallelCluster doesn't validate that the source and destination are properly configured.

Update policy: If this setting is changed, the update is not allowed.

`FileCacheSettings`

Note

You must define `FileCacheSettings` if `FileCache` is specified for `StorageType`.

(Optional) The settings for a File Cache.

```
FileCacheSettings:
  FileCacheId: string
```

Update policy: If this setting is changed, the update is not allowed.

`FileCacheSettings` properties

`FileCacheId` (**Required**, `String`)

Specifies the File Cache ID of an existing File Cache.

> **Note**
>
> - File Cache doesn't support AWS Batch schedulers.
> - Support for File Cache is added in AWS ParallelCluster version 3.7.0.
> - The file system must be associated to a security group that allows inbound and outbound TCP traffic through port 988.

Make sure that traffic is allowed between the cluster and file system by doing one of the following:

- Configure the security groups of the File Cache to allow the traffic to and from the CIDR or prefix list of cluster subnets.

  > **Note**
  >
  > AWS ParallelCluster validates that ports are open and that the CIDR or prefix list is configured. AWS ParallelCluster doesn't validate the content of CIDR block or prefix list.

- Set custom security groups for cluster nodes by using `SlurmQueues` / `Networking` / `SecurityGroups` and `HeadNode` / `Networking` / `SecurityGroups`. The custom security groups must be configured to allow traffic between the cluster and the file system.

> **Note**
>
> If all cluster nodes use custom security groups, AWS ParallelCluster only validates that the ports are open. AWS ParallelCluster doesn't validate that the source and destination are properly configured.

Update policy: If this setting is changed, the update is not allowed.

`Iam` section

(Optional) Specifies IAM properties for the cluster.

```
Iam:
  Roles:
    LambdaFunctionsRole: string
  PermissionsBoundary: string
  ResourcePrefix: string
```

Update policy: This setting can be changed during an update.

`Iam` properties

`PermissionsBoundary` (**Optional**, `String`)

The ARN of the IAM policy to use as permissions boundary for all roles created by AWS ParallelCluster. For more information, see Permissions boundaries for IAM entities in the *IAM User Guide*. The format is `arn:${Partition}:iam::${Account}:policy/${PolicyName}`.

Update policy: This setting can be changed during an update.

`Roles` (**Optional**)

Specifies settings for the IAM roles used by the cluster.

Update policy: This setting can be changed during an update.

`LambdaFunctionsRole` (**Optional**, `String`)

The ARN of the IAM role to use for AWS Lambda. This overrides the default role attached to all Lambda functions backing AWS CloudFormation custom resources. Lambda needs

to be configured as the principal allowed to assume the role. This will not override the role of Lambda functions used for AWS Batch. The format is `arn:${Partition}:iam::${Account}:role/${RoleName}`.

Update policy: This setting can be changed during an update.

`ResourcePrefix` (**Optional**)

Specifies a path or name prefix for IAM resources that are created by AWS ParallelCluster.

The resource prefix must follow the naming rules specified by IAM:

- A name can contain up to 30 characters.
- A name can only be a string with no slash (/) characters.
- A path can be up to 512 characters.
- A path must start and end with a slash (/). It can contain multiple slashes (/) between the start and end slashes (/).
- You can combine the path and name `/path/name`.

Specify a name.

```
Iam:
  ResourcePrefix: my-prefix
```

Specify a path.

```
Iam:
  ResourcePrefix: /org/dept/team/project/user/
```

Specify a path and name.

```
Iam:
  ResourcePrefix: /org/dept/team/project/user/my-prefix
```

If you specify `/my-prefix`, an error is returned.

```
Iam:
  ResourcePrefix: /my-prefix
```

A configuration error is returned. A path must have two /s. A prefix by itself can't have /s.

Update policy: If this setting is changed, the update is not allowed.

`LoginNodes` section

> **Note**
>
> Support for `LoginNodes` is added in AWS ParallelCluster version 3.7.0.

(Optional) Specifies the configuration for the login nodes pool.

```
LoginNodes:
  Pools:
    - Name: string
      Count: integer
      InstanceType: string
      GracetimePeriod: integer
      Image:
        CustomAmi: string
      Ssh:
        KeyName: string
        AllowedIps: string
      Networking:
        SubnetIds:
          - string
        SecurityGroups:
          - string
        AdditionalSecurityGroups:
          - string
      Dcv:
        Enabled: boolean
        Port: integer
        AllowedIps: string
      CustomActions:
        OnNodeStart:
          Sequence:
            - Script: string
              Args:
                - string
          Script: string
          Args:
            - string
```

```
        OnNodeConfigured:
          Sequence:
            - Script: string
              Args:
                - string
          Script: string
          Args:
            - string
        OnNodeUpdated:
          Sequence:
            - Script: string
              Args:
                - string
          Script: string
          Args:
            - string
      Iam:
        InstanceRole: string
        InstanceProfile: string
        AdditionalIamPolicies:
          - Policy: string
```

Update policy: The login nodes in the cluster must be stopped for this setting to be changed for an update.

`LoginNodes` properties

`Pools` properties

Defines groups of login nodes that have the same resource configuration. Starting with AWS ParallelCluster 3.11.0 up to 10 pools can be specified.

```
Pools:
  - Name: string
    Count: integer
    InstanceType: string
    GracetimePeriod: integer
    Image:
      CustomAmi: string
    Ssh:
      KeyName: string
      AllowedIps: string
    Networking:
```

```
      SubnetIds:
        - string
      SecurityGroups:
        - string
      AdditionalSecurityGroups:
        - string
    Dcv:
      Enabled: boolean
      Port: integer
      AllowedIps: string
    CustomActions:
      OnNodeStart:
        Sequence:
          - Script: string
            Args:
              - string
        Script: string
        Args:
          - string
      OnNodeConfigured:
        Sequence:
          - Script: string
            Args:
              - string
        Script: string
        Args:
          - string
      OnNodeUpdated:
        Sequence:
          - Script: string
            Args:
              - string
        Script: string
        Args:
          - string
    Iam:
      InstanceRole: string
      InstanceProfile: string
      AdditionalIamPolicies:
        - Policy: string
```

Update policy: Login node pools can be added, but removing a pool requires all login nodes in the cluster are stopped.

`Name` (**Required** `String`)

Specifies the name of the `LoginNodes` pool. This is used to tag the `LoginNodes` resources.

Update policy: If this setting is changed, the update is not allowed.

> **Note**
>
> Starting with AWS ParallelCluster version 3.11.0, the update policy is: The login nodes in the pool must be stopped for this setting to be changed for an update.

`Count` (**Required** `Integer`)

Specifies the number of login nodes to keep active.

Update policy: This setting can be changed during an update.

`InstanceType` (**Required** `String`)

Specifies the Amazon EC2 instance type that's used for the login node. The architecture of the instance type must be the same as the architecture used for Slurm `InstanceType` setting.

Update policy: This setting can be changed if the login nodes pool is stopped.

> **Note**
>
> Starting with AWS ParallelCluster version 3.11.0, the update policy is: The login nodes in the pool must be stopped for this setting to be changed for an update.

`GracetimePeriod` (**Optional** `Integer`)

Specifies the minimum amount of time in minutes that elapse between the notification to the logged in user that a login node is to be decommissioned and the actual stop event. Valid values for `GracetimePeriod` are from 3 up to 120 minutes. The default is 10 minutes.

> **Note**
>
> The triggering event involves interactions between multiple AWS services. Sometimes, network latency and propagation of the information might take some time so the grace time period may take longer than expected due to internal delays in AWS services.

Update policy: This setting can be changed during an update.

`Image` (**Optional**)

Defines the image configuration for the login nodes.

```
Image:
  CustomAmi: String
```

`CustomAmi` (**Optional** `String`)

Specifies the custom AMI used to provision the login nodes. If not specified, the value defaults to the one specified in the HeadNode section.

Update policy: If this setting is changed, the update is not allowed.

`Ssh` (**Optional**)

Defines the `ssh` configuration for the login nodes.

```
Ssh:
  KeyName: string
  AllowedIps: string
```

> **Note**
>
> Starting with AWS ParallelCluster version 3.11.0, the update policy is: The login nodes in the pool must be stopped for this setting to be changed for an update.

`KeyName` (**Optional** `String`)

Specifies the `ssh` key used to log in into the login nodes. If not specified, the value defaults to the one specified in the HeadNode section.

Update policy: The login nodes in the pool must be stopped for this setting to be changed for an update.

`AllowedIps` (**Optional** `String`)

Specifies the CIDR-formatted IP range or a prefix list id for SSH connections to login nodes in the pool. The default is the AllowedIps defined in the head node configuration, or `0.0.0.0/0` if not specified.HeadNode section.

Update policy: The login nodes in the pool must be stopped for this setting to be changed for an update.

> **Note**
>
> Support for AllowedIps for login nodes is added in AWS ParallelCluster version 3.11.0.

`Networking` (**Required**)

```
Networking:
  SubnetIds:
    - string
  SecurityGroups:
    - string
  AdditionalSecurityGroups:
    - string
```

> **Note**
>
> Starting with AWS ParallelCluster version 3.11.0, the update policy is: The login nodes in the pool must be stopped for this setting to be changed for an update.

`SubnetIds` (**Required** `[String]`)

The ID of existing subnet that you provision the login nodes pool in. You can only define one subnet.

Update policy: If this setting is changed, the update is not allowed.

`SecurityGroups` (**Optional** `[String]`)

A list of security groups to use for the login nodes pool. If no security groups are specified, AWS ParallelCluster creates security groups for you.

Update policy: If this setting is changed, the update is not allowed.

`AdditionalSecurityGroups` (**Optional** `[String]`)

A list of additional security groups to use for the login nodes pool.

Update policy: If this setting is changed, the update is not allowed.

`Dcv` (**Optional**)

Defines configuration settings for the NICE DCV server that runs on the login nodes. For more information, see Connect to the head and login nodes through Amazon DCV

```
Dcv:
  Enabled: boolean
  Port: integer
  AllowedIps: string
```

Important

By default, the NICE DCV port setup by AWS ParallelCluster is open to all IPv4 addresses. You can connect to a NICE DCV port only if you have the URL for the NICE DCV session and connect to the NICE DCV session within 30 seconds of when the URL is returned from pcluster dcv-connect. Use the `AllowedIps` setting to further restrict access to the NICE DCV port with a CIDR-formatted IP range and use the Port setting to set a nonstandard port.

Update policy: If this setting is changed, the update is not allowed.

Note

Support for DCV on login nodes is added in AWS ParallelCluster version 3.11.0.

`Enabled` (**Required** `Boolean`)

Specifies whether NICE DCV is enabled on the login nodes in the pool. The default value is `false`.

Update policy: If this setting is changed, the update is not allowed.

> **Note**
>
> NICE DCV automatically generates a self-signed certificate that's used to secure traffic between the NICE DCV client and NICE DCV server that runs on the login node. To configure your own certificate, see Amazon DCV HTTPS certificate.

`Port` (**Optional** `Integer`)

Specifies the port for NICE DCV. The default value is 8443.

Update policy: If this setting is changed, the update is not allowed.

`AllowedIps` (**Optional** `String`)

Specifies the CIDR-formatted IP range for connections to NICE DCV. This setting is used only when AWS ParallelCluster creates the security group. The default value is `0.0.0.0/0`, which allows access from any Internet address.

Update policy: If this setting is changed, the update is not allowed.

`CustomActions` (**Optional**)

Specifies the custom scripts to run on the login nodes.

```
CustomActions:
  OnNodeStart:
    Sequence:
      - Script: string
        Args:
          - string
    Script: string
    Args:
      - string
  OnNodeConfigured:
    Sequence:
      - Script: string
```

```
        Args:
          - string
    Script: string
    Args:
      - string
  OnNodeUpdated:
    Sequence:
      - Script: string
        Args:
          - string
    Script: string
    Args:
      - string
```

> **Note**
>
> Support for custom actions on login nodes is added in AWS ParallelCluster version 3.11.0.

`OnNodeStart` (**Optional**)

Specifies single script or a sequence of scripts to run on the login nodes before any node deployment bootstrap action is started. For more information, see Custom bootstrap actions.

`Sequence` (**Optional**)

List of scripts to run. AWS ParallelCluster runs the scripts in the same order as they are listed in the configuration file, starting with the first.

`Script` (**Required** `String`)

Specifies the file to use. The file path can start with `https://` or `s3://`.

`Args` (**Optional** `[String]`)

List of arguments to pass to the script.

Update policy: If this setting is changed, the update is not allowed.

`Script` (**Required** `String`)

Specifies the file to use for a single script. The file path can start with `https://` or `s3://`.

`Args` (**Optional** `[String]`)

List of arguments to pass to the single script.

`OnNodeConfigured` (**Optional**)

Specifies single script or a sequence of scripts to run on the login nodes after the node bootstrap processes are complete. For more information, see Custom bootstrap actions.

`Sequence` (**Optional**)

List of scripts to run. AWS ParallelCluster runs the scripts in the same order as they are listed in the configuration file, starting with the first.

`Script` (**Required** `String`)

Specifies the file to use. The file path can start with `https://` or `s3://`.

`Args` (**Optional** `[String]`)

List of arguments to pass to the script.

Update policy: If this setting is changed, the update is not allowed.

`Script` (**Required** `String`)

Specifies the file to use for a single script. The file path can start with `https://` or `s3://`.

`Args` (**Optional** `[String]`)

List of arguments to pass to the single script.

Update policy: If this setting is changed, the update is not allowed.

`OnNodeUpdated` (**Optional**)

Specifies single script or a sequence of scripts to run after the head node update is completed and the scheduler and shared storage are aligned with the latest cluster configuration changes. For more information, see Custom bootstrap actions.

`Sequence` (**Optional**)

List of scripts to run. AWS ParallelCluster runs the scripts in the same order as they are listed in the configuration file, starting with the first.

`Script` (**Required** `String`)

Specifies the file to use. The file path can start with `https://` or `s3://`.

`Args` (**Optional** `[String]`)

List of arguments to pass to the script.

`Script` (**Required** `String`)

Specifies the file to use for a single script. The file path can start with `https://` or `s3://`.

`Args` (**Optional** `[String]`)

List of arguments to pass to the single script.

Update policy: If this setting is changed, the update is not allowed.

> **Note**
>
> AWS ParallelCluster doesn't support including both a single script and `Sequence` for the same custom action.

`Iam` (**Optional**)

Specifies either an instance role or an instance profile to use on the login nodes to override the default instance role or instance profile for the cluster.

```
Iam:
  InstanceRole: string
  InstanceProfile: string
  AdditionalIamPolicies:
    - Policy: string
```

> **Note**
>
> Starting with AWS ParallelCluster version 3.11.0, the update policy is: The login nodes in the pool must be stopped for this setting to be changed for an update.

`InstanceProfile` (**Optional** `String`)

Specifies an instance profile to override the default login node instance profile. You can't specify both `InstanceProfile` and `InstanceRole`. The format is `arn:Partition:iam::Account:instance-profile/`*`InstanceProfileName`*. If this is specified, the `InstanceRole` and `AdditionalIamPolicies` settings can't be specified.

Update policy: If this setting is changed, the update is not allowed.

`InstanceRole` (**Optional** `String`)

Specifies an instance role to override the default login node instance role. You can't specify both `InstanceProfile` and `InstanceRole`. The format is `arn:Partition:iam::Account:role/RoleName`. If this is specified, the `InstanceProfile` and `AdditionalIamPolicies` settings can't be specified.

Update policy: If this setting is changed, the update is not allowed.

`AdditionalIamPolicies` (**Optional**)

```
AdditionalIamPolicies:
  - Policy: string
```

An IAM policy Amazon Resource Name (ARN).

Specifies a list of Amazon Resource Names (ARNs) of IAM policies for Amazon EC2. This list is attached to the root role used for the login node in addition to the permissions that are required by AWS ParallelCluster.

An IAM policy name and its ARN are different. Names can't be used.

If this is specified, the `InstanceProfile` and `InstanceRole` settings can't be specified. We recommend that you use `AdditionalIamPolicies` because `AdditionalIamPolicies` are added to the permissions that AWS ParallelCluster requires, and the `InstanceRole` must include all required permissions. The required permissions often change from release to release as features are added.

There's no default value.

Update policy: If this setting is changed, the update is not allowed.

`Policy` (**Required** `[String]`)

Update policy: If this setting is changed, the update is not allowed.

`Monitoring` section

(Optional) Specifies the monitoring settings for the cluster.

```
Monitoring:
  Logs:
    CloudWatch:
      Enabled: boolean
      RetentionInDays: integer
      DeletionPolicy: string
    Rotation:
      Enabled: boolean
  Dashboards:
    CloudWatch:
      Enabled: boolean
  DetailedMonitoring: boolean
  Alarms:
   Enabled: boolean
```

Update policy: This setting is not analyzed during an update.

`Monitoring` properties

`Logs` (**Optional**)

The log settings for the cluster.

Update policy: If this setting is changed, the update is not allowed.

`CloudWatch` (**Optional**)

The CloudWatch Logs settings for the cluster.

Update policy: If this setting is changed, the update is not allowed.

`Enabled` (**Required**, `Boolean`)

If `true`, cluster logs are streamed to CloudWatch Logs. The default value is `true`.

Update policy: If this setting is changed, the update is not allowed.

`RetentionInDays` (**Optional**, `Integer`)

The number of days to retain the log events in CloudWatch Logs. The default value is 180. The supported values are 0, 1, 3, 5, 7, 14, 30, 60, 90, 120, 150, 180, 365, 400, 545, 731, 1827, and 3653. A value of 0 will use the default CloudWatch log retention setting, i.e. never expire.

Update policy: This setting can be changed during an update.

`DeletionPolicy` (**Optional**, `String`)

Indicates whether to delete log events on CloudWatch Logs when the cluster is deleted. The possible values are `Delete` and `Retain`. The default value is `Retain`.

Update policy: This setting can be changed during an update.

`Rotation` (**Optional**)

The log rotation settings for the cluster.

Update policy: If this setting is changed, the update is not allowed.

`Enabled` (**Required**, `Boolean`)

If `true`, log rotation is enabled. The default is `true`. When a AWS ParallelCluster configured log file reaches a certain size, it is rotated and a single backup is maintained. For more information, see AWS ParallelCluster configured log rotation.

Update policy: If this setting is changed, the update is not allowed.

`Dashboards` (**Optional**)

The dashboard settings for the cluster.

Update policy: This setting can be changed during an update.

`CloudWatch` (**Optional**)

The CloudWatch dashboard settings for the cluster.

Update policy: This setting can be changed during an update.

`Enabled` (**Required**, `Boolean`)

If `true`, the CloudWatch dashboard is enabled. The default value is `true`.

Update policy: This setting can be changed during an update.

`DetailedMonitoring` (**Optional**, `Boolean`)

If set to `true`, detailed monitoring is enabled for the compute fleet Amazon EC2 instances. When enabled, the Amazon EC2 console displays graphs for monitoring the instances at 1 minute intervals. There are added costs when this feature is enabled. The default is `false`.

For more information, see Enable or turn off detailed monitoring for your instances in the *Amazon EC2 User Guide for Linux Instances*.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

> **Note**
>
> `DetailedMonitoring` is added starting with AWS ParallelCluster version 3.6.0.

`Alarms` (**Optional**)

CloudWatch Alarms for the cluster.

Update policy: This setting can be changed during an update.

`Enabled` (**Optional**)

If `true`, the CloudWatch Alarms for the cluster will be created. The default value is `true`.

Update policy: This setting can be changed during an update.

> **Note**
>
> Starting with AWS ParallelCluster version 3.8.0, the following alarms are created for the Head Node: Amazon EC2 Health Check, CPU/Memory/Disk utilization and a composite alarm that includes all the others.

Tags section

(Optional), Array Defines the tags that are used by AWS CloudFormation and propagated to all the cluster resources. For more information, see AWS CloudFormation resource tag in the *AWS CloudFormation User Guide*.

```
Tags:
  - Key: string
    Value: string
```

Update policy: If this setting is changed, the update is not allowed.

Tags properties

Key (**Required**, String)

Defines the name of the tag.

Update policy: If this setting is changed, the update is not allowed.

Value (**Required**, String)

Defines the value of the tag.

Update policy: If this setting is changed, the update is not allowed.

AdditionalPackages section

(Optional) Used to identify additional packages to install.

```
AdditionalPackages:
  IntelSoftware:
    IntelHpcPlatform: boolean
```

Update policy: If this setting is changed, the update is not allowed.

IntelSoftware

(Optional) Defines the configuration for Intel select solutions.

```
IntelSoftware:
  IntelHpcPlatform: boolean
```

Update policy: If this setting is changed, the update is not allowed.

`IntelSoftware` properties

`IntelHpcPlatform` (**Optional**, `Boolean`)

If `true`, indicates that the End user license agreement for Intel Parallel Studio is accepted. This causes Intel Parallel Studio to be installed on the head node and shared with the compute nodes. This adds several minutes to the time it takes the head node to bootstrap.

Update policy: If this setting is changed, the update is not allowed.

`DirectoryService` section

> **Note**
>
> Support for `DirectoryService` was added in AWS ParallelCluster version 3.1.1.

(Optional) The directory service settings for a cluster that supports multiple user access.

AWS ParallelCluster manages permissions that support multiple user access to clusters with an Active Directory (AD) over Lightweight Directory Access Protocol (LDAP) supported by the System Security Services Daemon (SSSD). For more information, see What is AWS Directory Service? in the *AWS Directory Service Administration Guide*.

We recommend that you use LDAP over TLS/SSL (abbreviated LDAPS for short) to ensure that any potentially sensitive information is transmitted over encrypted channels.

```
DirectoryService:
  DomainName: string
  DomainAddr: string
  PasswordSecretArn: string
  DomainReadOnlyUser: string
  LdapTlsCaCert: string
  LdapTlsReqCert: string
  LdapAccessFilter: string
  GenerateSshKeysForUsers: boolean
  AdditionalSssdConfigs: dict
```

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`DirectoryService` properties

> **Note**
>
> If you plan to use AWS ParallelCluster in a single subnet with no internet access, see AWS ParallelCluster in a single subnet with no internet access for additional requirements.

`DomainName` (**Required**, `String`)

The Active Directory (AD) domain that you use for identity information.

`DomainName` accepts both the Fully Qualified Domain Name (FQDN) and LDAP Distinguished Name (DN) formats.

- FQDN example: `corp.example.com`
- LDAP DN example: `DC=corp,DC=example,DC=com`

This property corresponds to the sssd-ldap parameter that's called `ldap_search_base`.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`DomainAddr` (**Required**, `String`)

The URI or URIs that point to the AD domain controller that's used as the LDAP server. The URI corresponds to the SSSD-LDAP parameter that's called `ldap_uri`. The value can be a comma separated string of URIs. To use LDAP, you must add `ldap://` to the beginning of the each URI.

Example values:

```
ldap://192.0.2.0,ldap://203.0.113.0          # LDAP
ldaps://192.0.2.0,ldaps://203.0.113.0        # LDAPS without support for certificate
 verification
ldaps://abcdef01234567890.corp.example.com  # LDAPS with support for certificate
 verification
192.0.2.0,203.0.113.0                        # AWS ParallelCluster uses LDAPS by
 default
```

If you use LDAPS with certificate verification, the URIs must be hostnames.

If you use LDAPS without certificate verification or LDAP, URIs can be hostnames or IP addresses.

Use LDAP over TLS/SSL (LDAPS) to avoid transmission of passwords and other sensitive information over unencrypted channels. If AWS ParallelCluster doesn't find a protocol, it adds `ldaps://` to the beginning of each URI or hostname.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`PasswordSecretArn` (**Required**, `String`)

The Amazon Resource Name (ARN) of the AWS Secrets Manager secret that contains the `DomainReadOnlyUser` plaintext password. The content of the secret corresponds to SSSD-LDAP parameter that's called `ldap_default_authtok`.

> **Note**
>
> When you use the AWS Secrets Manager console to create a secret, be sure to select "Other type of secret", select plaintext, and only include the password text in the secret. For more information on how to use AWS Secrets Manager to create a secret refer to Create an AWS Secrets Manager Secret

The LDAP client uses the password to authenticate to the AD domain as a `DomainReadOnlyUser` when it requests identity information.

If the user has the permission to `DescribeSecret`, `PasswordSecretArn` is validated. `PasswordSecretArn` is valid if the specified secret exists. If the user IAM policy doesn't include `DescribeSecret`, `PasswordSecretArn` isn't validated and a warning message is displayed. For more information, see Base AWS ParallelCluster `pcluster` user policy.

When the value of the secret changes, the cluster *isn't* automatically updated. To update the cluster for the new secret value, you must stop the compute fleet with the the section called "`pcluster update-compute-fleet`" command and then run the following command from within the head node.

```
$ sudo /opt/parallelcluster/scripts/directory_service/
update_directory_service_password.sh
```

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`DomainReadOnlyUser` (**Required**, `String`)

The identity that's used to query the AD domain for identity information when authenticating cluster user logins. It corresponds to SSSD-LDAP parameter that's called `ldap_default_bind_dn`. Use your AD identity information for this value.

Specify the identity in the form required by the specific LDAP client that's on the node:

- MicrosoftAD:

```
cn=ReadOnlyUser,ou=Users,ou=CORP,dc=corp,dc=example,dc=com
```

- SimpleAD:

```
cn=ReadOnlyUser,cn=Users,dc=corp,dc=example,dc=com
```

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`LdapTlsCaCert` (**Optional**, `String`)

The absolute path to a certificates bundle that contains the certificates for every certification authority in the certification chain that issued a certificate for the domain controllers. It corresponds to the SSSD-LDAP parameter that's called `ldap_tls_cacert`.

A certificate bundle is a file that's composed of the concatenation of distinct certificates in PEM format, also known as DER Base64 format in Windows. It is used to verify the identity of the AD domain controller that acts as the LDAP server.

AWS ParallelCluster isn't responsible for initial placement of certificates onto nodes. As the cluster administrator, you can configure the certificate in the head node manually after the cluster is created or you can use a bootstrap script. Alternatively, you can use an Amazon Machine Image (AMI) that includes the certificate configured on the head node.

Simple AD doesn't provide LDAPS support. To learn how to integrate a Simple AD directory with AWS ParallelCluster, see How to configure an LDAPS endpoint for Simple AD in the *AWS Security Blog*.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`LdapTlsReqCert` (**Optional**, `String`)

Specifies what checks to perform on server certificates in a TLS session. It corresponds to SSSD-LDAP parameter that's called `ldap_tls_reqcert`.

Valid values: `never`, `allow`, `try`, `demand`, and `hard`.

`never`, `allow`, and `try` enable connections to proceed even if problems with certificates are found.

`demand` and `hard` enable communication to continue if no problems with certificates are found.

If the cluster administrator uses a value that doesn't require the certificate validation to succeed, a warning message is returned to the administrator. For security reasons, we recommend that you don't disable certificate verification.

The default value is `hard`.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`LdapAccessFilter` (**Optional**, `String`)

Specifies a filter to limit directory access to a subset of users. This property corresponds to the SSSD-LDAP parameter that's called `ldap_access_filter`. You can use it to limit queries to an AD that supports a large number of users.

This filter can block user access to the cluster. However, it doesn't impact the discoverability of blocked users.

If this property is set, the SSSD parameter `access_provider` is set to `ldap` internally by AWS ParallelCluster and must not be modified by `DirectoryService` / `AdditionalSssdConfigs` settings.

If this property is omitted and customized user access isn't specified in `DirectoryService` / `AdditionalSssdConfigs`, all users in the directory can access the cluster.

Examples:

```
"!(cn=SomeUser*)"  # denies access to every user with an alias that starts with
 "SomeUser"
"(cn=SomeUser*)"   # allows access to every user with alias that starts with
 "SomeUser"
```

```
"memberOf=cn=TeamOne,ou=Users,ou=CORP,dc=corp,dc=example,dc=com" # allows access
 only to users in group "TeamOne".
```

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`GenerateSshKeysForUsers` (**Optional**, `Boolean`)

Defines whether AWS ParallelCluster generates an SSH key for cluster users immediately after their initial authentication on the head node.

If set to `true`, an SSH key is generated and saved to *`USER_HOME_DIRECTORY`*`/.ssh/id_rsa`, if it doesn't exist, for every user after their first authentication on the head node.

For a user that has not yet been authenticated on the head node, first authentication can happen in the following cases:

- The user logs into the head node for the first time with her or his own password.
- In the head node, a sudoer switches to the user for the first time: `su` *`USERNAME`*
- In the head node, a sudoer runs a command as the user for the first time: `su -u` *`USERNAME COMMAND`*

Users can use the SSH key for subsequent logins to the cluster head node and compute nodes. With AWS ParallelCluster, password logins to cluster compute nodes are disabled by design. If a user hasn't logged into the head node, SSH keys aren't generated and the user won't be able to log in to compute nodes.

The default is `true`.

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`AdditionalSssdConfigs` (**Optional**, `Dict`)

A dictionary of key-value pairs that contain SSSD parameters and values to write to the SSSD config file on cluster instances. For a full description of the SSSD configuration file, see the on-instance man pages for SSSD and related configuration files.

The SSSD parameters and values must be compatible with AWS ParallelCluster's SSSD configuration as described in the following list.

- `id_provider` is set to `ldap` internally by AWS ParallelCluster and must not be modified.

- `access_provider` is set to `ldap` internally by AWS ParallelCluster when `DirectoryService` / `LdapAccessFilter` is specified, and this setting must not be modified.

  If `DirectoryService` / `LdapAccessFilter` is omitted, its `access_provider` specification is omitted also. For example, if you set `access_provider` to `simple` in `AdditionalSssdConfigs`, then `DirectoryService` / `LdapAccessFilter` must not be specified.

The following configuration snippets are examples of valid configurations for `AdditionalSssdConfigs`.

This example enables debug level for SSSD logs, restricts the search base to a specific organizational unit, and disables credentials caching.

```
DirectoryService:
  ...
  AdditionalSssdConfigs:
    debug_level: "0xFFF0"
    ldap_search_base: OU=Users,OU=CORP,DC=corp,DC=example,DC=com
    cache_credentials: False
```

This example specifies the configuration of an SSSD `simple` `access_provider`. Users from the `EngineeringTeam` are provided access to the directory. `DirectoryService` / `LdapAccessFilter` must not be set in this case.

```
DirectoryService:
  ...
  AdditionalSssdConfigs:
    access_provider: simple
    simple_allow_groups: EngineeringTeam
```

Update policy: The compute fleet must be stopped for this setting to be changed for an update.

`DeploymentSettings` section

> **Note**
>
> `DeploymentSettings` is added starting with AWS ParallelCluster version 3.4.0.

(Optional) Specifies the deployment settings configuration.

```
DeploymentSettings:
  LambdaFunctionsVpcConfig:
    SecurityGroupIds:
      - string
    SubnetIds:
      - string
  DisableSudoAccessForDefaultUser: Boolean
  DefaultUserHome: string # 'Shared' or 'Local'
```

`DeploymentSettings` properties

`LambdaFunctionsVpcConfig`

(Optional) Specifies the AWS Lambda functions VPC configurations. For more information, see AWS Lambda VPC configuration in AWS ParallelCluster.

```
LambdaFunctionsVpcConfig:
  SecurityGroupIds:
    - string
  SubnetIds:
    - string
```

`LambdaFunctionsVpcConfig` properties

`SecurityGroupIds` (**Required**, `[String]`)

The list of Amazon VPC security group IDs that are attached to the Lambda functions.

Update policy: If this setting is changed, the update is not allowed.

`SubnetIds` (**Required**, `[String]`)

The list of subnet IDs that are attached to the Lambda functions.

Update policy: If this setting is changed, the update is not allowed.

> **Note**
>
> The subnets and security groups must be in the same VPC.

DisableSudoAccessForDefaultUser property

> **Note**
>
> This config Option is only supported with Slurm Clusters.

(Optional) If `True` , the sudo privileges of the default User will be disabled. This applies to all the nodes in the cluster.

```
# Main DeploymentSettings section in config yaml(applies to HN, CF and LN)
DeploymentSettings:
  DisableSudoAccessForDefaultUser: True
```

To update the value of `DisableSudoAccessForDefaultUser`, you must stop the compute fleet and all login nodes.

Update policy: The compute fleet and login nodes must be stopped for this setting to be changed for an update.

DefaultUserHome property

When set to `Shared`, the cluster will use the default setup and share the default user's directory across the cluster by `/home/<default user>`.

When set to `Local`, the head node, login nodes, and compute nodes will each have a separate local default user directory stored in `local/home/<default user>`.

Build image configuration files

AWS ParallelCluster version 3 uses YAML 1.1 files for build image configuration parameters. Please confirm that indentation is correct to reduce configuration errors. For more information, see the YAML 1.1 spec at https://yaml.org/spec/1.1/.

These configuration files are used to define how your custom AWS ParallelCluster AMIs are built using EC2 Image Builder. Custom AMI building processes are triggered using the `pcluster build-image` command. For some example configuration files, see https://github.com/aws/aws-parallelcluster/tree/release-3.0/cli/tests/pcluster/schemas/test_imagebuilder_schema/test_imagebuilder_schema.

Topics

Build image configuration file properties

`Region` (**Optional**, `String`)

Specifies the AWS Region for the `build-image` operation. For example, `us-east-2`.

`CustomS3Bucket` (**Optional**, `String`)

Specifies the name of an Amazon S3 bucket that is created in your AWS account to store resources that are used by the custom AMI build process and to export logs. The info used by the image is in the custom bucket for image config. AWS ParallelCluster maintains one Amazon S3 bucket in each AWS Region that you create clusters in. By default, these Amazon S3 buckets are named `parallelcluster-hash-v1-DO-NOT-DELETE`.

`Build` section

(Required) Specifies the configuration in which the image will be built.

```
Build:
  Imds:
    ImdsSupport: string
  InstanceType: string
  SubnetId: string
  ParentImage: string
  Iam:
    InstanceRole: string
```

```
    InstanceProfile: string
    CleanupLambdaRole: string
    AdditionalIamPolicies:
      - Policy: string
    PermissionsBoundary: string
  Components:
    - Type: string
      Value: string
  Tags:
    - Key: string
      Value: string
  SecurityGroupIds:
    - string
  UpdateOsPackages:
    Enabled: boolean
  Installation:
    NvidiaSoftware:
      Enabled: boolean
    LustreClient:
      Enabled: boolean
```

`Build` properties

`InstanceType` (**Required**, `String`)

Specifies the instance type for the instance used to build the image.

`SubnetId` (**Optional**, `String`)

Specifies the ID of an existing subnet in which to provision the instance to build the image. The provided subnet requires internet access. Note that you might need to Modify the IP addressing attributes of your subnet if the build fails.

> **Warning**
>
> `pcluster build-image` uses the default VPC. If the default VPC has been deleted, perhaps by using AWS Control Tower or AWS Landing Zone, then the subnet ID must be specified.

When you specify the SubnetId, it is recommended to specify the SecurityGroupIds property as well. If you leave SecurityGroupIds out, AWS ParallelCluster will use default security groups

or rely on the default behavior within the specified subnet. When you use both, you gain these advantages:

- Granular control: When you explicitly define both you ensure the instances launched during the image build process are placed in the correct subnet and have the precise network access you need for your build components and any required services (like access to S3 for build scripts).
- Security best practices: When you define appropriate security groups this helps restrict network access to only necessary ports and services, which enhances the security of your build environment.
- Avoiding potential issues: If you rely solely on defaults this might result in security groups that are too open or too restrictive, which can lead to problems during the build process.

`ParentImage` (**Required**, `String`)

Specifies the base image. The parent image can be either a non AWS ParallelCluster AMI or an official AWS ParallelCluster AMI for the same version. You can't use a AWS ParallelCluster official or custom AMI from a different version of AWS ParallelCluster. The format must either be the ARN of a image `arn:`*`Partition`*`:imagebuilder:`*`Region`*`:`*`Account`*`:image/`*`ImageName`*`/`*`ImageVersion`* or an AMI ID `ami-12345678`.

`SecurityGroupIds` (**Optional**, `[String]`)

Specifies the list of security group IDs for the image.

`Imds`

`Imds` properties

(Optional) Specifies the Amazon EC2 ImageBuilder build and test instance metadata service (IMDS) settings.

```
Imds:
  ImdsSupport: string
```

`ImdsSupport` (**Optional**, `String`)

Specifies which IMDS versions are supported in the Amazon EC2 ImageBuilder build and test instances. Supported values are `v2.0` and `v1.0`. The default value is `v2.0`.

If `ImdsSupport` is set to `v1.0`, both IMDSv1 and IMDSv2 are supported.

If `ImdsSupport` is set to `v2.0`, only IMDSv2 is supported.

For more information, see Use IMDSv2 in the *Amazon EC2 User Guide for Linux instances*.

Update policy: If this setting is changed, the update is not allowed.

> **Note**
>
> Starting with AWS ParallelCluster version 3.7.0, the `ImdsSupport` default value is `v2.0`. We recommend that you set `ImdsSupport` to `v2.0` and replace IMDSv1 with IMDSv2 in your custom actions calls.
> Support for `Imds` / `ImdsSupport` is added with AWS ParallelCluster version 3.3.0.

`Iam`

`Iam` properties

(**Optional**) Specifies the IAM resources for the image build.

```
Iam:
  InstanceRole: string
  InstanceProfile: string
  CleanupLambdaRole: string
  AdditionalIamPolicies:
    - Policy: string
  PermissionsBoundary: string
```

`InstanceProfile` (**Optional**, `String`)

Specifies an instance profile to override the default instance profile for the EC2 Image Builder instance. `InstanceProfile` and `InstanceRole` and `AdditionalIamPolicies` cannot be specified together. The format is `arn:`*`Partition`*`:iam::`*`Account`*`:instance-profile/`*`InstanceProfileName`*.

`InstanceRole` (**Optional**, `String`)

Specifies an instance role to override the default instance role for the EC2 Image Builder instance. `InstanceProfile` and `InstanceRole` and `AdditionalIamPolicies` cannot be specified together. The format is `arn:`*`Partition`*`:iam::`*`Account`*`:role/`*`RoleName`*.

`CleanupLambdaRole` (**Optional**, `String`)

The ARN of the IAM role to use for the AWS Lambda function that backs the AWS CloudFormation custom resource that removes build artifacts on build completion. Lambda needs to be configured as the principal allowed to assume the role. The format is `arn:Partition:iam::Account:role/RoleName`.

`AdditionalIamPolicies` (**Optional**)

Specifies additional IAM policies to attach to the EC2 Image Builder instance used to produce the custom AMI.

```
AdditionalIamPolicies:
  - Policy: string
```

`Policy` (**Optional**, `[String]`)

List of IAM policies. The format is `arn:Partition:iam::Account:policy/PolicyName`.

`PermissionsBoundary` (**Optional**, `String`)

The ARN of the IAM policy to use as permissions boundary for all roles created by AWS ParallelCluster. For more information on IAM permissions boundaries please refer to Permissions boundaries for IAM entities in the *IAM User Guide*. The format is `arn:Partition:iam::Account:policy/PolicyName`.

`Components`

`Components` properties

(**Optional**) Specifies Amazon EC2 ImageBuilder components to use during the AMI build process in addition to the ones provided by default by AWS ParallelCluster. Such components can be used to customize the AMI build process. For more information, see AWS ParallelCluster AMI customization.

```
Components:
  - Type: string
    Value: string
```

`Type` (**Optional**, `String`)

Specifies the type of the type-value pair for the component. Type can be `arn` or `script`.

`Value` (**Optional**, `String`)

Specifies the value of the type-value pair for the component. When type is `arn`, this is the ARN of a EC2 Image Builder component. When type is `script`, this is the https or s3 link that points to the script to use when you create the EC2 Image Builder component.

Tags

Tags properties

(**Optional**) Specifies the list of tags to be set in the resources used to build the AMI.

```
Tags:
  - Key: string
    Value: string
```

`Key` (**Optional**, `String`)

Defines the name of the tag.

`Value` (**Optional**, `String`)

Defines the value of the tag.

UpdateOsPackages

UpdateOsPackages properties

(**Optional**) Specifies whether the operating system is updated before installing AWS ParallelCluster software stack.

```
UpdateOsPackages:
  Enabled: boolean
```

`Enabled` (**Optional**, `Boolean`)

If `true`, the OS is updated and rebooted before installing the AWS ParallelCluster software. The default is `false`.

> **Note**
>
> When `UpdateOsPackages` is enabled, all available OS packages are updated, including the kernel. As a customer, it is your responsibility to verify that the update is compatible with the AMI dependencies that aren't included in the update.
> For example, suppose you want to build an AMI for AWS ParallelCluster version X.0 that's shipped with kernel version Y.0 and some component version Z.0. Suppose the available update includes updated kernel version Y.1 without updates to component Z.0. Before you enable `UpdateOsPackages`, it's your responsibility to verify that component Z.0 supports kernel Y.1.

`Installation`

`Installation` properties

(Optional) Specifies additional software to be installed on the image.

```
Installation:
  NvidiaSoftware:
    Enabled: boolean
  LustreClient:
    Enabled: boolean
```

`NvidiaSoftware` properties (**Optional**)

Specifies the Nvidia Software to be installed.

```
NvidiaSoftware:
    Enabled: boolean
```

`Enabled` (**Optional**, `boolean`)

If `true`, the Nvidia GPU driver and CUDA will be installed. The default is `false`.

`LustreClient` properties (**Optional**)

Specifies that the Amazon FSx Lustre client will be installed.

```
LustreClient:
```

```
    Enabled: boolean
```

`Enabled` (**Optional**, `boolean`)

If `true`, the Lustre client will be installed. The default is `true`.

`Image` section

(Optional) Defines the image properties for the image build.

```
Image:
  Name: string
  RootVolume:
    Size: integer
    Encrypted: boolean
    KmsKeyId: string
  Tags:
    - Key: string
      Value: string
```

`Image` properties

Name (**Optional**, `String`)

Specifies the name of the AMI. If not specified, the name used when calling the `pcluster build-image` command is used.

`Tags`

`Tags` properties

(**Optional**) Specifies key-value pairs for the image.

```
Tags:
  - Key: string
    Value: string
```

Key (**Optional**, `String`)

Defines the name of the tag.

`Value` (**Optional**, `String`)

Defines the value of the tag.

`RootVolume`

`RootVolume` properties

(**Optional**) Specifies properties of the root volume for the image.

```
RootVolume:
  Size: integer
  Encrypted: boolean
  KmsKeyId: string
```

`Size` (**Optional**, `Integer`)

Specifies the size of the root volume for the image, in GiB. The default size is the size of the `ParentImage` plus 27 GiB.

`Encrypted` (**Optional**, `Boolean`)

Specifies if the volume is encrypted. The default value is `false`.

`KmsKeyId` (**Optional**, `String`)

Specifies the ARN of the AWS KMS key used to encrypt the volume. The format is "`arn:`*`Partition`*`:kms:`*`Region`*`:`*`Account`*`:key/`*`KeyId`*.

`DeploymentSettings` section

> **Note**
>
> `DeploymentSettings` is added starting with AWS ParallelCluster version 3.4.0.

(Optional) Specifies the deployment settings configuration.

```
DeploymentSettings:
  LambdaFunctionsVpcConfig:
    SecurityGroupIds:
```

```
      - string
    SubnetIds:
      - string
```

`DeploymentSettings` properties

`LambdaFunctionsVpcConfig`

(Optional) Specifies the AWS Lambda functions VPC configurations. For more information, see AWS Lambda VPC configuration in AWS ParallelCluster.

```
LambdaFunctionsVpcConfig:
  SecurityGroupIds:
    - string
  SubnetIds:
    - string
```

`LambdaFunctionsVpcConfig` properties

`SecurityGroupIds` (**Required**, `[String]`)

The list of Amazon VPC security group IDs that are attached to the Lambda functions.

Update policy: If this setting is changed, the update is not allowed.

`SubnetIds` (**Required**, `[String]`)

The list of subnet IDs that are attached to the Lambda functions.

Update policy: If this setting is changed, the update is not allowed.

> **Note**
>
> The subnets and security groups must be in the same VPC.

AWS ParallelCluster API reference

This section provides descriptions, syntax, and usage examples for each of the AWS ParallelCluster API actions.

Topics

buildImage

Create a custom AWS ParallelCluster image in an AWS Region.

Topics

Request syntax

```
POST /v3/images/custom
{
  "imageConfiguration": "string",
  "imageId": "string",
  "dryrun": boolean,
  "region": "string",
  "rollbackOnFailure": boolean,
  "supressValidators": [ "string" ],
  "validationFailureLevel": "string"
}
```

Request body

imageConfiguration

The image configuration as a YAML document.

Type: string

Required: Yes

imageId

The ID of the image to build.

Type: string

Required: Yes

dryrun

If set to `true`, only perform request validation without creating any resource. Use this parameter to validate the image configuration. The default is `false`.

Type: boolean

Required: No

region

The AWS Region in which you run the command to build the image.

Type: string

Required: No

rollbackOnFailure

If set to `true`, image stack rollback occurs if the image fails to create. The default is `false`.

Type: boolean

Required: No

suppressValidators

Identify one or more configuration validators to suppress.

Type: list of strings

Format: `(ALL|type:[A-Za-z0-9]+)`

Required: No

validationFailureLevel

The minimum validation level that causes image build to fail. The default is `ERROR`.

Type: string

Valid values: `INFO | WARNING | ERROR`

Required: No

Response syntax

```
{
  "image": {
    "imageId": "string",
    "ec2AmiInfo": {
      "amiId": "string"
    },
    "region": "string",
    "version": "string",
    "cloudformationStackArn": "string",
    "imageBuildStatus": "BUILD_IN_PROGRESS",
    "cloudformationStackStatus": "CREATE_IN_PROGRESS"
  },
```

```
  "validationMessages": [
    {
      "id": "string",
      "type": "string",
      "level": "INFO",
      "message": "string"
    }
  ]
}
```

Response body

image

imageId

The ID of the image.

Type: string

cloudformationStackArn

The Amazon Resource Name (ARN) of the main CloudFormation stack.

Type: string

cloudformationStackStatus

The CloudFormation stack status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | ROLLBACK_IN_PROGRESS | ROLLBACK_FAILED | ROLLBACK_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_ROLLBACK_IN_PROGRESS | UPDATE_ROLLBACK_FAILED | UPDATE_ROLLBACK_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_ROLLBACK_COMPLETE`

ec2AmiInfo

ami_id

The Amazon EC2 AMI ID.

Type: string

imageBuildStatus

The image build status.

Type: string

Valid values: `BUILD_IN_PROGRESS | BUILD_FAILED | BUILD_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE`

region

The AWS Region in which the image is built.

Type: string

version

The AWS ParallelCluster version that's used to build the image.

Type: string

validationMessages

A list of messages with a validation level lower than `validationFailureLevel`. The list of messages is collected during configuration validation.

id

The validator ID.

Type: string

level

The validation level.

Type: string

Valid values: `INFO | WARNING | ERROR`

message

A validation message.

Type: string

type

The type of validator.

Type: string

Example

Python

Request

```
$ build_image(custom-image-id, custom-image-config.yaml)
```

200 Response

```
{
  "image": {
    "cloudformation_stack_arn": "arn:aws:cloudformation:us-
east-1:123456789012:stack/custom-image-id/711b76b0-af81-11ec-a29f-0ee549109f1f",
    "cloudformation_stack_status": "CREATE_IN_PROGRESS",
    "image_build_status": "BUILD_IN_PROGRESS",
    "image_id": "custom-image-id",
    "region": "us-east-1",
    "version": "3.2.1"
  }
}
```

createCluster

Create a managed cluster in an AWS Region.

Topics

Request syntax

```
POST /v3/clusters
{
  "clusterName": "string",
  "clusterConfiguration": "string",
  "dryrun": boolean,
  "region": "string",
  "rollbackOnFailure", boolean,
  "suppressValidators": [ "string" ],
  "validationFailureLevel": "string"
}
```

Request body

clusterConfiguration

The cluster configuration as a YAML document.

Type: string

Required: Yes

clusterName

The name of the cluster to create.

The name must start with an alphabetical character. The name can have up to 60 characters. If Slurm accounting is enabled, the name can have up to 40 characters.

Type: string

Required: Yes

dryrun

If set to `true`, only perform request validation but do not create any resource. Use this parameter to validate the cluster configuration. The default is `false`.

Type: boolean

Required: No

region

The AWS Region that the cluster is in.

Type: string

Required: No

rollbackOnFailure

If set to `true`, cluster stack rollback occurs if the cluster fails to create. The default is `true`.

Type: boolean

Required: No

suppressValidators

Identify one or more configuration validators to suppress.

Type: list of strings

Format: `(ALL|type:[A-Za-z0-9]+)`

Required: No

validationFailureLevel

The minimum validation level that causes cluster create to fail. The default is `ERROR`.

Type: string

Valid values: `INFO | WARNING | ERROR`

Required: No

Response syntax

```
{
  "cluster": {
    "clusterName": "string",
    "region": "string",
    "version": "string",
    "cloudformationStackArn": "string",
```

```
    "cloudformationStackStatus": "CREATE_IN_PROGRESS",
    "clusterStatus": "CREATE_IN_PROGRESS",
    "scheduler": {
      "type": "string",
      "metadata": {
        "name": "string",
        "version": "string"
      }
    }
  },
  "validationMessages": [
    {
      "id": "string",
      "type": "string",
      "level": "INFO",
      "message": "string"
    }
  ]
}
```

Response body

clusterName

The name of cluster.

Type: string

cloudformationStackArn

The Amazon Resource Name (ARN) of the main CloudFormation stack.

Type: string

cloudformationStackStatus

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | ROLLBACK_IN_PROGRESS | ROLLBACK_FAILED | ROLLBACK_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_ROLLBACK_IN_PROGRESS | UPDATE_ROLLBACK_FAILED | UPDATE_ROLLBACK_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_ROLLBACK_COMPLETE`

clusterStatus

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_FAILED`

region

The AWS Region that the cluster is created in.

Type: string

scheduler

metadata

The scheduler metadata

name

The name of the scheduler.

Type: string

version

The scheduler version.

Type: string

type

The scheduler type.

Type: string

version

The AWS ParallelCluster version that's used to create the cluster.

Type: string

validation_messages

A list of messages with a validation level lower than `validationFailureLevel`. The list of messages is collected during configuration validation.

id

The ID of the validator.

Type: string

level

Type: string

Valid values: `INFO | WARNING | ERROR`

message

A validation message.

Type: string

type

The type of the validator.

Type: string

Example

Python

Request

```
$ create_cluster(cluster_name_3x, cluster-config.yaml)
```

200 Response

```
{
  "cluster": {
    "cloudformation_stack_arn": "arn:aws:cloudformation:us-
east-1:123456789012:stack/cluster-3x/e0462730-50b5-11ed-99a3-0a5ddc4a34c7",
    "cloudformation_stack_status": "CREATE_IN_PROGRESS",
    "cluster_name": "cluster-3x",
    "cluster_status": "CREATE_IN_PROGRESS",
    "region": "us-east-1",
    "scheduler": {
      "type": "slurm"
    },
```

```
    "version": "3.2.1"
  }
}
```

deleteCluster

Initiate deleting a cluster.

Topics

Request syntax

```
DELETE /v3/clusters/{clusterName}
{
  "region": "string"
}
```

Request body

clusterName

The name of the cluster.

Type: string

Required: Yes

region

The AWS Region in which the cluster is deleted.

Type: string

Required: No

Response syntax

```
{
   "cluster": {
       "clusterName": "string",
       "region": "string",
       "version": "string",
       "cloudformationStackArn": "string",
       "cloudformationStackStatus": "DELETE_IN_PROGRESS",
       "clusterStatus": "DELETE_IN_PROGRESS",
       "scheduler": {
           "type": "string",
           "metadata": {
               "name": "string",
               "version": "string"
           }
        }
    }
}
```

Response body

cluster

A list of cluster instances.

clusterName

The name of a cluster.

Type: string

cloudformationStackArn

The Amazon Resource Name (ARN) of the main CloudFormation stack.

Type: string

cloudformationStackStatus

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | ROLLBACK_IN_PROGRESS | ROLLBACK_FAILED | ROLLBACK_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE |`

UPDATE_IN_PROGRESS | UPDATE_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_ROLLBACK_IN_PROGRESS | UPDATE_ROLLBACK_FAILED | UPDATE_ROLLBACK_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_ROLLBACK_COMPLETE

clusterStatus

Type: string

Valid values: CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_FAILED

region

The AWS Region in which the cluster is created.

Type: string

scheduler

metadata

The scheduler metadata.

name

The name of the scheduler.

Type: string

version

The scheduler version.

Type: string

type

The scheduler type.

Type: string

version

The AWS ParallelCluster version that's used to create the cluster.

Type: string

Example

Python

Request

```
$ delete_cluster(cluster_name_3x)
```

200 Response

```
{
  "cluster": {
    "cloudformation_stack_arn": "arn:aws:cloudformation:us-
east-1:123456789012:stack/cluster_name_3x/16b49540-aee5-11ec-8e18-0ac1d712b241",
    "cloudformation_stack_status": "DELETE_IN_PROGRESS",
    "cluster_name": "cluster_name_3x",
    "cluster_status": "DELETE_IN_PROGRESS",
    "region": "us-east-1",
    "version": "3.2.1"
  }
}
```

deleteClusterInstances

Initiate the forced termination of all cluster compute nodes. This action doesn't support AWS Batch clusters.

Topics

Request syntax

```
DELETE /v3/clusters/{clusterName}/instances
{
  "force": boolean,
```

```
  "region": "string"
}
```

Request body

clusterName

The name of the cluster.

Type: string

Required: Yes

force

If set to `true`, force the deletion when the cluster with the given name isn't found. The default is `false`.

Type: boolean

Required: No

region

The AWS Region that the cluster is in.

Type: string

Required: No

Response body

None

Example

Python

Request

```
$ delete_cluster_instances(cluster_name_3x)
```

200 Response

None

deleteImage

Initiate the deletion of the custom AWS ParallelCluster image.

Topics

Request syntax

```
DELETE /v3/images/custom/{imageId}
{
  "force": boolean,
  "region": "string"
}
```

Request body

imageId

The ID of the image.

Type: string

Required: Yes

force

If set to `true`, force the AMI delete. Use this parameter if there are instances that use the AMI or if the AMI is shared. The default is `false`.

Type: boolean

Required: No

region

The AWS Region in which the image was created.

Type: string

Required: No

Response syntax

```
{
  "image": {
    "imageId": "string",
    "ec2AmiInfo": {
      "amiId": "string"
    },
    "region": "string",
    "version": "string",
    "cloudformationStackArn": "string",
    "imageBuildStatus": "DELETE_IN_PROGRESS",
    "cloudformationStackStatus": "DELETE_IN_PROGRESS"
  }
}
```

Response body

image

cloudformationStackArn

The Amazon resource name (ARN) of the main CloudFormation stack.

Type: string

cloudformationStackStatus

The CloudFormation stack status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | ROLLBACK_IN_PROGRESS | ROLLBACK_FAILED | ROLLBACK_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE |`

UPDATE_IN_PROGRESS | UPDATE_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_ROLLBACK_IN_PROGRESS | UPDATE_ROLLBACK_FAILED | UPDATE_ROLLBACK_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_ROLLBACK_COMPLETE

ec2AmiInfo

amiId

The Amazon EC2 AMI ID.

Type: string

imageBuildStatus

The image build status.

Type: string

Valid values: BUILD_IN_PROGRESS | BUILD_FAILED | BUILD_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE

imageId

The ID of the image.

Type: string

region

The AWS Region in which the image is created.

Type: string

version

The AWS ParallelCluster version that's used to build the image.

Type: string

Example

Python

Request

```
$ delete_image(custom-image-id)
```

200 Response

```
{
  "image": {
    "image_build_status": "DELETE_IN_PROGRESS",
    "image_id": "custom-image-id",
    "region": "us-east-1",
    "version": "3.2.1"
  }
}
```

describeCluster

Get detailed information about an existing cluster.

Topics

Request syntax

```
GET /v3/clusters/{clusterName}
{
  "region": "string"
}
```

Request body

clusterName

The name of the cluster.

Type: string

Required: Yes

region

The AWS Region that the cluster is in.

Type: string

Required: No

Response syntax

> **Note**
>
> `failureReason` has changed to `failures` starting with AWS ParallelCluster version 3.5.0.

```
{
  "clusterName": "string",
  "region": "string",
  "version": "string",
  "cloudFormationStackStatus": "CREATE_IN_PROGRESS",
  "clusterStatus": "CREATE_IN_PROGRESS",
  "scheduler": {
    "type": "string",
    "metadata": {
      "name": "string",
      "version": "string"
    }
  },
  "cloudformationStackArn": "string",
  "creationTime": "2019-08-24T14:15:22Z",
  "lastUpdatedTime": "2019-08-24T14:15:22Z",
  "clusterConfiguration": {
    "url": "string"
  },
  "computeFleetStatus": "START_REQUESTED",
  "tags": [
    {
```

```
      "key": "string",
      "value": "string"
    }
  ],
  "headNode": {
    "instanceId": "string",
    "instanceType": "string",
    "launchTime": "2019-08-24T14:15:22Z",
    "privateIpAddress": "string",
    "publicIpAddress": "string",
    "state": "pending"
  },
  "failures": [
    {
      "failureCode": "string",
      "failureReason": "string"
    }
  ],
  "loginNodes": {
    "status": "string",
    "address": "string",
    "poolName": "string",
    "scheme": "string",
    "healthyNodes": integer,
    "unhealthyNodes": integer
  }
}
```

Response body

clusterName

The name of the cluster.

Type: string

cloudformationStackArn

The Amazon Resource Name (ARN) of the main CloudFormation stack.

Type: string

cloudformationStackStatus

The CloudFormation stack status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | ROLLBACK_IN_PROGRESS | ROLLBACK_FAILED | ROLLBACK_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_ROLLBACK_IN_PROGRESS | UPDATE_ROLLBACK_FAILED | UPDATE_ROLLBACK_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_ROLLBACK_COMPLETE`

clusterConfiguration

url

The URL of the cluster configuration file.

Type: string

clusterStatus

The cluster status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_FAILED`

computeFleetStatus

The compute fleet status.

Type: string

Valid values: `START_REQUESTED | STARTING | RUNNING | PROTECTED | STOP_REQUESTED | STOPPING | STOPPED | UNKNOWN | ENABLED | DISABLED`

creationTime

Timestamp for when the cluster was created.

Type: datetime

lastUpdatedTime

Timestamp for when the cluster was last updated.

Type: datetime

region

The AWS Region in which the cluster is created.

Type: string

tags

The list of tags that are associated with the cluster.

key

Tag name.

Type: string

tag

Tag value.

Type: string

version

The AWS ParallelCluster version that's used to create the cluster.

Type: string

failures

The list of failures when the cluster stack is in `CREATE_FAILED` status.

failureCode

The failure code when the cluster stack is in `CREATE_FAILED` status.

Type: string

failureReason

The reason for the failure when the cluster stack is in `CREATE_FAILED` status.

Type: string

head_node

The cluster head node.

instanceId

The Amazon EC2 instance ID.

Type: string

instanceType

The Amazon EC2 instance type.

Type: string

launchTime

The time when the Amazon EC2 instance was launched.

Type: datetime

privateIpAddress

The cluster private IP address.

Type: string

publicIpAddress

The cluster public IP address.

Type: string

state

The head node instance status.

Type: string

Valid values: `pending | running | shutting-down | terminated | stopping | stopped`

scheduler

metadata

The scheduler metadata.

name

The name of the scheduler.

Type: string

version

The scheduler version.

Type: string

loginNodes

status

The login node status.

Type: string

Valid values: `PENDING | FAILED | ACTIVE`

address

The login node address.

Type: string

poolName

The login node pool name.

Type: string

scheme

The login node scheme.

Type: string

healthyNodes

The number of healthy nodes.

Type: integer

unhealthyNodes

The number of unhealthy nodes.

Type: integer

type

The scheduler type.

Type: string

Example

Python

Request

```
$ describe_cluster(cluster_name_3x)
```

200 Response


```
{
  "cloud_formation_stack_status": "CREATE_COMPLETE",
  "cloudformation_stack_arn": "arn:aws:cloudformation:us-east-1:123456789012:stack/
cluster_name_3x/16b49540-aee5-11ec-8e18-0ac1d712b241",
  "cluster_configuration": {
    "url": "https://parallelcluster-...."
  },
  "cluster_name": "cluster_name_3x",
  "cluster_status": "CREATE_COMPLETE",
  "compute_fleet_status": "RUNNING",
  "creation_time": datetime.datetime(2022, 3, 28, 22, 19, 9, 661000,
 tzinfo=tzlocal()),
  "head_node": {
    "instance_id": "i-abcdef01234567890",
    "instance_type": "t2.micro",
    "launch_time": datetime.datetime(2022, 3, 28, 22, 21, 56, tzinfo=tzlocal()),
    "private_ip_address": "172.31.56.3",
    "public_ip_address": "107.23.100.164",
    "state": "running"
  },
  "last_updated_time": datetime.datetime(2022, 3, 28, 22, 19, 9, 661000,
 tzinfo=tzlocal()),
  "region": "us-east-1",
  "tags": [
    {
      "key": "parallelcluster:version", "value": "3.2.1"
```

```
    }
  ],
  "version": "3.2.1"
}
```

describeClusterInstances

Describe the instances that belong to a cluster.

Topics

Request syntax

```
GET /v3/clusters/{clusterName}/instances
{
  "nextToken": "string",
  "nodeType": "string",
  "queueName": "string",
  "region": "string"
}
```

Request body

clusterName

The name of the cluster.

Type: string

Required: Yes

nextToken

The token for the next set of results.

Type: string

Required: No

nodeType

Filter the instances by node type.

Type: string

Valid values: `HeadNode`, `ComputeNode`, `LoginNode`

Required: No

queueName

Filter the instances by queue name.

Type: string

Required: No

region

The AWS Region that the cluster is in.

Type: string

Required: No

Response syntax

```
{
  "nextToken": "string",
  "instances": [
    {
      "instanceId": "string",
      "instanceType": "string",
      "launchTime": "2019-08-24T14:15:22Z",
      "privateIpAddress": "string",
      "publicIpAddress": "string",
      "state": "pending",
      "nodeType": "HeadNode",
      "queueName": "string",
```

```
      "poolName": "string"
    }
  ]
}
```

Response body

instances

The list of cluster instances.

instanceId

The Amazon EC2 instance ID.

Type: string

instanceType

The Amazon EC2 instance type.

Type: string

launchTime

The time when the Amazon EC2 instance was launched.

Type: datetime

nodeType

The node type.

Type: string

Valid values: `HeadNode`, `ComputeNode`, `LoginNode`

publicIpAddress

The cluster public IP address.

Type: string

queueName

The name of the queue in which the Amazon EC2 instance is backing a node.

Type: string

state

The node Amazon EC2 instance status.

Type: string

Valid values: `pending | running | shutting-down | terminated | stopping | stopped`

nextToken

A token that can be used to retrieve the next set of results, or `null` if there are no additional results.

Type: string

Example

Python

Request

```
$ describe_cluster_instances(cluster_name_3x)
```

200 Response

```
{
  "instances": [
    {
      "instance_id": "i-abcdef01234567890",
      "instance_type": "t2.micro",
      "launch_time": datetime.datetime(2022, 3, 30, 14, 2, 7, tzinfo=tzlocal()),
      "node_type": "HeadNode",
      "private_ip_address": "192.0.2.5",
      "public_ip_address": "198.51.100.180",
      "state": "running"
    }
  ]
}
```

describeComputeFleet

Describe the status of the compute fleet.

Topics

Request syntax

```
GET /v3/clusters/{clusterName}/computefleet
{
  "region": "string"
}
```

Request body

clusterName

The name of the cluster.

Type: string

Required: Yes

region

The AWS Region that the cluster is in.

Type: string

Required: No

Response syntax

```
{
```

```
  "status": "START_REQUESTED",
  "lastStatusUpdatedTime": "2019-08-24T14:15:22Z"
}
```

Response body

status

Type: string

Valid values: START_REQUESTED | STARTING | RUNNING | PROTECTED | STOP_REQUESTED | STOPPING | STOPPED | UNKNOWN | ENABLED | DISABLED

lastStatusUpdatedTime

The timestamp representing the last status update time.

Type: datetime

Example

Python

Request

```
$ describe_compute_fleet(cluster_name_3x)
```

200 Response

```
{
  "last_status_updated_time": datetime.datetime(2022, 3, 28, 22, 27, 14,
 tzinfo=tzlocal()),
  "status": "RUNNING"
}
```

describeImage

Get detailed information about an existing image.

Topics

Request syntax

```
GET /v3/images/custom/{imageId}
{
  "region": "string"
}
```

Request body

imageId

The ID of the image.

Type: string

Required: Yes

region

The AWS Region in which the image was created.

Type: string

Required: No

Response syntax

```
{
  "imageId": "string",
  "region": "string",
  "version": "string",
  "imageBuildStatus": "BUILD_IN_PROGRESS",
  "imageBuildLogsArn": "string",
```

```
  "cloudformationStackStatus": "CREATE_IN_PROGRESS",
  "cloudformationStackStatusReason": "string",
  "cloudformationStackArn": "string",
  "creationTime": "2019-08-24T14:15:22Z",
  "cloudformationStackCreationTime": "2019-08-24T14:15:22Z",
  "cloudformationStackTags": [
    {
      "key": "string",
      "value": "string"
    }
  ],
  "imageConfiguration": {
    "url": "string"
  },
  "imagebuilderImageStatus": "PENDING",
  "imagebuilderImageStatusReason": "string",
  "ec2AmiInfo": {
    "amiId": "string",
    "tags": [
      {
        "key": "string",
        "value": "string"
      }
    ],
    "amiName": "string",
    "architecture": "string",
    "state": "PENDING",
    "description": "string"
  }
}
```

Response body

imageId

The ID of the image to retrieve detailed information for.

Type: string

imageBuildStatus

The image build status.

Type: string

Valid values: `BUILD_IN_PROGRESS | BUILD_FAILED | BUILD_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE`

imageConfiguration

url

The URL of the image configuration file.

Type: string

region

The AWS Region in which the image is created.

Type: string

version

The AWS ParallelCluster version that's used to build the image.

Type: string

cloudformationStackArn

The Amazon Resource Name (ARN) of the main CloudFormation stack.

Type: string

cloudformationStackCreationTime

The time when the CloudFormation stack was created.

Type: datetime

cloudformationStackStatus

The CloudFormation stack status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | ROLLBACK_IN_PROGRESS | ROLLBACK_FAILED | ROLLBACK_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_ROLLBACK_IN_PROGRESS | UPDATE_ROLLBACK_FAILED | UPDATE_ROLLBACK_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_ROLLBACK_COMPLETE`

cloudformationStackStatusReason

The reason for the CloudFormation stack status.

Type: string

cloudformationStackTags

The list of tags for the CloudFormation stack.

key

The tag name.

Type: string

value

The tag value.

Type: string

creationTime

The times when the image was created.

Type: datetime

ec2AmiInfo

amiId

The Amazon EC2 AMI ID.

Type: string

amiName

The Amazon EC2 AMI name.

Type: string

architecture

The Amazon EC2 AMI architecture.

Type: string

state

The state of the Amazon EC2 AMI.

Type: string

Valid values: `PENDING | AVAILABLE | INVALID | DEREGISTERED | TRANSIENT | FAILED | ERROR`

tags

List of Amazon EC2 AMI Tags.

key

Tag name.

Type: string

value

Tag value.

Type: string

imagebuilderImageStatus

The ImageBuilder Image status.

Type: string

Valid values: `PENDING | CREATING | BUILDING | TESTING | DISTRIBUTING | INTEGRATING | AVAILABLE | CANCELLED | FAILED | DEPRECATED | DELETED`

imagebuilderImageStatusReason

The reason the ImageBuilder Image has that status.

Type: string

imageBuildLogsArn

The Amazon Resource Name (ARN) of the logs for the image build process.

Type: string

Example

Python

Request

```
$ describe_image(custom-image-id)
```

200 Response

```
{
  "cloudformation_stack_arn": "arn:aws:cloudformation:us-east-1:123456789012:stack/
custom-image-id/6accc570-b080-11ec-845e-0e2dc6386985",
  "cloudformation_stack_creation_time": datetime.datetime(2022, 3, 30, 23, 23, 33,
 731000, tzinfo=tzlocal()),
  "cloudformation_stack_status": "CREATE_IN_PROGRESS",
  "cloudformation_stack_tags": [
    {
      "key": "parallelcluster:version", "value": "3.2.1"
    },
    {
      "key": "parallelcluster:image_name",
      "value": 'custom-image-id"
    },
    {
      "key": "parallelcluster:custom-image-id",
      "value": "custom-image-id"
    },
    {
      "key": 'parallelcluster:amzn-s3-demo-bucket",
      "value": 'amzn-s3-demo-bucket"
    },
    {
      "key": "parallelcluster:s3_image_dir",
      "value": "parallelcluster/3.2.1/images/custom-image-id-1234567890abcdef0"
    },
    {
      "key": "parallelcluster:build_log",
      "value": "arn:aws:logs:us-east-1:123456789012:log-group:/aws/imagebuilder/
ParallelClusterImage-custom-image-id"
    },
    {
      "key": "parallelcluster:build_config",
      "value": "s3://amzn-s3-demo-bucket/parallelcluster/3.2.1/images/custom-image-
id-1234567890abcdef0/configs/image-config.yaml"
    }
  ],
  "image_build_logs_arn": "arn:aws:logs:us-east-1:123456789012:log-group:/aws/
imagebuilder/ParallelClusterImage-alinux2-image",
```

```
  "image_build_status": "BUILD_IN_PROGRESS",
  "image_configuration": {
    "url": "https://amzn-s3-demo-bucket.s3.amazonaws.com/parallelcluster/3.2.1/
images/custom-image-id-1234567890abcdef0/configs/image-config.yaml?..."
  },
  "image_id": 'custom-image-id',
  "imagebuilder_image_status": "PENDING",
  "region": "us-east-1",
  "version": "3.2.1"
}
```

getClusterLogEvents

Retrieve the events that are associated with a log stream.

Topics

Request syntax

```
GET /v3/clusters/{clusterName}/logstreams/{logStreamName}
{
  "endTime": datetime,
  "limit": float,
  "nextToken": "string",
  "region": "string",
  "startFromHead": boolean,
  "startTime": datetime
}
```

Request body

clusterName

The name of the cluster.

Type: string

Required: Yes

logStreamName

The name of the log stream.

Type: string

Required: Yes

endTime

The end of the time range, expressed in ISO 8601 format. Events with a timestamp equal to or later than this time are not included.

Type: datetime

Format: `2021-01-01T20:00:00Z`

Required: No

limit

The maximum number of log events returned. If you don't specify a value, the maximum is the number of log events that can fit in a response size of 1 MB, or up to 10,000 log events.

Type: float

Required: No

nextToken

The token for the next set of results.

Type: string

Required: No

region

The AWS Region that the cluster is in.

Type: string

Required: No

startFromHead

If set to `true`, the earliest log events are returned first. If the value is `false`, the latest log events are returned first. The default is `false`.

Type: boolean

Required: No

startTime

The start of the time range, expressed in ISO 8601 format. Events with a timestamp equal to this time or later than this time are included.

Type: datetime

Format: `2021-01-01T20:00:00Z`

Required: No

Response syntax

```
{
  "nextToken": "string",
  "prevToken": "string",
  "events": [
    {
      "timestamp": "2019-08-24T14:15:22Z",
      "message": "string"
    }
  ]
}
```

Response body

events

List of filtered events.

message

The event message.

Type: string

timestamp

The event timestamp.

Type: datetime

nextToken

A token that can be used to retrieve the next set of results, or `null` if there are no additional results.

Type: string

prevToken

A token that can be used to retrieve the previous set of results, or `null` if there are no additional results.

Type: string

Example

Python

Request

```
$ get_cluster_log_events(cluster_name_3x, log_stream_name=ip-192-0-2-26.i-abcdef01234567890.cfn-init)
```

200 Response

```
"events": [
  {
    "message": "2022-09-22 16:40:15,127 [DEBUG] CloudFormation client initialized
 with endpoint https://cloudformation.us-east-1.amazonaws.com",
    "timestamp": "2022-09-22T16:40:15.127Z"
  },
  {
    "message": "2022-09-22 16:40:15,127 [DEBUG] Describing resource
 HeadNodeLaunchTemplate in stack cluster_name_3x",
    "timestamp": "2022-09-22T16:40:15.127Z"
  },
  ...
```

```
]
```

getClusterStackEvents

Retrieve the events that are associated with the stack for a cluster.

> **Note**
>
> Starting in version 3.6.0, AWS ParallelCluster uses nested stacks to create the resources associated with queues and compute resources. The `GetClusterStackEvents` API and the `pcluster get-cluster-stack-events` command only return the cluster main stack events. You can view the cluster stack events, including those related to queues and compute resources, in the CloudFormation console.

Topics

Request syntax

```
GET /v3/clusters/{clusterName}/stackevents
{
  "nextToken": "string",
  "region": "string"
}
```

Request body

clusterName

The name of the cluster.

Type: string

Required: Yes

nextToken

The token for the next set of results.

Type: string

Required: No

region

The AWS Region that the cluster is in.

Type: string

Required: No

Response syntax

```
{
  "nextToken": "string",
  "events": [
    {
      "stackId": "string",
      "eventId": "string",
      "stackName": "string",
      "logicalResourceId": "string",
      "physicalResourceId": "string",
      "resourceType": "string",
      "timestamp": "2019-08-24T14:15:22Z",
      "resourceStatus": "CREATE_IN_PROGRESS",
      "resourceStatusReason": "string",
      "resourceProperties": "string",
      "clientRequestToken": "string"
    }
  ]
}
```

Response body

events

List of filtered events.

clientRequestToken

The token passed to the action that generated this event.

Type: string

eventId

The unique ID of this event.

Type: string

logicalResourceId

The logical name of the resource specified in the template.

Type: string

physicalResourceId

The name or unique identifier that's associated with the physical instance of the resource.

Type: string

resourceProperties

A BLOB of the properties that are used to create the resource.

Type: string

resourceStatus

The resource status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | DELETE_SKIPPED | UPDATE_IN_PROGRESS | UPDATE_FAILED | UPDATE_COMPLETE | IMPORT_FAILED | IMPORT_COMPLETE | IMPORT_IN_PROGRESS | IMPORT_ROLLBACK_IN_PROGRESS | IMPORT_ROLLBACK_FAILED | IMPORT_ROLLBACK_COMPLETE`

resourceStatusReason

A success or failure message that's associated with the resource.

Type: string

resourceType

The type of resource.

Type: string

stackId

The unique ID name of the instance of the stack.

Type: string

stackName

The name that's associated with a stack.

Type: string

timestamp

The time when the status was updated.

Type: datetime

nextToken

A token that can be used to retrieve the next set of results, or `null` if there are no additional results.

Type: string

Example

Python

Request

```
$ get_cluster_stack_events(cluster_name_3x)
```

200 Response

```
{
```

```
  "events": [
    {
      "event_id": "590b3820-b081-11ec-985e-0a7af5751497",
      "logical_resource_id": "cluster_name_3x",
      "physical_resource_id": "arn:aws:cloudformation:us-east-1:123456789012:stack/
cluster_name_3x/11a59710-b080-11ec-b8bd-129def1380e9",
      "resource_status": "CREATE_COMPLETE",
      "resource_type": "AWS::CloudFormation::Stack",
      "stack_id": "arn:aws:cloudformation:us-east-1:123456789012:stack/
cluster_name_3x/11a59710-b080-11ec-b8bd-129def1380e9",
      "stack_name": "cluster_name_3x",
      "timestamp": datetime.datetime(2022, 3, 30, 23, 30, 13, 268000,
 tzinfo=tzlocal())
    },
    ...
  ]
}
```

getImageLogEvents

Retrieve the events that are associated with an image build.

Topics

Request syntax

```
GET /v3/images/custom/{imageId}/logstreams/{logStreamName}
{
  "endTime": datetime,
  "limit": float,
  "nextToken": "string",
  "region": "string",
  "startFromHead": boolean,
  "startTime": datetime
```

```
}
```

Request body

imageId

The ID of the image.

Type: string

Required: Yes

logStreamName

The name of the logstream.

Type: string

Required: Yes

endTime

The end of the time range, expressed in ISO 8601 format. Events with a timestamp equal to or later than this time aren't included.

Type: datetime

Format: `2021-01-01T20:00:00Z`

Required: No

limit

The maximum number of log events returned. If you don't specify a value, the maximum is as many log events as can fit in a response size of 1 MB, up to 10,000 log events.

Type: float

Required: No

nextToken

The token for the next set of results.

Type: string

Required: No

region

The AWS Region that the image in.

Type: string

Required: No

startFromHead

If set to `true`, return the earliest log events first. If set to false, return the latest log events first. The default is `false`.

Type: boolean

Required: No

startTime

The start of the time range, expressed in ISO 8601 format. Events with a timestamp equal to this time or later than this time are included.

Type: datetime

Format: `2021-01-01T20:00:00Z`

Required: No

Response syntax

```
{
  "nextToken": "string",
  "prevToken": "string",
  "events": [
    {
      "timestamp": "2019-08-24T14:15:22Z",
      "message": "string"
    }
  ]
}
```

Response body

events

A list of filtered events.

message

The event message.

Type: string

timestamp

The event timestamp.

Type: datetime

nextToken

A token that can be used to retrieve the next set of results, or `null` if there are no additional results.

Type: string

prevToken

A token that can be used to retrieve the previous set of results, or `null` if there are no additional results.

Type: string

Example

Python

Request

```
$ get_image_log_events(image_id, log_stream_name=3.2.1/1)
```

200 Response

```
"events": [
  {
```

```
    "message": "ExecuteBash: STARTED EXECUTION",
    "timestamp": 2022-04-05T15:51:20.228Z"
  },
  {
    "message": "ExecuteBash: Created temporary directory: /tmp/1234567890abcdef0",
    "timestamp": "2022-04-05T15:51:20.228Z"
  },
  ...
]
```

getImageStackEvents

Retrieve the events that are associated with the stack for an image build.

Topics

Request syntax

```
GET /v3/images/custom/{imageId}/stackevents
{
  "nextToken": "string",
  "region": "string"
}
```

Request body

imageId

The ID of the image.

Type: string

Required: Yes

nextToken

The token for the next set of results.

Type: string

Required: No

region

The AWS Region that the image is in.

Type: string

Required: No

Response syntax

```
{
  "nextToken": "string",
  "events": [
    {
      "stackId": "string",
      "eventId": "string",
      "stackName": "string",
      "logicalResourceId": "string",
      "physicalResourceId": "string",
      "resourceType": "string",
      "timestamp": "2019-08-24T14:15:22Z",
      "resourceStatus": "CREATE_IN_PROGRESS",
      "resourceStatusReason": "string",
      "resourceProperties": "string",
      "clientRequestToken": "string"
    }
  ]
}
```

Response body

events

A list of filtered events.

clientRequestToken

The token passed to the action that generated this event.

Type: string

eventId

The unique ID of this event.

Type: string

logicalResourceId

The logical name of the resource specified in the template.

Type: string

physicalResourceId

The name or unique identifier that's associated with the physical instance of the resource.

Type: string

resourceProperties

A BLOB of the properties that are used to create the resource.

Type: string

resourceStatus

The resource status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | DELETE_SKIPPED | UPDATE_IN_PROGRESS | UPDATE_FAILED | UPDATE_COMPLETE | IMPORT_FAILED | IMPORT_COMPLETE | IMPORT_IN_PROGRESS | IMPORT_ROLLBACK_IN_PROGRESS | IMPORT_ROLLBACK_FAILED | IMPORT_ROLLBACK_COMPLETE`

resourceStatusReason

A success or failure message that's associated with the resource.

Type: string

resourceType

The type of resource.

Type: string

stackId

The unique ID name of the instance of the stack.

Type: string

stackName

The name that's associated with a stack.

Type: string

timestamp

The time when the status was updated.

Type: datetime

nextToken

A token that can be used to retrieve the next set of results, or `null` if there are no additional results.

Type: string

Example

Python

Request

```
$ get_image_stack_events(image_id)
```

200 Response

```
{
  'events': [
    {
```

```
      'event_id': 'ParallelClusterImage-
CREATE_IN_PROGRESS-2022-03-30T23:26:33.499Z',
      'logical_resource_id': 'ParallelClusterImage',
      'physical_resource_id': 'arn:aws:imagebuilder:us-east-1:123456789012:image/
parallelclusterimage-alinux2-image/3.2.1/1',
      'resource_properties': {
        "InfrastructureConfigurationArn":"arn:aws:imagebuilder:us-
east-1:123456789012:infrastructure-configuration/parallelclusterimage-6accc570-
b080-11ec-845e-0e2dc6386985",
        "ImageRecipeArn":"arn:aws:imagebuilder:us-east-1:123456789012:image-recipe/
parallelclusterimage-alinux2-image/3.2.1",
        "DistributionConfigurationArn":"arn:aws:imagebuilder:us-
east-1:123456789012:distribution-configuration/parallelclusterimage-6accc570-
b080-11ec-845e-0e2dc6386985",
        "EnhancedImageMetadataEnabled":"false",
        "Tags": {
          "parallelcluster:image_name":"alinux2-
image","parallelcluster:image_id":"alinux2-image"
        }
      },
      'resource_status': 'CREATE_IN_PROGRESS',
      'resource_status_reason': 'Resource creation Initiated',
      'resource_type': 'AWS::ImageBuilder::Image',
      'stack_id': 'arn:aws:cloudformation:us-east-1:123456789012:stack/alinux2-
image/6accc570-b080-11ec-845e-0e2dc6386985',
      'stack_name': 'alinux2-image',
      'timestamp': datetime.datetime(2022, 3, 30, 23, 26, 33, 499000,
 tzinfo=tzlocal())
    },
    ...
  ]
}
```

listClusters

Retrieve a list of existing clusters.

Topics

- Request syntax
- Request body
- Response syntax

Request syntax

```
GET /v3/clusters
{
  "clusterStatus": "string",
  "nextToken": "string",
  "region": "string"
}
```

Request body

clusterStatus

Filter by cluster status. The default is all clusters.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | UPDATE_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_FAILED`

Required: no

nextToken

The token for the next set of results.

Type: string

Required: No

region

The AWS Region of the clusters.

Type: string

Required: No

Response syntax

```
{
  "nextToken": "string",
  "clusters": [
    {
      "clusterName": "string",
      "region": "string",
      "version": "string",
      "cloudformationStackArn": "string",
      "cloudformationStackStatus": "CREATE_IN_PROGRESS",
      "clusterStatus": "CREATE_IN_PROGRESS",
      "scheduler": {
        "type": "string",
        "metadata": {
          "name": "string",
          "version": "string"
        }
      }
    }
  ]
}
```

Response body

clusters

cloudformationStackArn

The Amazon Resource Name (ARN) of the main CloudFormation stack.

Type: string

cloudformationStackStatus

The CloudFormation stack status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | ROLLBACK_IN_PROGRESS | ROLLBACK_FAILED | ROLLBACK_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE_CLEANUP_IN_PROGRESS`

`| UPDATE_COMPLETE | UPDATE_ROLLBACK_IN_PROGRESS | UPDATE_ROLLBACK_FAILED | UPDATE_ROLLBACK_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_ROLLBACK_COMPLETE`

clusterName

The name of the cluster.

Type: string

clusterStatus

The cluster status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_FAILED`

scheduler

metadata

The scheduler metadata.

name

The name of the scheduler.

Type: string

version

The scheduler version.

Type: string

type

The type of scheduler.

Type: string

region

The AWS Region in which the cluster is created.

Type: string

version

The AWS ParallelCluster version that's used to create the cluster.

Type: string

nextToken

A token that can be used to retrieve the next set of results, or `null` if there are no additional results.

Type: string

Example

Python

Request

```
$ list_clusters()
```

200 Response

```
{
  'clusters':
   [
      {
        'cloudformation_stack_arn': 'arn:aws:cloudformation:us-
east-1:123456789012:stack/cluster_name_3x/16b49540-aee5-11ec-8e18-0ac1d712b241',
        'cloudformation_stack_status': 'CREATE_COMPLETE',
        'cluster_name': 'cluster_name_3x',
        'cluster_status': 'CREATE_COMPLETE',
        'region': 'us-east-1',
        'version': '3.2.1'
      },
      ...
   ]
}
```

listClusterLogStreams

Retrieve the list of log streams that are associated with a cluster.

Topics

Request syntax

```
GET /v3/clusters/{clusterName}/logstreams
{
  "filters": [ "string" ],
  "nextToken": "string",
  "region": "string"
}
```

Request body

clusterName

The name of the cluster.

Type: string

Required: Yes

filters

Filter the log streams.

Accepted filters are:

- private-dns-name: The short form of the private DNS name of the instance (e.g. ip-10-0-0-101).
- node-type: Valid value: `HeadNode`.

Type: Array of strings unique

Format: `Name=a,Values=1 Name=b,Values=2,3`

Required: No

nextToken

The token for the next set of results.

Type: string

Required: No

region

The AWS Region that the cluster is in.

Type: string

Required: No

Response syntax

```
{
  "nextToken": "string",
  "logStreams": [
    {
      "logStreamName": "string",
      "creationTime": "2019-08-24T14:15:22Z",
      "firstEventTimestamp": "2019-08-24T14:15:22Z",
      "lastEventTimestamp": "2019-08-24T14:15:22Z",
      "lastIngestionTime": "2019-08-24T14:15:22Z",
      "uploadSequenceToken": "string",
      "logStreamArn": "string"
    }
  ]
}
```

Response body

logStreams

A list of log streams.

creationTime

The time when the stream was created.

Type: datetime

firstEventTimestamp

The time of the first event of the stream.

Type: datetime

lastEventTimestamp

The time of the last event of the stream. The lastEventTime value updates on an eventual consistency basis. It typically updates in less than an hour from ingestion, but in rare situations might take longer.

Type: datetime

lastIngestionTime

The last ingestion time.

Type: datetime

logStreamArn

The Amazon Resource Name (ARN) of the log stream.

Type: string

logStreamName

Name of the log stream.

Type: string

uploadSequenceToken

The sequence token.

Type: string

nextToken

A token that can be used to retrieve the next set of results, or `null` if there are no additional results.

Type: string

Example

Python

Request

```
$ list_cluster_log_streams(cluster_name_3x)
```

200 Response

```
{
  'log_streams': [
      {
        'creation_time': datetime.datetime(2022, 3, 30, 14, 7, 34, 354000,
 tzinfo=tzlocal()),
        'first_event_timestamp': datetime.datetime(2022, 3, 30, 14, 6, 41, 444000,
 tzinfo=tzlocal()),
        'last_event_timestamp': datetime.datetime(2022, 3, 30, 14, 25, 55, 462000,
 tzinfo=tzlocal()),
        'last_ingestion_time': datetime.datetime(2022, 3, 30, 14, 49, 50, 62000,
 tzinfo=tzlocal()),
        'log_stream_arn': 'arn:aws:logs:us-east-1:123456789012:log-group:/aws/
parallelcluster/cluster_name_3x:log-stream:ip-192-0-2-26.i-abcdef01234567890.cfn-
init',
        'log_stream_name': 'ip-192-0-2-26.i-abcdef01234567890.cfn-init',
         ...
        'upload_sequence_token': '####'
      },
      ...
   ]
}
```

listImageLogStreams

Retrieve the list of log streams that's associated with an image.

Topics

- Request syntax
- Request body
- Response syntax

Request syntax

```
GET /v3/images/custom/{imageId}/logstreams
{
  "nextToken": "string",
  "region": "string"
}
```

Request body

imageId

The ID of the image.

Type: string

Required: Yes

nextToken

The token for the next set of results.

Type: string

Required: No

region

The AWS Region that the image is in.

Type: string

Required: No

Response syntax

```
{
  "nextToken": "string",
  "logStreams": [
```

```
    {
      "logStreamName": "string",
      "creationTime": "2019-08-24T14:15:22Z",
      "firstEventTimestamp": "2019-08-24T14:15:22Z",
      "lastEventTimestamp": "2019-08-24T14:15:22Z",
      "lastIngestionTime": "2019-08-24T14:15:22Z",
      "uploadSequenceToken": "string",
      "logStreamArn": "string"
    }
  ]
}
```

Response body

logStreams

A list of log streams.

creationTime

The time when the stream was created.

Type: datetime

firstEventTimestamp

The time of the first event in the stream.

Type: datetime

lastEventTimestamp

The time of the last event of the stream. The lastEventTime value updates on an eventual consistency basis. It typically updates in less than an hour from ingestion, but in rare situations might take longer.

Type: datetime

lastIngestionTime

The last ingestion time.

Type: datetime

logStreamArn

The Amazon Resource Name (ARN) of the log stream.

Type: string

logStreamName

The name of the log stream.

Type: string

uploadSequenceToken

The sequence token.

Type: string

next_token

A token that can be used to retrieve the next set of results, or `null` if there are no additional results.

Type: string

Example

Python

Request

```
$ list_image_log_streams(custom-image-id)
```

200 Response

```
{
  'log_streams': [
    {
      'creation_time': datetime.datetime(2022, 3, 29, 20, 29, 24, 875000,
 tzinfo=tzlocal()),
      'first_event_timestamp': datetime.datetime(2022, 3, 29, 20, 29, 24, 775000,
 tzinfo=tzlocal()),
      'last_event_timestamp': datetime.datetime(2022, 3, 29, 20, 38, 23, 944000,
 tzinfo=tzlocal()),
      'last_ingestion_time': datetime.datetime(2022, 3, 29, 20, 51, 56, 26000,
 tzinfo=tzlocal()),
      'log_stream_arn': 'arn:aws:logs:us-east-1:123456789012:log-group:/aws/
imagebuilder/ParallelClusterImage-alinux2-image:log-stream:3.2.1/1',
```

```
      'log_stream_name': '3.2.1/1',
      'upload_sequence_token': '####'
    },
    ...
  ]
}
```

listImages

Retrieve the list of existing custom images.

Topics

Request syntax

```
GET /images/custom
{
  "imageStatus": "string",
  "nextToken": "string",
  "region": "string"
}
```

Request body

imageStatus

Filter images by the status provided.

Type: string

Valid values: `AVAILABLE | PENDING | FAILED`

Required: Yes

nextToken

The token for the next set of results.

Type: string

Required: No

region

The AWS Region that images are in.

Type: string

Required: No

Response syntax

```
{
  "nextToken": "string",
  "images": [
    {
      "imageId": "string",
      "ec2AmiInfo": {
        "amiId": "string"
      },
      "region": "string",
      "version": "string",
      "cloudformationStackArn": "string",
      "imageBuildStatus": "BUILD_IN_PROGRESS",
      "cloudformationStackStatus": "CREATE_IN_PROGRESS"
    }
  ]
}
```

Response body

images

A list of images.

cloudformationStackArn

The Amazon Resource Name (ARN) of the main CloudFormation stack.

Type: string

cloudformationStackStatus

The CloudFormation stack status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | ROLLBACK_IN_PROGRESS | ROLLBACK_FAILED | ROLLBACK_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_ROLLBACK_IN_PROGRESS | UPDATE_ROLLBACK_FAILED | UPDATE_ROLLBACK_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_ROLLBACK_COMPLETE`

ec2AmiInfo

ami_id

The Amazon EC2 AMI ID.

Type: string

imageBuildStatus

The image build status.

Valid values: `BUILD_IN_PROGRESS | BUILD_FAILED | BUILD_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE`

Type: string

imageId

The ID of the image.

Type: string

region

The AWS Region in which the image is created.

Type: string

version

The AWS ParallelCluster version that's used to build the image.

Type: string

nextToken

A token that can be used to retrieve the next set of results, or `null` if there are no additional results.

Type: string

Example

Python

Request

```
$ list_images("AVAILABLE")
```

200 Response

```
{
  'images': [
    {
      'ec2_ami_info': {
        'ami_id': 'ami-abcdef01234567890'
      },
      'image_build_status': 'BUILD_COMPLETE',
      'image_id': 'custom-image',
      'region': 'us-east-1',
      'version': '3.2.1'
    }
  ]
}
```

listOfficialImages

Retrieve the list of AWS ParallelCluster official images.

Topics

- Request syntax
- Request body

Request syntax

```
GET /v3/images/official
{
  "architecture": "string",
  "os": "string",
  "region": "string"
}
```

Request body

architecture

Filter by architecture. The default is no filtering.

Type: string

Valid values: `x86_64 | arm64`

Required: No

os

Filter by OS distribution. The default is no filtering.

Type: string

Valid values: `alinux2 | alinux2023 | ubuntu2404 | ubuntu2204 | ubuntu2004 | rhel8 | rhel9`

Required: No

region

The AWS Region in which official images are listed.

Type: string

Required: No

Response syntax

```
{
  "images": [
    {
      "architecture": "string",
      "amiId": "string",
      "name": "string",
      "os": "string",
      "version": "string"
    }
  ]
}
```

Response body

images

amiId

The ID of the AMI.

Type: string

architecture

The AMI architecture.

Type: string

name

The name of the AMI.

Type: string

os

The AMI operating system.

Type: string

version

The AWS ParallelCluster version.

Type: string

Example

Python

Request

```
$ list_official_images()
```

200 Response

```
{
  'images': [
    {
      'ami_id': 'ami-015cfeb4e0d6306b2',
      'architecture': 'x86_64',
      'name': 'aws-parallelcluster-3.2.1-ubuntu-2004-lts-hvm-x86_64-202202261505 '
      '2022-02-26T15-08-34.759Z',
      'os': 'ubuntu2004',
      'version': '3.2.1'
    },
    ...
  ]
}
```

updateCluster

Update the cluster.

Topics

Request syntax

```
PUT /v3/clusters/{clusterName}
```

```
{
  "clusterConfiguration": "string",
  "dryrun": boolean,
  "forceUpdate": boolean,
  "region": "string",
  "suppressValidators": "string",
  "validationFailureLevel": "string"
}
```

Request body

clusterConfiguration

The cluster configuration as a YAML document.

Required: Yes

clusterName

The name of the cluster.

Type: string

Required: Yes

dryrun

If set to `true`, only perform request validation without creating any resource. Use this parameter to validate the cluster configuration and update requirements. The default is `false`.

Type: boolean

Required: No

forceUpdate

If set to `true`, ignore the update validation errors and force the update. The default is `false`.

Type: boolean

Required: No

region

The AWS Region that the cluster is in.

Type: string

Required: No

suppressValidators

Identifies one or more configuration validators to suppress.

Type: string

Format: `(ALL|type:[A-Za-z0-9]+)`

Required: No

Example valid values: `currentValue`, `requestedValue`, `message`

validationFailureLevel

The minimum validation level to cause the update to fail.

Type: string

Valid values: `INFO | WARNING | ERROR`

Required: No

Response syntax

```
{
  "cluster": {
    "clusterName": "string",
    "region": "string",
    "version": "string",
    "cloudformationStackArn": "string",
    "cloudformationStackStatus": "UPDATE_IN_PROGRESS",
    "clusterStatus": "UPDATE_IN_PROGRESS",
    "scheduler": {
      "type": "string",
      "metadata": {
        "name": "string",
        "version": "string"
      }
    }
  },
```

```
  "validationMessages": [
    {
      "id": "string",
      "type": "string",
      "level": "INFO",
      "message": "string"
    }
  ],
  "changeSet": [
    {
      "parameter": "string",
      "currentValue": "string",
      "requestedValue": "string"
    }
  ]
}
```

Response body

changeSet

The change set for the cluster update.

currentValue

The current value of the parameter to be updated.

Type: string

parameter

The parameter to be updated.

Type: string

requestedValue

The requested value for the parameter to be updated.

Type: string

cluster

cloudformationStackArn

The Amazon Resource Name (ARN) of the main CloudFormation stack.

Type: string

cloudformationStackStatus

The CloudFormation stack status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | ROLLBACK_IN_PROGRESS | ROLLBACK_FAILED | ROLLBACK_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_ROLLBACK_IN_PROGRESS | UPDATE_ROLLBACK_FAILED | UPDATE_ROLLBACK_COMPLETE_CLEANUP_IN_PROGRESS | UPDATE_ROLLBACK_COMPLETE`

clusterName

The name of cluster.

Type: string

clusterStatus

The cluster status.

Type: string

Valid values: `CREATE_IN_PROGRESS | CREATE_FAILED | CREATE_COMPLETE | DELETE_IN_PROGRESS | DELETE_FAILED | DELETE_COMPLETE | UPDATE_IN_PROGRESS | UPDATE_COMPLETE | UPDATE_FAILED`

region

The AWS Region in which the cluster is created.

Type: string

scheduler

metadata

The scheduler metadata.

name

The name of the scheduler.

Type: string

version

The scheduler version.

Type: string

type

The scheduler type.

Type: string

version

AWS ParallelCluster version that's used to create the cluster.

Type: string

validationMessages

A list of messages with a validation level lower than `validationFailureLevel`. The list of messages is collected during configuration validation.

id

The ID of the validator.

Type: string

level

The validation level.

Type: string

Valid values: `INFO | WARNING | ERROR`

message

The validation message.

Type: string

type

The type of the validator.

Type: string

Example

Python

Request

```
$ update_cluster(cluster_name_3x, path/config-file.yaml)
```

200 Response

```
{
  'change_set': [
    {
      'current_value': '10',
      'parameter':
 'Scheduling.SlurmQueues[queue1].ComputeResources[t2micro].MaxCount',
      'requested_value': '15'
    }
  ],
  'cluster': {
    'cloudformation_stack_arn': 'arn:aws:cloudformation:us-
east-1:123456789012:stack/test-api-cluster/e0462730-50b5-11ed-99a3-0a5ddc4a34c7',
    'cloudformation_stack_status': 'UPDATE_IN_PROGRESS',
    'cluster_name': 'cluster-3x',
    'cluster_status': 'UPDATE_IN_PROGRESS',
    'region': 'us-east-1',
    'scheduler': {
      'type': 'slurm'
    },
    'version': '3.2.1'
  }
}
```

updateComputeFleet

Update the status of the cluster compute fleet.

Topics

Request syntax

```
PATCH /v3/clusters/{clusterName}/computefleet
{
  "status": "string",
  "region": "string"
}
```

Request body

clusterName

The name of the cluster.

Type: string

Required: Yes

status

The compute fleet status.

Type: string

Valid values: `START_REQUESTED | STOP_REQUESTED | ENABLED | DISABLED`

Required: Yes

region

The AWS Region that the cluster is in.

Type: string

Required: No

Response syntax

```
{
  "status": "START_REQUESTED",
  "lastStatusUpdatedTime": "2019-08-24T14:15:22Z"
}
```

Response body

status

The compute fleet status.

Type: string

Valid values: `START_REQUESTED | STARTING | RUNNING | PROTECTED | STOP_REQUESTED | STOPPING | STOPPED | UNKNOWN | ENABLED | DISABLED`

lastStatusUpdatedTime

The timestamp that represents the last status update time.

Type: datetime

Example

Python

Request

```
$ update_compute_fleet(cluster_name_3x,  "START_REQUESTED")
```

200 Response

```
{
  'last_status_updated_time': datetime.datetime(2022, 3, 28, 22, 27, 14,
 tzinfo=tzlocal()),
```

```
  'status': 'START_REQUESTED'
}
```

AWS ParallelCluster Python library API

Starting with AWS ParallelCluster version 3.5.0, you can access AWS ParallelCluster with the AWS ParallelCluster Python library. You can access the AWS ParallelCluster library in your `pcluster` environment or from within an AWS Lambda runtime. Learn how to access the AWS ParallelCluster API by using the AWS ParallelCluster Python library. The AWS ParallelCluster Python library offers the same functionality that the AWS ParallelCluster API delivers.

The AWS ParallelCluster Python library operations and parameters mirror those of the API parameters when converted to `snake_case` with no capital letters.

Topics

AWS ParallelCluster Python library authorization

Specify credentials by using any of the standard ways that are valid for boto3. For more information, see the boto3 documentation.

Install the AWS ParallelCluster Python library

1. Install `pcluster` CLI version 3.5.0 or later by following the instructions given in Setting up AWS ParallelCluster.

2. Import the `pcluster` module and start using the library, as shown in the following example:

```
import pcluster.lib as pc
pc.create_cluster(cluster_name="mycluster", cluster_configuration="config.yaml"
```

Cluster API operations

Topics

`list_clusters`

```
list_clusters(region, next_token, cluster_status)
```

Retrieve the list of existing clusters.

Parameters:

`region`

Lists clusters deployed to a given AWS Region.

`next_token`

The token for the next set of results.

`cluster_status`

Filters by cluster status. The default is to list all clusters.

Valid values: `CREATE_IN_PROGRESS` | `CREATE_FAILED` | `CREATE_COMPLETE` | `DELETE_IN_PROGRESS` | `DELETE_FAILED` | `UPDATE_IN_PROGRESS` | `UPDATE_COMPLETE` | `UPDATE_FAILED`

create_cluster

```
create_cluster(cluster_name, cluster_configuration, region, suppress_validators,
 validation_failure_level, dry_run, rollback_on_failure, wait)
```

Create a cluster in a given Region.

Parameters:

cluster_name (required)

The cluster name.

cluster_configuration (required)

The cluster configuration as a Python data type.

region

The cluster AWS Region.

suppress_validators

Identifies one or more cluster configuration validators to suppress.

Format: `(ALL | type:[A-Za-z0-9]+)`

validation_failure_level

The minimum validation level that causes the cluster creation to fail. The default is `ERROR`.

Valid values: `INFO` | `WARNING` | `ERROR`.

dry_run

Performs the request validation without creating any resources. You can use this to validate the cluster configuration. The default is `False`.

rollback_on_failure

If set to `True`, AWS ParallelCluster automatically initiates a cluster stack rollback on failures. The default is `True`.

wait

If set to `True`, AWS ParallelCluster waits for the operation to complete. The default is `False`.

delete_cluster

```
delete_cluster(cluster_name, region, wait)
```

Delete a cluster in a given Region.

Parameters:

cluster_name (required)

The cluster name.

region

The cluster AWS Region.

wait

If set to `True`, waits for the operation to complete. The default is `False`.

describe_cluster

```
describe_cluster(cluster_name, region)
```

Get detailed information about an existing cluster.

Parameters:

cluster_name (required)

The cluster name.

region

The cluster AWS Region.

update_cluster

```
update_cluster(cluster_name, cluster_configuration, suppress_validators,
 validation_failure_level, region, force_update, dry_run, wait)
```

Update a cluster in a given Region.

Parameters:

`cluster_name` (required)

The cluster name.

`cluster_configuration` (required)

The cluster configuration as a Python data type.

`suppress_validators`

Identifies one or more cluster configuration validators to suppress.

Format: `(ALL | type:[A-Za-z0-9]+)`

`validation_failure_level`

The minimum validation level that causes the cluster update to fail. The default is `ERROR`.

Valid values: `INFO` | `WARNING` | `ERROR`

`region`

The cluster AWS Region.

`dry_run`

Performs the request validation without creating or updating any resources. You can use this to validate the cluster configuration. The default is `False`.

`force_update`

If set to `True`, forces the update by ignoring the update validation errors. The default is `False`.

`wait`

If set to `True`, waits for the operation to complete. The default is `False`.

Compute fleet API operations

Topics

- describe_compute_fleet
- update_compute_fleet
- delete_cluster_instances
- describe_cluster_instances

describe_compute_fleet

```
describe_compute_fleet(cluster_name, region)
```

Describe the status of a cluster compute fleet for a given cluster.

Parameters:

cluster_name (required)

The cluster name.

region

Describes the compute fleet status for a cluster deployed to a given AWS Region.

update_compute_fleet

```
update_compute_fleet(cluster_name, status, region)
```

Update the status of the cluster compute fleet.

Parameters:

cluster_name (required)

The cluster name.

status (required)

The status to update to.

Valid values: START_REQUESTED | STOP_REQUESTED | ENABLED | DISABLED

`region`

The cluster AWS Region.

`delete_cluster_instances`

```
delete_cluster_instances(cluster_name, region, force)
```

Initiate the forced termination of all cluster compute nodes. This action doesn't support AWS Batch clusters.

Parameters:

`cluster_name` (required)

The cluster name.

`region`

The cluster AWS Region.

`force`

If set to `True`, forces deletion when the cluster with the given `cluster_name` isn't found. The default is `False`.

`describe_cluster_instances`

```
describe_cluster_instances(cluster_name, region, next_token, node_type, queue_name)
```

Describe a cluster's instances.

Parameters:

`cluster_name` (required)

The cluster name.

`region`

The cluster AWS Region.

`next_token`

The token for the next set of results.

`node_type`

Filters the instances by `node_type`.

Valid values: `HeadNode` | `ComputeNode`

`queue_name`

Filters the instances by queue name.

Cluster and stack log operations

Topics

- list_cluster_log_streams
- get_cluster_log_events
- get_cluster_stack_events

`list_cluster_log_streams`

```
list_cluster_log_streams(cluster_name, region, filters, next_token)
```

List log streams for a given cluster.

Parameters:

`cluster_name` (required)

The cluster name.

`region`

The cluster AWS Region.

`filters`

Filters the cluster log streams.

Format: `'Name=a,Values=1 Name=b,Values=2,3'`

Accepted filters:

code-dns-name

The short form of the private DNS name of the instance; for example, `ip-10-0-0-101`.

node-type

The node type.

Valid values: `HeadNode`

`next_token`

The token for the next set of results.

`get_cluster_log_events`

```
get_cluster_log_events(cluster_name, log_stream_name, region, next_token,
 start_from_head, limit, start_time, end_time)
```

Get log events for a given cluster and log stream.

Parameters:

`cluster_name` (required)

The cluster name.

`log_stream_name` (required)

The log stream name.

`region`

The cluster AWS Region.

`next_token`

The token for the next set of results.

`start_from_head`

If set to `True`, AWS ParallelCluster returns the earliest log events first. If set to `False`, it returns the latest log events first. The default is `False`.

`limit`

The maximum number of log events returned. If you don't specify a value, the maximum is the number of logs that can fit in a response size of 1 MB, up to 10,000 log events.

`start_time`

The start of the time range for log events, expressed in ISO 8601 format; for example, `'2021-01-01T20:00:00Z'`. Events with a timestamp equal to, or later than, this time are included.

`end_time`

The end of the time range for log events, expressed in ISO 8601 format; for example, `'2021-01-01T20:00:00Z'`. Events with a timestamp equal to, or later than, this time are not included.

`get_cluster_stack_events`

```
get_cluster_stack_events(cluster_name, region, next_token)
```

Get stack events for a given cluster.

Parameters:

`cluster_name` (required)

The cluster name.

`region`

The cluster AWS Region.

`next_token`

The token for the next set of results.

Image API operations

Topics

- delete_image
- describe_image

list_images

```
list_images(image_status, region, next_token)
```

Retrieve the list of existing images.

Parameters:

image_status (required)

Filters by image status.

Valid values: AVAILABLE | PENDING | FAILED

region

Lists images built in a given AWS Region.

next_token

The token for the next set of results.

build_image

```
build_image(image_configuration, image_id, suppress_validators,
 validation_failure_level, dry_run, rollback_on_failure, region)
```

Create a custom AWS ParallelCluster image in a given Region.

Parameters:

image_configuration (required)

The image configuration as Python data.

image_id (required)

The image ID.

`suppress_validators`

Identifies one or more image configuration validators to suppress.

Format: `(ALL | type:[A-Za-z0-9]+)`

`validation_failure_level`

The minimum validation level that causes the image creation to fail. The default is `ERROR`.

Valid values: `INFO` | `WARNING` | `ERROR`

`dry_run`

If set to `True`, AWS ParallelCluster performs the request validation without creating any resources. You can use this to validate the image configuration. The default is `False`.

`rollback_on_failure`

If set to `True`, AWS ParallelCluster automatically initiates an image stack rollback on failures. The default is `False`.

`region`

The image AWS Region.

`delete_image`

```
delete_image(image_id, region, force)
```

Delete an image in a given Region.

Parameters:

`image_id` (required)

The image ID.

`region`

The image AWS Region.

`force`

If set to `True`, AWS ParallelCluster forces deletion if instances are using the AMI or if the AMI is shared. The default is `False`.

`describe_image`

```
describe_image(image_id, region)
```

Get detailed information about an existing image.

Parameters:

`image_id` (required)

The image ID.

`region`

The image AWS Region.

Image and stack log operations

Topics

- list_image_log_streams
- get_image_log_events
- get_image_stack_events
- list_official_images

`list_image_log_streams`

```
list_image_log_streams(image_id, region, next_token)
```

List log streams for an image.

Parameters:

`image_id` (required)

The image ID.

`region`

The image AWS Region.

next_token

The token for the next set of results.

get_image_log_events

```
get_image_log_events(image_id, log_stream_name, region, next_token, start_from_head,
 limit, start_time, end_time)
```

Get log events for a given image and log stream.

Parameters:

image_id (required)

The image ID.

log_stream_name (required)

The log stream name.

region

The image AWS Region.

next_token

The token for the next set of results.

start_from_head

If set to `True`, AWS ParallelCluster returns the earliest log events first. If set to `False`, it returns the latest log events first. The default is `False`.

limit

The maximum number of log events returned. If you don't specify a value, the maximum is the number of logs that can fit in a response size of 1 MB, up to 10,000 log events.

start_time

The start of the time range for log events, expressed in ISO 8601 format; for example, `'2021-01-01T20:00:00Z'`. Events with a timestamp equal to, or later than, this time are included.

`end_time`

The end of the time range for log events, expressed in ISO 8601 format; for example, `'2021-01-01T20:00:00Z'`. Events with a timestamp equal to, or later than, this time are not included.

`get_image_stack_events`

```
get_image_stack_events(image_id, region, next_token)
```

Get stack events for a given image.

Parameters:

`image_id` (required)

The image ID.

`region`

The image AWS Region.

`next_token`

The token for the next set of results.

`list_official_images`

```
list_official_images(region,os, architecture)
```

Retrieve the list of official AWS ParallelCluster images.

Parameters:

`region`

The image AWS Region.

`os`

Filters by operating system distribution. The default is no filtering.

`architecture`

Filters by architecture. The default is no filtering.

Example

Topics

- Create a cluster

Create a cluster

When you run the following example script, with the given inputs stored in your environment, you create a cluster. The cluster configuration is created as a Python data type based on the cluster configuration documentation.

```
import os
import pprint
import pcluster.lib as pc
pp = pprint.PrettyPrinter()

HEAD_NODE_SUBNET = os.environ["HEAD_NODE_SUBNET"]
COMPUTE_NODE_SUBNET = os.environ["HEAD_NODE_SUBNET"]
KEY_NAME = os.environ["KEY_NAME"]
CONFIG = {'Image': {'Os': 'alinux2'},
          'HeadNode': {'InstanceType': 't2.large',
                       'Networking': {'SubnetId': HEAD_NODE_SUBNET},
                       'Ssh': {'KeyName': KEY_NAME}},

          'Scheduling': {'Scheduler': 'slurm',
                         'SlurmQueues':
                         [{'Name': 'queue0',
                           'ComputeResources':
                           [{'Name': 'queue0-i0', 'InstanceType': 't2.micro',
                             'MinCount': 0, 'MaxCount': 10}],
                           'Networking': {'SubnetIds': [COMPUTE_NODE_SUBNET]}}]}}


pp.pprint(pc.create_cluster(cluster_name="mycluster", cluster_configuration=CONFIG))
```

Output:

```
{'cluster': {'cloudformationStackArn': 'arn:aws:cloudformation:us-
east-2:123456789012:stack/mycluster/00000000-aaaa-1111-999-000000000000',
             'cloudformationStackStatus': 'CREATE_IN_PROGRESS',
             'clusterName': 'mycluster',
             'clusterStatus': 'CREATE_IN_PROGRESS',
             'region': 'us-east-2',
             'scheduler': {'type': 'slurm'},
             'version': '3.13.2'}}
```

AWS Lambda for the AWS ParallelCluster Python library

You can deploy a Lambda layer and runtime to access to the AWS ParallelCluster Python library. We host AWS ParallelCluster zip files that you can use by entering the link to the zip file as described in the following steps. Lambda uses the zip files to prepare the runtime environment to support access to the Python library. The AWS ParallelCluster Python library is added with AWS ParallelCluster version 3.5.0. You can only use the library for versions 3.5.0 and later.

The hosted zip file URL is in the format: `s3://`*`aws-region-id`*`-aws-parallelcluster/parallelcluster/`*`3.13.2`*`/layers/aws-parallelcluster/lambda-layer.zip`. (Replace *`3.13.2`* with the AWS ParallelCluster version you want to use in the following step.)

Get started accessing the AWS ParallelCluster Python library with AWS Lambda

Create a Lambda layer

1. Log in to the AWS Management Console and navigate to the AWS Lambda console.
2. In the navigation pane, select **Layers**, then **Create layer**.
3. Enter a name for your layer and select **Upload a file from Amazon S3**.
4. Enter the URL to the zip file: s3://*aws-region-id*-aws-parallelcluster/parallelcluster/*3.13.2*/layers/aws-parallelcluster/lambda-layer.zip.
5. For **Compatible architectures**, choose the **x86_64** architecture.
6. For **Compatible runtimes**, choose the **Python 3.12** runtime.
7. Choose **Create**.

Use your Lambda layer

1. In the Lambda console navigation pane, select **Functions**, then **Create function**.

2. Enter a name for your function.
3. For **Runtime**, choose the **Python 3.12** runtime.
4. For **Architecture**, choose the **x86_64** architecture.
5. Choose **Create function**.
6. After the function is created, choose **Layers** and select **Add a layer**.
7. Select **Custom layers** and choose the layer you created in previous steps.
8. Choose the layer version.
9. Choose **Add**.
10. Your Lambda needs permissions to manage clusters created with AWS ParallelCluster. Create a Lambda role with the permissions listed in Base AWS ParallelCluster `pcluster` user policy.

You can now access AWS ParallelCluster from the Python library as described in AWS ParallelCluster Python library API.

Tutorials on how to use AWS ParallelCluster

The following tutorials show you how to get started with AWS ParallelCluster version 3, and provide best practice guidance for some common tasks.

When using the AWS ParallelCluster command line interface (CLI) or API, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

Topics

Running your first job on AWS ParallelCluster

This tutorial walks you through running your first Hello World job on AWS ParallelCluster

When using the AWS ParallelCluster command line interface (CLI) or API, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

Prerequisites

- AWS ParallelCluster is installed.
- The AWS CLI is installed and configured.
- You have an Amazon EC2 key pair.
- You have an IAM role with the permissions required to run the `pcluster` CLI.

Verifying your installation

First, we verify that AWS ParallelCluster is correctly, including the Node.js dependency, installed and configured.

```
$ node --version
v16.8.0
$ pcluster version
{
  "version": "3.13.2"
}
```

This returns the running version of AWS ParallelCluster.

Creating your first cluster

Now it's time to create your first cluster. Because the workload for this tutorial isn't performance intensive, we can use the default instance size of `t2.micro`. (For production workloads, you choose an instance size that best fits your needs.) Let's call your cluster `hello-world`.

```
$ pcluster create-cluster \
    --cluster-name hello-world \
    --cluster-configuration hello-world.yaml
```

> **Note**
>
> The AWS Region to use must be specified for most `pcluster` commands. If it's not specified in the `AWS_DEFAULT_REGION` environment variable, or the `region` setting in the `[default]` section of the `~/.aws/config` file, then the `--region` parameter must be provided on the `pcluster` command line.

If the output gives you a message about configuration, you need to run the following to configure AWS ParallelCluster:

```
$ pcluster configure --config hello-world.yaml
```

If the pcluster create-cluster command succeeds, you see output similar to the following:

```
{
  "cluster": {
    "clusterName": "hello-world",
    "cloudformationStackStatus": "CREATE_IN_PROGRESS",
    "cloudformationStackArn": "arn:aws:cloudformation:xxx:stack/xxx",
    "region": "...",
    "version": "...",
    "clusterStatus": "CREATE_IN_PROGRESS"
  }
}
```

You monitor the creation of the cluster using:

```
$ pcluster describe-cluster --cluster-name hello-world
```

The `clusterStatus` reports "CREATE_IN_PROGRESS" while the cluster is being created. The `clusterStatus` transitions to "CREATE_COMPLETE" when the cluster is created successfully. The output also provides us with the `publicIpAddress` and `privateIpAddress` of our head node.

Logging into your head node

Use your OpenSSH pem file to log into your head node.

```
$ pcluster ssh --cluster-name hello-world -i /path/to/keyfile.pem
```

After you log in, run the command `sinfo` to verify that your compute nodes are set up and configured.

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
queue1*      up   infinite     10  idle~ queue1-dy-queue1t2micro-[1-10]
```

The output shows that we have one queue in our cluster, with up to ten nodes.

Running your first job using Slurm

Next, we create a job that sleeps for a little while and then outputs its own hostname. Create a file called `hellojob.sh`, with the following contents.

```
#!/bin/bash
sleep 30
echo "Hello World from $(hostname)"
```

Next, submit the job using `sbatch`, and verify that it runs.

```
$ sbatch hellojob.sh
Submitted batch job 2
```

Now, you can view your queue and check the status of the job. The provisioning of a new Amazon EC2 instance is started in the background. You can monitor the status of the cluster instances with the `sinfo` command.

```
$ squeue
             JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
                 2    queue1 hellojob ec2-user CF       3:30      1 queue1-dy-
queue1t2micro-1
```

The output shows that the job has been submitted to `queue1`. Wait 30 seconds for the job to finish, and then run `squeue` again.

```
$ squeue
             JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
```

Now that there are no jobs in the queue, we can check for output in our current directory.

```
$ ls -l
total 8
-rw-rw-r-- 1 ec2-user ec2-user 57 Sep  1 14:25 hellojob.sh
-rw-rw-r-- 1 ec2-user ec2-user 43 Sep  1 14:30 slurm-2.out
```

In the output, we see a "`out`" file. We can see output from our job:

```
$ cat slurm-2.out
```

```
Hello World from queue1-dy-queue1t2micro-1
```

The output also shows that our job ran successfully on instance `queue1-dy-queue1t2micro-1`.

In the cluster you just created, only the home directory is shared among all nodes of the cluster.

To learn more about creating and using clusters, see Best practices.

If your application requires shared software, libraries, or data, consider the following options:

- Build a AWS ParallelCluster enabled custom AMI that includes your software as described in Building a custom AWS ParallelCluster AMI.
- Use the StorageSettings option in the AWS ParallelCluster configuration file to specify a shared filesystem and store your installed software in the specified mount location.
- Use Custom bootstrap actions to automate the bootstrap procedure of each node of your cluster.

Building a custom AWS ParallelCluster AMI

When using the AWS ParallelCluster command line interface (CLI) or API, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

> **Important**
>
> If you build a custom AMI, you must repeat the steps that you used to create your custom AMI with each new AWS ParallelCluster release.

Before reading further, we recommend that you first review the Custom bootstrap actions section. Determine if the modifications that you want to make can be scripted and supported with future AWS ParallelCluster releases.

Even though building a custom AMI in general isn't ideal, there are specific scenarios where building a custom AMI for AWS ParallelCluster is necessary. This tutorial covers how to build a custom AMI for these scenarios.

Prerequisites

- AWS ParallelCluster is installed.

- The AWS CLI is installed and configured.
- You have an Amazon EC2 key pair.
- You have an IAM role with the permissions required to run the `pcluster` CLI and build images.

How to customize the AWS ParallelCluster AMI

There are two ways to build a custom AWS ParallelCluster AMI. One of these two methods is to build a new AMI using the AWS ParallelCluster CLI. Another method requires that you to make manual modifications to build a new AMI that's available under your AWS account.

Build a custom AWS ParallelCluster AMI

If you have a customized AMI and software, you can apply the changes that are needed by AWS ParallelCluster on top of it. AWS ParallelCluster relies on the EC2 Image Builder service to build customized AMIs. For more information, see the Image Builder User Guide.

Key points:

- The process takes about 1 hour. This time can vary if there are additional `Build` / `Components` to be installed at build time.
- The AMI is tagged with the versions of the main components. These include the kernel, scheduler, and EFA driver. A subset of the component versions are also reported in the AMI description.
- Starting from AWS ParallelCluster 3.0.0, a new set of CLI commands can be used to manage the lifecycle of images. This includes `build-image`, `list-images`, `describe-image`, and `delete-image`.
- This method is repeatable. You can re-run it to keep AMIs updated (for example, OS updates), and then use them when you update an existing cluster.

> **Note**
>
> If you use this method in the AWS China Partition, you might get network errors. For example, you might see these errors from the `pcluster build-image` command when it downloads packages from GitHub or from an OS repository. If this happens, we recommend that you use one of the following alternative methods:
>
> 1. Follow the Modify an AWS ParallelCluster AMI approach that bypasses this command.

2. Build the image in another Partition and Region, such as `us-east-1`, and then store/restore it to move it to the China Region. For more information, see Store and restore an AMI using S3 in the *Amazon EC2 User Guide*.

Steps:

1. Configure your AWS account credentials so that the AWS ParallelCluster client can make calls to AWS API operations on your behalf. For a list of the required permissions, see AWS Identity and Access Management permissions in AWS ParallelCluster.
2. Create a basic *build image* configuration file. To do this, specify the `InstanceType` to be used to build the image and the `ParentImage`. These are used as the starting point to create the AMI. For more information about optional build parameters, see Image Configuration.

```
Build:
 InstanceType: <BUILD_INSTANCE_TYPE>
 ParentImage: <BASE_AMI_ID>
```

3. Use the CLI command `pcluster build-image` to build an AWS ParallelCluster AMI starting from the AMI that you provide as the base.

```
$ pcluster build-image --image-id IMAGE_ID --image-configuration IMAGE_CONFIG.yaml --
region REGION
    {
 "image": {
   "imageId": "IMAGE_ID",
   "imageBuildStatus": "BUILD_IN_PROGRESS",
   "cloudformationStackStatus": "CREATE_IN_PROGRESS",
   "cloudformationStackArn": "arn:aws:cloudformation:us-east-1:123456789012:stack/
IMAGE_ID/abcd1234-ef56-gh78-ij90-1234abcd5678",
   "region": "us-east-1",
   "version": "3.13.2"
 }
}
```

> **Warning**
>
> `pcluster build-image` uses the default VPC. If you delete the default VPC using AWS Control Tower or AWS Landing Zone, the subnet ID must be specified in the image configuration file. For more information, see `SubnetId`.

For a list of other parameters, see the `pcluster build-image` command reference page. The results of the preceding command are as follows:

- A CloudFormation stack is created based on the image configuration. The stack includes all of the EC2 Image Builder resources required for the build.
- The created resources include the official Image Builder AWS ParallelCluster components that custom Image Builder components can be added to. For more information, see Create a custom component with Image Builder in the *EC2 Image Builder User Guide*.
- EC2 Image Builder launches a build instance, applies the AWS ParallelCluster cookbook, installs the AWS ParallelCluster software stack, and performs necessary configuration tasks. The AWS ParallelCluster cookbook is used to build and bootstrap AWS ParallelCluster.
- The instance is stopped and a new AMI is created from it.
- Another instance is launched from the newly created AMI. During the test phase, EC2 Image Builder runs tests that are defined in the Image Builder components.
- If the build is successful, the stack is deleted. If the build fails, the stack is retained and available for inspection.

4. You can monitor the status of the build process by running the following command. After the build completes, you can run it to retrieve the AMI ID given in the response.

```
$ pcluster describe-image --image-id IMAGE_ID --region REGION

# BEFORE COMPLETE
{
 "imageConfiguration": {
   "url": "https://parallelcluster-1234abcd5678efgh-v1-do-not-
delete.s3.amazonaws.com/parallelcluster/3.13.2/images/IMAGE_ID-abcd1234efgh5678/
configs/image-config.yaml?...",
 },
 "imageId": "IMAGE_ID",
 "imagebuilderImageStatus": "BUILDING",
 "imageBuildStatus": "BUILD_IN_PROGRESS",
 "cloudformationStackStatus": "CREATE_IN_PROGRESS",
```

```
 "cloudformationStackArn": "arn:aws:cloudformation:us-east-1:123456789012:stack/
IMAGE_ID/abcd1234-ef56-gh78-ij90-1234abcd5678",
 "region": "us-east-1",
 "version": "3.13.2",
 "cloudformationStackTags": [
   {
     "value": "3.13.2",
     "key": "parallelcluster:version"
   },
   {
     "value": "IMAGE_ID",
     "key": "parallelcluster:image_name"
   },
   ...
 ],
 "imageBuildLogsArn": "arn:aws:logs:us-east-1:123456789012:log-group:/aws/
imagebuilder/ParallelClusterImage-IMAGE_ID",
 "cloudformationStackCreationTime": "2022-04-05T21:36:26.176Z"
}

# AFTER COMPLETE
{
 "imageConfiguration": {
   "url": "https://parallelcluster-1234abcd5678efgh-v1-do-not-delete.s3.us-
east-1.amazonaws.com/parallelcluster/3.13.2/images/IMAGE_ID-abcd1234efgh5678/configs/
image-config.yaml?Signature=..."
 },
 "imageId": "IMAGE_ID",
 "imageBuildStatus": "BUILD_COMPLETE",
 "region": "us-east-1",
 "ec2AmiInfo": {
     "amiName": "IMAGE_ID 2022-04-05T21-39-24.020Z",
     "amiId": "ami-1234stuv5678wxyz",
     "description": "AWS ParallelCluster AMI for alinux2,
 kernel-4.14.238-182.422.amzn2.x86_64, lustre-2.10.8-5.amzn2.x86_64,
 efa-1.13.0-1.amzn2.x86_64, dcv-2021.1.10598-1.el7.x86_64, slurm-20-11-8-1",
     "state": "AVAILABLE",
     "tags": [
      {
        "value": "2021.3.11591-1.el7.x86_64",
        "key": "parallelcluster:dcv_version"
      },
      ...
     ],
```

```
    "architecture": "x86_64"
  },
  "version": "3.13.2"
}
```

5. To create your cluster, enter the AMI ID in the `CustomAmi` field in your cluster configuration.

Troubleshooting and monitoring AMI creation process

Image creation completes in about an hour. You can monitor the process by running the `pcluster describe-image` command or log retrieval commands.

```
$ pcluster describe-image --image-id IMAGE_ID --region REGION
```

The `build-image` command creates a CloudFormation stack with all the Amazon EC2 resources that are required to build the image, and launches the EC2 Image Builder process.

After running the `build-image` command, it's possible to retrieve CloudFormation stack events by using `pcluster get-image-stack-events`. You can filter results with the `--query` parameter to see the latest events. For more information, see Filtering AWS CLI output in the *AWS Command Line Interface User Guide*.

```
$ pcluster get-image-stack-events --image-id IMAGE_ID --region REGION --query
 "events[0]"
{
 "eventId": "ParallelClusterImage-CREATE_IN_PROGRESS-2022-04-05T21:39:24.725Z",
 "physicalResourceId": "arn:aws:imagebuilder:us-east-1:123456789012:image/
parallelclusterimage-IMAGE_ID/3.13.2/1",
 "resourceStatus": "CREATE_IN_PROGRESS",
 "resourceStatusReason": "Resource creation Initiated",
 "resourceProperties": "{\"InfrastructureConfigurationArn\":
\"arn:aws:imagebuilder:us-east-1:123456789012:infrastructure-configuration/
parallelclusterimage-abcd1234-ef56-gh78-ij90-1234abcd5678\",\"ImageRecipeArn\":
\"arn:aws:imagebuilder:us-east-1:123456789012:image-recipe/parallelclusterimage-
IMAGE_ID/3.13.2\",\"DistributionConfigurationArn\":\"arn:aws:imagebuilder:us-
east-1:123456789012:distribution-configuration/parallelclusterimage-abcd1234-ef56-
gh78-ij90-1234abcd5678\",\"Tags\":{\"parallelcluster:image_name\":\"IMAGE_ID\",
\"parallelcluster:image_id\":\"IMAGE_ID\"}}",
 "stackId": "arn:aws:cloudformation:us-east-1:123456789012:stack/IMAGE_ID/abcd1234-
ef56-gh78-ij90-1234abcd5678",
 "stackName": "IMAGE_ID",
 "logicalResourceId": "ParallelClusterImage",
```

```
 "resourceType": "AWS::ImageBuilder::Image",
 "timestamp": "2022-04-05T21:39:24.725Z"
}
```

After about 15 minutes, the stack events appear in the log event entry related to Image Builder creation. You can now list image log streams and monitor the Image Builder steps by using pcluster list-image-log-streams and pcluster get-image-log-events commands.

```
$ pcluster list-image-log-streams --image-id IMAGE_ID --region REGION \
    --query 'logStreams[*].logStreamName'

 "3.13.2/1"
]

$ pcluster get-image-log-events --image-id IMAGE_ID --region REGION \
--log-stream-name 3.13.2/1 --limit 3
{
 "nextToken": "f/36295977202298886557255241372854078762600452615936671762",
 "prevToken": "b/36295977196879805474012299949460899222346900769983430672",
 "events": [
   {
     "message": "ExecuteBash: FINISHED EXECUTION",
     "timestamp": "2022-04-05T22:13:26.633Z"
   },
   {
     "message": "Document arn:aws:imagebuilder:us-east-1:123456789012:component/
parallelclusterimage-test-abcd1234-ef56-gh78-ij90-1234abcd5678/3.13.2/1",
     "timestamp": "2022-04-05T22:13:26.741Z"
   },
   {
     "message": "TOE has completed execution successfully",
     "timestamp": "2022-04-05T22:13:26.819Z"
   }
 ]
}
```

Continue to check with the describe-image command until you see the BUILD_COMPLETE status.

```
$ pcluster describe-image --image-id IMAGE_ID --region REGION
{
 "imageConfiguration": {
```

```
   "url": "https://parallelcluster-1234abcd5678efgh-v1-do-not-delete.s3.us-
east-1.amazonaws.com/parallelcluster/3.13.2/images/IMAGE_ID-abcd1234efgh5678/configs/
image-config.yaml?Signature=..."
 },
 "imageId": "IMAGE_ID",
 "imageBuildStatus": "BUILD_COMPLETE",
 "region": "us-east-1",
 "ec2AmiInfo": {
     "amiName": "IMAGE_ID 2022-04-05T21-39-24.020Z",
     "amiId": "ami-1234stuv5678wxyz",
     "description": "AWS ParallelCluster AMI for alinux2,
 kernel-4.14.238-182.422.amzn2.x86_64, lustre-2.10.8-5.amzn2.x86_64,
 efa-1.13.0-1.amzn2.x86_64, dcv-2021.1.10598-1.el7.x86_64, slurm-20-11-8-1",
     "state": "AVAILABLE",
     "tags": [
      {
        "value": "2021.3.11591-1.el7.x86_64",
        "key": "parallelcluster:dcv_version"
      },
      ...
     ],
   "architecture": "x86_64"
 },
 "version": "3.13.2"
}
```

If you need to troubleshoot a custom AMI creation issue, create an archive of the image logs as described in following steps.

It's possible to archive the logs in an Amazon S3 bucket or in a local file, depending on the `--output` parameter.

```
$ pcluster export-image-logs --image-id IMAGE_ID --region REGION \
--bucket BUCKET_NAME --bucket-prefix BUCKET_FOLDER
{
 "url": "https://BUCKET_NAME.s3.us-east-1.amazonaws.com/BUCKET-FOLDER/IMAGE_ID-
logs-202209071136.tar.gz?AWSAccessKeyId=..."
}

$ pcluster export-image-logs --image-id IMAGE_ID \
--region REGION --bucket BUCKET_NAME --bucket-prefix BUCKET_FOLDER --output-file /tmp/
archive.tar.gz
{
```

```
  "path": "/tmp/archive.tar.gz"
}
```

The archive contains the CloudWatch Logs Streams related to the Image Builder process and AWS CloudFormation stack events. The command might take several minutes to run.

Managing Custom AMIs

Starting from AWS ParallelCluster 3.0.0, a new set of commands has been added in the CLI to build, monitor, and manage the image lifecycle. For more information about the commands, see pcluster commands.

Modify an AWS ParallelCluster AMI

This method consists of modifying an official AWS ParallelCluster AMI by adding customization on top of it. The base AWS ParallelCluster AMIs are updated with new releases. These AMIs have all of the components that are required for AWS ParallelCluster to function when it's installed and configured. You can start with one of these as your base.

Key points:

- This method is faster than the `build-image` command. However, it's a manual process and not automatically repeatable.
- With this method, you don't have access to the log retrieval and image lifecycle management commands that are available through the CLI.

Steps:

New Amazon EC2 console

1. Find the AMI that corresponds to the specific AWS Region that you use. To find it, use the `pcluster list-official-images` command with the `--region` parameter to select the specific AWS Region and `--os` and `--architecture` parameters to filter for the desired AMI with the OS and architecture that you want to use. From the output, retrieve the Amazon EC2 Image ID.
2. Sign in to the AWS Management Console and open the Amazon EC2 console at https://console.aws.amazon.com/ec2/.
3. In the navigation pane, choose **Images**, and then **AMIs**. Search for the retrieved EC2 Image ID, select the AMI, and choose **Launch instance from AMI**.

4. Scroll down and choose your **Instance type**.
5. Choose your **Key pair** and **Launch Instance**.
6. Log in to your instance using the OS user and your SSH key.
7. Manually customize your instance to meet your requirements.
8. Run the following command to prepare your instance for AMI creation.

```
sudo /usr/local/sbin/ami_cleanup.sh
```

9. From the console, choose **Instance state** and **Stop instance**.

   Navigate to **Instances**, choose the new instance, select **Instance state**, and **Stop instance**.

10. Create a new AMI from the instance using the Amazon EC2 console or AWS CLI create-image.

    From the Amazon EC2 console

    a. Choose **Instances** in the navigation pane.

    b. Choose the instance that you created and modified.

    c. In **Actions**, choose **Image** and then **Create image**.

    d. Choose **Create Image**.

11. Enter the new AMI ID in the `CustomAmi` field in your cluster configuration and create a cluster.

Old Amazon EC2 console

1. Find the AWS ParallelCluster AMI that corresponds to the specific AWS Region that you use. To find it you can use the `pcluster list-official-images` command with the `--region` parameter to select the specific AWS Region and `--os` and `--architecture` parameters to filter for the desired AMI with the OS and architecture that you want to use. From the output you can retrieve the Amazon EC2 Image ID.
2. Sign in to the AWS Management Console and open the Amazon EC2 console at https://console.aws.amazon.com/ec2/.
3. In the navigation pane, choose **Images**, and then **AMIs**. Set the filter for **Public images** and search for the retrieved EC2 Image ID, select the AMI, and choose **Launch**.
4. Choose your instance type and select **Next: Configure Instance Details** or **Review and Launch** to launch your instance.

5. Choose **Launch**, select your **Key pair**, and **Launch Instances**.
6. Log into your instance using the OS user and your SSH key. For more information, navigate to **Instances**, select the new instance and **Connect**.
7. Manually customize your instance to meet your requirements.
8. Run the following command to prepare your instance for AMI creation:

```
sudo /usr/local/sbin/ami_cleanup.sh
```

9. From the Amazon EC2 console, choose **Instances** in the navigation pane, select your new instance and choose **Actions**, **Instance State** and **Stop**.
10. Create a new AMI from the instance using the Amazon EC2 console or AWS CLI create-image.

    From the Amazon EC2 console

    a. Choose **Instances** in the navigation pane.

    b. Choose the instance you created and modified.

    c. In **Actions**, choose **Image** and then **Create Image**.

    d. Choose **Create Image**.
11. Enter the new AMI ID in the `CustomAmi` field in your cluster configuration and create a cluster.

Integrating Active Directory

In this tutorial, you create a multiple user environment. This environment includes an AWS ParallelCluster that's integrated with an AWS Managed Microsoft AD (Active Directory) at `corp.example.com`. You configure an `Admin` user to manage the directory, a `ReadOnly` user to read the directory, and a `user000` user to log into the cluster. You can use either the automated path or the manual path to create the networking resources, an Active Directory (AD), and the Amazon EC2 instance that you use to configure the AD. Regardless of the path, the infrastructure that you create is pre-configured to integrate AWS ParallelCluster using one of the following methods:

- LDAPS with certificate verification (recommended as the most secure option)
- LDAPS without certificate verification
- LDAP

LDAP by itself *doesn't* provide encryption. To ensure secure transmission of potentially sensitive information, we strongly recommend that you use LDAPS (LDAP over TLS/SSL) for clusters integrated with ADs. For more information, see Enable server-side LDAPS using AWS Managed Microsoft AD in the AWS Directory Service *Administration Guide*.

After you create these resources, proceed to configure and create your cluster integrated with your Active Directory (AD). After the cluster is created, log in as the user you created. For more information about the configuration that you create in this tutorial, see Multiple user access to clusters and the `DirectoryService` configuration section.

This tutorial covers how to create an environment that supports multiple user access to clusters. This tutorial doesn't cover how you create and use an AWS Directory Service AD. The steps that you take to set up an AWS Managed Microsoft AD in this tutorial are provided for testing purposes only. They *aren't* provided to replace the official documentation and best practices you can find at AWS Managed Microsoft AD and Simple AD in the *AWS Directory Service Administration Guide*.

Note

Directory user passwords expire according to the directory password policy property definitions. To reset directory passwords with AWS ParallelCluster, see How to reset a user password and expired passwords.

Note

The directory domain controller IP addresses can change due to domain controller changes and directory maintenance. If you chose the automated quick create method to create the directory infrastructure, you must manually align the load balancer in front of the directory controllers when the directory IP addresses change. If you use the quick create method, the directory IP addresses aren't automatically aligned with the load balancers.

When using the AWS ParallelCluster command line interface (CLI) or API, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

Prerequisites

- AWS ParallelCluster is installed.

- The AWS CLI is installed and configured.
- You have an Amazon EC2 key pair.
- You have an IAM role with the permissions required to run the `pcluster` CLI.

As you go through the tutorial, replace *`inputs highlighted in red`*, such as *`region-id`* and *`d-abcdef01234567890`*, with your own names and IDs. Replace *`0123456789012`* with your AWS account number.

Create the AD infrastructure

Choose the *Automated* tab to create the Active Directory (AD) infrastructure with an AWS CloudFormation quick create template.

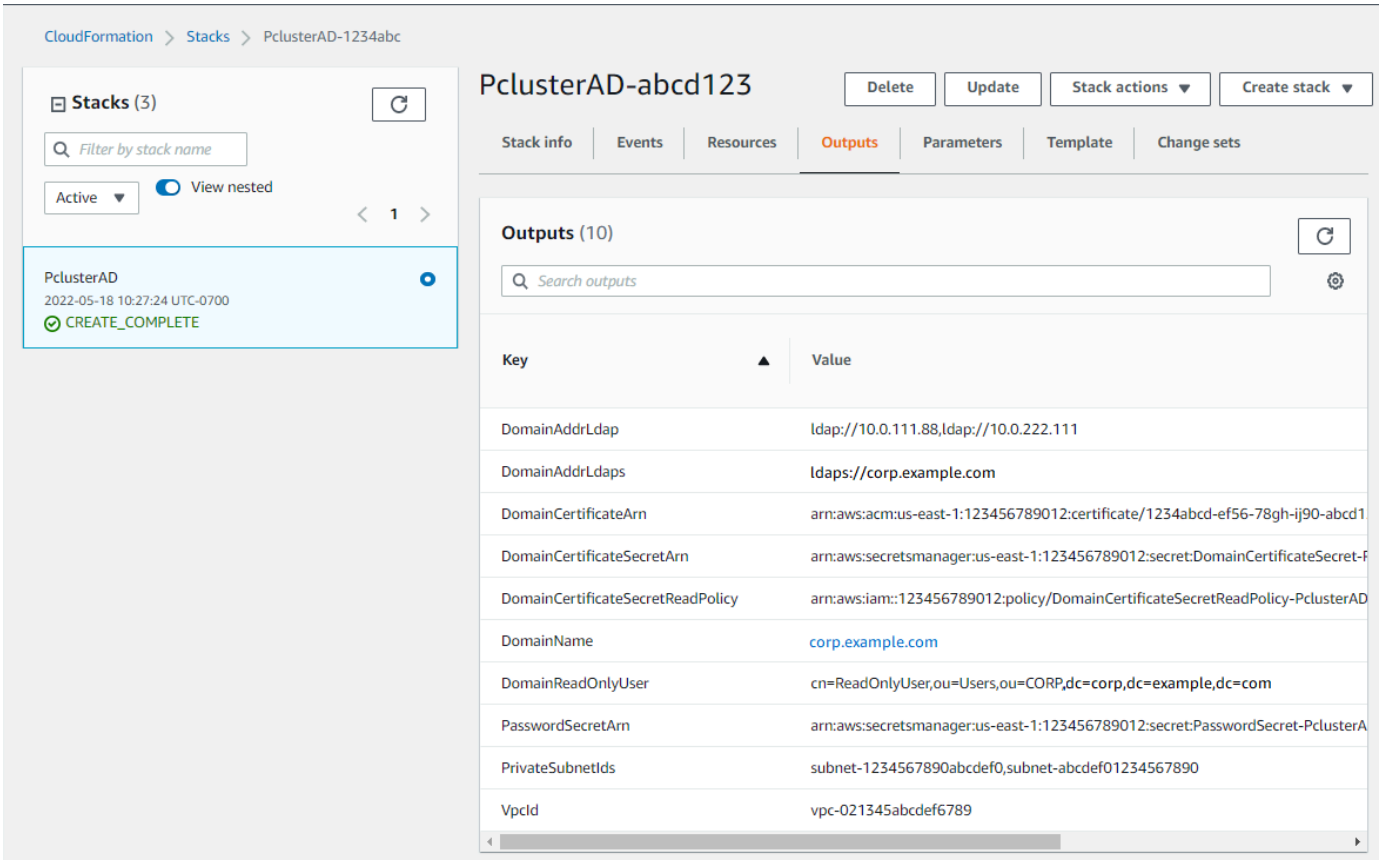
Choose the *Manual* tab to manually create the AD infrastructure.

Automated

1. Sign in to the AWS Management Console.
2. Open CloudFormation Quick Create (region us-east-1) to create the following resources in the CloudFormation console:
   - A VPC with two subnets and routing for public access, if no VPC is specified.
   - An AWS Managed Microsoft AD.
   - An Amazon EC2 instance that's joined to the AD that you can use to manage the directory.
3. In the **Quick create stack** page **Parameters** section, enter passwords for the following parameters:
   - **AdminPassword**
   - **ReadOnlyPassword**
   - **UserPassword**

   Make note of the passwords. You use them later on in this tutorial.
4. For **DomainName**, enter **`corp.example.com`**
5. For **Keypair**, enter the name of an Amazon EC2 key pair.
6. Check the boxes to acknowledge each of the access capabilities at the bottom of the page.
7. Choose **Create stack**.

8. After the CloudFormation stack has reached the `CREATE_COMPLETE` state, choose the **Outputs** tab of the stack. Make a note of the output resource names and IDs because you need to use them in later steps. The outputs provide the information that's needed to create the cluster.



9. To complete the exercises (Optional) Manage AD users and groups, you need the directory ID. Choose **Resources** and scroll down to make note of the directory ID.
10. Continue at (Optional) Manage AD users and groups or Create the cluster.

Manual

Create a VPC for the directory service with two subnets in different Availability Zones and an AWS Managed Microsoft AD.

Create the AD

Note

- The directory and domain name is `corp.example.com`. The short name is `CORP`.

- Change the `Admin` password in the script.
- The Active Directory (AD) takes at least 15 minutes to create.

Use the following Python script to create the VPC, subnets, and AD resources in your local AWS Region. Save this file as `ad.py` and run it.

```
import boto3
import time
from pprint import pprint

vpc_name = "PclusterVPC"
ad_domain = "corp.example.com"
admin_password = "asdfASDF1234"

ec2 = boto3.client("ec2")
ds = boto3.client("ds")
region = boto3.Session().region_name

# Create the VPC, Subnets, IGW, Routes
vpc = ec2.create_vpc(CidrBlock="10.0.0.0/16")["Vpc"]
vpc_id = vpc["VpcId"]
time.sleep(30)
ec2.create_tags(Resources=[vpc_id], Tags=[{"Key": "Name", "Value": vpc_name}])
subnet1 = ec2.create_subnet(VpcId=vpc_id, CidrBlock="10.0.0.0/17",
 AvailabilityZone=f"{region}a")["Subnet"]
subnet1_id = subnet1["SubnetId"]
time.sleep(30)
ec2.create_tags(Resources=[subnet1_id], Tags=[{"Key": "Name", "Value": f"{vpc_name}/
subnet1"}])
ec2.modify_subnet_attribute(SubnetId=subnet1_id, MapPublicIpOnLaunch={"Value": True})
subnet2 = ec2.create_subnet(VpcId=vpc_id, CidrBlock="10.0.128.0/17",
 AvailabilityZone=f"{region}b")["Subnet"]
subnet2_id = subnet2["SubnetId"]
time.sleep(30)
ec2.create_tags(Resources=[subnet2_id], Tags=[{"Key": "Name", "Value": f"{vpc_name}/
subnet2"}])
ec2.modify_subnet_attribute(SubnetId=subnet2_id, MapPublicIpOnLaunch={"Value": True})
igw = ec2.create_internet_gateway()["InternetGateway"]
ec2.attach_internet_gateway(InternetGatewayId=igw["InternetGatewayId"], VpcId=vpc_id)
route_table = ec2.describe_route_tables(Filters=[{"Name": "vpc-id", "Values":
 [vpc_id]}])["RouteTables"][0]
```

```
ec2.create_route(RouteTableId=route_table["RouteTableId"],
 DestinationCidrBlock="0.0.0.0/0", GatewayId=igw["InternetGatewayId"])
ec2.modify_vpc_attribute(VpcId=vpc_id, EnableDnsSupport={"Value": True})
ec2.modify_vpc_attribute(VpcId=vpc_id, EnableDnsHostnames={"Value": True})

# Create the Active Directory
ad = ds.create_microsoft_ad(
    Name=ad_domain,
    Password=admin_password,
    Description="ParallelCluster AD",
    VpcSettings={"VpcId": vpc_id, "SubnetIds": [subnet1_id, subnet2_id]},
    Edition="Standard",
)
directory_id = ad["DirectoryId"]

# Wait for completion
print("Waiting for the directory to be created...")
directories = ds.describe_directories(DirectoryIds=[directory_id])
["DirectoryDescriptions"]
directory = directories[0]
while directory["Stage"] in {"Requested", "Creating"}:
    time.sleep(3)
    directories = ds.describe_directories(DirectoryIds=[directory_id])
["DirectoryDescriptions"]
    directory = directories[0]

dns_ip_addrs = directory["DnsIpAddrs"]

pprint({"directory_id": directory_id,
        "vpc_id": vpc_id,
        "subnet1_id": subnet1_id,
        "subnet2_id": subnet2_id,
        "dns_ip_addrs": dns_ip_addrs})
```

The following is example output from the Python script.

```
{
  "directory_id": "d-abcdef01234567890",
  "dns_ip_addrs": ["192.0.2.254", "203.0.113.237"],
  "subnet1_id": "subnet-021345abcdef6789",
  "subnet2_id": "subnet-1234567890abcdef0",
  "vpc_id": "vpc-021345abcdef6789"
}
```

Make a note of the output resource names and IDs. You use them in later steps.

After the script completes, continue to the next step.

Create an Amazon EC2 instance

New Amazon EC2 console

1. Sign in to the AWS Management Console.
2. If you don't have a role with the policies listed in step 4 attached, open the IAM console at https://console.aws.amazon.com/iam/. Otherwise, skip to step 5.
3. Create the `ResetUserPassword` policy, replacing the red highlighted content with your AWS Region ID, Account ID, and the directory ID from the output of the script you ran to create the AD.

   ResetUserPassword

   ```
   {
          "Statement": [
           {
               "Action": [
                   "ds:ResetUserPassword"
               ],
               "Resource": "arn:aws:ds:region-id:123456789012:directory/d-
   abcdef01234567890",
               "Effect": "Allow"
           }
       ]
   }
   ```

4. Create an IAM role with the following policies attached.
   - AWS managed policy AmazonSSMManagedInstanceCore
   - AWS managed policy AmazonSSMDirectoryServiceAccess
   - ResetUserPassword policy
5. Open the Amazon EC2 console at https://console.aws.amazon.com/ec2/.
6. In the **Amazon EC2 Dashboard**, choose **Launch Instance**.
7. In **Application and OS Images**, select a recent Amazon Linux 2 AMI.
8. For **Instance type**, choose t2.micro.

9. For **Key pair**, choose a key pair.
10. For **Network settings**, choose **Edit**.
11. For **VPC**, select the directory VPC.
12. Scroll down and select **Advanced details**.
13. In **Advanced details**, **Domain join directory**, choose **`corp.example.com`**.
14. For **IAM Instance profile**, choose the role you created in step 1 or a role with policies listed in step 4 attached.
15. In **Summary** choose **Launch instance**.
16. Make note of the Instance ID (for example, i-1234567890abcdef0) and wait for the instance to finish launching.
17. After the instance has launched, continue to the next step.

Old Amazon EC2 console

1. Sign in to the AWS Management Console.
2. If you don't have a role with the policies listed in step 4 attached, open the IAM console at https://console.aws.amazon.com/iam/. Otherwise, skip to step 5.
3. Create the `ResetUserPassword` policy. Replace the red highlighted content with your AWS Region ID, AWS account ID, and the directory ID from the output of the script you ran to create the Active Directory (AD).

   ResetUserPassword

   ```
   {
           "Statement": [
               {
                   "Action": [
                       "ds:ResetUserPassword"
                   ],
                   "Resource": "arn:aws:ds:region-id:123456789012:directory/d-
   abcdef01234567890",
                   "Effect": "Allow"
               }
           ]
       }
   ```

4. Create an IAM role with the following policies attached.

- AWS managed policy AmazonSSMManagedInstanceCore
- AWS managed policy AmazonSSMDirectoryServiceAccess
- ResetUserPassword policy

5. Open the Amazon EC2 console at https://console.aws.amazon.com/ec2/.
6. In the **Amazon EC2 Dashboard**, choose **Launch Instance**.
7. In **Application and OS Images**, select a recent Amazon Linux 2 AMI.
8. For **Instance type**, choose t2.micro.
9. For **Key pair**, choose a key pair.
10. In **Network settings**, choose **Edit**.
11. In **Network settings**, **VPC**, select the directory VPC.
12. Scroll down and select **Advanced details**.
13. In **Advanced details**, **Domain join directory**, choose **`corp.example.com`**.
14. In **Advanced details**, **Instance profile**, choose the role that you created in step 1 or a role with the policies that are listed in step 4 attached.
15. In **Summary** choose **Launch instance**.
16. Make note of the Instance ID (for example, i-1234567890abcdef0) and wait for the instance to finish launching.
17. After the instance has launched, continue to the next step.

Join your instance to the AD

1. **Connect to your instance and join the AD realm as `admin`.**

   Run the following commands to connect to the instance.

```
$ INSTANCE_ID="i-1234567890abcdef0"
```

```
$ PUBLIC_IP=$(aws ec2 describe-instances \
--instance-ids $INSTANCE_ID \
--query "Reservations[0].Instances[0].PublicIpAddress" \
--output text)
```

```
$ ssh -i ~/.ssh/keys/keypair.pem ec2-user@$PUBLIC_IP
```

2. **Install necessary software and join the realm.**

```
$ sudo yum -y install sssd realmd oddjob oddjob-mkhomedir adcli samba-common samba-common-tools krb5-workstation openldap-clients policycoreutils-python
```

3. **Replace the admin password with your admin password.**

```
$ ADMIN_PW="asdfASDF1234"
```

```
$ echo $ADMIN_PW | sudo realm join -U Admin corp.example.com
Password for Admin:
```

If the preceding has succeeded, you're joined to the realm and can proceed to the next step.

Add users to the AD

1. **Create the ReadOnlyUser and an additional user.**

In this step, you use adcli and openldap-clients tools that you installed in a preceding step.

```
$ echo $ADMIN_PW | adcli create-user -x -U Admin --domain=corp.example.com --display-name=ReadOnlyUser ReadOnlyUser
```

```
$ echo $ADMIN_PW | adcli create-user -x -U Admin --domain=corp.example.com --display-name=user000 user000
```

2. **Verify the users are created:**

The directory DNS IP addresses are outputs of the Python script.

```
$ DIRECTORY_IP="192.0.2.254"
```

```
$ ldapsearch -x -h $DIRECTORY_IP -D Admin -w $ADMIN_PW -b
 "cn=ReadOnlyUser,ou=Users,ou=CORP,dc=corp,dc=example,dc=com"
```

```
$ ldapsearch -x -h $DIRECTORY_IP -D Admin -w $ADMIN_PW -b
 "cn=user000,ou=Users,ou=CORP,dc=corp,dc=example,dc=com"
```

By default, when you create a user with the `ad-cli`, the user is disabled.

3. **Reset and activate the user passwords from your local machine:**

   Log out of your Amazon EC2 instance.

   > **Note**
   >
   > - `ro-p@ssw0rd` is the password of `ReadOnlyUser`, retrieved from AWS Secrets Manager.
   > - `user-p@ssw0rd` is the password of a cluster user that's provided when you connect (`ssh`) to the cluster.

   The `directory-id` is an output of the Python script.

   ```
   $ DIRECTORY_ID="d-abcdef01234567890"
   ```

   ```
   $ aws ds reset-user-password \
   --directory-id $DIRECTORY_ID \
   --user-name "ReadOnlyUser" \
   --new-password "ro-p@ssw0rd" \
   --region "region-id"
   ```

   ```
   $ aws ds reset-user-password \
   --directory-id $DIRECTORY_ID \
   --user-name "user000" \
   --new-password "user-p@ssw0rd" \
   --region "region-id"
   ```

4. **Add the password to a Secrets Manager secret.**

   Now that you created a `ReadOnlyUser` and set the password, store it in a secret that AWS ParallelCluster uses for validating logins.

   Use Secrets Manager to create a new secret to hold the password for the `ReadOnlyUser` as the value. The secret value format must be plain text only (not JSON format). Make note of the secret ARN for future steps.

```
$ aws secretsmanager create-secret --name "ADSecretPassword" \
--region region_id \
--secret-string "ro-p@ssw0rd" \
--query ARN \
--output text
arn:aws:secretsmanager:region-id:123456789012:secret:ADSecretPassword-1234
```

LDAPS with certificate verification (recommended) setup

Make a note of resource IDs. You use them in steps later on.

1. **Generate domain certificate, locally.**

```
$ PRIVATE_KEY="corp-example-com.key"
CERTIFICATE="corp-example-com.crt"
printf ".\n.\n.\n.\ncorp.example.com\n.\n" | openssl req -x509 -sha256 -nodes -
newkey rsa:2048 -keyout $PRIVATE_KEY -days 365 -out $CERTIFICATE
```

2. **Store the certificate to Secrets Manager to make it retrievable from within the cluster later on.**

```
$ aws secretsmanager create-secret --name example-cert \
  --secret-string file://$CERTIFICATE \
  --region region-id
{
  "ARN": "arn:aws:secretsmanager:region-id:123456789012:secret:example-
cert-123abc",
  "Name": "example-cert",
  "VersionId": "14866070-092a-4d5a-bcdd-9219d0566b9c"
}
```

3. Add the following policy to the IAM role that you created to join the Amazon EC2 instance to the AD domain.

   PutDomainCertificateSecrets

```
{
    "Statement": [
        {
            "Action": [
```

```
                "secretsmanager:PutSecretValue"
            ],
            "Resource": [
                "arn:aws:secretsmanager:region-id:123456789012:secret:example-
cert-123abc"
            ],
            "Effect": "Allow"
        }
    ]
}
```

4. **Import the certificate to AWS Certificate Manager (ACM).**

```
$ aws acm import-certificate --certificate fileb://$CERTIFICATE \
  --private-key fileb://$PRIVATE_KEY \
  --region region-id
{
  "CertificateArn": "arn:aws:acm:region-
id:123456789012:certificate/343db133-490f-4077-b8d4-3da5bfd89e72"
}
```

5. **Create and the load balancer that is put in front of the Active Directory endpoints.**

```
$ aws elbv2 create-load-balancer --name CorpExampleCom-NLB \
  --type network \
  --scheme internal \
  --subnets subnet-1234567890abcdef0 subnet-021345abcdef6789 \
  --region region-id
{
  "LoadBalancers": [
    {
      "LoadBalancerArn": "arn:aws:elasticloadbalancing:region-
id:123456789012:loadbalancer/net/CorpExampleCom-NLB/3afe296bf4ba80d4",
      "DNSName": "CorpExampleCom-NLB-3afe296bf4ba80d4.elb.region-id.amazonaws.com",
      "CanonicalHostedZoneId": "Z2IFOLAFXWLO4F",
      "CreatedTime": "2022-05-05T12:56:55.988000+00:00",
      "LoadBalancerName": "CorpExampleCom-NLB",
      "Scheme": "internal",
      "VpcId": "vpc-021345abcdef6789",
      "State": {
        "Code": "provisioning"
       },
       "Type": "network",
```

```
      "AvailabilityZones": [
        {
          "ZoneName": "region-idb",
          "SubnetId": "subnet-021345abcdef6789",
          "LoadBalancerAddresses": []
        },
        {
          "ZoneName": "region-ida",
          "SubnetId": "subnet-1234567890abcdef0",
          "LoadBalancerAddresses": []
        }
      ],
      "IpAddressType": "ipv4"
    }
  ]
}
```

6. **Create the target group that's targeting the Active Directory endpoints.**

```
$ aws elbv2 create-target-group --name CorpExampleCom-Targets --protocol TCP \
  --port 389 \
  --target-type ip \
  --vpc-id vpc-021345abcdef6789 \
  --region region-id
{
  "TargetGroups": [
    {
      "TargetGroupArn": "arn:aws:elasticloadbalancing:region-
id:123456789012:targetgroup/CorpExampleCom-Targets/44577c583b695e81",
      "TargetGroupName": "CorpExampleCom-Targets",
      "Protocol": "TCP",
      "Port": 389,
      "VpcId": "vpc-021345abcdef6789",
      "HealthCheckProtocol": "TCP",
      "HealthCheckPort": "traffic-port",
      "HealthCheckEnabled": true,
      "HealthCheckIntervalSeconds": 30,
      "HealthCheckTimeoutSeconds": 10,
      "HealthyThresholdCount": 3,
      "UnhealthyThresholdCount": 3,
      "TargetType": "ip",
      "IpAddressType": "ipv4"
    }
  ]
```

```
}
```

7. **Register the Active Directory (AD) endpoints into the target group.**

```
$ aws elbv2 register-targets --target-group-arn
 arn:aws:elasticloadbalancing:region-id:123456789012:targetgroup/CorpExampleCom-
Targets/44577c583b695e81 \
  --targets Id=192.0.2.254,Port=389 Id=203.0.113.237,Port=389 \
  --region region-id
```

8. **Create the LB listener with the certificate.**

```
$ aws elbv2 create-listener --load-balancer-arn
 arn:aws:elasticloadbalancing:region-id:123456789012:loadbalancer/net/
CorpExampleCom-NLB/3afe296bf4ba80d4 \
  --protocol TLS \
  --port 636 \
  --default-actions
 Type=forward,TargetGroupArn=arn:aws:elasticloadbalancing:region-
id:123456789012:targetgroup/CorpExampleCom-Targets/44577c583b695e81 \
  --ssl-policy ELBSecurityPolicy-TLS-1-2-2017-01 \
  --certificates CertificateArn=arn:aws:acm:region-
id:123456789012:certificate/343db133-490f-4077-b8d4-3da5bfd89e72 \
  --region region-id
  "Listeners": [
  {
    "ListenerArn": "arn:aws:elasticloadbalancing:region-id:123456789012:listener/
net/CorpExampleCom-NLB/3afe296bf4ba80d4/a8f9d97318743d4b",
    "LoadBalancerArn": "arn:aws:elasticloadbalancing:region-
id:123456789012:loadbalancer/net/CorpExampleCom-NLB/3afe296bf4ba80d4",
    "Port": 636,
    "Protocol": "TLS",
    "Certificates": [
      {
        "CertificateArn": "arn:aws:acm:region-
id:123456789012:certificate/343db133-490f-4077-b8d4-3da5bfd89e72"
       }
     ],
     "SslPolicy": "ELBSecurityPolicy-TLS-1-2-2017-01",
     "DefaultActions": [
       {
         "Type": "forward",
```

```
          "TargetGroupArn": "arn:aws:elasticloadbalancing:region-
id:123456789012:targetgroup/CorpExampleCom-Targets/44577c583b695e81",
          "ForwardConfig": {
            "TargetGroups": [
              {
                 "TargetGroupArn": "arn:aws:elasticloadbalancing:region-
id:123456789012:targetgroup/CorpExampleCom-Targets/44577c583b695e81"
               }
             ]
           }
         }
       ]
     }
  ]
}
```

9. **Create the hosted zone to make the domain discoverable within the cluster VPC.**

```
$ aws route53 create-hosted-zone --name corp.example.com \
  --vpc VPCRegion=region-id,VPCId=vpc-021345abcdef6789 \
  --caller-reference "ParallelCluster AD Tutorial"
{
  "Location": "https://route53.amazonaws.com/2013-04-01/hostedzone/
Z09020002B5MZQNXMSJUB",
  "HostedZone": {
    "Id": "/hostedzone/Z09020002B5MZQNXMSJUB",
    "Name": "corp.example.com.",
    "CallerReference": "ParallelCluster AD Tutorial",
    "Config": {
         "PrivateZone": true
    },
    "ResourceRecordSetCount": 2
  },
  "ChangeInfo": {
    "Id": "/change/C05533343BF3IKSORW1TQ",
    "Status": "PENDING",
    "SubmittedAt": "2022-05-05T13:21:53.863000+00:00"
  },
  "VPC": {
    "VPCRegion": "region-id",
    "VPCId": "vpc-021345abcdef6789"
  }
}
```

10. **Create a file that's named `recordset-change.json` with the following content. `HostedZoneId` is the canonical hosted zone ID of the load balancer.**

```
{
  "Changes": [
    {
      "Action": "CREATE",
      "ResourceRecordSet": {
        "Name": "corp.example.com",
        "Type": "A",
        "Region": "region-id",
        "SetIdentifier": "example-active-directory",
        "AliasTarget": {
          "HostedZoneId": "Z2IFOLAFXWLO4F",
          "DNSName": "CorpExampleCom-NLB-3afe296bf4ba80d4.elb.region-
id.amazonaws.com",
          "EvaluateTargetHealth": true
        }
      }
    }
  ]
}
```

11. **Submit the recordset change to the hosted zone, this time using the hosted zone ID.**

```
$ aws route53 change-resource-record-sets --hosted-zone-id Z09020002B5MZQNXMSJUB \
  --change-batch file://recordset-change.json
{
  "ChangeInfo": {
    "Id": "/change/C0137926I56R3GC7XW2Y",
    "Status": "PENDING",
    "SubmittedAt": "2022-05-05T13:40:36.553000+00:00"
  }
}
```

12. **Create a policy document `policy.json` with the following content.**

    JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
```

```
        {
            "Action": [
                "secretsmanager:GetSecretValue"
            ],
            "Resource": [
                "arn:aws:secretsmanager:us-
east-1:123456789012:secret:example-cert-abc123"
            ],
            "Effect": "Allow"
        }
    ]
}
```

13. **Create a policy document that is named `policy.json` with the following content.**

```
$ aws iam create-policy --policy-name ReadCertExample \
  --policy-document file://policy.json
{
  "Policy": {
    "PolicyName": "ReadCertExample",
    "PolicyId": "ANPAUUXUVBC42VZSI4LDY",
    "Arn": "arn:aws:iam::123456789012:policy/ReadCertExample-efg456",
    "Path": "/",
    "DefaultVersionId": "v1",
    "AttachmentCount": 0,
    "PermissionsBoundaryUsageCount": 0,
    "IsAttachable": true,
    "CreateDate": "2022-05-05T13:42:18+00:00",
    "UpdateDate": "2022-05-05T13:42:18+00:00"
  }
}
```

14. Continue to follow the steps at (Optional) Manage AD users and groups or Create the cluster.

(Optional) Manage AD users and groups

In this step, you manage users and groups from an Amazon EC2 Amazon Linux 2 instance that's joined to the Active Delivery (AD) domain.

If you followed the *automated* path, restart and log in to the AD joined instance that was created as part of the automation.

If you followed the *manual* path, restart and log in to the instance that you created and joined to the AD in preceding steps.

In these steps, you use the adcli and openldap-clients tools that were installed in the instance as part of a preceding step.

Log in to an Amazon EC2 instance that is joined to the AD domain

1. From the Amazon EC2 console, select the untitled Amazon EC2 instance that was created in previous steps. The instance state might be **Stopped**.
2. If the instance state is **Stopped**, choose **Instance state** and then **Start instance**.
3. After the status checks pass, select the instance and choose **Connect** and SSH in to the instance.

Manage users and groups when logged into an Amazon EC2 Amazon Linux 2 instance that's joined the AD

When you run the `adcli` commands with the `-U "Admin"` option, you're prompted to enter the AD `Admin` password. You include the AD `Admin` password as part of the `ldapsearch` commands.

1. **Create a user.**

```
$ adcli create-user "clusteruser" --domain "corp.example.com" -U "Admin"
```

2. **Set a user password.**

```
$ aws --region "region-id" ds reset-user-password --directory-id "d-
abcdef01234567890" --user-name "clusteruser" --new-password "new-p@ssw0rd"
```

3. **Create a group.**

```
$ adcli create-group "clusterteam" --domain "corp.example.com" -U "Admin"
```

4. **Add a user to a group.**

```
$ adcli add-member "clusterteam" "clusteruser" --domain "corp.example.com" -U
 "Admin"
```

5. **Describe users and groups.**

   Describe all users.

```
$ ldapsearch "(&(objectClass=user))" -x -h "192.0.2.254" -b
 "DC=corp,DC=example,DC=com" -D
 "CN=Admin,OU=Users,OU=CORP,DC=corp,DC=example,DC=com" -w "p@ssw0rd"
```

   Describe a specific user.

```
$ ldapsearch "(&(objectClass=user)(cn=clusteruser))"
 -x -h "192.0.2.254" -b "DC=corp,DC=example,DC=com" -D
 "CN=Admin,OU=Users,OU=CORP,DC=corp,DC=example,DC=com" -w "p@ssw0rd"
```

   Describe all users with a name pattern.

```
$ ldapsearch "(&(objectClass=user)(cn=user*))" -x -h "192.0.2.254" -b
 "DC=corp,DC=example,DC=com" -D
 "CN=Admin,OU=Users,OU=CORP,DC=corp,DC=example,DC=com" -w "p@ssw0rd"
```

   Describe all users that are part of a specific group.

```
$ ldapsearch "(&(objectClass=user)
(memberOf=CN=clusterteam,OU=Users,OU=CORP,DC=corp,DC=example,DC=com))"
 -x -h "192.0.2.254" -b "DC=corp,DC=example,DC=com" -D
 "CN=Admin,OU=Users,OU=CORP,DC=corp,DC=example,DC=com" -w "p@ssw0rd"
```

   Describe all groups

```
$ ldapsearch "objectClass=group" -x -h "192.0.2.254" -b "DC=corp,DC=example,DC=com"
 -D "CN=Admin,OU=Users,OU=CORP,DC=corp,DC=example,DC=com" -w "p@ssw0rd"
```

   Describe a specific group

```
$ ldapsearch "(&(objectClass=group)(cn=clusterteam))"
 -x -h "192.0.2.254" -b "DC=corp,DC=example,DC=com" -D
 "CN=Admin,OU=Users,OU=CORP,DC=corp,DC=example,DC=com" -w "p@ssw0rd"
```

6. **Remove a user from a group.**

```
$ adcli remove-member "clusterteam" "clusteruser" --domain "corp.example.com" -U
 "Admin"
```

7. **Delete a user.**

```
$ adcli delete-user "clusteruser" --domain "corp.example.com" -U "Admin"
```

8. **Delete a group.**

```
$ adcli delete-group "clusterteam" --domain "corp.example.com" -U "Admin"
```

Create the cluster

If you haven't exited the Amazon EC2 instance, do so now.

The environment is set up to create a cluster that can authenticate users against the Active Directory (AD).

Create a simple cluster configuration and provide the settings relevant to connecting to the AD. For more information, see the `DirectoryService` section.

Choose one of the following cluster configurations and copy it to a file that's named `ldaps_config.yaml`, `ldaps_nocert_config.yaml`, or `ldap_config.yaml`.

We recommend that you choose the LDAPS configuration with certificate verification. If you choose this configuration, you must also copy the bootstrap script to a file that's named `active-directory.head.post.sh`. And, you must store it in an Amazon S3 bucket as indicated in the configuration file.

LDAPS with certificate verification configuration (recommended)

> **Note**
>
> **The following components must be changed.**
>
> - KeyName: One of your Amazon EC2 keypairs.

- `SubnetId / SubnetIds`: One of the subnet IDs provided in the output of the CloudFormation quick create stack (automated tutorial) or python script (manual tutorial).
- `Region`: The Region where you created the AD infrastructure.
- `DomainAddr`: This IP address is one of the DNS addresses of your AD service.
- `PasswordSecretArn`: The Amazon Resource Name (ARN) of the secret that contains the password for the `DomainReadOnlyUser`.
- `BucketName`: The name of the bucket that holds the bootstrap script.
- `AdditionalPolicies / Policy`: The Amazon Resource Name (ARN) of the read domain certification policy ReadCertExample.
- `CustomActions / OnNodeConfigured / Args`: The Amazon Resource Name (ARN) of secret that holds the domain certification policy.

For better security posture, we suggest to use the `HeadNode / Ssh / AllowedIps` configuration to limit the SSH access to the head node.

```
Region: region-id
Image:
  Os: alinux2
HeadNode:
  InstanceType: t2.micro
  Networking:
    SubnetId: subnet-abcdef01234567890
  Ssh:
    KeyName: keypair
  Iam:
    AdditionalIamPolicies:
      - Policy: arn:aws:iam::123456789012:policy/ReadCertExample
    S3Access:
      - BucketName: amzn-s3-demo-bucket
        EnableWriteAccess: false
        KeyName: bootstrap/active-directory/active-directory.head.post.sh
  CustomActions:
    OnNodeConfigured:
      Script: s3://amzn-s3-demo-bucket/bootstrap/active-directory/active-
directory.head.post.sh
      Args:
```

```
        - arn:aws:secretsmanager:region-id:123456789012:secret:example-cert-123abc
        - /opt/parallelcluster/shared/directory_service/domain-certificate.crt
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue0
      ComputeResources:
        - Name: queue0-t2-micro
          InstanceType: t2.micro
          MinCount: 1
          MaxCount: 10
      Networking:
        SubnetIds:
          - subnet-abcdef01234567890
DirectoryService:
  DomainName: corp.example.com
  DomainAddr: ldaps://corp.example.com
  PasswordSecretArn: arn:aws:secretsmanager:region-
id:123456789012:secret:ADSecretPassword-1234
  DomainReadOnlyUser: cn=ReadOnlyUser,ou=Users,ou=CORP,dc=corp,dc=example,dc=com
  LdapTlsCaCert: /opt/parallelcluster/shared/directory_service/domain-certificate.crt
  LdapTlsReqCert: hard
```

Bootstrap script

After you create the bootstrap file and before you upload it to your S3 bucket, run `chmod +x active-directory.head.post.sh` to give AWS ParallelCluster run permission.

```
#!/bin/bash
set -e

CERTIFICATE_SECRET_ARN="$1"
CERTIFICATE_PATH="$2"

[[ -z $CERTIFICATE_SECRET_ARN ]] && echo "[ERROR] Missing CERTIFICATE_SECRET_ARN" &&
 exit 1
[[ -z $CERTIFICATE_PATH ]] && echo "[ERROR] Missing CERTIFICATE_PATH" && exit 1

source /etc/parallelcluster/cfnconfig
REGION="${cfn_region:?}"

mkdir -p $(dirname $CERTIFICATE_PATH)
```

```
aws secretsmanager get-secret-value --region $REGION --secret-id
 $CERTIFICATE_SECRET_ARN --query SecretString --output text > $CERTIFICATE_PATH
```

LDAPS without certificate verification configuration

Note

The following components must be changed.

- `KeyName`: One of your Amazon EC2 keypairs.
- `SubnetId / SubnetIds`: One of the subnet IDs that's in the output of the CloudFormation quick create stack (automated tutorial) or python script (manual tutorial).
- `Region`: The Region where you created the AD infrastructure.
- `DomainAddr`: This IP address is one of the DNS addresses of your AD service.
- `PasswordSecretArn`: The Amazon Resource Name (ARN) of the secret that contains the password for the `DomainReadOnlyUser`.

For better security posture, we suggest to use the HeadNode/Ssh/AllowedIps configuration to limit the SSH access to the head node.

```
Region: region-id
Image:
  Os: alinux2
HeadNode:
  InstanceType: t2.micro
  Networking:
    SubnetId: subnet-abcdef01234567890
  Ssh:
    KeyName: keypair
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue0
      ComputeResources:
        - Name: queue0-t2-micro
          InstanceType: t2.micro
```

```
          MinCount: 1
          MaxCount: 10
      Networking:
        SubnetIds:
          - subnet-abcdef01234567890
DirectoryService:
  DomainName: corp.example.com
  DomainAddr: ldaps://corp.example.com
  PasswordSecretArn: arn:aws:secretsmanager:region-
id:123456789012:secret:ADSecretPassword-1234
  DomainReadOnlyUser: cn=ReadOnlyUser,ou=Users,ou=CORP,dc=corp,dc=example,dc=com
  LdapTlsReqCert: never
```

LDAP configuration

> **Note**
>
> **The following components must be changed.**
>
> - `KeyName`: One of your Amazon EC2 keypairs.
> - `SubnetId` / `SubnetIds`: One of the subnet IDs provided in the output of the CloudFormation quick create stack (automated tutorial) or python script (manual tutorial).
> - `Region`: The Region where you created the AD infrastructure.
> - `DomainAddr`: This IP address is one of the DNS addresses of your AD service.
> - `PasswordSecretArn`: The Amazon Resource Name (ARN) of the secret that contains the password for the `DomainReadOnlyUser`.
>
> For better security posture, we suggest to use the HeadNode/Ssh/AllowedIps configuration to limit the SSH access to the head node.

```
Region: region-id
Image:
  Os: alinux2
HeadNode:
  InstanceType: t2.micro
  Networking:
```

```
    SubnetId: subnet-abcdef01234567890
  Ssh:
    KeyName: keypair
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue0
      ComputeResources:
        - Name: queue0-t2-micro
          InstanceType: t2.micro
          MinCount: 1
          MaxCount: 10
      Networking:
        SubnetIds:
          - subnet-abcdef01234567890
DirectoryService:
  DomainName: dc=corp,dc=example,dc=com
  DomainAddr: ldap://192.0.2.254,ldap://203.0.113.237
  PasswordSecretArn: arn:aws:secretsmanager:region-
id:123456789012:secret:ADSecretPassword-1234
  DomainReadOnlyUser: cn=ReadOnlyUser,ou=Users,ou=CORP,dc=corp,dc=example,dc=com
  AdditionalSssdConfigs:
    ldap_auth_disable_tls_never_use_in_production: True
```

Create your cluster with the following command.

```
$ pcluster create-cluster --cluster-name "ad-cluster" --cluster-configuration "./
ldaps_config.yaml"
{
  "cluster": {
    "clusterName": "pcluster",
    "cloudformationStackStatus": "CREATE_IN_PROGRESS",
    "cloudformationStackArn": "arn:aws:cloudformation:region-id:123456789012:stack/ad-
cluster/1234567-abcd-0123-def0-abcdef0123456",
    "region": "region-id",
    "version": 3.13.2,
    "clusterStatus": "CREATE_IN_PROGRESS"
  }
}
```

Connect to the cluster as a user

You can determine the status of the cluster with the following commands.

```
$ pcluster describe-cluster -n ad-cluster --region "region-id" --query "clusterStatus"
```

The output is as follows.

```
"CREATE_IN_PROGRESS" / "CREATE_COMPLETE"
```

When the status reaches "CREATE_COMPLETE", log in with the created user name and password.

```
$ HEAD_NODE_IP=$(pcluster describe-cluster -n "ad-cluster" --region "region-id" --query
 headNode.publicIpAddress | xargs echo)
```

```
$ ssh user000@$HEAD_NODE_IP
```

You can log in without the password by providing the SSH key that was created for the new user at /home/user000@HEAD_NODE_IP/.ssh/id_rsa.

If the ssh command succeeded, you have successfully connected to the cluster as a user that's authenticated to use the Active Directory (AD).

Clean up

1. **From your local machine, delete the cluster.**

   ```
   $ pcluster delete-cluster --cluster-name "ad-cluster" --region "region-id"
   {
     "cluster": {
       "clusterName": "ad-cluster",
       "cloudformationStackStatus": "DELETE_IN_PROGRESS",
       "cloudformationStackArn": "arn:aws:cloudformation:region-id:123456789012:stack/
   ad-cluster/1234567-abcd-0123-def0-abcdef0123456",
       "region": "region-id",
       "version": "3.13.2",
       "clusterStatus": "DELETE_IN_PROGRESS"
     }
   }
   ```

2. **Check the progress of the cluster being deleted.**

   ```
   $ pcluster describe-cluster --cluster-name "ad-cluster" --region "region-id" --
   query "clusterStatus"
   ```

```
"DELETE_IN_PROGRESS"
```

After the cluster is successfully deleted, proceed to the next step.

Automated

Delete the Active Directory resources

1. From https://console.aws.amazon.com/cloudformation/.
2. In the navigation pane, choose **Stacks**.
3. From the list of stacks, choose the AD stack (for example, `pcluster-ad`).
4. Choose **Delete**.

Manual

1. **Delete the Amazon EC2 instance.**
   a. From https://console.aws.amazon.com/ec2/, choose **Instances** in the navigation pane.
   b. From the list of instances, choose the instance that you created to add users to the directory.
   c. Choose **Instance state**, then **Terminate instance**.
2. **Delete the hosted zone.**
   a. Create a `recordset-delete.json` with the following content. In this example, HostedZoneId is the canonical hosted zone ID of the load balancer.

```
{
  "Changes": [
    {
      "Action": "DELETE",
      "ResourceRecordSet": {
        "Name": "corp.example.com",
        "Type": "A",
        "Region": "region-id",
        "SetIdentifier": "pcluster-active-directory",
        "AliasTarget": {
          "HostedZoneId": "Z2IFOLAFXWLO4F",
```

```
          "DNSName": "CorpExampleCom-NLB-3afe296bf4ba80d4.elb.region-
id.amazonaws.com",
          "EvaluateTargetHealth": true
        }
      }
    }
  ]
}
```

b. Submit the recordset change to the hosted zone using the hosted zone ID.

```
$ aws route53 change-resource-record-sets --hosted-zone-
id Z09020002B5MZQNXMSJUB \
  --change-batch file://recordset-delete.json
{
 "ChangeInfo": {
     "Id": "/change/C04853642A0TH2TJ5NLNI",
     "Status": "PENDING",
     "SubmittedAt": "2022-05-05T14:25:51.046000+00:00"
 }
}
```

c. Delete the hosted zone.

```
$ aws route53 delete-hosted-zone --id Z09020002B5MZQNXMSJUB
{
 "ChangeInfo": {
     "Id": "/change/C0468051QFABTVHMDEG9",
     "Status": "PENDING",
     "SubmittedAt": "2022-05-05T14:26:13.814000+00:00"
 }
}
```

3. **Delete the LB listener.**

```
$ aws elbv2 delete-listener \
  --listener-arn arn:aws:elasticloadbalancing:region-id:123456789012:listener/net/
CorpExampleCom-NLB/3afe296bf4ba80d4/a8f9d97318743d4b --region region-id
```

4. **Delete the target group.**

```
$ aws elbv2 delete-target-group \
```

```
  --target-group-arn arn:aws:elasticloadbalancing:region-
id:123456789012:targetgroup/CorpExampleCom-Targets/44577c583b695e81 --
region region-id
```

5. **Delete the load balancer.**

```
$ aws elbv2 delete-load-balancer \
  --load-balancer-arn arn:aws:elasticloadbalancing:region-
id:123456789012:loadbalancer/net/CorpExampleCom-NLB/3afe296bf4ba80d4 --
region region-id
```

6. **Delete the policy that the cluster uses to read the certificate from Secrets Manager.**

```
$ aws iam delete-policy --policy-arn arn:aws:iam::123456789012:policy/
ReadCertExample
```

7. **Delete the secret that contains the domain certificate.**

```
$ aws secretsmanager delete-secret \
  --secret-id arn:aws:secretsmanager:region-id:123456789012:secret:example-
cert-123abc \
  --region region-id
{
 "ARN": "arn:aws:secretsmanager:region-id:123456789012:secret:example-cert-123abc",
 "Name": "example-cert",
 "DeletionDate": "2022-06-04T16:27:36.183000+02:00"
}
```

8. **Delete the certificate from ACM.**

```
$ aws acm delete-certificate \
  --certificate-arn arn:aws:acm:region-
id:123456789012:certificate/343db133-490f-4077-b8d4-3da5bfd89e72 --region region-id
```

9. **Delete the Active Directory (AD) resources.**

   a. Get the following resource IDs from the output of the python script `ad.py`:

      - AD ID
      - AD subnet IDs
      - AD VPC ID

   b. Delete the directory by running the following command.

```
$ aws ds delete-directory --directory-id d-abcdef0123456789 --region region-id
{
   "DirectoryId": "d-abcdef0123456789"
}
```

c. List the security groups in the VPC.

```
$ aws ec2 describe-security-groups --filters '[{"Name":"vpc-id","Values":
["vpc-07614ade95ebad1bc"]}]' --region region-id
```

d. Delete the custom security group.

```
$ aws ec2 delete-security-group --group-id sg-021345abcdef6789 --region region-
id
```

e. Delete the subnets.

```
$ aws ec2 delete-subnet --subnet-id subnet-1234567890abcdef --region region-id
```

```
$ aws ec2 delete-subnet --subnet-id subnet-021345abcdef6789 --region region-id
```

f. Describe internet gateway.

```
$ aws ec2 describe-internet-gateways \
  --filters Name=attachment.vpc-id,Values=vpc-021345abcdef6789 \
  --region region-id
{
  "InternetGateways": [
    {
      "Attachments": [
        {
          "State": "available",
          "VpcId": "vpc-021345abcdef6789"
        }
      ],
      "InternetGatewayId": "igw-1234567890abcdef",
      "OwnerId": "123456789012",
      "Tags": []
    }
  ]
}
```

g. Detach the internet gateway.

```
$ aws ec2 detach-internet-gateway \
  --internet-gateway-id igw-1234567890abcdef \
  --vpc-id vpc-021345abcdef6789 \
  --region region-id
```

h. Delete the internet gateway.

```
$ aws ec2 delete-internet-gateway \
  --internet-gateway-id igw-1234567890abcdef \
  --region region-id
```

i. Delete the VPC.

```
$ aws ec2 delete-vpc \
  --vpc-id vpc-021345abcdef6789 \
  --region region-id
```

j. Delete the secret that contains the `ReadOnlyUser` password.

```
$ aws secretsmanager delete-secret \
  --secret-id arn:aws:secretsmanager:region-
id:123456789012:secret:ADSecretPassword-1234" \
  --region region-id
```

Configuring shared storage encryption with an AWS KMS key

Learn how to set up a customer managed AWS KMS key to encrypt and protect your data in the cluster file storage systems that are configured for AWS ParallelCluster.

When using the AWS ParallelCluster command line interface (CLI) or API, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

AWS ParallelCluster supports following shared storage configuration options:

- `SharedStorage` / `EbsSettings` / `KmsKeyId`
- `SharedStorage` / `EfsSettings` / `KmsKeyId`
- `SharedStorage` / `FsxLustreSettings` / `KmsKeyId`

You can use these options to provide a customer managed AWS KMS key for Amazon EBS, Amazon EFS, and FSx for Lustre shared storage system encryption. To use them, you must create and configure an IAM policy for the following:

- `HeadNode` / `Iam` / `AdditionalIamPolicies` / `Policy`
- `Scheduler` / `SlurmQueues` / `Iam` / `AdditionalIamPolicies` / `Policy`

Prerequisites

- AWS ParallelCluster is installed.
- The AWS CLI is installed and configured.
- You have an Amazon EC2 key pair.
- You have an IAM role with the permissions that are required to run the `pcluster` CLI.

Topics

Create the policy

In this tutorial, you will create a policy for configuring shared storage encryption with an AWS KMS key.

Create a policy.

1. Go to the IAM Console: https://console.aws.amazon.com/iam/home.
2. Choose **Policies**.
3. Choose **Create policy**.
4. Choose the **JSON** tab and paste in the following policy. Make sure to replace all occurrences of *123456789012* with your AWS account ID and the key Amazon Resource Name (ARN) and AWS Region with that of your own.

   JSON

```
{
    "Version": "2012-10-17",
```

```
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "kms:DescribeKey",
                "kms:ReEncrypt*",
                "kms:CreateGrant",
                "kms:Decrypt"
            ],
            "Resource": [
                "arn:aws:kms:us-east-1:123456789012:key/abcd1234-ef56-gh78-
ij90-abcd1234efgh5678"
            ]
        }
    ]
}
```

5. For this tutorial, name the policy `ParallelClusterKmsPolicy`, and then choose **Create Policy**.
6. Make a note of the policy ARN. You need it to configure your cluster.

Configure and create the cluster

The following is an example cluster configuration that includes an Amazon Elastic Block Store shared file system with encryption.

```
Region: eu-west-1
Image:
  Os: alinux2
HeadNode:
  InstanceType: t2.micro
  Networking:
    SubnetId: subnet-abcdef01234567890
  Ssh:
    KeyName: my-ssh-key
  Iam:
    AdditionalIamPolicies:
      - Policy: arn:aws:iam::123456789012:policy/ParallelClusterKmsPolicy
Scheduling:
  Scheduler: slurm
  SlurmQueues:
```

```
    - Name: q1
      ComputeResources:
        - Name: t2micro
          InstanceType: t2.micro
          MinCount: 0
          MaxCount: 10
      Networking:
        SubnetIds:
          - subnet-abcdef01234567890
      Iam:
        AdditionalIamPolicies:
          - Policy: arn:aws:iam::123456789012:policy/ParallelClusterKmsPolicy
SharedStorage:
  - MountDir: /shared/ebs1
    Name: shared-ebs1
    StorageType: Ebs
    EbsSettings:
      Encrypted: True
      KmsKeyId: abcd1234-ef56-gh78-ij90-abcd1234efgh5678
```

Replace the items in red text with your own values. Then, create a cluster that uses your AWS KMS key to encrypt your data in Amazon EBS.

The configuration is similar for Amazon EFS and FSx for Lustre file systems.

The Amazon EFS `SharedStorage` configuration is as follows.

```
...
SharedStorage:
  - MountDir: /shared/efs1
    Name: shared-efs1
    StorageType: Efs
    EfsSettings:
      Encrypted: True
      KmsKeyId: abcd1234-ef56-gh78-ij90-abcd1234efgh5678
```

The FSx for Lustre `SharedStorage` configuration is as follows.

```
...
SharedStorage:
  - MountDir: /shared/fsx1
    Name: shared-fsx1
```

```
    StorageType: FsxLustre
    FsxLustreSettings:
      StorageCapacity: 1200
      DeploymentType: PERSISTENT_1
      PerUnitStorageThroughput: 200
      KmsKeyId: abcd1234-ef56-gh78-ij90-abcd1234efgh5678
```

Running jobs in a multiple queue mode cluster

This tutorial covers how to run your first "Hello World" job on AWS ParallelCluster with multiple queue mode.

When using the AWS ParallelCluster command line interface (CLI) or API, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

Prerequisites

- AWS ParallelCluster is installed.
- The AWS CLI is installed and configured.
- You have an Amazon EC2 key pair.
- You have an IAM role with the permissions that are required to run the `pcluster` CLI.

Configure your cluster

First, verify that AWS ParallelCluster is correctly installed by running the following command.

```
$ pcluster version
```

For more information about `pcluster version`, see `pcluster version`.

This command returns the running version of AWS ParallelCluster.

Next, run `pcluster configure` to generate a basic configuration file. Follow all the prompts that follow this command.

```
$ pcluster configure --config multi-queue-mode.yaml
```

For more information about the `pcluster configure` command, see `pcluster configure`.

After you complete this step, a basic configuration file named `multi-queue-mode.yaml` appears. This file contains a basic cluster configuration.

In the next step, you modify your new configuration file and launch a cluster with multiple queues.

> **Note**
>
> Some instances that this tutorial uses aren't free-tier eligible.

For this tutorial, modify your configuration file to match the following configuration. The items that are highlighted in red represent your configuration file values. Keep your own values.

```
Region: region-id
Image:
 Os: alinux2
HeadNode:
 InstanceType: c5.xlarge
 Networking:
   SubnetId: subnet-abcdef01234567890
 Ssh:
   KeyName: yourkeypair
Scheduling:
 Scheduler: slurm
 SlurmQueues:
 - Name: spot
   ComputeResources:
   - Name: c5xlarge
     InstanceType: c5.xlarge
     MinCount: 1
     MaxCount: 10
   - Name: t2micro
     InstanceType: t2.micro
     MinCount: 1
     MaxCount: 10
   Networking:
     SubnetIds:
     - subnet-abcdef01234567890
 - Name: ondemand
   ComputeResources:
   - Name: c52xlarge
```

```
      InstanceType: c5.2xlarge
      MinCount: 0
      MaxCount: 10
    Networking:
      SubnetIds:
      - subnet-021345abcdef6789
```

Create your cluster

Create a cluster that's named `multi-queue-cluster` based on your configuration file.

```
$ pcluster create-cluster --cluster-name multi-queue-cluster --cluster-configuration
 multi-queue-mode.yaml
{
 "cluster": {
   "clusterName": "multi-queue-cluster",
   "cloudformationStackStatus": "CREATE_IN_PROGRESS",
   "cloudformationStackArn": "arn:aws:cloudformation:eu-west-1:123456789012:stack/
multi-queue-cluster/1234567-abcd-0123-def0-abcdef0123456",
   "region": "eu-west-1",
   "version": "3.13.2",
   "clusterStatus": "CREATE_IN_PROGRESS"
 }
}
```

For more information about the `pcluster create-cluster` command, see [pcluster create-cluster].

To check the status of the cluster, run the following command.

```
$ pcluster list-clusters
{
 "cluster": {
   "clusterName": "multi-queue-cluster",
   "cloudformationStackStatus": "CREATE_IN_PROGRESS",
   "cloudformationStackArn": "arn:aws:cloudformation:eu-west-1:123456789012:stack/
multi-queue-cluster/1234567-abcd-0123-def0-abcdef0123456",
   "region": "eu-west-1",
   "version": "3.13.2",
   "clusterStatus": "CREATE_IN_PROGRESS"
 }
}
```

When the cluster is created, the `clusterStatus` field shows `CREATE_COMPLETE`.

Log in to the head node

Use your private SSH key file to log in to the head node.

```
$ pcluster ssh --cluster-name multi-queue-cluster -i ~/path/to/yourkeyfile.pem
```

For more information about `pcluster ssh`, see `pcluster ssh`.

After logging in, run the `sinfo` command to verify that your scheduler queues are set up and configured.

For more information about `sinfo`, see sinfo in the *Slurm documentation*.

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
spot*        up   infinite     18  idle~ spot-dy-c5xlarge-[1-9],spot-dy-t2micro-[1-9]
spot*        up   infinite      2  idle  spot-st-c5xlarge-1,spot-st-t2micro-1
ondemand     up   infinite     10  idle~ ondemand-dy-c52xlarge-[1-10]
```

The output shows that you have one `t2.micro` and one `c5.xlarge` compute node in the `idle` state that are available in your cluster.

Other nodes are all in the power saving state, indicated by the ~ suffix in node state, with no Amazon EC2 instances backing them. The default queue is indicated by a * suffix after its queue name. `spot` is your default job queue.

Run job in multiple queue mode

Next, try to run a job to sleep for a while. The job later outputs its own hostname. Make sure that this script can be run by the current user.

```
$ tee <<EOF hellojob.sh
#!/bin/bash
sleep 30
echo "Hello World from \$(hostname)"
EOF

$ chmod +x hellojob.sh
$ ls -l hellojob.sh
```

```
-rwxrwxr-x 1 ec2-user ec2-user 57 Sep 23 21:57 hellojob.sh
```

Submit the job using the `sbatch` command. Request two nodes for this job with the `-N 2` option, and verify that the job submits successfully. For more information about `sbatch`, see sbatch in the *Slurm documentation*.

```
$ sbatch -N 2 --wrap "srun hellojob.sh"
Submitted batch job 1
```

You can view your queue and check the status of the job with the `squeue` command. Because you didn't specify a specific queue, the default queue (`spot`) is used. For more information about `squeue`, see squeue in the *Slurm documentation*.

```
$ squeue
JOBID PARTITION     NAME     USER  ST       TIME  NODES NODELIST(REASON)
   1      spot     wrap ec2-user  R       0:10      2 spot-st-c5xlarge-1,spot-st-
t2micro-1
```

The output shows that the job is currently in a running state. Wait for the job to finish. This takes about 30 seconds. Then, run `squeue` again.

```
$ squeue
JOBID PARTITION     NAME     USER          ST       TIME  NODES NODELIST(REASON)
```

Now that the jobs in the queue have all finished, look for the output file that's named `slurm-1.out` in your current directory.

```
$ cat slurm-1.out
Hello World from spot-st-t2micro-1
Hello World from spot-st-c5xlarge-1
```

The output shows that the job ran successfully on the `spot-st-t2micro-1` and `spot-st-c5xlarge-1` nodes.

Now submit the same job by specifying constraints for specific instances with the following commands.

```
$ sbatch -N 3 -p spot -C "[c5.xlarge*1&t2.micro*2]" --wrap "srun hellojob.sh"
Submitted batch job 2
```

You used these parameters for `sbatch`:

- `-N 3`– requests three nodes.
- `-p spot`– submits the job to the `spot` queue. You can also submit a job to the ondemand queue by specifying `-p ondemand`.
- `-C "[c5.xlarge*1&t2.micro*2]"`– specifies the specific node constraints for this job. This requests one `c5.xlarge` node and two `t2.micro` nodes to be used for this job.

Run the `sinfo` command to view the nodes and queues. Queues in AWS ParallelCluster are called partitions in Slurm.

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
spot*        up   infinite      1  alloc# spot-dy-t2micro-1
spot*        up   infinite     17  idle~ spot-dy-c5xlarge-[2-10],spot-dy-t2micro-[2-9]
spot*        up   infinite      1  mix   spot-st-c5xlarge-1
spot*        up   infinite      1  alloc spot-st-t2micro-1
ondemand     up   infinite     10  idle~ ondemand-dy-c52xlarge-[1-10]
```

The nodes are powering up. This is indicated by the # suffix on the node state. Run the squeue command to view information about the jobs in the cluster.

```
$ squeue
JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
   2      spot     wrap ec2-user CF       0:04      3 spot-dy-c5xlarge-1,spot-dy-
t2micro-1,spot-st-t2micro-1
```

Your job is in the CF (CONFIGURING) state, waiting for instances to scale up and join the cluster.

After about three minutes, the nodes are available and the job enters the R (RUNNING) state.

```
$ squeue
JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
   2      spot     wrap ec2-user  R       0:07      3 spot-dy-t2micro-1,spot-st-
c5xlarge-1,spot-st-t2micro-1
```

The job finishes, and all three nodes are in the `idle` state.

```
$ squeue
```

```
JOBID PARTITION     NAME     USER ST       TIME  NODES NODELIST(REASON)
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
spot*        up   infinite     17  idle~ spot-dy-c5xlarge-[1-9],spot-dy-t2micro-[2-9]
spot*        up   infinite      3  idle  spot-dy-t2micro-1,spot-st-c5xlarge-1,spot-st-
t2micro-1
ondemand     up   infinite     10  idle~ ondemand-dy-c52xlarge-[1-10]
```

Then, after no jobs remain in the queue, check for `slurm-2.out` in your local directory.

```
$ cat slurm-2.out
Hello World from spot-st-t2micro-1
Hello World from spot-dy-t2micro-1
Hello World from spot-st-c5xlarge-1
```

This is the final state of the cluster.

```
$ sinfo
PARTITION AVAIL  TIMELIMIT  NODES  STATE NODELIST
spot*        up   infinite     17  idle~ spot-dy-c5xlarge-[1-9],spot-dy-t2micro-[2-9]
spot*        up   infinite      3  idle  spot-dy-t2micro-1,spot-st-c5xlarge-1,spot-st-
t2micro-1
ondemand     up   infinite     10  idle~ ondemand-dy-c52xlarge-[1-10]
```

After logging off of the cluster, you can clean up by running `pcluster delete-cluster`. For more information, see `pcluster list-clusters` and `pcluster delete-cluster`.

```
$ pcluster list-clusters
{
 "clusters": [
   {
     "clusterName": "multi-queue-cluster",
     "cloudformationStackStatus": "CREATE_COMPLETE",
     "cloudformationStackArn": "arn:aws:cloudformation:eu-west-1:123456789012:stack/
multi-queue-cluster/1234567-abcd-0123-def0-abcdef0123456",
     "region": "eu-west-1",
     "version": "3.1.4",
     "clusterStatus": "CREATE_COMPLETE"
   }
 ]
}
$ pcluster delete-cluster -n multi-queue-cluster
```

```
{
 "cluster": {
   "clusterName": "multi-queue-cluster",
   "cloudformationStackStatus": "DELETE_IN_PROGRESS",
   "cloudformationStackArn": "arn:aws:cloudformation:eu-west-1:123456789012:stack/
multi-queue-cluster/1234567-abcd-0123-def0-abcdef0123456",
   "region": "eu-west-1",
   "version": "3.1.4",
   "clusterStatus": "DELETE_IN_PROGRESS"
 }
}
```

Using the AWS ParallelCluster API

In this tutorial, you build and test the API with Amazon API Gateway and an AWS ParallelCluster CloudFormation template. Then, you use the example client available on GitHub to use the API. For more information about using the API, see the AWS ParallelCluster API.

For more information, see Create a custom component with Image Builder in the *EC2 Image Builder User Guide*.

When using the AWS ParallelCluster command line interface (CLI) or API, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

Prerequisites

- The AWS CLI is installed and configured in your compute environment.
- AWS ParallelCluster is installed in a virtual environment. For more information, see Install AWS ParallelCluster in a virtual environment (recommended).
- You have an Amazon EC2 key pair.
- You have an IAM role with the permissions that are required to run the `pcluster` CLI.

Step 1: Build the API with Amazon API Gateway

Stay in your home user directory and activate your virtual environment:

1. Install a helpful JSON command line processor.

```
$ sudo yum groupinstall -y "Development Tools"
 sudo yum install -y jq python3-devel
```

2. Run the following command to get your AWS ParallelCluster version and assign it to an environment variable.

```
$ PCLUSTER_VERSION=$(pcluster version | jq -r '.version')
 echo "export PCLUSTER_VERSION=${PCLUSTER_VERSION}" |tee -a ~/.bashrc
```

3. Create an environment variable and assign your Region ID to it.

```
$ export AWS_DEFAULT_REGION="us-east-1"
 echo "export AWS_DEFAULT_REGION=${AWS_DEFAULT_REGION}" |tee -a ~/.bashrc
```

4. Run the following commands to deploy the API.

```
API_STACK_NAME="pc-api-stack"
 echo "export API_STACK_NAME=${API_STACK_NAME}" |tee -a ~/.bashrc
```

```
aws cloudformation create-stack \
   --region ${AWS_DEFAULT_REGION} \
   --stack-name ${API_STACK_NAME} \
   --template-url https://${AWS_DEFAULT_REGION}-aws-parallelcluster.s3.
${AWS_DEFAULT_REGION}.amazonaws.com/parallelcluster/${PCLUSTER_VERSION}/api/
parallelcluster-api.yaml \
   --capabilities CAPABILITY_NAMED_IAM CAPABILITY_AUTO_EXPAND \
   --parameters ParameterKey=EnableIamAdminAccess,ParameterValue=true

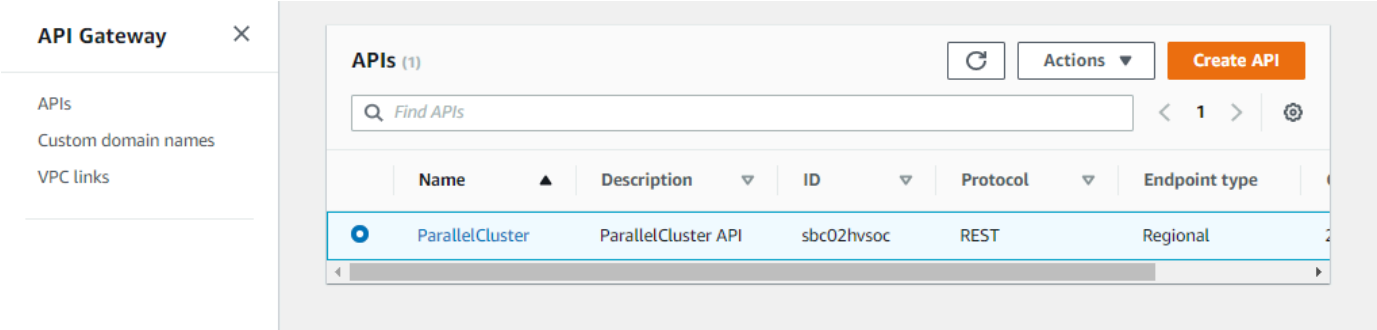
    {
       "StackId": "arn:aws:cloudformation:us-east-1:123456789012:stack/my-api-
stack/abcd1234-ef56-gh78-ei90-1234abcd5678"
    }
```

After the process completes, proceed to the next step.

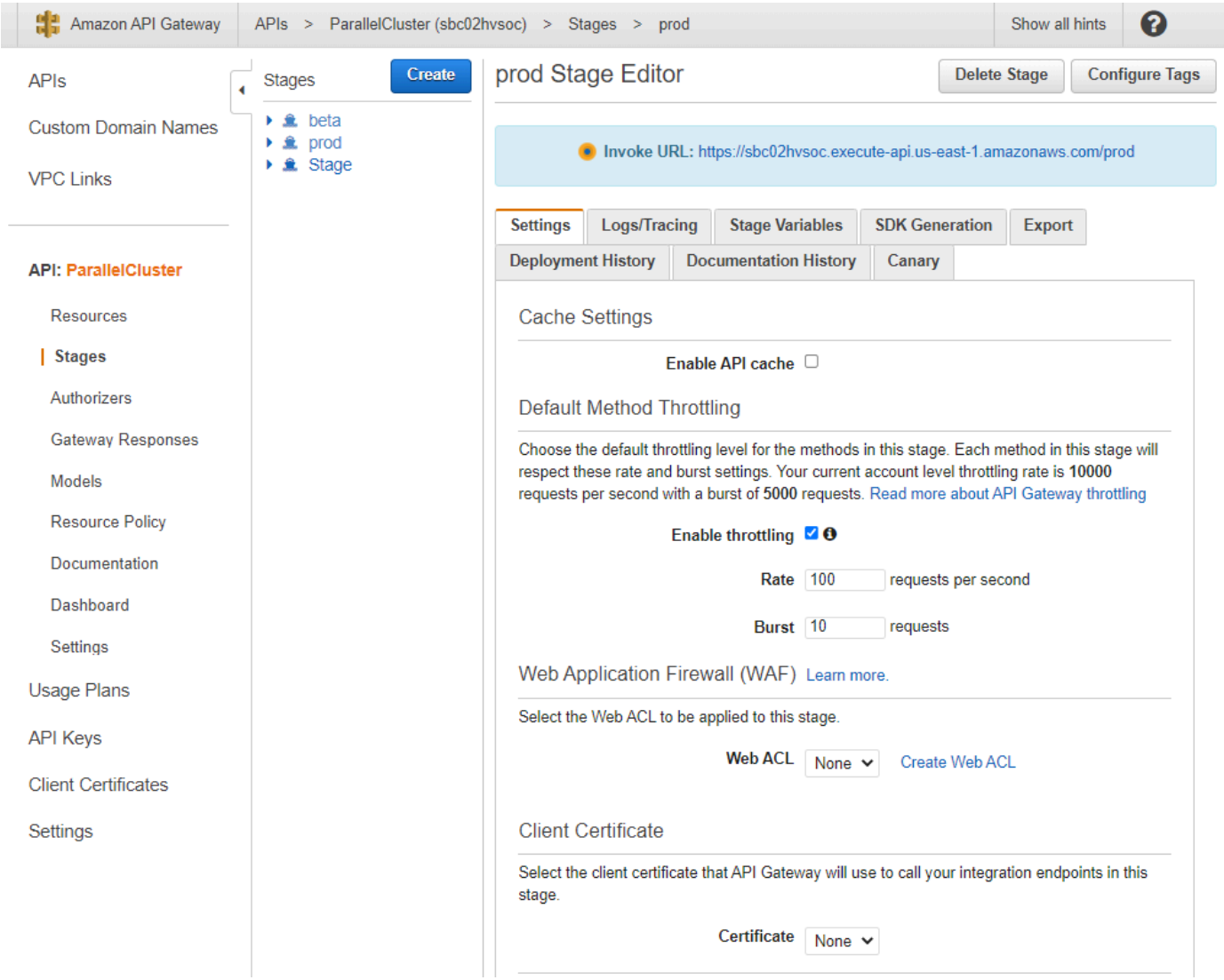
Step 2: Test the API in the Amazon API Gateway console

1. Sign in to the AWS Management Console.
2. Navigate to the Amazon API Gateway console.

3. Choose your API deployment.



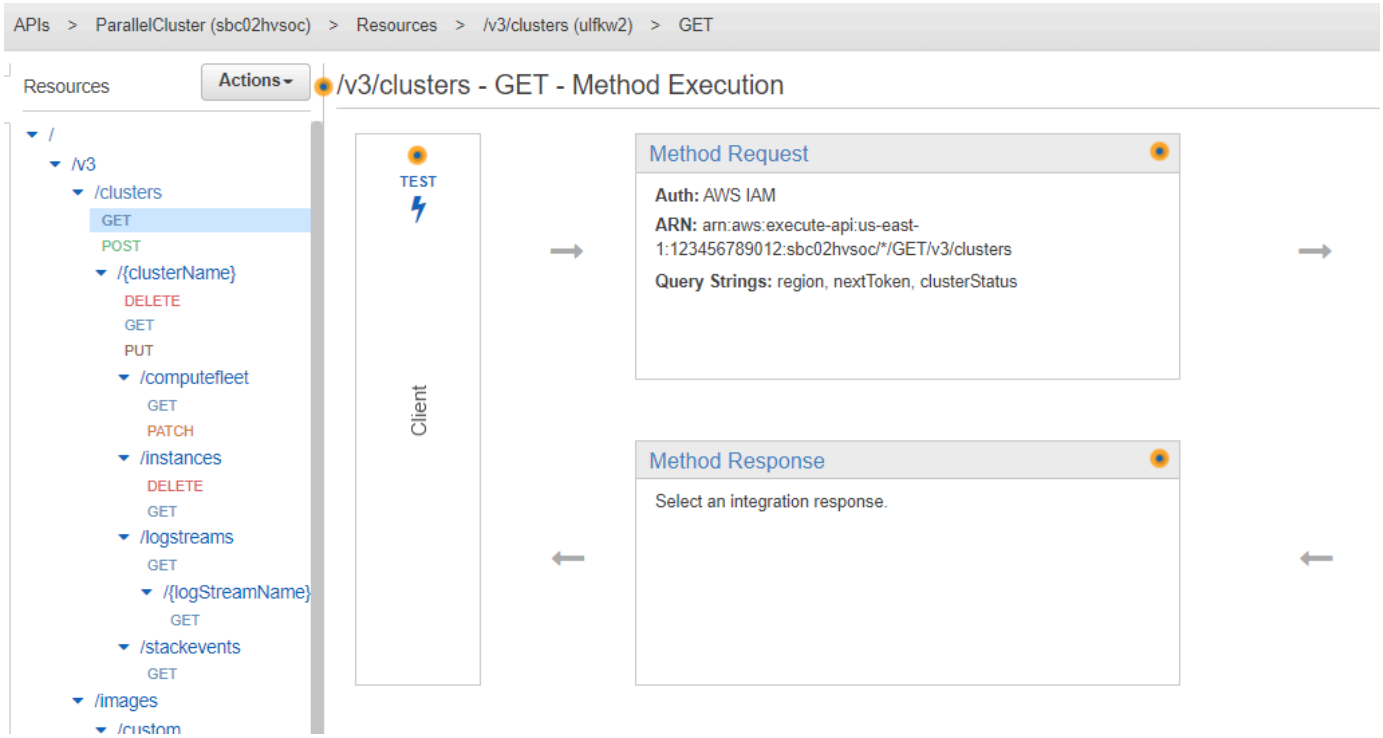
4. Choose **Stages** and select a stage.



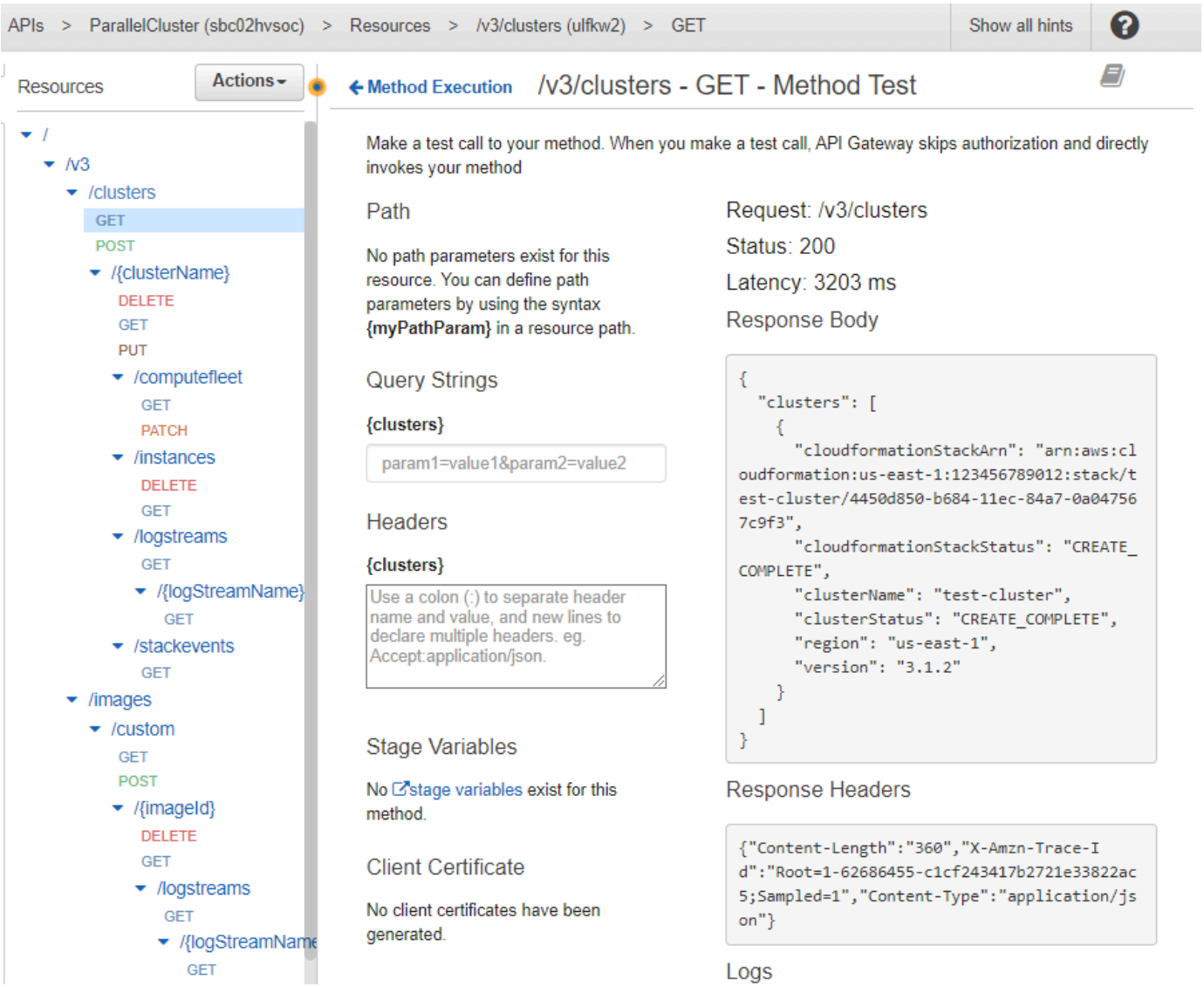
5. Note the URL that API Gateway provides for accessing or invoking your API. It's highlighted in blue.

6. Choose **Resources**, and select **`GET`** under **`/clusters`**.

7. Choose the **TEST** icon and then scroll down and choose **TEST** icon.



The response to your `/clusters GET` appears.



Step 3: Prepare and test an example client to invoke the API

Clone the AWS ParallelCluster source code, `cd` to the `api` directory, and install the Python client libraries.

1.
```
$ git clone -b v${PCLUSTER_VERSION} https://github.com/aws/aws-parallelcluster aws-parallelcluster-v${PCLUSTER_VERSION}
 cd aws-parallelcluster-v${PCLUSTER_VERSION}/api
```

```
$ pip3 install client/src
```

2. Navigate back to your home user directory.
3. Export the API Gateway base URL that the client uses when running.

```
$ export PCLUSTER_API_URL=$( aws cloudformation describe-stacks
 --stack-name ${API_STACK_NAME} --query 'Stacks[0].Outputs[?
OutputKey==`ParallelClusterApiInvokeUrl`].OutputValue' --output text )
 echo "export PCLUSTER_API_URL=${PCLUSTER_API_URL}" |tee -a ~/.bashrc
```

4. Export a cluster name that the client uses to create a cluster.

```
$ export CLUSTER_NAME="test-api-cluster"
 echo "export CLUSTER_NAME=${CLUSTER_NAME}" |tee -a ~/.bashrc
```

5. Run the following commands to store the credentials that the example client uses to access the API.

```
$ export PCLUSTER_API_USER_ROLE=$( aws cloudformation describe-
stacks --stack-name ${API_STACK_NAME} --query 'Stacks[0].Outputs[?
OutputKey==`ParallelClusterApiUserRole`].OutputValue' --output text )
 echo "export PCLUSTER_API_USER_ROLE=${PCLUSTER_API_USER_ROLE}" |tee -a ~/.bashrc
```

Step 4: Copy client code script and run cluster tests

1. Copy the following example client code to `test_pcluster_client.py` in your home user directory. The client code makes requests to do the following:
   - Create the cluster.
   - Describe the cluster.
   - List the clusters.
   - Describe the compute fleet.
   - Describe the cluster instances.

```
# Copyright 2021 Amazon.com, Inc. or its affiliates. All Rights Reserved.
# SPDX-License-Identifier: MIT-0
#
# Permission is hereby granted, free of charge, to any person obtaining a copy of
 this
```

```
# software and associated documentation files (the "Software"), to deal in the
 Software
# without restriction, including without limitation the rights to use, copy,
 modify,
# merge, publish, distribute, sublicense, and/or sell copies of the Software, and
 to
# permit persons to whom the Software is furnished to do so.
#
# THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR
 IMPLIED,
# INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A
# PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR
 COPYRIGHT
# HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION
# OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE
# SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.
#
#  Author: Evan F. Bollig (Github: bollig)

import time, datetime
import os
import pcluster_client
from pprint import pprint
from pcluster_client.api import (
    cluster_compute_fleet_api,
    cluster_instances_api,
    cluster_operations_api
)
from pcluster_client.model.create_cluster_request_content import
 CreateClusterRequestContent
from pcluster_client.model.cluster_status import ClusterStatus
region=os.environ.get("AWS_DEFAULT_REGION")

# Defining the host is optional and defaults to http://localhost
# See configuration.py for a list of all supported configuration parameters.
configuration = pcluster_client.Configuration(
    host = os.environ.get("PCLUSTER_API_URL")
)
cluster_name=os.environ.get("CLUSTER_NAME")

# Enter a context with an instance of the API client
with pcluster_client.ApiClient(configuration) as api_client:
    cluster_ops = cluster_operations_api.ClusterOperationsApi(api_client)
    fleet_ops = cluster_compute_fleet_api.ClusterComputeFleetApi(api_client)
```

```
    instance_ops = cluster_instances_api.ClusterInstancesApi(api_client)

    # Create cluster
    build_done = False
    try:
        with open('cluster-config.yaml', encoding="utf-8") as f:
            body = CreateClusterRequestContent(cluster_name=cluster_name,
cluster_configuration=f.read())
            api_response = cluster_ops.create_cluster(body, region=region)
    except pcluster_client.ApiException as e:
        print("Exception when calling create_cluster: %s\n" % e)
        build_done = True
    time.sleep(60)

    # Confirm cluster status with describe_cluster
    while not build_done:
        try:
            api_response = cluster_ops.describe_cluster(cluster_name,
region=region)
            pprint(api_response)
            if api_response.cluster_status == ClusterStatus('CREATE_IN_PROGRESS'):
                print('. . . working . . .', end='', flush=True)
                time.sleep(60)
            elif api_response.cluster_status == ClusterStatus('CREATE_COMPLETE'):
                print('READY!')
                build_done = True
            else:
                print('ERROR!!!!')
                build_done = True
        except pcluster_client.ApiException as e:
            print("Exception when calling describe_cluster: %s\n" % e)

    # List clusters
    try:
        api_response = cluster_ops.list_clusters(region=region)
        pprint(api_response)
    except pcluster_client.ApiException as e:
        print("Exception when calling list_clusters: %s\n" % e)

    # DescribeComputeFleet
    try:
        api_response = fleet_ops.describe_compute_fleet(cluster_name,
region=region)
        pprint(api_response)
```

```
    except pcluster_client.ApiException as e:
        print("Exception when calling compute fleet: %s\n" % e)

    # DescribeClusterInstances
    try:
        api_response = instance_ops.describe_cluster_instances(cluster_name,
 region=region)
        pprint(api_response)
    except pcluster_client.ApiException as e:
        print("Exception when calling describe_cluster_instances: %s\n" % e)
```

2. Create a cluster configuration.

```
$ pcluster configure --config cluster-config.yaml
```

3. The API Client library automatically detects configuration details from your environment variables (for example, AWS_ACCESS_KEY_ID, AWS_SECRET_ACCESS_KEY, or AWS_SESSION_TOKEN) or $HOME/.aws. The following command switches your current IAM role to the designated ParallelClusterApiUserRole.

```
$  eval $(aws sts assume-role --role-arn ${PCLUSTER_API_USER_ROLE} --role-
session-name ApiTestSession | jq -r '.Credentials | "export AWS_ACCESS_KEY_ID=
\(.AccessKeyId)\nexport AWS_SECRET_ACCESS_KEY=\(.SecretAccessKey)\nexport
 AWS_SESSION_TOKEN=\(.SessionToken)\n"')
```

Error to watch for:

If you see an error similar to the following, you already assumed the ParallelClusterApiUserRole and your AWS_SESSION_TOKEN has expired.

```
An error occurred (AccessDenied) when calling the AssumeRole operation:
User: arn:aws:sts::XXXXXXXXXXXX:assumed-role/ParallelClusterApiUserRole-XXXXXXXX-
XXXX-XXXX-XXXX-XXXXXXXXXXXX/ApiTestSession
is not authorized to perform: sts:AssumeRole on resource:
 arn:aws:iam::XXXXXXXXXXXX:role/ParallelClusterApiUserRole-XXXXXXXX-XXXX-XXXX-XXXX-
XXXXXXXXXXXX
```

Drop the role and then re-run the `aws sts assume-role` command to use the ParallelClusterApiUserRole.

```
$ unset AWS_SESSION_TOKEN
```

```
unset AWS_SECRET_ACCESS_KEY
unset AWS_ACCESS_KEY_ID
```

To provide your current user permissions for API access, you must expand the Resource Policy.

4. Run the following command to test the example client.

```
$ python3 test_pcluster_client.py
{'cluster_configuration': 'Region: us-east-1\n'
                          'Image:\n'
                          '  Os: alinux2\n'
                          'HeadNode:\n'
                          '  InstanceType: t2.micro\n'
                          '  Networking . . . :\n'
                          '    SubnetId: subnet-1234567890abcdef0\n'
                          '  Ssh:\n'
                          '    KeyName: adpc\n'
                          'Scheduling:\n'
                          '  Scheduler: slurm\n'
                          '  SlurmQueues:\n'
                          '  - Name: queue1\n'
                          '    ComputeResources:\n'
                          '    - Name: t2micro\n'
                          '      InstanceType: t2.micro\n'
                          '      MinCount: 0\n'
                          '      MaxCount: 10\n'
                          '    Networking . . . :\n'
                          '      SubnetIds:\n'
                          '      - subnet-1234567890abcdef0\n',
 'cluster_name': 'test-api-cluster'}
{'cloud_formation_stack_status': 'CREATE_IN_PROGRESS',
 'cloudformation_stack_arn': 'arn:aws:cloudformation:us-east-1:123456789012:stack/
test-api-cluster/abcd1234-ef56-gh78-ij90-1234abcd5678',
 'cluster_configuration': {'url': 'https://parallelcluster-021345abcdef6789-v1-do-
not-delete...},
 'cluster_name': 'test-api-cluster',
 'cluster_status': 'CREATE_IN_PROGRESS',
 'compute_fleet_status': 'UNKNOWN',
 'creation_time': datetime.datetime(2022, 4, 28, 16, 18, 47, 972000,
 tzinfo=tzlocal()),
 'last_updated_time': datetime.datetime(2022, 4, 28, 16, 18, 47, 972000,
 tzinfo=tzlocal()),
 'region': 'us-east-1',
 'tags': [{'key': 'parallelcluster:version', 'value': '3.1.3'}],
```

```
 'version': '3.1.3'}
        .
        .
        .
. . . working . . . {'cloud_formation_stack_status': 'CREATE_COMPLETE',
 'cloudformation_stack_arn': 'arn:aws:cloudformation:us-east-1:123456789012:stack/
test-api-cluster/abcd1234-ef56-gh78-ij90-1234abcd5678',
 'cluster_configuration': {'url': 'https://parallelcluster-021345abcdef6789-v1-do-
not-delete...},
 'cluster_name': 'test-api-cluster',
 'cluster_status': 'CREATE_COMPLETE',
 'compute_fleet_status': 'RUNNING',
 'creation_time': datetime.datetime(2022, 4, 28, 16, 18, 47, 972000,
 tzinfo=tzlocal()),
 'head_node': {'instance_id': 'i-abcdef01234567890',
               'instance_type': 't2.micro',
               'launch_time': datetime.datetime(2022, 4, 28, 16, 21, 46,
 tzinfo=tzlocal()),
               'private_ip_address': '172.31.27.153',
               'public_ip_address': '52.90.156.51',
               'state': 'running'},
 'last_updated_time': datetime.datetime(2022, 4, 28, 16, 18, 47, 972000,
 tzinfo=tzlocal()),
 'region': 'us-east-1',
 'tags': [{'key': 'parallelcluster:version', 'value': '3.1.3'}],
 'version': '3.1.3'}
READY!
```

Step 5: Copy client code script and delete cluster

1. Copy the following example client code to `delete_cluster_client.py`. The client code makes a request to delete the cluster.

```
# Copyright 2021 Amazon.com, Inc. or its affiliates. All Rights Reserved.
# SPDX-License-Identifier: MIT-0
#
# Permission is hereby granted, free of charge, to any person obtaining a copy of
 this
# software and associated documentation files (the "Software"), to deal in the
 Software
# without restriction, including without limitation the rights to use, copy,
 modify,
```

```
# merge, publish, distribute, sublicense, and/or sell copies of the Software, and
 to
# permit persons to whom the Software is furnished to do so.
#
# THE SOFTWARE IS PROVIDED "AS IS", WITHOUT WARRANTY OF ANY KIND, EXPRESS OR
 IMPLIED,
# INCLUDING BUT NOT LIMITED TO THE WARRANTIES OF MERCHANTABILITY, FITNESS FOR A
# PARTICULAR PURPOSE AND NONINFRINGEMENT. IN NO EVENT SHALL THE AUTHORS OR
 COPYRIGHT
# HOLDERS BE LIABLE FOR ANY CLAIM, DAMAGES OR OTHER LIABILITY, WHETHER IN AN ACTION
# OF CONTRACT, TORT OR OTHERWISE, ARISING FROM, OUT OF OR IN CONNECTION WITH THE
# SOFTWARE OR THE USE OR OTHER DEALINGS IN THE SOFTWARE.
#
#  Author: Evan F. Bollig (Github: bollig)

import time, datetime
import os
import pcluster_client
from pprint import pprint
from pcluster_client.api import (
    cluster_compute_fleet_api,
    cluster_instances_api,
    cluster_operations_api
)
from pcluster_client.model.create_cluster_request_content import
 CreateClusterRequestContent
from pcluster_client.model.cluster_status import ClusterStatus
region=os.environ.get("AWS_DEFAULT_REGION")

# Defining the host is optional and defaults to http://localhost
# See configuration.py for a list of all supported configuration parameters.
configuration = pcluster_client.Configuration(
    host = os.environ.get("PCLUSTER_API_URL")
)
cluster_name=os.environ.get("CLUSTER_NAME")

# Enter a context with an instance of the API client
with pcluster_client.ApiClient(configuration) as api_client:
    cluster_ops = cluster_operations_api.ClusterOperationsApi(api_client)

    # Delete the cluster
    gone = False
    try:
        api_response = cluster_ops.delete_cluster(cluster_name, region=region)
```

```
    except pcluster_client.ApiException as e:
        print("Exception when calling delete_cluster: %s\n" % e)
    time.sleep(60)

    # Confirm cluster status with describe_cluster
    while not gone:
        try:
            api_response = cluster_ops.describe_cluster(cluster_name,
 region=region)
            pprint(api_response)
            if api_response.cluster_status == ClusterStatus('DELETE_IN_PROGRESS'):
                print('. . . working . . .', end='', flush=True)
                time.sleep(60)
        except pcluster_client.ApiException as e:
            gone = True
            print("DELETE COMPLETE or Exception when calling describe_cluster: %s
\n" % e)
```

2. Run the following command to delete the cluster.

```
$ python3 delete_cluster_client.py
{'cloud_formation_stack_status': 'DELETE_IN_PROGRESS',
'cloudformation_stack_arn': 'arn:aws:cloudformation:us-east-1:123456789012:stack/
test-api-cluster/abcd1234-ef56-gh78-ij90-1234abcd5678',
'cluster_configuration': {'url': 'https://parallelcluster-021345abcdef6789-v1-do-
not-delete...},
'cluster_name': 'test-api-cluster',
'cluster_status': 'DELETE_IN_PROGRESS',
'compute_fleet_status': 'UNKNOWN',
'creation_time': datetime.datetime(2022, 4, 28, 16, 50, 47, 943000,
 tzinfo=tzlocal()),
'head_node': {'instance_id': 'i-abcdef01234567890',
              'instance_type': 't2.micro',
              'launch_time': datetime.datetime(2022, 4, 28, 16, 53, 48,
 tzinfo=tzlocal()),
              'private_ip_address': '172.31.17.132',
              'public_ip_address': '34.201.100.37',
              'state': 'running'},
'last_updated_time': datetime.datetime(2022, 4, 28, 16, 50, 47, 943000,
 tzinfo=tzlocal()),
'region': 'us-east-1',
'tags': [{'key': 'parallelcluster:version', 'value': '3.1.3'}],
'version': '3.1.3'}
       .
```

```
        .
        .
. . . working . . . {'cloud_formation_stack_status': 'DELETE_IN_PROGRESS',
'cloudformation_stack_arn': 'arn:aws:cloudformation:us-east-1:123456789012:stack/
test-api-cluster/abcd1234-ef56-gh78-ij90-1234abcd5678',
'cluster_configuration': {'url': 'https://parallelcluster-021345abcdef6789-v1-do-
not-delete...},
'cluster_name': 'test-api-cluster',
'cluster_status': 'DELETE_IN_PROGRESS',
'compute_fleet_status': 'UNKNOWN',
'creation_time': datetime.datetime(2022, 4, 28, 16, 50, 47, 943000,
 tzinfo=tzlocal()),
'last_updated_time': datetime.datetime(2022, 4, 28, 16, 50, 47, 943000,
 tzinfo=tzlocal()),
'region': 'us-east-1',
'tags': [{'key': 'parallelcluster:version', 'value': '3.1.3'}],
'version': '3.1.3'}
. . . working . . . DELETE COMPLETE or Exception when calling describe_cluster:
 (404)
Reason: Not Found
        .
        .
        .
HTTP response body: {"message":"Cluster 'test-api-cluster' does not exist or
 belongs to an incompatible ParallelCluster major version."}
```

3. After you are finished testing, unset the environment variables.

```
$ unset AWS_SESSION_TOKEN
unset AWS_SECRET_ACCESS_KEY
unset AWS_ACCESS_KEY_ID
```

Step 6: Clean up

You can use the AWS Management Console or AWS CLI to delete your API.

1. From the AWS CloudFormation console, choose the API stack and then choose **Delete**.
2. Run the following command if using the AWS CLI.

   Using AWS CloudFormation.

```
$ aws cloudformation delete-stack --stack-name ${API_STACK_NAME}
```

Creating a cluster with Slurm accounting

Learn how to configure and create a cluster with Slurm accounting. For more information, see Slurm accounting with AWS ParallelCluster.

When using the AWS ParallelCluster command line interface (CLI) or API, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

In this tutorial, you use a CloudFormation quick-create template (us-east-1) to create an Amazon Aurora for MySQL serverless database. The template instructs CloudFormation to create all the necessary components to deploy an Amazon Aurora serverless database on the same VPC as the cluster. The template also creates a basic networking and security configuration for the connection between the cluster and the database.

Note

Starting with version 3.3.0, AWS ParallelCluster supports Slurm accounting with the cluster configuration parameter SlurmSettings / Database.

Note

The quick-create template serves as an example. This template doesn't cover all possible use cases for a Slurm accounting database server. It's your responsibility to create a database server with the configuration and capacity appropriate for your production workloads.

Prerequisites:

- AWS ParallelCluster is installed.
- The AWS CLI is installed and configured.
- You have an Amazon EC2 key pair.

- You have an IAM role with the permissions that are required to run the `pcluster` CLI.
- The region that you deploy the quick-create template in supports Amazon Aurora MySQL serverless v2. For more information, see Aurora Serverless v2 with Aurora MySQL.

Step 1: Create the VPC and subnets for AWS ParallelCluster

To use the provided CloudFormation template for the Slurm accounting database, you must have the VPC for the cluster ready. You can do this manually or as part of the Configure and create a cluster with the AWS ParallelCluster command line interface procedure. If you already used AWS ParallelCluster, you might have a VPC ready for the deployment of the cluster and the database server.

Step 2: Create the database stack

Use the CloudFormation quick-create template(us-east-1) to create a database stack for Slurm accounting. The template requires following inputs:

- Database server credentials, specifically the admin user name and password.
- Sizing of the Amazon Aurora serverless cluster. This depends on the expected cluster loading.
- Networking parameters, specifically the target VPC and subnets or CIDR blocks for the creation of the subnets.

Select appropriate credentials and size for your database server. For the networking options, you're required to use the same VPC that the AWS ParallelCluster cluster is deployed to. You can create the subnets for the database and pass them as input to the template. Or, provide two disjoint CIDR blocks for the two subnets and let the CloudFormation template create the two subnets for CIDR blocks. Make sure that the CIDR blocks don't overlap with existing subnets. If the CIDR blocks overlap with existing subnets, the stack fails to be created.

The database server takes several minutes to be created.

Step 3: Create a cluster with Slurm accounting enabled

The provided CloudFormation template generates a CloudFormation stack with some defined outputs. From the AWS Management Console, you can view the outputs in the **Outputs** tab in the CloudFormation stack view. To enable the Slurm accounting, some of these outputs must be used in the AWS ParallelCluster cluster configuration file:

- `DatabaseHost`: Used for the `SlurmSettings` / `Database` / `Uri` cluster config parameter.
- `DatabaseAdminUser`: Used for the `SlurmSettings` / `Database` / `UserName` cluster configuration parameter value.
- `DatabaseSecretArn`: Used for the `SlurmSettings` / `Database` / `PasswordSecretArn` cluster config parameter.
- `DatabaseClientSecurityGroup`: This is the security group that's attached to the head node of the cluster that's defined in the `HeadNode` / `Networking` / `SecurityGroups` configuration parameter.

Update your cluster configuration file `Database` parameters with the output values. Use the `pcluster` CLI to create the cluster.

```
$ pcluster create-cluster -n cluster-3.x -c path/to/cluster-config.yaml
```

After the cluster is created, you can start using Slurm accounting commands such as `sacctmgr` or `sacct`.

Creating a cluster with an external Slurmdbd accounting

Learn how to configure and create a cluster with external Slurmdbd accounting. For more information, see Slurm accounting with AWS ParallelCluster.

When using the AWS ParallelCluster command line interface (CLI) or API, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

The AWS ParallelCluster UI is built on a serverless architecture and you can use it within the AWS Free Tier category for most cases. For more information, see AWS ParallelCluster UI costs.

In this tutorial, you use a AWS CloudFormation quick-create template to create the necessary components to deploy a Slurmdbd instance on the same VPC as the cluster. The template creates a basic networking and security configuration for the connection between the cluster and the database.

Note

Starting with `version 3.10.0`, AWS ParallelCluster supports external Slurmdbd with the cluster configuration parameter `SlurmSettings / ExternelSlurmdbd`.

Note

The quick-create template serves as an example. This template doesn't cover all possible use cases. It's your responsibility to create an external Slurmdbd with the configuration and capacity appropriate for your production workloads.

Prerequisites:

- AWS ParallelCluster is installed.
- The AWS CLI is installed and configured.
- You have an Amazon Elastic Compute Cloud key pair.
- You have an AWS Identity and Access Management role with the permissions that are required to run the `pcluster` CLI.
- You have a Slurm accounting database. To step through a tutorial of creating Slurm accounting database, follow steps 1 and 2 in Create the Slurm Accounting Database stack.

Step 1: Create the Slurmdbd stack

In this tutorial, use a CloudFormation quick-create template (`us-east-1`) to create a Slurmdbd stack. The template requires following inputs:

Networking

- **VPCId**: The VPC ID to launch the Slurmdbd instance.
- **SubnetId**: The Subnet ID to launch the Slurmdbd instance.
- **PrivatePrefix**: The CIDR prefix of the VPC.
- **PrivateIp**: A secondary private IP to assign to the Slurmdbd instance.

Database connection

- **DBMSClientSG**: The security group to be attach to the Slurmdbd instance. This security group should allows connections between the database server and the Slurmdbd instance.
- **DBMSDatabaseName**: The name of the database.
- **DBMSUsername**: The username to the database.
- **DBMSPasswordSecretArn**: The secret containing the password to the database.
- **DBMSUri**: The URI of the database server.

Instance settings

- **InstanceType**: An instance type to use for the slurmdbd instance.
- **KeyName**: An Amazon EC2 key pair to use for the slurmdbd instance.

Slurmdbd settings

- **AMIID**: An AMI of the Slurmdbd instance. The AMI should be a ParallelCluster AMI. The version of the ParallelCluster AMI determines the version of Slurmdbd.
- **MungeKeySecretArn**: The secret containing the munge key to use for authenticating communications between Slurmdbd and clusters.
- **SlurmdbdPort**: A port number that the slurmdbd uses.
- **EnableSlurmdbdSystemService**: Enables slurmdbd as system service and have it run when an instance launches.

> **Warning**
>
> If the database was created by a different version of SlurmDB, do not use Slurmdbd as a system service.
> If the database contains a large number of entries, the Slurm Database Daemon (SlurmDBD) may require tens of minutes to update the database and be unresponsive during this time interval.
> Before upgrading SlurmDB, make a backup of the database. For more information, see the Slurm documentation.

Step 2: Create a cluster with external Slurmdbd enabled

The provided AWS CloudFormation template generates a AWS CloudFormation stack with some defined outputs.

From the AWS Management Console, view the **Outputs** tab in the AWS CloudFormation stack to review the entities created. To enable the Slurm accounting, some of these outputs must be used in the AWS ParallelCluster configuration file:

- **SlurmdbdPrivateIp**: Used for the SlurmSettings / ExternalSlurmdbd / Host cluster config parameter.
- **SlurmdbdPort**: Used for the SlurmSettings / ExternalSlurmdbd / Port cluster configuration parameter value.
- **AccountingClientSecurityGroup**: This is the security group that's attached to the head node of the cluster that's defined in the HeadNode / Networking / AdditionalSecurityGroups configuration parameter.

Additional, from the **Parameters** tab in the AWS CloudFormation stack view:

- **MungeKeySecretArn**: Used for the SlurmSettings / MungeKeySecretArn cluster configuration parameter value.

Update your cluster configuration file database parameters with the output values. Use the pcluster AWS CLI to create the cluster.

```
$  pcluster create-cluster -n cluster-3.x-c path/to/cluster-config.yaml
```

After the cluster is created, you can start using Slurm accounting commands such as `sacctmgr` or `sacct`.

> **Warning**
>
> Traffic between `ParallelCluster` and the external SlurmDB is not encrypted. It is recommended to run the cluster and the external SlurmDB in a trusted network.

Reverting to a previous AWS Systems Manager document version

Learn how to revert to a previous AWS Systems Manager document version. For more information about SSM documents, see AWS Systems Manager Documents in the *AWS Systems Manager User Guide*.

When using the AWS ParallelCluster command line interface (CLI) or API, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

Prerequisites:

- An AWS account with permissions to manage SSM documents.
- The AWS CLI is installed and configured.

Revert to a previous SSM document version

1. In your terminal, run the following command to get the list of existing SSM documents that you own.

   ```
   $ aws ssm list-documents --document-filter "key=Owner,value=Self"
   ```

2. Revert an SSM document to a previous version. In this example, we revert to a previous version of the `SessionManagerRunShell` document. You can use the SSM `SessionManagerRunShell` document to customize every SSM shell session that you initiate.

   a. Find the `DocumentVersion` parameter for `SessionManagerRunShell` by running the following command:

   ```
   $ aws ssm describe-document --name "SSM-SessionManagerRunShell"
   {
       "Document": {
           "Hash": "...",
           "HashType": "Sha256",
           "Name": "SSM-SessionManagerRunShell",
           "Owner": "123456789012",
           "CreatedDate": "2023-02-20T19:04:32.390000+00:00",
           "Status": "Active",
   ```

```
        "DocumentVersion": "1",
        "Parameters": [
            {
                "Name": "linuxcmd",
                "Type": "String",
                "Description": "The command to run on connection...",
                "DefaultValue": "if [ -d '/opt/parallelcluster' ]; then
 source /opt/parallelcluster/cfnconfig; sudo su - $cfn_cluster_user; fi; /bin/
bash"
            }
        ],
        "PlatformTypes": [
            "Windows",
            "Linux",
            "MacOS"
        ],
        "DocumentType": "Session",
        "SchemaVersion": "1.0",
        "LatestVersion": "2",
        "DefaultVersion": "1",
        "DocumentFormat": "JSON",
        "Tags": []
    }
}
```

The latest version is 2.

b. Revert to the previous version by running the following command:

```
$ aws ssm delete-document --name "SSM-SessionManagerRunShell" --document-
version 2
```

3. Verify that the document version has been reverted by running the `describe-document` command again:

```
$ aws ssm describe-document --name "SSM-SessionManagerRunShell"
{
    "Document": {
        "Hash": "...",
        "HashType": "Sha256",
        "Name": "SSM-SessionManagerRunShell",
        "Owner": "123456789012",
        "CreatedDate": "2023-02-20T19:04:32.390000+00:00",
```

```
        "Status": "Active",
        "DocumentVersion": "1",
        "Parameters": [
            {
                "Name": "linuxcmd",
                "Type": "String",
                "Description": "The command to run on connection...",
                "DefaultValue": "if [ -d '/opt/parallelcluster' ]; then source /
opt/parallelcluster/cfnconfig; sudo su - $cfn_cluster_user; fi; /bin/bash"
            }
        ],
        "PlatformTypes": [
            "Windows",
            "Linux",
            "MacOS"
        ],
        "DocumentType": "Session",
        "SchemaVersion": "1.0",
        "LatestVersion": "1",
        "DefaultVersion": "1",
        "DocumentFormat": "JSON",
        "Tags": []
    }
}
```

The latest version is 1.

Creating a cluster with AWS CloudFormation

Learn how to create a cluster with an AWS ParallelCluster CloudFormation custom resource. For more information, see AWS CloudFormation custom resource.

When using AWS ParallelCluster, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

Prerequisites:

- The AWS CLI is installed and configured.
- An Amazon EC2 key pair.
- An IAM role with the permissions that are required to run the `pcluster` CLI.

Cluster creation with a CloudFormation quick-create stack

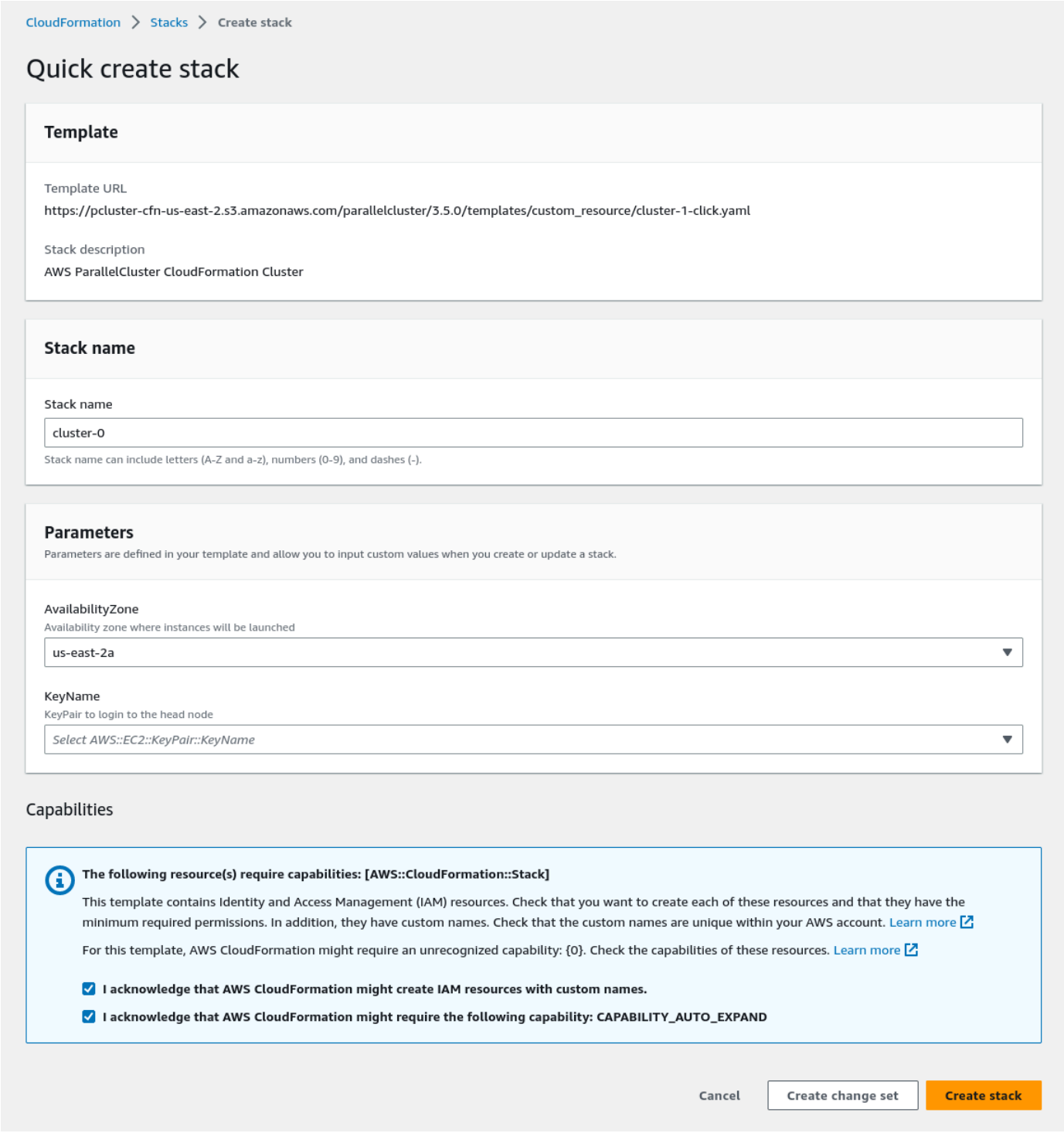
In this tutorial, you use a quick-create stack to deploy a CloudFormation template that creates a cluster and the following AWS resources:

- A root CloudFormation stack created by using a CloudFormation quick-create stack.
- Nested CloudFormation stacks that include default policies, default VPC set up, and a custom resource provider.
- An example AWS ParallelCluster cluster stack and a cluster that you can log in to and run jobs.

Create a cluster with AWS CloudFormation

1. Sign in to the AWS Management Console.
2. Open the CloudFormation quick-create link to create the following resources in the CloudFormation console:

   - A nested CloudFormation stack with a VPC with a public and private subnet for running the cluster head node and compute nodes, respectively.
   - A nested CloudFormation stack with an AWS ParallelCluster custom resource for managing the cluster.
   - A nested CloudFormation stack with the default policies for managing the cluster.
   - A root CloudFormation stack for the nested stacks.
   - An AWS ParallelCluster cluster with the Slurm scheduler and a defined number of compute nodes.



3. In the **Quick create stack Parameters** section, enter values for the following parameters:

   a. For **KeyName**, enter the name of your Amazon EC2 key pair.

   b. For **AvailabilityZone**, choose an AZ for your cluster nodes, for example, `us-east-1a`.

4. Check the boxes to acknowledge each of the access capabilities at the bottom of the page.
5. Choose **Create stack**.
6. Wait for the CloudFormation stack to reach the `CREATE_COMPLETE` state.

Cluster creation with the AWS CloudFormation Command Line Interface (CLI)

In this tutorial, you use the AWS Command Line Interface (CLI) for CloudFormation to deploy a CloudFormation template that creates a cluster.

Create the following AWS resources:

- A root CloudFormation stack created by using a CloudFormation quick-create stack.
- Nested CloudFormation stacks that include default policies, default VPC setup, and a custom resource provider.
- An example AWS ParallelCluster cluster stack and a cluster that you can log in to and run jobs.

Replace *`inputs highlighted in red`*, such as *`keypair`*, with your own values.

Create a cluster with AWS CloudFormation

1. Create a CloudFormation template named `cluster_template.yaml` with the following content:

```
AWSTemplateFormatVersion: '2010-09-09'
Description: >
  AWSParallelCluster CloudFormation Template

Parameters:
  KeyName:
    Description: KeyPair to login to the head node
    Type: AWS::EC2::KeyPair::KeyName

  AvailabilityZone:
    Description: Availability zone where instances will be launched
    Type: AWS::EC2::AvailabilityZone::Name
    Default: us-east-2a

Mappings:
```

```
  ParallelCluster:
    Constants:
      Version: 3.13.2

Resources:
  PclusterClusterProvider:
    Type: AWS::CloudFormation::Stack
    Properties:
      TemplateURL: !Sub
        - https://${AWS::Region}-aws-parallelcluster.s3.${AWS::Region}.
${AWS::URLSuffix}/parallelcluster/${Version}/templates/custom_resource/cluster.yaml
        - { Version: !FindInMap [ParallelCluster, Constants, Version] }

  PclusterVpc:
    Type: AWS::CloudFormation::Stack
    Properties:
      Parameters:
        PublicCIDR: 10.0.0.0/24
        PrivateCIDR: 10.0.16.0/20
        AvailabilityZone: !Ref AvailabilityZone
      TemplateURL: !Sub
        - https://${AWS::Region}-aws-parallelcluster.s3.${AWS::Region}.
${AWS::URLSuffix}/parallelcluster/${Version}/templates/networking/public-private-
${Version}.cfn.json
        - { Version: !FindInMap [ParallelCluster, Constants, Version ] }

  PclusterCluster:
    Type: Custom::PclusterCluster
    Properties:
      ServiceToken: !GetAtt [ PclusterClusterProvider , Outputs.ServiceToken ]
      ClusterName: !Sub 'c-${AWS::StackName}'
      ClusterConfiguration:
        Image:
          Os: alinux2
        HeadNode:
          InstanceType: t2.medium
          Networking:
            SubnetId: !GetAtt [ PclusterVpc , Outputs.PublicSubnetId ]
          Ssh:
            KeyName: !Ref KeyName
        Scheduling:
          Scheduler: slurm
          SlurmQueues:
          - Name: queue0
```

```
            ComputeResources:
            - Name: queue0-cr0
              InstanceType: t2.micro
            Networking:
              SubnetIds:
              -  !GetAtt [ PclusterVpc , Outputs.PrivateSubnetId ]
Outputs:
  HeadNodeIp:
    Description: The Public IP address of the HeadNode
    Value: !GetAtt [ PclusterCluster, headNode.publicIpAddress ]
```

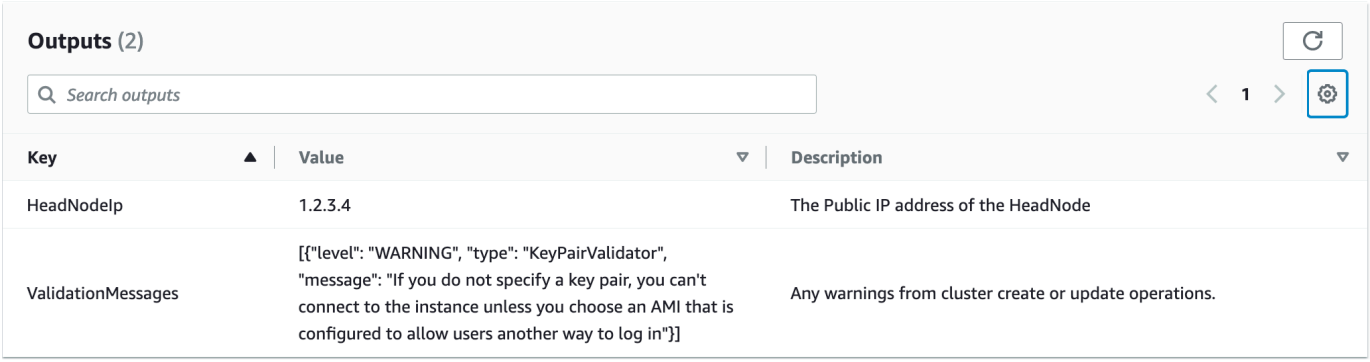
2. Run the following AWS CLI command to deploy the CloudFormation stack for cluster creation and management.

```
$ aws cloudformation deploy --template-file ./cluster_template.yaml \
  --stack-name mycluster \
  --parameter-overrides KeyName=keypair \
                        AvailabilityZone=us-east-2b \
  --capabilities CAPABILITY_NAMED_IAM CAPABILITY_AUTO_EXPAND
```

View CloudFormation cluster output

View the CloudFormation cluster output to obtain useful cluster details. The added `ValidationMessages` property provides access to validation messages from cluster create and update operations.

1. Navigate to the CloudFormation console and select the stack that includes your AWS ParallelCluster custom resource.
2. Choose **Stack details**, and select the **Outputs** tab.



Validation messages might be truncated. For more information about how to retrieve logs, see AWS ParallelCluster troubleshooting.

Access your cluster

Access the cluster.

ssh into the cluster head node

1. After the CloudFormation stack deployment is complete, obtain the IP address of the head node with the following command:

```
$ HEAD_NODE_IP=$(aws cloudformation describe-stacks --stack-name=mycluster --query
 "Stacks|[0].Outputs[?OutputKey=='HeadNodeIp']|[0].OutputValue" --output=text)
```

   You can also retrieve the head node IP address from **HeadNodeIp** parameter in the cluster stack **Outputs** tab in the CloudFormation console.

   You can find the head node IP address here because it was added in the `Outputs` section of the cluster CloudFormation template, specifically for this example cluster.

2. Connect to the cluster head node by running the following command:

```
$ ssh -i keyname.pem ec2-user@$HEAD_NODE_IP
```

Clean up

Delete the cluster.

1. Run the following AWS CLI command to delete the CloudFormation stack and cluster.

```
$ aws cloudformation delete-stack --stack-name=mycluster
```

2. Check the stack delete status by running the following command.

```
$ aws cloudformation describe-stacks --stack-name=mycluster
```

Deploy ParallelCluster API with Terraform

In this tutorial, you will define a simple Terraform project to deploy a ParallelCluster API.

Prerequisites

- Terraform v1.5.7+ is installed.
- IAM role with the permissions to deploy the ParallelCluster API. See the section called “Required permissions”.

Define a Terraform project

In this tutorial, you will define a Terraform project.

1. Create a directory called `my-pcluster-api`.

   All files that you create will be within this directory.
2. Create the file `provider.tf` to configure the AWS provider.

```
provider "aws" {
  region  = var.region
  profile = var.profile
}
```

3. Create the file `main.tf` to define the resources using the ParallelCluster module.

```
module "parallelcluster_pcluster_api" {
  source = "aws-tf/parallelcluster/aws//modules/pcluster_api"
  version = "1.1.0"

  region                = var.region
  api_stack_name        = var.api_stack_name
  api_version           = var.api_version

  parameters = {
    EnableIamAdminAccess = "true"
  }
}
```

4. Create the file `variables.tf` to define the variables that can be injected for this project.

```
variable "region" {
  description = "The region the ParallelCluster API is deployed in."
  type        = string
  default     = "us-east-1"
}

variable "profile" {
  type        = string
  description = "The AWS profile used to deploy the clusters."
  default     = null
}

variable "api_stack_name" {
  type        = string
  description = "The name of the CloudFormation stack used to deploy the
 ParallelCluster API."
  default     = "ParallelCluster"
}

variable "api_version" {
  type        = string
  description = "The version of the ParallelCluster API."
}
```

5. Create the file `terraform.tfvars` to set arbitrary values for the variables.

   The file below deploys a ParallelCluster API 3.11.1 in `us-east-1` using the stack name `MyParallelClusterAPI-3111`. You'll be able to reference this ParallelCluster API deployment using its stack name.

   > **Note**
   >
   > The `api_version` assignment in the following code can be replaced with any supported AWS ParallelCluster version.

   ```
   region = "us-east-1"
   api_stack_name = "MyParallelClusterAPI-3111"
   api_version = "3.11.1"
   ```

6. Create the file `outputs.tf` to define the outputs returned by this project.

```
output "pcluster_api_stack_outputs" {
  value = module.parallelcluster_pcluster_api.stack_outputs
}
```

The project directory is:

```
my-pcluster-api
### main.tf - Terraform entrypoint to define the resources using the
 ParallelCluster module.
### outputs.tf - Defines the outputs returned by Terraform.
### providers.tf - Configures the AWS provider.
### terraform.tfvars - Set the arbitrary values for the variables, i.e. region,
 PCAPI version, PCAPI stack name
### variables.tf - Defines the variables, e.g. region, PCAPI version, PCAPI stack
 name.
```

Deploy the API

To deploy the API, run the standard Terraform commands in order.

1. Build the project:

   ```
   terraform init
   ```

2. Define the deployment plan:

   ```
   terraform plan -out tfplan
   ```

3. Deploy the plan:

   ```
   terraform apply tfplan
   ```

Required permissions

You need the following permissions to deploy the ParallelCluster API with Terraform:

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": [
                "cloudformation:DescribeStacks",
                "cloudformation:GetTemplate"
            ],
            "Resource": "arn:aws:cloudformation:us-east-1:111122223333:stack/*",
            "Effect": "Allow",
            "Sid": "CloudFormationRead"
        },
        {
            "Action": [
                "cloudformation:CreateStack",
                "cloudformation:DeleteStack",
                "cloudformation:CreateChangeSet"
            ],
            "Resource": "arn:aws:cloudformation:us-east-1:111122223333:stack/
MyParallelClusterAPI*",
            "Effect": "Allow",
            "Sid": "CloudFormationWrite"
        },
        {
            "Action": [
                "cloudformation:CreateChangeSet"
            ],
            "Resource": [
                "arn:aws:cloudformation:us-east-1:111122223333:aws:transform/
Include",
                "arn:aws:cloudformation:us-east-1:111122223333:aws:transform/
Serverless-2016-10-31"
            ],
            "Effect": "Allow",
            "Sid": "CloudFormationTransformWrite"
        },
        {
            "Action": [
                "s3:GetObject"
            ],
            "Resource": [
```

```
                "arn:aws:s3:us-east-1:111122223333:*-aws-parallelcluster/
parallelcluster/*/api/ParallelCluster.openapi.yaml",
                "arn:aws:s3:us-east-1:111122223333:*-aws-parallelcluster/
parallelcluster/*/layers/aws-parallelcluster/lambda-layer.zip"
            ],
            "Effect": "Allow",
            "Sid": "S3ParallelClusterArtifacts"
        },
        {
            "Action": [
                "iam:CreateRole",
                "iam:DeleteRole",
                "iam:GetRole",
                "iam:CreatePolicy",
                "iam:DeletePolicy",
                "iam:GetPolicy",
                "iam:GetRolePolicy",
                "iam:AttachRolePolicy",
                "iam:DetachRolePolicy",
                "iam:PutRolePolicy",
                "iam:DeleteRolePolicy",
                "iam:ListPolicyVersions"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/*",
                "arn:aws:iam::111122223333:policy/*"
            ],
            "Effect": "Allow",
            "Sid": "IAM"
        },
        {
            "Action": [
                "iam:PassRole"
            ],
            "Resource": [
                "arn:aws:iam::111122223333:role/ParallelClusterLambdaRole-*",
                "arn:aws:iam::111122223333:role/APIGatewayExecutionRole-*"
            ],
            "Effect": "Allow",
            "Sid": "IAMPassRole"
        },
        {
            "Action": [
                "lambda:CreateFunction",
```

```
                "lambda:DeleteFunction",
                "lambda:GetFunction",
                "lambda:PublishLayerVersion",
                "lambda:DeleteLayerVersion",
                "lambda:GetLayerVersion",
                "lambda:TagResource",
                "lambda:UntagResource"
            ],
            "Resource": [
                "arn:aws:lambda:us-east-1:111122223333:layer:PCLayer-*",
                "arn:aws:lambda:us-east-1:111122223333:function:*-
ParallelClusterFunction-*"
            ],
            "Effect": "Allow",
            "Sid": "Lambda"
        },
        {
            "Action": [
                "logs:CreateLogGroup",
                "logs:DeleteLogGroup",
                "logs:DescribeLogGroups",
                "logs:PutRetentionPolicy",
                "logs:TagLogGroup",
                "logs:UntagLogGroup"
            ],
            "Resource": [
                "arn:aws:logs:us-east-1:111122223333:log-group:/aws/lambda/*-
ParallelClusterFunction-*"
            ],
            "Effect": "Allow",
            "Sid": "Logs"
        },
        {
            "Action": [
                "apigateway:DELETE",
                "apigateway:GET",
                "apigateway:PATCH",
                "apigateway:POST",
                "apigateway:PUT",
                "apigateway:UpdateRestApiPolicy"
            ],
            "Resource": [
                "arn:aws:apigateway:us-east-1::/restapis",
                "arn:aws:apigateway:us-east-1::/restapis/*",
```

```
                "arn:aws:apigateway:us-east-1::/tags/*"
            ],
            "Effect": "Allow",
            "Sid": "APIGateway"
        }
    ]
}
```

Creating a cluster with Terraform

When using AWS ParallelCluster, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see the section called “AWS services used by AWS ParallelCluster”.

Prerequisites

- Terraform v1.5.7+ is installed.
- the section called “AWS ParallelCluster API” v3.8.0+ is deployed in your account. See the section called “Deploy ParallelCluster API with Terraform”.
- IAM role with the permissions to invoke the ParallelCluster API. See [Required permissions]

Define a Terraform project

In this tutorial, you will define a simple Terraform project to deploy a cluster.

1. Create a directory called `my-clusters`.

   All files that you create will be within this directory.

2. Create the file `terraform.tf` to import the ParallelCluster provider.

```
terraform {
  required_version = ">= 1.5.7"
  required_providers {
    aws-parallelcluster = {
      source  = "aws-tf/aws-parallelcluster"
      version = "~> 1.0"
    }
  }
```

```
}
```

3. Create the file `providers.tf` to configure the ParallelCluster and AWS providers.

```
provider "aws" {
  region  = var.region
  profile = var.profile
}

provider "aws-parallelcluster" {
  region         = var.region
  profile        = var.profile
  api_stack_name = var.api_stack_name
  use_user_role  = true
}
```

4. Create the file `main.tf` to define the resources using the ParallelCluster module.

```
module "pcluster" {
  source  = "aws-tf/parallelcluster/aws"
  version = "1.1.0"

  region                = var.region
  api_stack_name        = var.api_stack_name
  api_version           = var.api_version
  deploy_pcluster_api   = false

  template_vars         = local.config_vars
  cluster_configs       = local.cluster_configs
  config_path           = "config/clusters.yaml"
}
```

5. Create the file `clusters.tf` to define multiple clusters as Terraform local variables.

> **Note**
>
> You can define multiple clusters within the `cluster_config` element. For every cluster, you can explicitly define the cluster properties within the local variables (see `DemoCluster01`) or reference an external file (see `DemoCluster02`).

To review the cluster properties that you can set within the configuration element, see the section called “Cluster configuration file”.

To review the options that you can set for cluster creation, see the section called “`pcluster create-cluster`”.

```
locals {
  cluster_configs = {
    DemoCluster01 : {
      region : local.config_vars.region
      rollbackOnFailure : false
      validationFailureLevel : "WARNING"
      suppressValidators : [
        "type:KeyPairValidator"
      ]
      configuration : {
        Region : local.config_vars.region
        Image : {
          Os : "alinux2"
        }
        HeadNode : {
          InstanceType : "t3.small"
          Networking : {
            SubnetId : local.config_vars.subnet
          }
          Iam : {
            AdditionalIamPolicies : [
              { Policy : "arn:aws:iam::aws:policy/AmazonSSMManagedInstanceCore" }
            ]
          }
        }
        Scheduling : {
          Scheduler : "slurm"
          SlurmQueues : [{
            Name : "queue1"
            CapacityType : "ONDEMAND"
            Networking : {
              SubnetIds : [local.config_vars.subnet]
            }
            Iam : {
              AdditionalIamPolicies : [
```

```
                { Policy : "arn:aws:iam::aws:policy/AmazonSSMManagedInstanceCore" }
              ]
            }
            ComputeResources : [{
              Name : "compute"
              InstanceType : "t3.small"
              MinCount : "1"
              MaxCount : "4"
            }]
          }]
          SlurmSettings : {
            QueueUpdateStrategy : "TERMINATE"
          }
        }
      }
    }
    DemoCluster02 : {
      configuration : "config/cluster_config.yaml"
    }
  }
}
```

6. Create the file `config/clusters.yaml` to define multiple clusters as YAML configuration.

```
DemoCluster03:
  region: ${region}
  rollbackOnFailure: true
  validationFailureLevel: WARNING
  suppressValidators:
    - type:KeyPairValidator
  configuration: config/cluster_config.yaml
DemoCluster04:
  region: ${region}
  rollbackOnFailure: false
  configuration: config/cluster_config.yaml
```

7. Create the file `config/cluster_config.yaml`, which is a standard ParallelCluster config file where Terraform variables can be injected.

   To review the cluster properties that you can set within the configuration element, see the section called “Cluster configuration file”.

```
Region: ${region}
```

```
Image:
 Os: alinux2
HeadNode:
 InstanceType: t3.small
 Networking:
   SubnetId: ${subnet}
 Iam:
   AdditionalIamPolicies:
     - Policy: arn:aws:iam::aws:policy/AmazonSSMManagedInstanceCore
Scheduling:
 Scheduler: slurm
 SlurmQueues:
   - Name: queue1
     CapacityType: ONDEMAND
     Networking:
       SubnetIds:
         - ${subnet}
     Iam:
       AdditionalIamPolicies:
         - Policy: arn:aws:iam::aws:policy/AmazonSSMManagedInstanceCore
     ComputeResources:
       - Name: compute
         InstanceType: t3.small
         MinCount: 1
         MaxCount: 5
 SlurmSettings:
   QueueUpdateStrategy: TERMINATE
```

8. Create the file `clusters_vars.tf` to define the variables that can be injected into cluster configurations.

   This file allows you to define dynamic values that can be used in cluster configurations, such as region and subnet.

   This example retrieves values directly from the project variables, but you may need to use custom logic to determine them.

```
locals {
  config_vars = {
    subnet = var.subnet_id
    region = var.cluster_region
  }
```

```
}
```

9. Create the file `variables.tf` to define the variables that can be injected for this project.

```
variable "region" {
  description = "The region the ParallelCluster API is deployed in."
  type        = string
  default     = "us-east-1"
}

variable "cluster_region" {
  description = "The region the clusters will be deployed in."
  type        = string
  default     = "us-east-1"
}

variable "profile" {
  type        = string
  description = "The AWS profile used to deploy the clusters."
  default     = null
}

variable "subnet_id" {
  type        = string
  description = "The id of the subnet to be used for the ParallelCluster
 instances."
}

variable "api_stack_name" {
  type        = string
  description = "The name of the CloudFormation stack used to deploy the
 ParallelCluster API."
  default     = "ParallelCluster"
}

variable "api_version" {
  type        = string
  description = "The version of the ParallelCluster API."
}
```

10. Create the file `terraform.tfvars` to set arbitrary values for the variables.

The file below deploys the clusters in `eu-west-1` within the subnet `subnet-123456789`, using the existing ParallelCluster API 3.11.1, which is already deployed in `us-east-1` with stack name `MyParallelClusterAPI-3111`.

```
region = "us-east-1"
api_stack_name = "MyParallelClusterAPI-3111"
api_version = "3.11.1"

cluster_region = "eu-west-1"
subnet_id = "subnet-123456789"
```

11. Create the file `outputs.tf` to define the outputs returned by this project.

```
output "clusters" {
  value = module.pcluster.clusters
}
```

The project directory is:

```
my-clusters
### config
#   ### cluster_config.yaml - Cluster configuration, where terraform variables can
 be injected..
#   ### clusters.yaml - File listing all the clusters to deploy.
### clusters.tf - Clusters defined as Terraform local variables.
### clusters_vars.tf - Variables that can be injected into cluster configurations.
### main.tf - Terraform entrypoint where the ParallelCluster module is configured.
### outputs.tf - Defines the cluster as a Terraform output.
### providers.tf - Configures the providers: ParallelCluster and AWS.
### terraform.tf - Import the ParallelCluster provider.
### terraform.tfvars - Defines values for variables, e.g. region, PCAPI stack name.
### variables.tf - Defines the variables, e.g. region, PCAPI stack name.
```

Deploy the cluster

To deploy the cluster, run the standard Terraform commands in order.

> **Note**
>
> This example assumes that you've already deployed the ParallelCluster API in your account.

1. Build the project:

```
terraform init
```

2. Define the deployment plan:

```
terraform plan -out tfplan
```

3. Deploy the plan:

```
terraform apply tfplan
```

Deploy the ParallelCluster API with clusters

If you haven't deployed the ParallelCluster API and you want to deploy it with the clusters, change the following files:

- `main.tf`

```
module "pcluster" {
  source  = "aws-tf/aws/parallelcluster"
  version = "1.0.0"

  region                = var.region
  api_stack_name        = var.api_stack_name
  api_version           = var.api_version
  deploy_pcluster_api   = true
  parameters = {
    EnableIamAdminAccess = "true"
  }

  template_vars         = local.config_vars
  cluster_configs       = local.cluster_configs
  config_path           = "config/clusters.yaml"
}
```

- `providers.tf`

```
provider "aws-parallelcluster" {
  region   = var.region
  profile  = var.profile
  endpoint = module.pcluster.pcluster_api_stack_outputs.ParallelClusterApiInvokeUrl
  role_arn = module.pcluster.pcluster_api_stack_outputs.ParallelClusterApiUserRole
}
```

Required permissions

You need the following permissions to deploy a cluster with Terraform:

- assume the ParallelCluster API role, which is in charge of interacting with the ParallelCluster API
- describe the AWS CloudFormation stack of the ParallelCluster API to verify it exists and retrieve its parameters and outputs

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": "sts:AssumeRole",
            "Resource": "arn:aws:sts::111122223333:role/PCAPIUserRole-*",
            "Effect": "Allow",
            "Sid": "AssumePCAPIUserRole"
        },
        {
            "Action": [
                "cloudformation:DescribeStacks"
            ],
            "Resource": "arn:aws:cloudformation:us-east-1:111122223333:stack/*",
            "Effect": "Allow",
            "Sid": "CloudFormation"
        }
    ]
}
```

Creating a custom AMI with Terraform

When using AWS ParallelCluster, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see the section called “AWS services used by AWS ParallelCluster”.

Prerequisites

- Terraform v1.5.7+ is installed.
- the section called “AWS ParallelCluster API” v3.8.0+ is deployed in your account. See the section called “Creating a cluster with Terraform”.
- IAM role with the permissions to invoke the ParallelCluster API. See the section called “Required permissions”.

Define a Terraform project

In this tutorial, you will define a simple Terraform project to deploy a ParallelCluster custom AMI.

1. Create a directory called `my-amis`.

   All files that you create will be within this directory.
2. Create the file `terraform.tf` to import the ParallelCluster provider.

```
terraform {
  required_version = ">= 1.5.7"
  required_providers {
    aws-parallelcluster = {
      source  = "aws-tf/aws-parallelcluster"
      version = "~> 1.0"
    }
  }
}
```

3. Create the file `providers.tf` to configure the ParallelCluster and AWS providers.

```
provider "aws" {
  region  = var.region
  profile = var.profile
}
```

```
provider "aws-parallelcluster" {
  region         = var.region
  profile        = var.profile
  api_stack_name = var.api_stack_name
  use_user_role  = true
}
```

4. Create the file `main.tf` to define the resources using the ParallelCluster module.

   To review the image properties that you can set within the `image_configuration` element, see the section called “Build image configuration files”.

   To review the options that you can set for image creation, for example `image_id` and `rollback_on_failure`, see the section called “`pcluster build-image`”.

```
data "aws-parallelcluster_list_official_images" "parent_image" {
  region = var.region
  os = var.os
  architecture = var.architecture
}

resource "aws-parallelcluster_image" "demo01" {
  image_id            = "demo01"
  image_configuration = yamlencode({
    "Build":{
      "InstanceType": "c5.2xlarge",
      "ParentImage": data.aws-
parallelcluster_list_official_images.parent_image.official_images[0].amiId,
      "UpdateOsPackages": {"Enabled": false}
    }
  })
  rollback_on_failure = false
}
```

5. Create the file `variables.tf` to define the variables that can be injected for this project.

```
variable "region" {
  description = "The region the ParallelCluster API is deployed in."
  type        = string
  default     = "us-east-1"
}

variable "profile" {
```

```
  type        = string
  description = "The AWS profile used to deploy the clusters."
  default     = null
}

variable "api_stack_name" {
  type        = string
  description = "The name of the CloudFormation stack used to deploy the
 ParallelCluster API."
  default     = "ParallelCluster"
}

variable "api_version" {
  type        = string
  description = "The version of the ParallelCluster API."
}

variable "os" {
  type        = string
  description = "The OS of the ParallelCluster image."
}

variable "architecture" {
  type        = string
  description = "The architecture of the ParallelCluster image."
}
```

6. Create the file `terraform.tfvars` to set your arbitrary values for the variables.

   With the file below deploy the custom AMI in `us-east-1` based on Amazon Linux 2 for x86_64 architecture, using the existing ParallelCluster API 3.11.1 which is already deployed in `us-east-1` with stack name `MyParallelClusterAPI-3111`.

   ```
   region = "us-east-1"
   api_stack_name = "MyParallelClusterAPI-3111"
   api_version = "3.11.1"

   os = "alinux2"
   architecture = "x86_64"
   ```

7. Create the file `outputs.tf` to define the outputs returned by this project.

   ```
   output "parent_image" {
   ```

```
  value = data.aws-
parallelcluster_list_official_images.parent_image.official_images[0]
}

output "custom_image" {
  value = aws-parallelcluster_image.demo01
}
```

The project directory is:

```
my-amis
### main.tf - Terraform entrypoint where the ParallelCluster module is configured.
### outputs.tf - Defines the cluster as a Terraform output.
### providers.tf - Configures the providers: ParallelCluster and AWS.
### terraform.tf - Import the ParallelCluster provider.
### terraform.tfvars - Defines values for variables, e.g. region, PCAPI stack name.
### variables.tf - Defines the variables, e.g. region, PCAPI stack name.
```

Deploy the AMI

To deploy the AMI, run the standard Terraform commands in order.

1. Build the project:

   ```
   terraform init
   ```

2. Define the deployment plan:

   ```
   terraform plan -out tfplan
   ```

3. Deploy the plan:

   ```
   terraform apply tfplan
   ```

Required permissions

You need the following permissions to deploy a custom AMI with Terraform:

- assume the ParallelCluster API role, which is in charge of interacting with the ParallelCluster API

- describe the AWS CloudFormation stack of the ParallelCluster API, to verify it exists and retrieve its parameters and outputs

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Action": "sts:AssumeRole",
            "Resource": "arn:aws:sts::111122223333:role/PCAPIUserRole-*",
            "Effect": "Allow",
            "Sid": "AssumePCAPIUserRole"
        },
        {
            "Action": [
                "cloudformation:DescribeStacks"
            ],
            "Resource": "arn:aws:cloudformation:us-east-1:111122223333:stack/*",
            "Effect": "Allow",
            "Sid": "CloudFormation"
        }
    ]
}
```

AWS ParallelCluster UI Integration with Identity Center

The goal of this tutorial is to demonstrate how to integrate AWS ParallelCluster UI with IAM Identity Center for a single sign-on solution that unifies users in Active Directory that can be shared with AWS ParallelCluster clusters.

When using AWS ParallelCluster, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

Prerequisites:

- An existing AWS ParallelCluster UI which can be installed following the instructions here.

- An existing Managed Active Directory, preferably one that you will also use for integrating with AWS ParallelCluster.

Enable IAM Identity Center

If you already have an identity center connected to the your AWS Managed Microsoft AD (Active Directory) it can be used and you can skip to the section **Adding your Application to IAM Identity Center**.

If you do not already have an identity center connected to an AWS Managed Microsoft AD, follow the steps below to set it up.

Enabling Identity Center

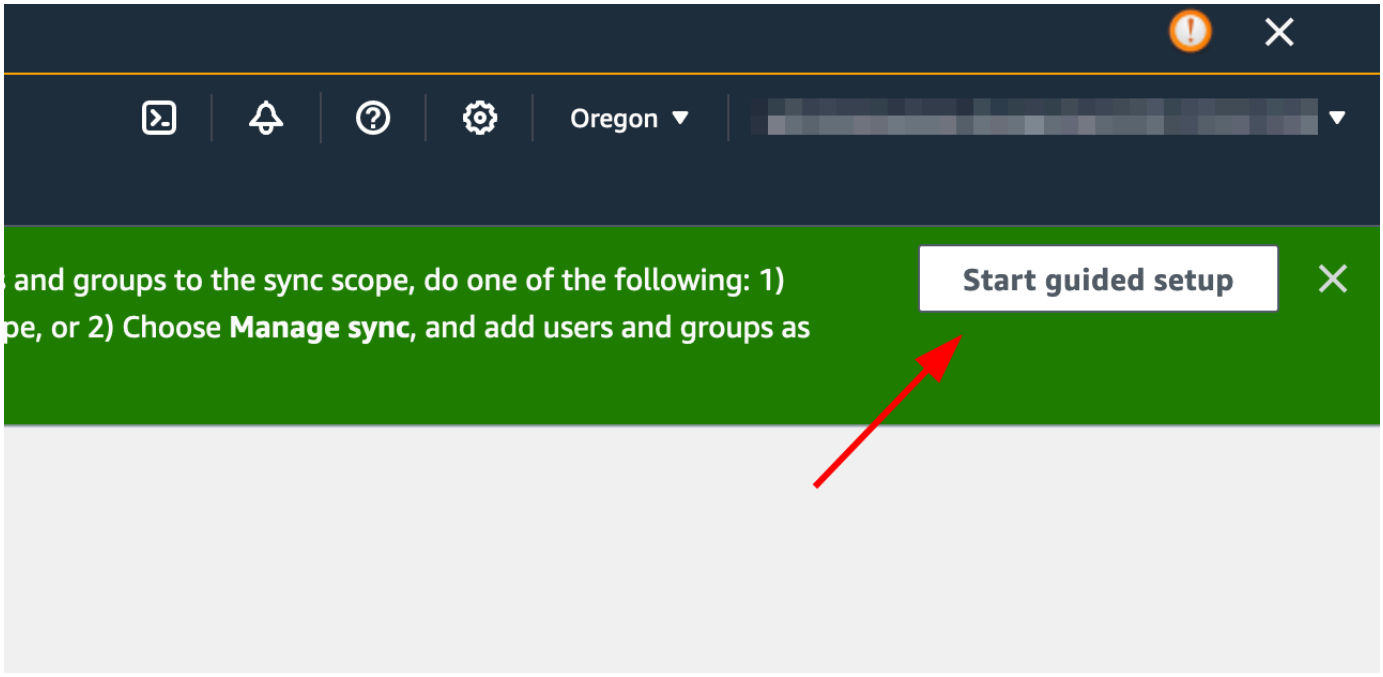
1. In the console, navigate to IAM Identity Center. (Make sure you are in the region in which you have your AWS Managed Microsoft AD.)
2. Click the **Enable** button, this may ask if you want to enable organizations, this is a requirement so you can select to enable it. **Note** : This will email the administrator of your account with a confirmation email that you should follow the link to confirm.

Connecting Identity Center to Managed AD

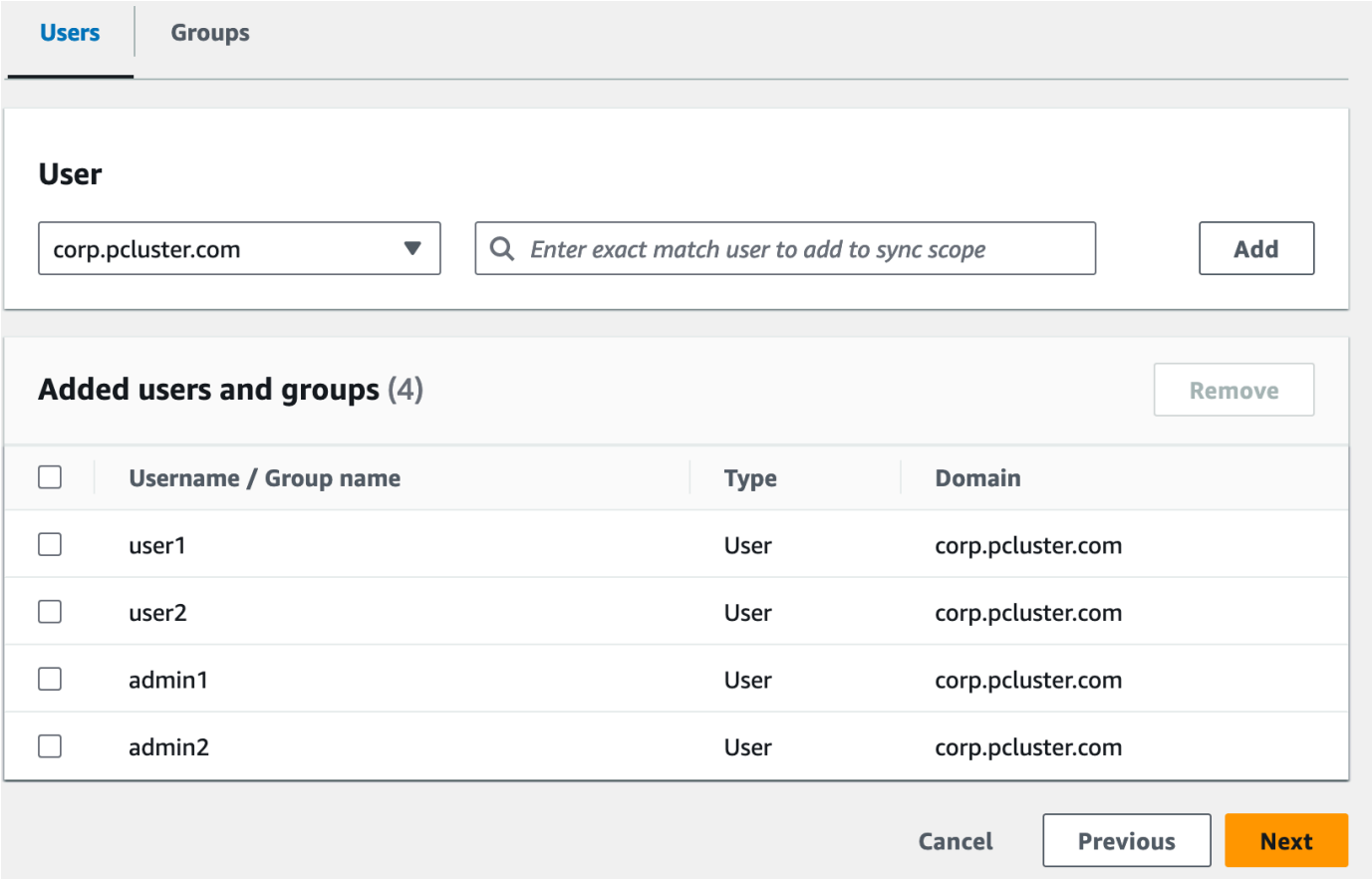
1. On the next page after enabling identity center you should see Recommended Set Up Steps, under Step 1, select **Choose Your Identity Source**.
2. In the Identity Source section, click on the **Actions** drop down menu (in the top right), then select **Change Identity Source**.
3. Select **Active Directory**.
4. Under **Existing Directories**, choose your directory.
5. Click Next.
6. Review your changes, scroll to the bottom, type ACCEPT into the text box to confirm, then click **Change Identity Source**.
7. Wait for the changes to complete, then you should see a green banner at the top.

Syncing users and groups to Identity Center

1. In the green banner click **Start Guided Setup** (button in the top right one)

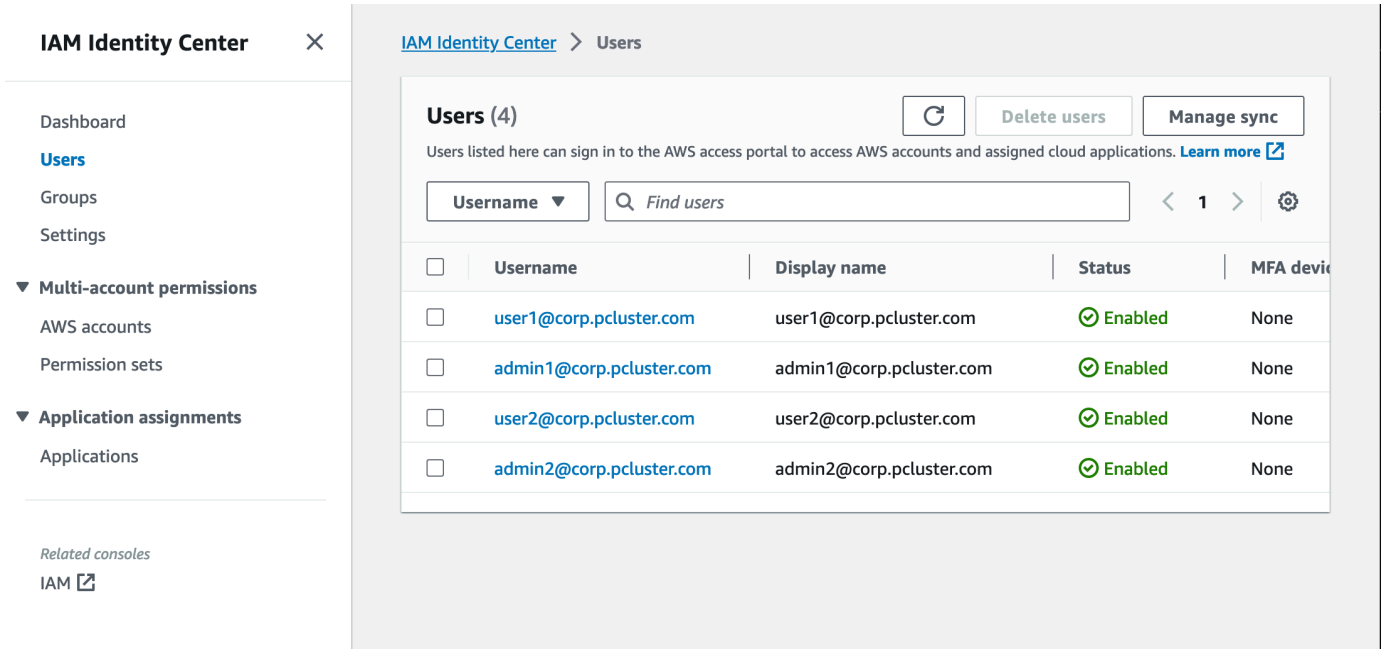


2. In the **Configure Attribute Mappings**, click **Next**
3. In the Configure sync scope section, type in the name of the users you want synced to identity center, then click **Add**
4. Once finished adding users and groups, click **Next**



5. Review your changes, then click **Save configuration**
6. If you see a warning in the next screen about users not being synced, select the **Resume sync button** in the top right.
7. Next, to enable users, In the **Users** tab on the left, select a user and then click **Enable user access** > **Enable user access**

   Note: You may need to select Resume sync if you have a warning banner at the top and then wait for users to sync (try the refresh button to see if they are synced yet).



Adding your Application to IAM Identity Center

Once you have synced your users with IAM Identity Center, you will need to add a new application. This configures which SSO enabled applications will be available from your IAM Identity Center portal. In this case, we will be adding AWS ParallelCluster UI as an application while IAM Identity Center will be the identity provider.

The next step will add the AWS ParallelCluster UI as an application in IAM Identity Center. AWS ParallelCluster UI is a web portal that helps the user to manage their clusters. For more information see AWS ParallelCluster UI.

Setting up the application in Identity Center

1. Under **IAM Identity Center** > **Applications** (found on the left menu bar, click on Applications)
2. Click **Add Application**
3. Select **Add custom SAML 2.0 application**
4. Click **Next**
5. Select the display name and description you would like to use (e.g. PCUI and AWS ParallelCluster UI)
6. Under **IAM Identity Center metadata**, copy the link for IAM Identity Center SAML metadata file and save for later, this will be used when configuring SSO on the web app

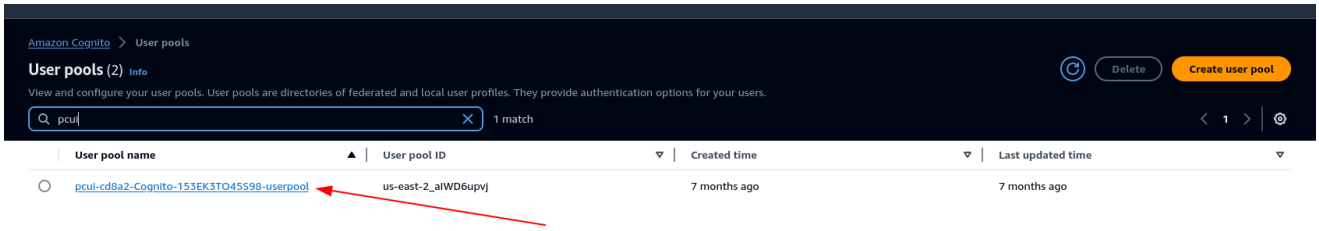
7. Under **Application properties**, in the Application start URL, put your PCUI address. This can be found by going to the CloudFormation console, selecting the stack that corresponds to PCUI (e.g. parallelcluster-ui) and going to the **Outputs** tab to find ParallelClusterUIUrl

   e.g. https://m2iwazsi1j.execute-api.us-east-1.amazonaws.com

8. Under **Application metadata**, choose **Manually type your metadata values**. Then provide the following values.

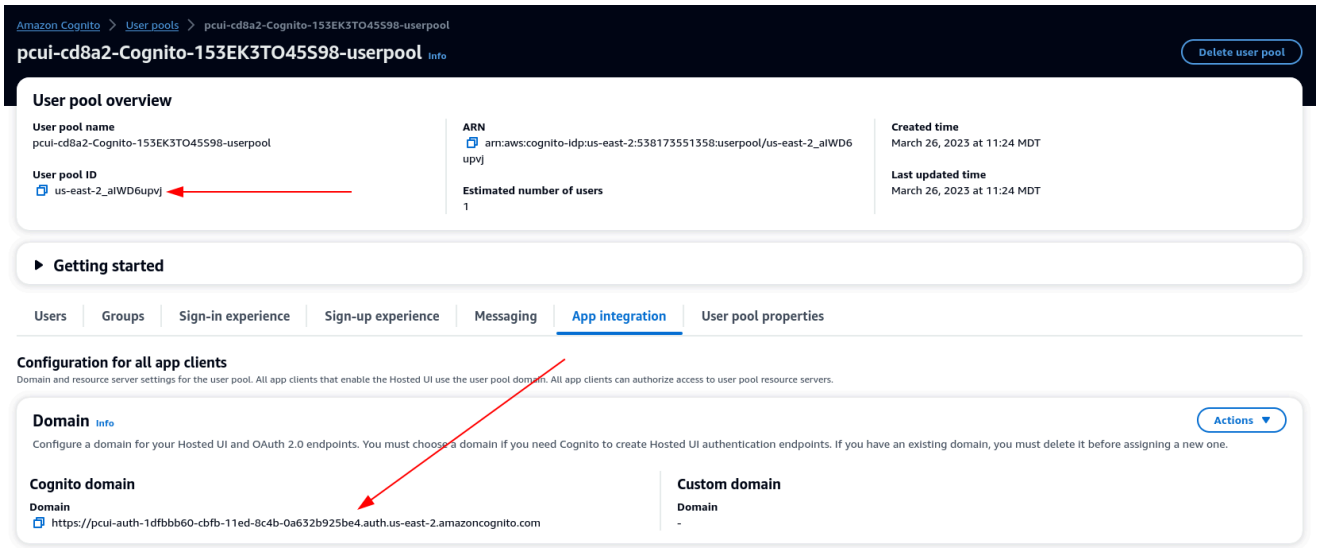
   a. **Important**: Make sure to replace the domain-prefix, region, and userpool-id values with information that's specific to your environment.

   b. The domain prefix, region and userpool-id can be obtained by opening the **Amazon Cognito** > **User pools console**



   c. Select the user pool that corresponds to PCUI (which will have a User pool name like pcui-cd8a2-Cognito-153EK3TO45S98-userpool)

   d. Navigate to **App Integration**



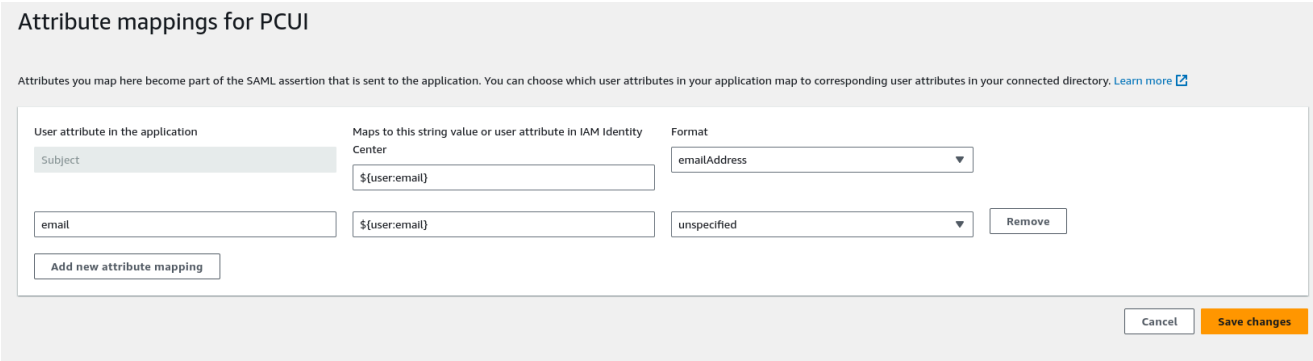
9. Application Assertion Consumer Service (ACS) URL: https://<domain-prefix>.auth.<region>.amazoncognito.com/saml2/idpresponse

   Application SAML audience: urn:amazon:cognito:sp:<userpool-id>

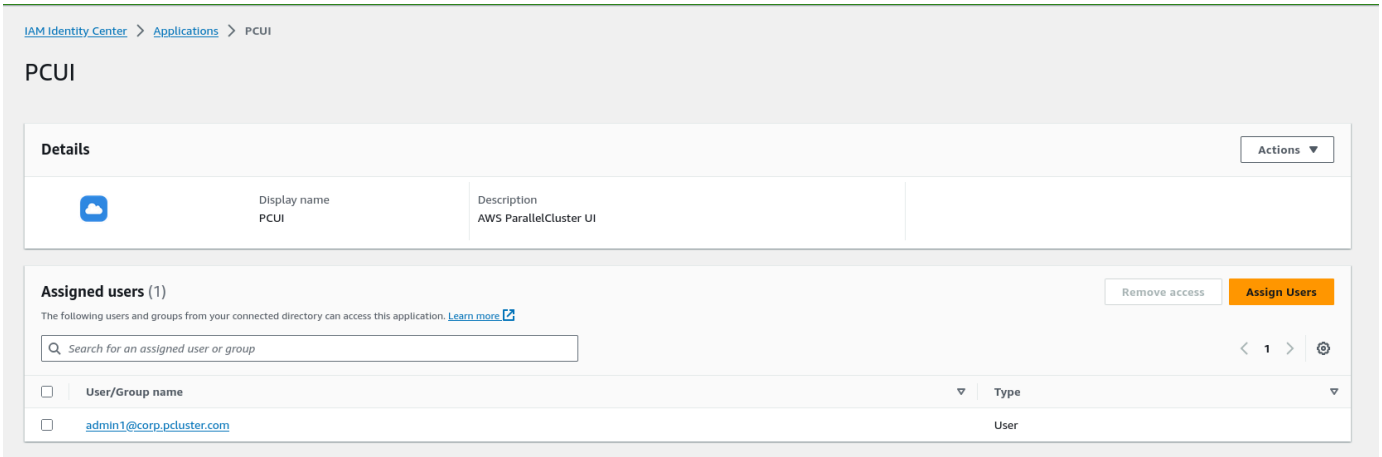
10. Choose **Submit**. Then, go to the **Details** page for the application that you added.
11. Select the **Actions** dropdown list and choose **Edit attribute mappings**. Then, provide the following attributes.

    a. User attribute in the application: **subject** (Note: **subject** is prefilled.) → Maps to this string value or user attribute in IAM Identity Center: **${user:email}**, Format: **emailAddress**

    b. User attribute in the application: **email** → Maps to this string value or user attribute in IAM Identity Center: **${user:email}**, Format: **unspecified**

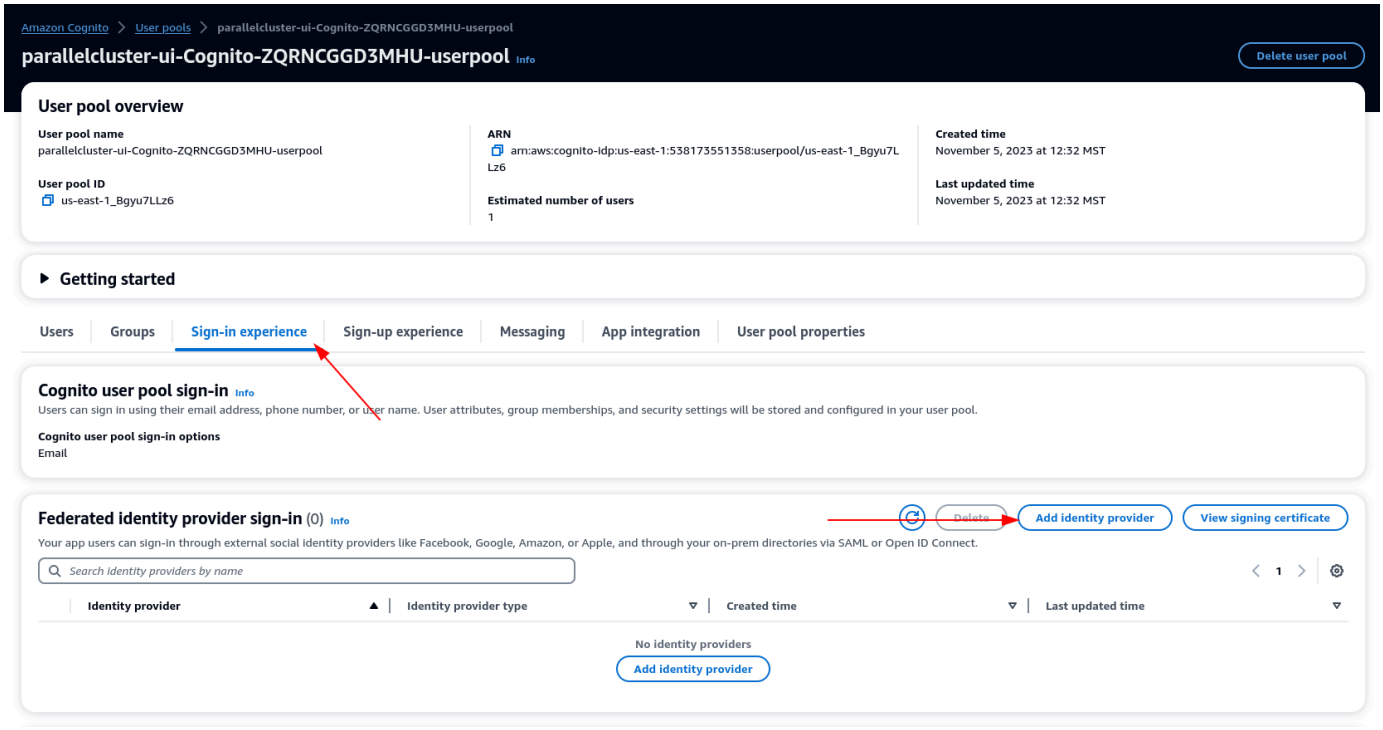


12. Save your changes.
13. Choose the **Assign Users** button and then assign your user to the application. These are the users in your Active Directory that will have access to the PCUI interface.

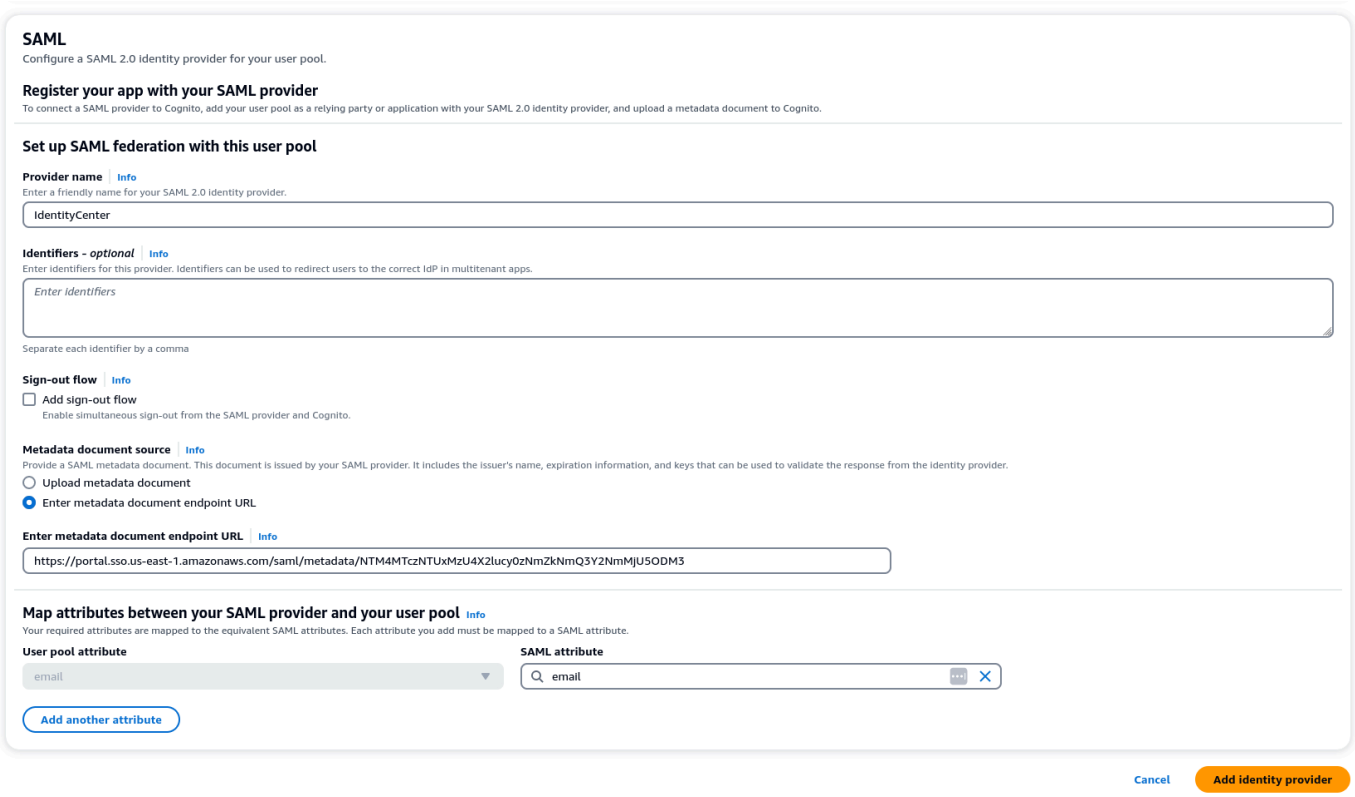


Configure IAM Identity Center as a SAML IdP in your user pool

1. In your user pool settings, select **Sign-in experience** > **Add identity provider**



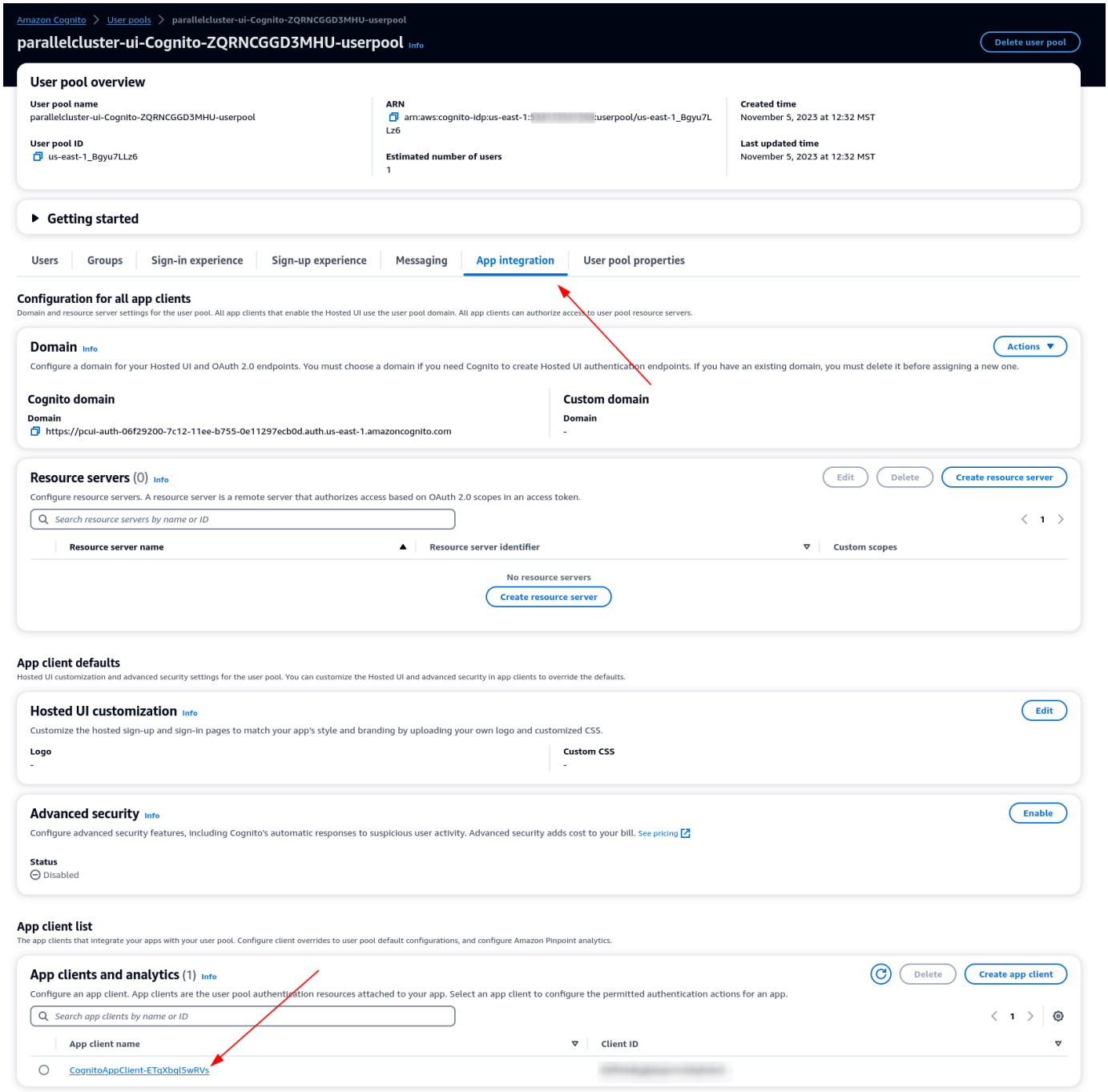
2. Choose a SAML IdP
3. For **Provider name** provide **IdentityCenter**
4. Under **Metadata document source** choose **Enter metadata document endpoint URL** and provide the URL copied during the Application setup of Identity Center
5. Under the **Attributes**, for email choose email



6. Select **Add identity provider**.

Integrate the IdP with the user pool app client

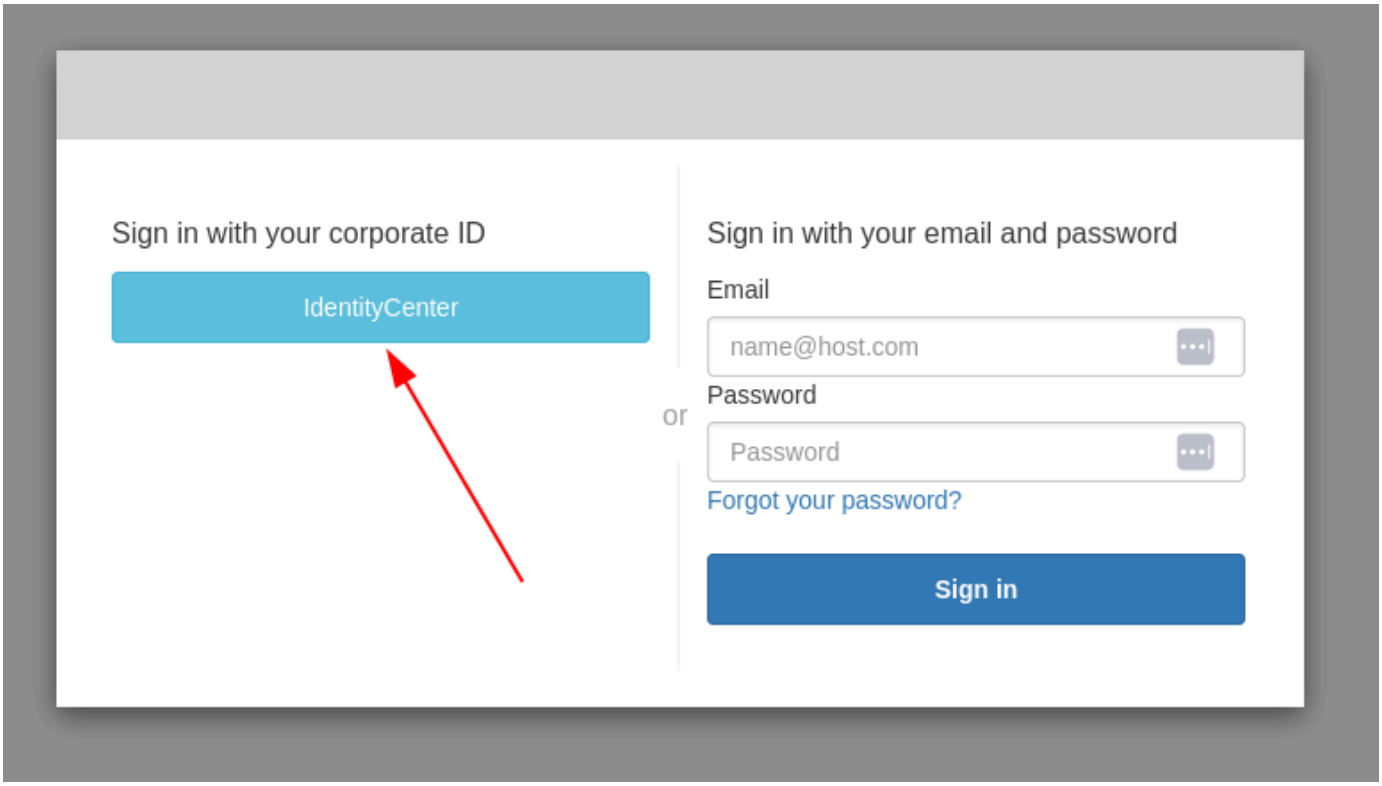
1. Next, under the **App Integration** section of your user pool, choose the client listed under **App client list**



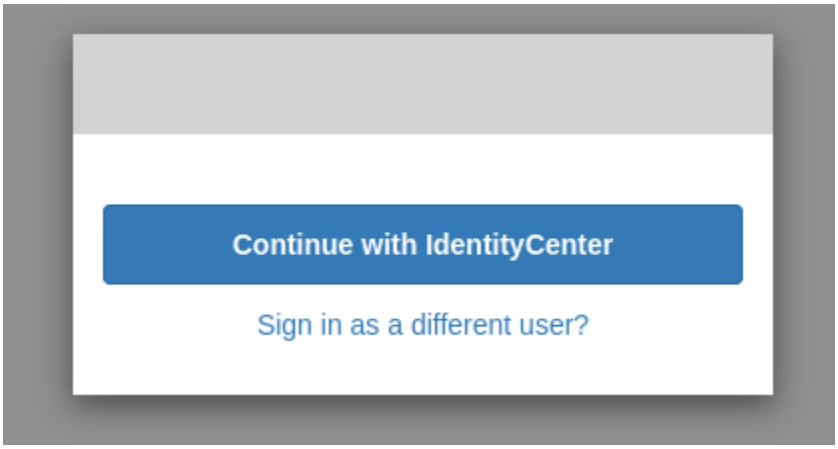
2. Under **Hosted UI** choose **Edit**
3. Under **Identity providers** choose **IdentityCenter** as well.
4. Choose **Save changes**

Validate your setup

1. Next we will validate the setup that we just created by logging in to PCUI. Sign in to your PCUI portal and you should now see an option to sign in with your Corporate ID:

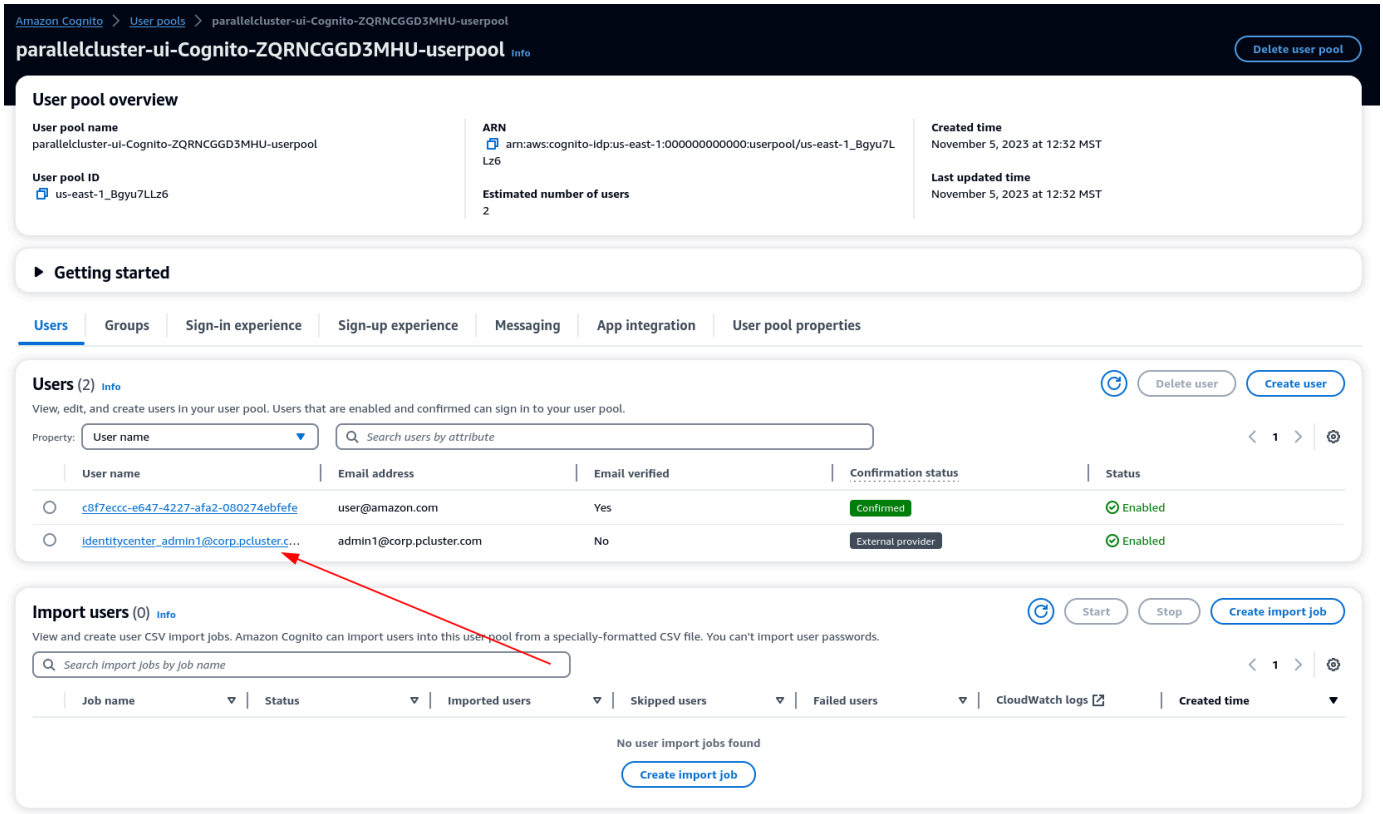


2. Clicking the **IdentityCenter** button should take you to the IAM Identity Center IdP login followed by a page with your applications on it which includes PCUI, open that application.
3. Once you get to the following screen, your user will have been added to the Cognito user pool.



Make your user an administrator

1. Now navigate to the **Amazon Cognito** > **User pools console** and select the newly created user which should have a prefix of identitycenter



2. Under **Group memberships** select **Add user to group**, choose **admin** and click **Add**.
3. Now when you click **Continue with IdentityCenter** you will be navigated to the AWS ParallelCluster UI page.

Running containerized jobs with Pyxis

Learn how to create a cluster that is able to run containerized jobs using Pyxis, which is a SPANK plugin to manage containerized jobs in SLURM. Containers in Pyxis are managed by Enroot, which is tool to turn traditional container/OS images into unprivileged sandboxes. For more information, see NVIDIA Pyxis and NVIDIA Enroot.

Note

- This feature is available with AWS ParallelCluster v3.11.1
- The scripts in this tutorial move (`mv`) some files, which deletes them from their original locations. If you want to keep copies of these files in their original locations, change the scripts to use the copy (`cp`) command instead.

When using AWS ParallelCluster, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

Prerequisites:

- The AWS CLI is installed and configured.
- An Amazon EC2 key pair..
- An IAM role with the permissions that are required to run the pcluster CLI.

Create the cluster

Starting with AWS ParallelCluster 3.11.1, all official AMIs comes with Pyxis and Enroot pre-installed. In particular, SLURM is recompiled with Pyxis support and Enroot is installed as a binary in the system. However, you must to configure them according to your specific needs. The folders used by Enroot and Pyxis will have a critical impact on cluster performance. For more information, see Pyxis documentation and Enroot documentation.

For your convenience, you can find sample configurations for both Pyxis, Enroot and SPANK within `/opt/parallelcluster/examples/`.

To deploy a cluster using the sample configurations we have provided, complete the following tutorial.

To create the cluster with sample configuration

Pyxis and Enroot must be configured on the head node by first creating the persistent and volatile directories for Enroot, then creating the runtime directory for Pyxis, and finally enabling Pyxis as SPANK plugin in the whole cluster.

1. Execute the below script as OnNodeConfigured custom action in the head node to configure Pyxis and Enroot on the head node.

```
#!/bin/bash
set -e

echo "Executing $0"

# Configure Enroot
```

```
ENROOT_PERSISTENT_DIR="/var/enroot"
ENROOT_VOLATILE_DIR="/run/enroot"

sudo mkdir -p $ENROOT_PERSISTENT_DIR
sudo chmod 1777 $ENROOT_PERSISTENT_DIR
sudo mkdir -p $ENROOT_VOLATILE_DIR
sudo chmod 1777 $ENROOT_VOLATILE_DIR
sudo mv /opt/parallelcluster/examples/enroot/enroot.conf /etc/enroot/enroot.conf
sudo chmod 0644 /etc/enroot/enroot.conf

# Configure Pyxis
PYXIS_RUNTIME_DIR="/run/pyxis"

sudo mkdir -p $PYXIS_RUNTIME_DIR
sudo chmod 1777 $PYXIS_RUNTIME_DIR

sudo mkdir -p /opt/slurm/etc/plugstack.conf.d/
sudo mv /opt/parallelcluster/examples/spank/plugstack.conf /opt/slurm/etc/
sudo mv /opt/parallelcluster/examples/pyxis/pyxis.conf /opt/slurm/etc/
plugstack.conf.d/
sudo -i scontrol reconfigure
```

2. Pyxis and Enroot must be configured on the compute fleet by creating the persistent and volatile directories for Enroot and the runtime directory for Pyxis. Execute the below script as OnNodeStart custom action in compute nodes to configure Pyxis and Enroot on the compute fleet.

```
#!/bin/bash
set -e

echo "Executing $0"

# Configure Enroot
ENROOT_PERSISTENT_DIR="/var/enroot"
ENROOT_VOLATILE_DIR="/run/enroot"
ENROOT_CONF_DIR="/etc/enroot"

sudo mkdir -p $ENROOT_PERSISTENT_DIR
sudo chmod 1777 $ENROOT_PERSISTENT_DIR
sudo mkdir -p $ENROOT_VOLATILE_DIR
sudo chmod 1777 $ENROOT_VOLATILE_DIR
```

```
sudo mkdir -p $ENROOT_CONF_DIR
sudo chmod 1777 $ENROOT_CONF_DIR
sudo mv /opt/parallelcluster/examples/enroot/enroot.conf /etc/enroot/enroot.conf
sudo chmod 0644 /etc/enroot/enroot.conf

# Configure Pyxis
PYXIS_RUNTIME_DIR="/run/pyxis"

sudo mkdir -p $PYXIS_RUNTIME_DIR
sudo chmod 1777 $PYXIS_RUNTIME_DIR

# In Ubuntu24.04 Apparmor blocks the creation of unprivileged user namespaces,
# which is required by Enroot. So to run Enroot, it is required to disable this
 restriction.
# See https://ubuntu.com/blog/ubuntu-23-10-restricted-unprivileged-user-namespaces
source /etc/os-release
if [ "${ID}${VERSION_ID}" == "ubuntu24.04" ]; then
    echo "kernel.apparmor_restrict_unprivileged_userns = 0" | sudo tee /etc/
sysctl.d/99-pcluster-disable-apparmor-restrict-unprivileged-userns.conf
    sudo sysctl --system
fi
```

Submit jobs

Now that Pyxis is configured in your cluster, you can submit containerized jobs using the sbatch and srun command, that are now enriched with container specific options.

```
# Submitting an interactive job
srun -N 2 --container-image docker://ubuntu:22.04 hostname

# Submitting a batch job
sbatch -N 2 --wrap='srun --container-image docker://ubuntu:22.04 hostname'
```

Creating a cluster with an EFA-enabled FSx Lustre

In this tutorial, you will create a cluster that uses an EFA-enabled FSx Lustre file system as shared storage. Using an FSx Lustre file system with EFA enabled can provide a boost in performance up to

8x. To verify if an EFA-enabled file system is what you need, look at Working with EFA-enabled file systems in the *FSx for Lustre User Guide*.

When you use AWS ParallelCluster, you only pay for the AWS resources that are created when you create or update AWS ParallelCluster images and clusters. For more information, see AWS services used by AWS ParallelCluster.

Requirements

- The AWS CLI is installed and configured.
- The ParallelCluster CLI is installed and configured.
- An Amazon EC2 key pair to log into the cluster.
- An IAM role with the permissions that are required to run the ParallelCluster CLI.

Create Security Groups

Create two security groups in the same VPC where the cluster and the file system will be deployed: one for the client running on cluster nodes and one for the file system.

```
# Create security group for the FSx client
aws ec2 create-security-group \
    --group-name Fsx-Client-SecurityGroup \
    --description "Allow traffic for the FSx Lustre client" \
    --vpc-id vpc-cluster \
    --region region

# Create security group for the FSx file system
aws ec2 create-security-group \
    --group-name Fsx-FileSystem-SecurityGroup \
    --description "Allow traffic for the FSx Lustre File System" \
    --vpc-id vpc-cluster \
    --region region
```

In the remainder of the tutorial, we will assume `sg-client` and `sg-file-system` are the security group ids of the client and file system, respectively.

Configure the security group for the client to allow all outbound traffic to the file system, as required by EFA.

```
# Allow all outbound traffic from the client to the file system
aws ec2 authorize-security-group-egress \
 --group-id sg-client \
 --protocol -1 \
 --port -1 \
 --source-group sg-file-system \
 --region region
```

Configure the security group for the file system to allow all inbound/outbound traffic within itself and all inbound traffic from the client, as required by EFA.

```
# Allow all inbound traffic within this security group
aws ec2 authorize-security-group-ingress \
    --group-id sg-file-system \
    --protocol -1 \
    --port -1 \
    --source-group sg-file-system \
    --region region

# Allow all outbound traffic within this security group
aws ec2 authorize-security-group-egress \
    --group-id sg-file-system \
    --protocol -1 \
    --port -1 \
    --source-group sg-file-system \
    --region region

# Allow all inbound traffic from the client
aws ec2 authorize-security-group-ingress \
    --group-id sg-file-system \
    --protocol -1 \
    --port -1 \
    --source-group sg-client \
    --region region

# Allow all outbound traffic to the client
aws ec2 authorize-security-group-egress \
    --group-id sg-file-system \
    --protocol -1 \
    --port -1 \
    --source-group sg-client \
    --region region
```

Create the file system

Create the file system within the same Availability Zone (AZ) where the compute nodes will be; and replace *subnet-compute-nodes* with its ID in the following code. This is required to allow EFA work with your file system. Note that, as part of the file system creation, we enable EFA using the EfaEnable property.

```
aws fsx create-file-system \
    --file-system-type LUSTRE \
    --storage-capacity 38400 \
    --storage-type SSD \
    --subnet-ids subnet-compute-nodes \
    --security-group-ids sg-file-system \
    --lustre-configuration
 DeploymentType=PERSISTENT_2,PerUnitStorageThroughput=125,EfaEnabled=true,MetadataConfiguration
 \
    --region region
```

Take note of the file system id returned by the previous command. In the remainder of the tutorial, replace *fs-id* with this file system id.

Create the cluster

1. Create the cluster with the following configurations set in the AWS ParallelCluster YAML configuration file:
   a. AMI based on a supported OS, such as Ubuntu 22.04.
   b. Compute nodes must use an EFA supported instance type having Nitro v4+, such as g6.16xlarge.
      - Compute nodes must be in the same AZ where the file system is.
      - Compute nodes must have Efa/Enabled set to true.
      - Compute nodes must run the configuration script `configure-efa-fsx-lustre-client.sh` as an OnNodeStart custom action. The script, provided in the FSx official documentation and offered in our public bucket for your convenience, is meant to configure the FSx Lustre client on compute nodes to let them use EFA.
2. Create a cluster configuration file `config.yaml`:

```
Region: region
Image:
```

```
  Os: ubuntu2204
HeadNode:
  InstanceType: c5.xlarge
  Networking:
    SubnetId: subnet-xxxxxxxxxx
    AdditionalSecurityGroups:
        - sg-client
  Ssh:
    KeyName: my-ssh-key
Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: q1
      ComputeResources:
        - Name: cr1
          Instances:
            - InstanceType: g6.16xlarge
          MinCount: 1
          MaxCount: 3
          Efa:
            Enabled: true
      Networking:
        SubnetIds:
          - subnet-xxxxxxxxxx # Subnet in the same AZ where the file system is
        AdditionalSecurityGroups:
          - sg-client
        PlacementGroup:
          Enabled: false
      CustomActions:
        OnNodeStart:
          Script: https://us-east-1-aws-parallelcluster.s3.us-east-1.amazonaws.com/
scripts/fsx-lustre-efa/configure-efa-fsx-lustre-client.sh
SharedStorage:
  - MountDir: /fsx
    Name: my-fsxlustre-efa-external
    StorageType: FsxLustre
    FsxLustreSettings:
      FileSystemId: fs-id
```

Then create a cluster using that configuration:

```
pcluster create-cluster \
    --cluster-name fsx-efa-tutorial \
```

```
    --cluster-configuration config.yaml \
    --region region
```

Validate FSx with EFA is working

To verify that Lustre network traffic is using EFA, use the Lustre `lnetctl` tool that can show the network traffic for a given network interface. To this aim, execute the following commands in a compute node:

```
# Take note of the number of packets flowing through the interface,
# which are specified in statistics:send_count and statistics:recv_count
sudo lnetctl net show --net efa -v

# Generate traffic to the file system
echo 'Hello World' > /fsx/hello-world.txt

# Take note of the number of packets flowing through the interface,
# which are specified in statistics:send_count and statistics:recv_count
sudo lnetctl net show --net efa -v
```

If the feature is working, the number of packets flowing through the interface is expected to increase.

AWS ParallelCluster troubleshooting

The following sections provide troubleshooting tips for issues that might occur while using AWS ParallelCluster. The AWS ParallelCluster community maintains a Wiki page that provides many troubleshooting tips on the AWS ParallelCluster GitHub Wiki. For a list of known issues, see Known issues.

Topics

Trying to create a cluster

When using AWS ParallelCluster version 3.5.0 and later to create a cluster, and a cluster creation failed with `--rollback-on-failure` set to `false`, use the `pcluster describe-cluster` CLI command to get status and failure information. In this case, the expected `clusterStatus` of the `pcluster describe-cluster` output is `CREATE_FAILED`. Check the `failures` section in the output to find the `failureCode` and `failureReason`. Then, in the following section, find the matching `failureCode` for additional troubleshooting help. For more information, see `pcluster describe-cluster`.

In the following sections, we recommend that you check the logs on the head node, such as the `/var/log/cfn-init.log` and `/var/log/chef-client.log` files. For more information about AWS ParallelCluster logs and how to view them, see Key logs for debugging and Retrieving and preserving logs.

If you don't have a `failureCode`, navigate to the AWS CloudFormation console to view the cluster stack. Check the `Status Reason` for the `HeadNodeWaitCondition` or failures on other resources to find additional failure details. For more information, see View AWS CloudFormation events on `CREATE_FAILED`. Check the `/var/log/cfn-init.log` and `/var/log/chef-client.log` files on the head node. If cluster creation fails because of head node creation failure and the cluster logs are not available in the cluster log group, you must retain the cluster on failure, specify `--rollback-on-failure` = `True` and retrieve the logs from within the head node itself.

`failureCode` is `OnNodeConfiguredExecutionFailure`

- **Why did it fail?**

  You provided a custom script in `OnNodeConfigured` of the head node section in the configuration to create a cluster. However, the custom script failed to run.

- **How to resolve?**

  Check the `/var/log/cfn-init.log` file to learn more about the failure and how to fix the issue in your custom script. Near the end of this log, you might see run information related to the `OnNodeConfigured` script after the `Running command runpostinstall` message.

`failureCode` is `OnNodeConfiguredDownloadFailure`

- **Why did it fail?**

  You provided a custom script in `OnNodeConfigured` of the head node section in the configuration to create a cluster. However, the custom script failed to download.

- **How to resolve?**

  Make sure that the URL is valid and that the access is correctly configured. For more information on the configuration of custom bootstrap scripts, see Custom bootstrap actions.

  Check the `/var/log/cfn-init.log`file. Near the end of this log, you might see run information related to `OnNodeConfigured` script processing, including downloading, after the `Running command runpostinstall` message.

`failureCode` is `OnNodeConfiguredFailure`

- **Why did it fail?**

  You provided a custom script in `OnNodeConfigured` of the head node section in the configuration to create a cluster. However, the use of the custom script failed in the cluster deployment. An immediate cause can't be determined and additional investigation is needed.

- **How to resolve?**

  Check the `/var/log/cfn-init.log`file. Near the end of this log, you might see run information related to `OnNodeConfigured` script processing after the `Running command runpostinstall` message.

`failureCode` is `OnNodeStartExecutionFailure`

- **Why did it fail?**

  You provided a custom script in `OnNodeStart` of the head node section in the configuration to create a cluster. However, the custom script failed to run.

- **How to resolve?**

Check the `/var/log/cfn-init.log` file to learn more about the failure and how to fix the issue in your custom script. Near the end of this log, you might see run information related to the `OnNodeStart` script after the `Running command runpreinstall` message.

`failureCode` is `OnNodeStartDownloadFailure`

- **Why did it fail?**

  You provided a custom script in `OnNodeStart` of the head node section in the configuration to create a cluster. However, the custom script failed to download.

- **How to resolve?**

  Make sure that the URL is valid and that the access is correctly configured. For more information on the configuration of custom bootstrap scripts, see Custom bootstrap actions.

  Check the `/var/log/cfn-init.log`file. Near the end of this log, you might see run information related to `OnNodeStart` script processing, including downloading, after the `Running command runpreinstall` message.

`failureCode` is `OnNodeStartFailure`

- **Why did it fail?**

  You provided a custom script in the `OnNodeStart` of the head node section in the configuration to create a cluster. However, the use of the custom script failed in the cluster deployment. An immediate cause can't be determined and additional investigation is needed.

- **How to resolve?**

  Check the `/var/log/cfn-init.log`file. Near the end of this log, you might see run information related to `OnNodeStart` script processing after the `Running command runpreinstall` message.

`failureCode` is `EbsMountFailure`

- **Why did it fail?**

The EBS volume defined in the cluster configuration failed to mount.

- **How to resolve?**

  Check the `/var/log/chef-client.log` file for failure details.

`failureCode` is `EfsMountFailure`

- **Why did it fail?**

  The Amazon EFS volume defined in the cluster configuration failed to mount.

- **How to resolve?**

  If you defined an existing Amazon EFS file system, make sure that traffic is allowed between the cluster and the file system. For more information, see `SharedStorage` / `EfsSettings` / `FileSystemId`.

  Check the `/var/log/chef-client.log` file for failure details.

`failureCode` is `FsxMountFailure`

- **Why did it fail?**

  The Amazon FSx file system defined in the cluster configuration failed to mount.

- **How to resolve?**

  If you defined an existing Amazon FSx file system, make sure that traffic is allowed between the cluster and the file system. For more information, see `SharedStorage` / `FsxLustreSettings` / `FileSystemId`.

  Check the `/var/log/chef-client.log` file for failure details.

`failureCode` is `RaidMountFailure`

- **Why did it fail?**

  The RAID volumes defined in the cluster configuration failed to mount.

- **How to resolve?**

Check the `/var/log/chef-client.log` file for failure details.

`failureCode` is `AmiVersionMismatch`

- **Why did it fail?**

  The AWS ParallelCluster version used to create the custom AMI is different than the AWS ParallelCluster version used to configure the cluster. In the CloudFormation console, view the cluster CloudFormation stack details and check the `Status Reason` for the `HeadNodeWaitCondition` to get additional details on the AWS ParallelCluster versions and the AMI. For more information, see View AWS CloudFormation events on `CREATE_FAILED`.

- **How to resolve?**

  Make sure the AWS ParallelCluster version used to create the custom AMI is the same AWS ParallelCluster version used to configure the cluster. You can change either the custom AMI version or the `pcluster` CLI version to make them the same.

`failureCode` is `InvalidAmi`

- **Why did it fail?**

  The custom AMI is invalid because it wasn't built using AWS ParallelCluster.

- **How to resolve?**

  Use the `pcluster build-image` command to create an AMI by making your AMI the parent image. For more information, see `pcluster build-image`.

`failureCode` is `HeadNodeBootstrapFailure` with `failureReason` Failed to set up the head node.

- **Why did it fail?**

  An immediate cause can't be determined and additional investigation is needed. For example, it could be that the cluster is in protected status, and this could be caused by a failure to provision the static compute fleet.

- **How to resolve?**

  Check the `/var/log/chef-client.log.` file for failure details.

  > **Note**
  >
  > If you see `RuntimeError` exception `Cluster state has been set to PROTECTED mode due to failures detected in static node provisioning`, the cluster is in protected status. For more information, see How to debug protected mode.

`failureCode` is `HeadNodeBootstrapFailure` with `failureReason` Cluster creation timed out.

- **Why did it fail?**

  By default, there is a 30 minute time limit for cluster creation to complete. If cluster creation hasn't completed within this time frame, the cluster creation fails with a timeout error. The cluster creation can timeout for different reasons. For example, timeout failures can be caused by a head node creation failure, a network issue, custom scripts that take too long to run in the head node, an error in a custom script that runs in compute nodes, or long wait times for compute node provisioning. An immediate cause can't be determined and additional investigation is needed.

- **How to resolve?**

  Check the `/var/log/cfn-init.log` and `/var/log/chef-client.log` files for failure details. For more information about AWS ParallelCluster logs and how to get them, see Key logs for debugging and Retrieving and preserving logs.

  You might discover the following in these logs.

  - **Seeing `Waiting for static fleet capacity provisioning` near the end of the `chef-client.log`**

    This indicates that the cluster creation timed out when waiting for static nodes to power up. For more information, see Seeing errors in compute node initializations.

  - **Seeing `OnNodeConfigured` or `OnNodeStart` head node script hasn't finished at the end of the `cfn-init.log`**

This indicates that the `OnNodeConfigured` or `OnNodeStart` custom script took a long time to run and caused a timeout error. Check your custom script for issues that might cause it to run for a long time. If your custom script requires a long time to run, consider changing the timeout limit by adding a `DevSettings` section to your cluster configuration file, as shown in the following example:

```
DevSettings:
  Timeouts:
    HeadNodeBootstrapTimeout: 1800 # default setting: 1800 seconds
```

- **Can't find the logs, or the head node wasn't created successfully**

  It's possible that the head node wasn't created successfully and the logs can't be found. In the CloudFormation console, view the cluster stack details to check for additional failure details.

`failureCode` is `HeadNodeBootstrapFailure` with `failureReason` Failed to bootstrap the head node.

- **Why did it fail?**

  An immediate cause can't be determined and additional investigation is needed.

- **How to resolve?**

  Check the `/var/log/cfn-init.log` and `/var/log/chef-client.log` files.

`failureCode` is `ResourceCreationFailure`

- **Why did it fail?**

  The creation of some resources failed during the cluster creation process. The failure can occur for various reasons. For example, resource creation failures can be caused by capacity issues or a misconfigured IAM policy.

- **How to resolve?**

  In the CloudFormation console, view the cluster stack to check for additional resource creation failure details.

`failureCode` is `ClusterCreationFailure`

- **Why did it fail?**

  An immediate cause can't be determined and additional investigation is needed.

- **How to resolve?**

  In the CloudFormation console, view the cluster stack and check the `Status Reason` for the `HeadNodeWaitCondition` to find additional failure details.

  Check the `/var/log/cfn-init.log` and `/var/log/chef-client.log` files.

Seeing `WaitCondition timed out...` in CloudFormation stack

For more information, see [`failureCode` is `HeadNodeBootstrapFailure` with `failureReason` `Cluster creation timed out.`.]

Seeing `Resource creation cancelled` in CloudFormation stack

For more information, see [`failureCode` is `ResourceCreationFailure`].

Seeing `Failed to run cfn-init...` or other errors in the AWS CloudFormation stack

Check the `/var/log/cfn-init.log` and `/var/log/chef-client.log` for additional failure details.

Seeing `chef-client.log` ends with `INFO: Waiting for static fleet capacity provisioning`

This is related to cluster creation timeout when waiting for static nodes to power up. For more information, see Seeing errors in compute node initializations.

Seeing `Failed to run preinstall or postinstall in cfn-init.log`

You have an `OnNodeConfigured` or `OnNodeStart` script in the cluster configuration `HeadNode` section. The script isn't working correctly. Check the `/var/log/cfn-init.log` file for custom script error details.

Seeing `This AMI was created with xxx, but is trying to be used with xxx...` in CloudFormation stack

For more information, see `failureCode` is `AmiVersionMismatch`.

Seeing `This AMI was not baked by AWS ParallelCluster...` in CloudFormation stack

For more information, see `failureCode` is `InvalidAmi`.

Seeing `pcluster create-cluster` command fails to run locally

Check the `~/.parallelcluster/pcluster-cli.log` in your local file system for failure details.

Additional support

Follow the troubleshooting guidance in Troubleshooting cluster deployment issues.

Check to see if your scenario is covered in GitHub Known Issues at AWS ParallelCluster on GitHub.

Trying to run a job

The following section provides possible troubleshooting solutions if you run in to issues while trying to run a job.

`srun` interactive job fails with error `srun: error: fwd_tree_thread: can't find address for <host>, check slurm.conf`

- **Why did it fail?**

You ran the `srun` command to submit a job, and then you increased the size of a queue by using the `pcluster update-cluster` command without restarting the Slurm daemons after the update completed.

Slurm organizes Slurm daemons in a tree hierarchy to optimize communication. This hierarchy is only updated when the daemons start.

Suppose you use `srun` to launch a job and then run the `pcluster update-cluster` command to increase the size of the queue. New compute nodes launch as part of the update. Then, Slurm queues your job to one of the new compute nodes. In this case, both the Slurm daemons and `srun` don't detect the new compute nodes. `srun` returns an error because it doesn't detect the new nodes.

- **How to resolve?**

  Restart the Slurm daemons on all of the compute nodes, and then use `srun` to submit your job. You can schedule the Slurm daemons restart by running the `scontrol reboot` command that restarts the compute nodes. For more information, see scontrol reboot in the Slurm documentation. You can also manually restart the Slurm daemons on the compute nodes by requesting a restart of the corresponding `systemd` services.

Job is stuck in CF state with squeue command

This might be an issue with dynamic nodes powering up. For more information, see Seeing errors in compute node initializations.

Running large scale jobs and seeing `nfsd: too many open connections, consider increasing the number of threads in /var/log/messages`

With a networked file system, when network limits are reached, I/O wait time also increases. This can result in soft lockups because the network is used to write data for both networking and I/O metrics.

With 5th generation instances, we use the ENA driver to expose packet counters. These counters count the packets shaped by AWS when the network reaches instance bandwidth limits. You can check these counters to see if they are greater than 0. If they are, then you have exceeded

your bandwidth limits. You can view these counters by running `ethtool -S eth0 | grep exceeded`.

Exceeding network limits is often a result of supporting too many NFS connections. This is one of the first things to check when you reach or exceed network limits.

For example, the following output shows dropped packages:

```
$ ethtool -S eth0 | grep exceeded
  bw_in_allowance_exceeded: 38750610
  bw_out_allowance_exceeded: 1165693
  pps_allowance_exceeded: 103
  conntrack_allowance_exceeded: 0
  linklocal_allowance_exceeded: 0
```

To avoid getting this message, consider changing the head node instance type to a more performant instance type. Consider moving your data storage to shared storage file systems that aren't exported as an NFS share, such a Amazon EFS or Amazon FSx. For more information, see Shared storage and the Best Practices at the AWS ParallelCluster Wiki on GitHub.

Running an MPI job

Enabling debug mode

To enable OpenMPI debug mode, see What controls does Open MPI have that aid in debugging.

To enable IntelMPI debug mode, see Other Environment Variables.

Seeing `MPI_ERRORS_ARE_FATAL` and `OPAL ERROR` in the job output

These error codes come from the MPI layer in your application. To learn how to get MPI debug logs from your application, see Enabling debug mode.

A possible cause for this error is that your application has been compiled for a specific MPI implementation, such as OpenMPI, and you are trying to run it with a different MPI implementation, such as IntelMPI. Make sure you are both compiling and running your application with the same MPI implementation.

Using `mpirun` with managed DNS disabled

For clusters created with SlurmSettings / Dns / DisableManagedDns and UseEc2Hostnames set to `true`, the Slurm node name isn't resolved by the DNS. Slurm can bootstrap MPI processes when

nodenames aren't enabled and if the MPI job is run in a Slurm context. We recommend following the guidance in the Slurm MPI User's Guide to run MPI jobs with Slurm.

Trying to update a cluster

The following section provides possible troubleshooting solutions to issues that might happen while you're trying to update a cluster.

`pcluster update-cluster` command fails to run locally

Check the `~/.parallelcluster/pcluster-cli.log` in your local file system for failure details.

Seeing `clusterStatus` is `UPDATE_FAILED` with `pcluster describe-cluster` command

If the cluster stack update rolled back, check the `/var/log/chef-client.log` file for error details.

Check to see if your issue is mentioned in GitHub Known Issues at AWS ParallelCluster on GitHub.

The cluster update timed out

This could be an issue related to `cfn-hup` not running. If the `cfn-hup` demon is terminated by an external cause, it's not restarted automatically. If `cfn-hup` isn't running, during a cluster update, the CloudFormation stack starts the update process as expected, but the update procedure isn't activated on the head node and the stack deployment eventually times out. For more information, see Troubleshooting a cluster update timeout when `cfn-hup` isn't running to troubleshoot and recover from the issue.

Trying to access storage

Learn about the troubleshooting tips for trying to access storage.

Using an external Amazon FSx for Lustre file system

Make sure that traffic is allowed between the cluster and file system. The file system must be associated with a security group that allows inbound and outbound TCP traffic through ports

988, 1021, 1022, and 1023. For more information about how to set up security groups, see FileSystemId.

Using an external Amazon Elastic File System file system

Make sure that traffic is allowed between the cluster and file system. The file system must be associated with a security group that allows inbound and outbound TCP traffic through ports 988, 1021, 1022, and 1023. For more information about how to set up security groups, see FileSystemId.

Trying to delete a cluster

If you get an error while trying to delete a cluster, the following sections provide troubleshooting tips for the common scenarios.

The `pcluster delete-cluster` command fails to run locally

Check the `~/.parallelcluster/pcluster-cli.log` file in your local file system.

The cluster stack fails to delete

If the cluster stack fails to delete, check the CloudFormation stack events message.

Check if your issue is mentioned in GitHub Known Issues at AWS ParallelCluster on GitHub.

Trying to upgrade the AWS ParallelCluster API stack

If you get an error such as `UPDATE_FAILED` when you try to upgrade the AWS ParallelCluster API stack, we recommend that you check for a solution in the **Known Issues** sections of the AWS ParallelCluster Wiki on GitHub. For example, see ParallelCluster API Stack Upgrade Fails for ECR resources, which identifies one possible issue and provides mitigation options.

Seeing errors in compute node initializations

The following sections provide troubleshooting tips for when you see errors in compute node initializations. This includes bootstrap errors, seeing errors in logs, and where to go if none of the scenarios apply to your specific situation.

Topics

Seeing `Node bootstrap error` in `clustermgtd.log`

The problem is related to compute nodes failing to bootstrap. For information on how to debug a cluster protected mode issue, see How to debug protected mode.

I configured on demand capacity reservations (ODCRs) or zonal Reserved Instances

ODCRs that include instances that have multiple network interfaces, such as P4d, P4de, and AWS Trainium (Trn)

In the cluster configuration file, check that the `HeadNode` is in a public subnet and that the compute nodes are in a private subnet.

ODCRs are targeted ODCRS

Seeing `Unable to read file '/opt/slurm/etc/pcluster/run_instances_overrides.json'.` even though I already have `/opt/slurm/etc/pcluster/run_instances_overrides.json` in place by following the instructions given in Launch instances with On-Demand Capacity Reservations (ODCR)

If you are using AWS ParallelCluster versions 3.1.1 to 3.2.1 with targeted ODCRs, and you are also using the run instances override JSON file, it's possible that you don't have the JSON file formatted correctly. You could see an error in `clustermgtd.log`, such as the following:

```
Unable to read file '/opt/slurm/etc/pcluster/run_instances_overrides.json'.
Using default: {} in  /var/log/parallelcluster/clustermgtd.
```

Validate that the JSON file format is correct by running the following:

```
$ echo /opt/slurm/etc/pcluster/run_instances_overrides.json | jq
```

Seeing `Found RunInstances parameters override.` in `clustermgtd.log` when cluster creation failed, or in `slurm_resume.log` when run job failed

If you are using run instances override JSON file, check that you correctly set the queue name and the compute resources name in the `/opt/slurm/etc/pcluster/run_instances_overrides.json` file.

Seeing `An error occurred (InsufficientInstanceCapacity)` in `slurm_resume.log` when I fail to a run job, or in `clustermgtd.log` when I fail to create a cluster

Using PG-ODCR (Placement Group ODCR)

When creating an ODCR with an associated placement group, the same placement group name must be used in the configuration file. Set the corresponding placement group name in the cluster configuration.

Using zonal Reserved Instances

If you are using zonal Reserved Instances with `PlacementGroup` / `Enabled` to `true` in the cluster configuration, you might see an error, such as the following:

```
We currently do not have sufficient trn1.32xlarge capacity in the Availability Zone
 you requested (us-east-1d). Our system will be working on provisioning additional
 capacity.
You can currently get trn1.32xlarge capacity by not specifying an Availability Zone in
 your request or choosing us-east-1a, us-east-1b, us-east-1c, us-east-1e, us-east-1f.
```

You might see this because the zonal Reserved Instances aren't placed in the same UC (or spine), which can cause insufficient capacity errors (ICEs) when using placement groups. You can check this case by disabling the `PlacementGroupGroup` setting in the cluster configuration to determine if the cluster can allocate the instances.

Seeing `An error occurred (VcpuLimitExceeded)` in `slurm_resume.log` when I fail to run a job, or in `clustermgtd.log`, when I fail to create a cluster

Check the vCPU limits on your account for the specific Amazon EC2 instance type that you are using. If you see zero or fewer vCPUs than you are requesting, request an increase for your limits. For information about how to view current limits and request new limits, see Amazon EC2 service quotas in the *Amazon EC2 User Guide*.

Seeing `An error occurred (InsufficientInstanceCapacity)` in `slurm_resume.log` when I fail to run a job, or in `clustermgtd.log`, when I fail to create a cluster

You are experiencing an insufficient capacity issue. Follow https://aws.amazon.com/premiumsupport/knowledge-center/ec2-insufficient-capacity-errors/ to troubleshoot the issue.

Seeing nodes are in `DOWN` state with `Reason (Code:InsufficientInstanceCapacity)...`

You are experiencing an insufficient capacity issue. Follow https://aws.amazon.com/premiumsupport/knowledge-center/ec2-insufficient-capacity-errors/ to troubleshoot the issue. For more information about AWS ParallelCluster's fast insufficient capacity fail-over mode, see Slurm cluster fast insufficient capacity fail-over.

Seeing `cannot change locale (en_US.utf-8) because it has an invalid name` in `slurm_resume.log`

This can occur if you have an unsuccessful `yum` installation process that left the locale settings in an inconsistent state. For example, this can be caused when a user terminates the install process.

To verify the cause, take the following actions:

- Run `su - pcluster-admin`.

  The shell shows an error, such as, `cannot change locale...no such file or directory`.
- Run `localedef --list`.

Returns an empty list or doesn't contain the default locale.

- Check the last `yum` command with `yum history` and `yum history info #ID`. Does the last ID have `Return-Code: Success`?

  If the last ID doesn't have `Return-Code: Success`, the post-install scripts might not have run successfully.

To fix the issue, try rebuilding the locale with `yum reinstall glibc-all-langpacks`. After the rebuild, `su - pcluster-admin` doesn't show an error or warning if the issue is fixed.

None of the previous scenarios apply to my situation

To troubleshoot compute node initialization issues, see Troubleshooting node initialization issues.

Check to see if your scenario is covered in GitHub Known Issues at AWS ParallelCluster on GitHub.

Troubleshooting cluster health metrics

Cluster health metrics are added to the AWS ParallelCluster Amazon CloudWatch dashboard starting with AWS ParallelCluster version 3.6.0. In the following sections, you can learn about the dashboard health metrics, and actions you can take to troubleshoot and resolve issues.

Topics

Seeing the Instance Provisioning Errors graph

If you see a non-zero value in the `Instance Provisioning Errors` graph, then it means that the Amazon EC2 instance for backing slurm nodes failed to launch on the `CreateFleet` or `RunInstance` API.

Seeing `IAMPolicyErrors`

- **What happened?**

A number of instances failed to launch, which is caused by insufficient permissions with error code `UnauthorizedOperation`.

- **How to resolve?**

  If you have a configured a custom `InstanceRole` or `InstanceProfile`, check your IAM policies and verify that you are using the correct credentials.

  Check the `clustermgtd` file for static node error details. Check the `slurm_resume.log` file for dynamic node error details. Use the details to learn more about the missing permissions that must be added.

Seeing `VcpuLimitErrors`

- **What happened?**

  AWS ParallelCluster failed to launch instances because it reached the vCPU limit on your AWS account for a specific Amazon EC2 instance type that you configured for cluster compute nodes.

- **How to resolve?**

  Check for the `VcpuLimitExceeded` error in the `clustermgtd` file for static nodes, and check in the `slurm_resume.log` file for dynamic nodes to get additional details. To resolve this issue, you can request an increase to your vCPU limits. For more information about how to view current limits and request new limits, see Amazon Elastic Compute Cloud service quotas in the *Amazon Elastic Compute Cloud User Guide for Linux Instances*.

Seeing `VolumeLimitErrors`

- **What happened?**

  You have reached your Amazon EBS volume limit on your AWS account, and AWS ParallelCluster is unable to launch instances with error code `InsufficientVolumeCapacity` or `VolumeLimitExceeded`.

- **How to resolve?**

  Check the `clustermgtd` file for static nodes, and check the `slurm_resume.log` file for dynamic nodes to get additional volume limit details. To resolve this issue, you can use a

different AWS Region, clean up existing volumes, or contact the AWS Support Center to submit a request to increase your Amazon EBS volume limit.

Seeing `InsufficientCapacityErrors`

- **What happened?**

  AWS ParallelCluster doesn't have sufficient capacity to launch Amazon EC2 instances to back nodes.

- **How to resolve?**

  Check the `clustermgtd` file for static nodes, and check the `slurm_resume.log` file for dynamic nodes to get insufficient capacity error details. To troubleshoot the issue, follow the guidance at https://aws.amazon.com/premiumsupport/knowledge-center/ec2-insufficient-capacity-errors/.

`OtherInstanceLaunchFailures`

- **What happened?**

  The Amazon EC2 instance for backing compute nodes failed to launch with the `CreateFleet` or `RunInstance` API.

- **How to resolve?**

  Check the `clustermgtd` file for static nodes, and check the `slurm_resume.log` file for dynamic nodes to get error details.

Seeing the Unhealthy Instance Errors graph

- **What happened?**

  A number of compute instances were launched but later terminated as unhealthy.

- **How to resolve?**

  For more information about troubleshooting unhealthy nodes, see **Troubleshooting unexpected node replacements and terminations**.

Seeing `InstanceBootstrapTimeoutError`

- **What happened?**

  An instance can't join the cluster within the `resume_timeout` (for dynamic nodes) or `node_replacement_timeout` (for static nodes). This can occur if the network isn't configured correctly for the compute nodes, or it can occur if custom scripts running on the compute node take too long to finish.

- **How to resolve?**

  For dynamic nodes, check the `clustermgtd` log (`/var/log/parallelcluster/clustermgtd`) for the compute node IP address and errors such as the following:

  ```
  Node bootstrap error: Resume timeout expires for node
  ```

  For static nodes, check the `clustermgtd` log (`/var/log/parallelcluster/clustermgtd`) for the compute node IP address and errors such as the following:

  ```
  Node bootstrap error: Replacement timeout expires for node ... in replacement.
  ```

  For additional details, check the `/var/log/cloud-init-output.log` file for errors. You can retrieve problematic compute node IP addresses from the `clustermgtd` and `slurm_resume` log files.

Seeing `EC2HealthCheckErrors`

- **What happened?**

  An instance failed an Amazon EC2 health check.

- **How to resolve?**

  For information about how to troubleshoot this issue, see Troubleshoot instances with failed status checks.

Seeing `ScheduledEventHealthCheckErrors`

- **What happened?**

An instance failed an Amazon EC2 scheduled event health check, and it's unhealthy.

- **How to resolve?**

  For information about how to troubleshoot this issue, see Scheduled events for your instances.

Seeing `NoCorrespondingInstanceErrors`

- **What happened?**

  AWS ParallelCluster can't find instances backing nodes. The nodes have likely self-terminated during bootstrap operations. `SlurmQueues` / `CustomActions` / `OnNodeStart` | `OnNodeConfigured` script, or network errors can produce `NoCorrespondingInstanceErrors`.

- **How to resolve?**

  For additional details, check the `/var/log/cloud-init-output.log` for the compute node.

Seeing the Compute Fleet Idle Time graph

Seeing a `MaxDynamicNodeIdleTime` that is significantly longer than the Idle Time Scaledown threshold

- **What happened?**

  Your instance isn't terminating properly. `MaxDynamicNodeIdleTime` shows the maximum time in seconds that a dynamic node, backed by an Amazon EC2 instance, is idle. The **Idle Time Scaledown** threshold is derived from the cluster configuration `ScaledownIdletime` parameter. When a compute node has been idle for more than **Idle Time Scaledown** seconds, Slurm powers down the node and AWS ParallelCluster terminates the backing instance. In this case, something is preventing the instance termination.

- **How to resolve?**

  For more information about this issue, see **Replacing, terminating, or powering down problematic instances and nodes** in Troubleshooting scaling issues.

Troubleshooting cluster deployment issues

If your cluster fails to be created and rolls back stack creation, you can look through the log files to diagnose the issue. The failure message likely looks like the following output:

```
$ pcluster create-cluster --cluster-name mycluster --region eu-west-1 \
 --cluster-configuration cluster-config.yaml
{
  "cluster": {
    "clusterName": "mycluster",
    "cloudformationStackStatus": "CREATE_IN_PROGRESS",
    "cloudformationStackArn": "arn:aws:cloudformation:eu-west-1:xxx:stack/
mycluster/1bf6e7c0-0f01-11ec-a3b9-024fcc6f3387",
    "region": "eu-west-1",
    "version": "3.13.2",
    "clusterStatus": "CREATE_IN_PROGRESS"
  }
}

$ pcluster describe-cluster --cluster-name mycluster --region eu-west-1
{
  "creationTime": "2021-09-06T11:03:47.696Z",
  ...
  "cloudFormationStackStatus": "ROLLBACK_IN_PROGRESS",
  "clusterName": "mycluster",
  "computeFleetStatus": "UNKNOWN",
  "cloudformationStackArn": "arn:aws:cloudformation:eu-west-1:xxx:stack/
mycluster/1bf6e7c0-0f01-11ec-a3b9-024fcc6f3387",
  "lastUpdatedTime": "2021-09-06T11:03:47.696Z",
  "region": "eu-west-1",
  "clusterStatus": "CREATE_FAILED"
}
```

Topics

- View AWS CloudFormation events on CREATE_FAILED
- Use the CLI to view log streams
- Re-create the failed cluster with rollback-on-failure

View AWS CloudFormation events on CREATE_FAILED

You can use the console or the AWS ParallelCluster CLI to view CloudFormation events on CREATE_FAILED errors to help find the root cause.

Topics

View events in the CloudFormation console

To see more information about what caused the "CREATE_FAILED" status, you can use the CloudFormation console.

View CloudFormation error messages from the console.

1. Log in to the AWS Management Console and navigate to https://console.aws.amazon.com/cloudformation.
2. Select the stack named *cluster_name*.
3. Choose the **Events** tab.
4. Check the **Status** for the resource that failed to create by scrolling through the list of resource events by **Logical ID**. If a subtask failed to create, work backwards to find the failed resource event.
5. As an example, if you see the following status message, you must use instance types that won't exceed your current vCPU limit or request more vCPU capacity.

```
2022-02-04 16:09:44 UTC-0800 HeadNode CREATE_FAILED You have requested more vCPU
 capacity than your current vCPU limit of 0 allows
    for the instance bucket that the specified instance type belongs to. Please
 visit http://aws.amazon.com/contact-us/ec2-request to request an adjustment to
 this limit.
    (Service: AmazonEC2; Status Code: 400; Error Code: VcpuLimitExceeded; Request
 ID: a9876543-b321-c765-d432-dcba98766789; Proxy: null).
```

Use the CLI to view and filter CloudFormation events on CREATE_FAILED

To diagnose the cluster creation issue, you can use the pcluster get-cluster-stack-events command by filtering for CREATE_FAILED status. For more information, see Filtering AWS CLI output in the *AWS Command Line Interface User Guide*.

```
$ pcluster get-cluster-stack-events --cluster-name mycluster --region eu-west-1 \
    --query 'events[?resourceStatus==`CREATE_FAILED`]'
  [
    {
      "eventId": "3ccdedd0-0f03-11ec-8c06-02c352fe2ef9",
      "physicalResourceId": "arn:aws:cloudformation:eu-west-1:xxx:stack/
mycluster/1bf6e7c0-0f02-11ec-a3b9-024fcc6f3387",
      "resourceStatus": "CREATE_FAILED",
      "resourceStatusReason": "The following resource(s) failed to create: [HeadNode].
 ",
      "stackId": "arn:aws:cloudformation:eu-west-1:xxx:stack/
mycluster/1bf6e7c0-0f02-11ec-a3b9-024fcc6f3387",
      "stackName": "mycluster",
      "logicalResourceId": "mycluster",
      "resourceType": "AWS::CloudFormation::Stack",
      "timestamp": "2021-09-06T11:11:51.780Z"
    },
    {
      "eventId": "HeadNode-CREATE_FAILED-2021-09-06T11:11:50.127Z",
      "physicalResourceId": "i-04e91cc1f4ea796fe",
      "resourceStatus": "CREATE_FAILED",
      "resourceStatusReason": "Received FAILURE signal with UniqueId
 i-04e91cc1f4ea796fe",
      "resourceProperties": "{\"LaunchTemplate\":{\"Version\":\"1\",\"LaunchTemplateId
\":\"lt-057d2b1e687f05a62\"}}",
      "stackId": "arn:aws:cloudformation:eu-west-1:xxx:stack/
mycluster/1bf6e7c0-0f02-11ec-a3b9-024fcc6f3387",
      "stackName": "mycluster",
      "logicalResourceId": "HeadNode",
      "resourceType": "AWS::EC2::Instance",
      "timestamp": "2021-09-06T11:11:50.127Z"
    }
  ]
```

In the previous example, the failure was in the head node setup.

Use the CLI to view log streams

To debug this kind of issue, you can list the log streams available from the head node with the `pcluster list-cluster-log-streams` by filtering for `node-type` and then analyzing the log streams content.

```
$ pcluster list-cluster-log-streams --cluster-name mycluster --region eu-west-1 \
--filters 'Name=node-type,Values=HeadNode'
{
  "logStreams": [
    {
      "logStreamArn": "arn:aws:logs:eu-west-1:xxx:log-group:/aws/parallelcluster/
mycluster-202109061103:log-stream:ip-10-0-0-13.i-04e91cc1f4ea796fe.cfn-init",
      "logStreamName": "ip-10-0-0-13.i-04e91cc1f4ea796fe.cfn-init",
      ...
    },
    {
      "logStreamArn": "arn:aws:logs:eu-west-1:xxx:log-group:/aws/parallelcluster/
mycluster-202109061103:log-stream:ip-10-0-0-13.i-04e91cc1f4ea796fe.chef-client",
      "logStreamName": "ip-10-0-0-13.i-04e91cc1f4ea796fe.chef-client",
      ...
    },
    {
      "logStreamArn": "arn:aws:logs:eu-west-1:xxx:log-group:/aws/parallelcluster/
mycluster-202109061103:log-stream:ip-10-0-0-13.i-04e91cc1f4ea796fe.cloud-init",
      "logStreamName": "ip-10-0-0-13.i-04e91cc1f4ea796fe.cloud-init",
      ...
    },
    ...
  ]
}
```

The two primary log streams that you can use to find initialization errors are the following:

- `cfn-init` is the log for the `cfn-init` script. First check this log stream. You're likely to see the `Command chef failed` error in this log. Look at the lines immediately before this line for more specifics connected with the error message. For more information, see cfn-init.
- `cloud-init` is the log for cloud-init. If you don't see anything in `cfn-init`, then try checking this log next.

You can retrieve the content of the log stream by using the pcluster get-cluster-log-events (note the `--limit 5` option to limit the number of retrieved events):

```
$ pcluster get-cluster-log-events --cluster-name mycluster \
  --region eu-west-1 --log-stream-name ip-10-0-0-13.i-04e91cc1f4ea796fe.cfn-init \
  --limit 5
{
  "nextToken": "f/36370880979637159565202782352491087067973952362220945409/s",
  "prevToken": "b/36370880752972385367337528725601470541902663176996585497/s",
  "events": [
    {
      "message": "2021-09-06 11:11:39,049 [ERROR] Unhandled exception during build:
 Command runpostinstall failed",
      "timestamp": "2021-09-06T11:11:39.049Z"
    },
    {
      "message": "Traceback (most recent call last):\n  File \"/opt/aws/bin/
cfn-init\", line 176, in <module>\n    worklog.build(metadata, configSets)\n
 File \"/usr/lib/python3.7/site-packages/cfnbootstrap/construction.py\", line
 135, in build\n    Contractor(metadata).build(configSets, self)\n  File \"/
usr/lib/python3.7/site-packages/cfnbootstrap/construction.py\", line 561, in
 build\n    self.run_config(config, worklog)\n  File \"/usr/lib/python3.7/
site-packages/cfnbootstrap/construction.py\", line 573, in run_config\n
 CloudFormationCarpenter(config, self._auth_config).build(worklog)\n  File \"/usr/
lib/python3.7/site-packages/cfnbootstrap/construction.py\", line 273, in build\n
   self._config.commands)\n  File \"/usr/lib/python3.7/site-packages/cfnbootstrap/
command_tool.py\", line 127, in apply\n    raise ToolError(u\"Command %s failed\" %
 name)",
      "timestamp": "2021-09-06T11:11:39.049Z"
    },
    {
      "message": "cfnbootstrap.construction_errors.ToolError: Command runpostinstall
 failed",
      "timestamp": "2021-09-06T11:11:39.049Z"
    },
    {
      "message": "2021-09-06 11:11:49,212 [DEBUG] CloudFormation client initialized
 with endpoint https://cloudformation.eu-west-1.amazonaws.com",
      "timestamp": "2021-09-06T11:11:49.212Z"
    },
    {
      "message": "2021-09-06 11:11:49,213 [DEBUG] Signaling resource HeadNode in stack
 mycluster with unique ID i-04e91cc1f4ea796fe and status FAILURE",
```

```
      "timestamp": "2021-09-06T11:11:49.213Z"
    }
  ]
}
```

In the previous example, the failure is caused by a `runpostinstall` failure, so it is strictly related to the content of the custom bootstrap script used in the `OnNodeConfigured` configuration parameter of the `CustomActions`.

Re-create the failed cluster with `rollback-on-failure`

AWS ParallelCluster creates cluster CloudWatch log streams in log groups. You can view these logs in the CloudWatch console **Custom Dashboards** or **Log groups**. For more information, see Integration with Amazon CloudWatch Logs and Amazon CloudWatch dashboard. If there are no log streams available, the failure might be caused by the `CustomActions` custom bootstrap script or an AMI-related issue. To diagnose the creation issue in this case, create the cluster again using `pcluster create-cluster`, including the `--rollback-on-failure` parameter set to `false`. Then, use SSH to view the cluster, as shown in the following:

```
$ pcluster create-cluster --cluster-name mycluster --region eu-west-1 \
   --cluster-configuration cluster-config.yaml --rollback-on-failure false
 {
   "cluster": {
     "clusterName": "mycluster",
     "cloudformationStackStatus": "CREATE_IN_PROGRESS",
     "cloudformationStackArn": "arn:aws:cloudformation:eu-west-1:xxx:stack/
mycluster/1bf6e7c0-0f01-11ec-a3b9-024fcc6f3387",
     "region": "eu-west-1",
     "version": "3.13.2",
     "clusterStatus": "CREATE_IN_PROGRESS"
   }
 }
 $ pcluster ssh --cluster-name mycluster
```

After you're logged into the head node, you should find three primary log files that you can use to find the error.

- `/var/log/cfn-init.log` is the log for the `cfn-init` script. First check this log. You're likely to see an error such as `Command chef failed` in this log. Look at the lines immediately before this line for more specifics connected with the error message. For more information, see cfn-init.

- `/var/log/cloud-init.log` is the log for cloud-init. If you don't see anything in `cfn-init.log`, then try checking this log next.
- `/var/log/cloud-init-output.log` is the output of commands that were run by cloud-init. This includes the output from `cfn-init`. In most cases, you don't need to look at this log to troubleshoot this type of issue.

Troubleshooting cluster deployment using Terraform

This section is relevant to clusters that were deployed using Terraform.

ParallelCluster API not found

The planning could fail because the ParallelCluster API cannot be found. In this case, the returned error would be something like:

```
Planning failed. Terraform encountered an error while generating this plan.

#
# Error: Unable to retrieve ParallelCluster API cloudformation stack.
#
#   with provider["registry.terraform.io/aws-tf/aws-parallelcluster"],
#   on providers.tf line 6, in provider "aws-parallelcluster":
#    6: provider "aws-parallelcluster" {
#
# operation error CloudFormation: DescribeStacks, https response error StatusCode: 400,
 RequestID: REQUEST_ID, api error ValidationError: Stack with id PCAPI_STACK_NAME does
 not exist
```

To solve this error, deploy the ParallelCluster API in the account where the clusters are going to be created. See the section called “Creating a cluster with Terraform”.

User not authorized to call ParallelCluster API

The planning could fail because the IAM role/user you assumed to deploy your Terraform project doesn't have permissions to interact with the ParallelCluster API. In this case, the returned error would be something like:

```
Planning failed. Terraform encountered an error while generating this plan.

# Error: 403 Forbidden
```

```
#
#   with
 module.parallelcluster_clusters.module.clusters[0].pcluster_cluster.managed_configs["DemoClust
#   on .terraform/modules/parallelcluster_clusters/modules/clusters/main.tf line 35, in
 resource "pcluster_cluster" "managed_configs":
#   35: resource "pcluster_cluster" "managed_configs" {
#
# {{"Message":"User: USER_ARN is not authorized to perform: execute-api:Invoke on
 resource: PC_API_REST_RESOURCE with an explicit deny"}
# }
```

To solve this error, configure the ParallelCluster Provider so that it uses the ParallelCluster API role to interact with the API.

```
provider "aws-parallelcluster" {
  region         = var.region
  profile        = var.profile
  api_stack_name = var.api_stack_name
  **use_user_role** **= true**
}
```

Troubleshooting scaling issues

This section is relevant to clusters that were installed using AWS ParallelCluster version 3.0.0 and later with the Slurm job scheduler. For more information about configuring multiple queues, see Configuration of multiple queues.

If one of your running clusters is experiencing issues, place the cluster in a `STOPPED` state by running the following command before you begin to troubleshoot. This prevents incurring any unexpected costs.

```
$ pcluster update-compute-fleet --cluster-name mycluster \
   --status STOP_REQUESTED
```

You can list the log streams available from the cluster nodes by using the `pcluster list-cluster-log-streams` command and filtering by using the `private-dns-name` of one of the failing nodes or the head node:

```
$ pcluster list-cluster-log-streams --cluster-name mycluster --region eu-west-1 \
 --filters 'Name=private-dns-name,Values=ip-10-0-0-101'
```

Then, you can retrieve the content of the log stream to analyze it by using the `pcluster get-cluster-log-events` command and passing the `--log-stream-name` corresponding to one of the key logs mentioned in the following section:

```
$ pcluster get-cluster-log-events --cluster-name mycluster \
--region eu-west-1 --log-stream-name ip-10-0-0-13.i-04e91cc1f4ea796fe.cfn-init
```

AWS ParallelCluster creates cluster CloudWatch log streams in log groups. You can view these logs in the CloudWatch console **Custom Dashboards** or **Log groups**. For more information, see Integration with Amazon CloudWatch Logs and Amazon CloudWatch dashboard.

Topics

Key logs for debugging

The following table provides an overview of the key logs for the head node:

- `/var/log/cfn-init.log` - This is the AWS CloudFormation init log. It contains all commands that were run when an instance was set up. Use it to troubleshoot initialization issues.
- `/var/log/chef-client.log` - This is the Chef client log. It contains all commands that were run through Chef/CINC. Use it to troubleshoot initialization issues.
- `/var/log/parallelcluster/slurm_resume.log` - This is a `ResumeProgram` log. It launches instances for dynamic nodes. Use it to troubleshoot dynamic nodes launch issues.
- `/var/log/parallelcluster/slurm_suspend.log` - This is the `SuspendProgram` log. It's called when instances are terminated for dynamic nodes. Use it to troubleshoot dynamic nodes termination issues. When you check this log, you should also check the `clustermgtd` log.

- `/var/log/parallelcluster/clustermgtd` - This is the `clustermgtd` log. It runs as the centralized daemon that manages most cluster operation actions. Use it to troubleshoot any launch, termination, or cluster operation issues.
- `/var/log/slurmctld.log` - This is the Slurm control daemon log. AWS ParallelCluster doesn't make scaling decisions. Rather, it only attempts to launch resources to satisfy the Slurm requirements. It's useful for scaling and allocation issues, job-related issues, and any scheduler-related launch and termination issues.
- `/var/log/parallelcluster/compute_console_output` - This log records the console output from a sample subset of static compute nodes that have unexpectedly terminated. Use this log if static compute nodes terminate and the compute node logs aren't available in CloudWatch. The `compute_console_output log` content you receive is the same when you use the Amazon EC2 console or AWS CLI to retrieve the instance console output.

These are the key logs for the compute nodes:

- `/var/log/cloud-init-output.log` - This is the cloud-init log. It contains all commands that were run when an instance was set up. Use it to troubleshoot initialization issues.
- `/var/log/parallelcluster/computemgtd` - This is the `computemgtd` log. It runs on each compute node to monitor the node in the uncommon event that `clustermgtd` daemon on the head node is offline. Use it to troubleshoot unexpected termination issues.
- `/var/log/slurmd.log` - This is the Slurm compute daemon log. Use it to troubleshoot initialization and compute failure issues.

Seeing `InsufficientInstanceCapacity` error in `slurm_resume.log` when I fail to run a job, or in `clustermgtd.log` when I fail to create a cluster

If the cluster uses a Slurm scheduler, you are experiencing an insufficient capacity issue. If there aren't enough instances available when an instance launch request is made, an `InsufficientInstanceCapacity` error is returned.

For static instance capacity, you can find the error in the `clustermgtd` log at `/var/log/parallelcluster/clustermgtd`.

For dynamic instance capacity, you can find the error in the `ResumeProgram` log at `/var/log/parallelcluster/slurm_resume.log`.

The message looks similar to the following example:

```
An error occurred (InsufficientInstanceCapacity) when calling the RunInstances/
CreateFleet operation...
```

Based on your use case, consider using one of the following methods to avoid getting these types of error messages:

- Disable the placement group if it's enabled. For more information, see Placement groups and instance launch issues.
- Reserve capacity for the instances and launch them with ODCR (On-Demand Capacity Reservations). For more information, see Launch instances with On-Demand Capacity Reservations (ODCR).
- Configure multiple compute resources with different instance types. If your workload doesn't require a specific instance type, you can leverage fast insufficient capacity fail over with multiple compute resources. For more information, see Slurm cluster fast insufficient capacity fail-over.
- Configure multiple instance types in the same compute resource, and leverage the multiple instance type allocation. For more information about configuring multiple instances, see Multiple instance type allocation with Slurm and `Scheduling` / `SlurmQueues` / `ComputeResources` / `Instances`.
- Move the queue to a different Availability Zone by changing the subnet ID in the cluster configuration `Scheduling` / `SlurmQueues` / `Networking` / `SubnetIds`.
- If your workload isn't tightly coupled, span the queue across different Availability Zones. For more information about configuring multiple subnets, see `Scheduling` / `SlurmQueues` / `Networking` / `SubnetIds`.

Troubleshooting node initialization issues

This section covers how you can troubleshoot node initialization issues. This includes issues where the node fails to launch, power up, or join a cluster.

Topics

- Head node
- Compute nodes

Head node

Applicable logs:

- `/var/log/cfn-init.log`
- `/var/log/chef-client.log`
- `/var/log/parallelcluster/clustermgtd`
- `/var/log/parallelcluster/slurm_resume.log`
- `/var/log/slurmctld.log`

Check the `/var/log/cfn-init.log` and `/var/log/chef-client.log` logs or corresponding log streams. These logs contain all the actions that were run when the head node was set up. Most errors that occur during setup should have error messages located in the `/var/log/chef-client.log` log. If `OnNodeStart` or `OnNodeConfigured` scripts are specified in the configuration of the cluster, double check that the script runs successfully through log messages.

When a cluster is created, the head node must wait for the compute nodes to join the cluster before it can join the cluster. Because of this, if the compute nodes fail to join the cluster, then the head node also fails. You can follow one of these sets of procedures, depending on the type of compute notes you use, to troubleshoot this type of issue:

Compute nodes

- Applicable logs:
  - `/var/log/cloud-init-output.log`
  - `/var/log/slurmd.log`
- If a compute node is launched, first check `/var/log/cloud-init-output.log`, which should contain the setup logs similar to the `/var/log/chef-client.log` log on the head node. Most errors that occur during setup should have error messages located at the `/var/log/cloud-init-output.log` log. If pre-install or post-install scripts are specified in cluster configuration, check that they ran successfully.
- If you're using a custom AMI with modification to the Slurm configuration, then there might be a Slurm-related error that prevents the compute node from joining the cluster. For scheduler-related errors, check the `/var/log/slurmd.log` log.

Dynamic compute nodes:

- Search the `ResumeProgram` log (`/var/log/parallelcluster/slurm_resume.log`) for your compute node name to see if `ResumeProgram` was ever called with the node. (If `ResumeProgram` wasn't ever called, you can check the `slurmctld` log (`/var/log/slurmctld.log`) to determine if Slurm ever tried to call `ResumeProgram` with the node).
- Note that incorrect permissions for `ResumeProgram` might cause `ResumeProgram` to fail silently. If you're using a custom AMI with modification to `ResumeProgram` setup, check that the `ResumeProgram` is owned by the `slurm` user and has the `744` (`rwxr--r--`) permission.
- If `ResumeProgram` is called, check to see if an instance is launched for the node. If no instance was launched, you can see an error message that describes the launch failure.
- If the instance is launched, then there might have been a problem during the setup process. You should see the corresponding private IP address and instance ID from the `ResumeProgram` log. Moreover, you can look at corresponding setup logs for the specific instance. For more information about troubleshooting a setup error with a compute node, see the next section.

Static compute nodes:

- Check the `clustermgtd` (`/var/log/parallelcluster/clustermgtd`) log to see if instances were launched for the node. If they weren't launched, there should be clear error message detailing the launch failure.
- If instance is launched, there's some issue during setup process. You should see the corresponding private IP address and instance ID from the `ResumeProgram` log. Moreover, you can look at the corresponding setup logs for the specific instance.

Compute nodes backed by Spot Instances:

- If it's the first time you use Spot Instances and the job remains in a PD (pending state), double check the `/var/log/parallelcluster/slurm_resume.log` file. You'll probably find an error like the following:

```
2022-05-20 13:06:24,796 - [slurm_plugin.common:add_instances_for_nodes] - ERROR -
 Encountered exception when launching instances for nodes (x1) ['spot-dy-t2micro-2']:
 An error occurred (AuthFailure.ServiceLinkedRoleCreationNotPermitted) when calling
 the RunInstances operation: The provided credentials do not have permission to
 create the service-linked role for Amazon EC2 Spot Instances.
```

When using Spot Instances, an `AWSServiceRoleForEC2Spot` service-linked role must exist in your account. To create this role in your account using the AWS CLI, run the following command:

```
$ aws iam create-service-linked-role --aws-service-name spot.amazonaws.com
```

For more information, see Working with Spot Instances in the AWS ParallelCluster User Guide and Service-linked role for Spot Instance requests in the *Amazon EC2 User Guide*.

Troubleshooting unexpected node replacements and terminations

This section continues to explore how you can troubleshoot node related issues, specifically when a node is replaced or terminated unexpectedly.

- **Applicable logs:**
  - `/var/log/parallelcluster/clustermgtd` (head node)
  - `/var/log/slurmctld.log` (head node)
  - `/var/log/parallelcluster/computemgtd` (compute node)

Nodes replaced or terminated unexpectedly

- Check in the `clustermgtd` log (`/var/log/parallelcluster/clustermgtd`) to see if `clustermgtd` replaced or terminated a node. Note that `clustermgtd` handles all normal node maintenance action.
- If `clustermgtd` replaced or terminated the node, there should be a message detailing why this action was taken on the node. If the reason is scheduler related (for example, because the node is in `DOWN`), check in `slurmctld` log for more information. If the reason is Amazon EC2 related, there should be informative message detailing the Amazon EC2 related issue that required the replacement.
- If `clustermgtd` didn't terminate the node, first check if this was an expected termination by Amazon EC2 , more specifically a spot termination. `computemgtd`, running on a compute node, can also terminate a node if `clustermgtd` is determined as unhealthy. Check `computemgtd` log (`/var/log/parallelcluster/computemgtd`) to see if `computemgtd` terminated the node.

Nodes failed

- Check in `slurmctld` log (`/var/log/slurmctld.log`) to see why a job or a node failed. Note that jobs are automatically re-queued if a node failed.
- If `slurm_resume` reports that node is launched and `clustermgtd`reports after several minutes that there's no corresponding instance in Amazon EC2 for that node, the node might fail during setup. To retrieve the log from a compute (`/var/log/cloud-init-output.log`), do the following steps:
  - Submit a job to let Slurm spin up a new node.
  - Wait for the compute node to start.
  - Modify the instance initiated shutdown behavior so that a failing compute node will be stopped rather than terminated.

    ```
    $ aws ec2 modify-instance-attribute \
        --instance-id i-1234567890abcdef0 \
        --instance-initiated-shutdown-behavior "{\"Value\": \"stop\"}"
    ```

  - Enable termination protection.

    ```
    $ aws ec2 modify-instance-attribute \
        --instance-id i-1234567890abcdef0 \
        --disable-api-termination
    ```

  - Tag the node to be easily identifiable.

    ```
    $ aws ec2 create-tags \
        --resources i-1234567890abcdef0 \
        --tags Key=Name,Value=QUARANTINED-Compute
    ```

  - Detach the node from the cluster by changing the `parallelcluster:cluster-name` tag.

    ```
    $ aws ec2 create-tags \
        --resources i-1234567890abcdef0 \
        --tags Key=parallelcluster:clustername,Value=QUARANTINED-ClusterName
    ```

  - Retrieve the console output from the node with this command.

    ```
    $ aws ec2 get-console-output --instance-id i-1234567890abcdef0 --output text
    ```

Replacing, terminating, or powering down problematic instances and nodes

- **Applicable logs:**
  - `/var/log/parallelcluster/clustermgtd` (head node)
  - `/var/log/parallelcluster/slurm_suspend.log` (head node)
- In most cases, `clustermgtd` handles all expected instance termination action. Check in the `clustermgtd` log to see why it failed to replace or terminate a node.
- For dynamic nodes failing `SlurmSettings` Properties, check in the `SuspendProgram` log to see if `SuspendProgram` was called by `slurmctld` with the specific node as argument. Note that `SuspendProgram` doesn't actually perform any action. Rather, it only logs when it's called. All instance termination and `NodeAddr` reset is done by `clustermgtd`. Slurm puts nodes back into a `POWER_SAVING` state after `SuspendTimeout` automatically.
- If compute nodes are failing continuously due to bootstrap failures, verify if they are being launched with Slurm cluster protected mode enabled. If protected mode isn't enabled, modify the protected mode settings to enable protected mode. Troubleshoot and fix the bootstrap script.

Queue (partition) `Inactive` status

If you run `sinfo` and the output shows queues with `AVAIL` status of `inact`, your cluster might have Slurm cluster protected mode enabled and the queue has been set to the `INACTIVE` state for a pre-defined period of time.

Troubleshooting other known node and job issues

Another type of known issue is that AWS ParallelCluster might fail to allocate jobs or make scaling decisions. With this type of issue, AWS ParallelCluster only launches, terminates, or maintains resources according to Slurm instructions. For these issues, check the `slurmctld` log to troubleshoot them.

Placement groups and instance launch issues

To get the lowest inter-node latency, use a *placement group*. A placement group ensures that your instances are on the same networking backbone. If there aren't enough instances available when a

request is made, an `InsufficientInstanceCapacity` error is returned. To reduce the possibility of receiving this error when using cluster placement groups, set the `SlurmQueues` / `Networking` / `PlacementGroup` / `Enabled` parameter to `false`.

For additional control over capacity access, consider launching instances with ODCR (On-Demand Capacity Reservations).

For more information, see Troubleshooting instance launch issues and Placement groups roles and limitations in the *Amazon EC2 User Guide for Linux Instances*.

Replacing directories

Some directories can't be replaced. If you're having issues replacing the directory, that might be the case. The following directories are shared between the nodes and can't be replaced.

- `/opt/intel` - This includes Intel MPI, Intel Parallel Studio, and related files.
- `/opt/slurm` - This includes Slurm Workload Manager and related files. (Conditional, only if `Scheduler: slurm`.)

Troubleshooting issues in Amazon DCV

Topics

- Logs for Amazon DCV
- Ubuntu Amazon DCV issues

Logs for Amazon DCV

The logs for Amazon DCV are written to files in the `/var/log/dcv/` directory. Reviewing these logs can help to troubleshoot issues.

The instance type should have at least 1.7 gibibytes (GiB) of RAM to run Amazon DCV. Nano and micro instance types don't have enough memory to run Amazon DCV.

AWS ParallelCluster creates Amazon DCV log streams in log groups. You can view these logs in the CloudWatch console **Custom Dashboards** or **Log groups**. For more information, see Integration with Amazon CloudWatch Logs and Amazon CloudWatch dashboard.

Ubuntu Amazon DCV issues

When running Gnome Terminal over a Amazon DCV session on Ubuntu, you might not automatically have access to the user environment that AWS ParallelCluster makes available through the login shell. The user environment provides environment modules such as openmpi or intelmpi, and other user settings.

Gnome Terminal's default settings prevent the shell from starting as a login shell. This means that shell profiles aren't automatically sourced and the AWS ParallelCluster user environment isn't loaded.

To properly source the shell profile and access the AWS ParallelCluster user environment, do one of the following:

- **Change the default terminal settings:**
  1. Choose the **Edit** menu in the Gnome terminal.
  2. Select **Preferences**, then **Profiles**.
  3. Choose **Command** and select **Run Command as login shell**.
  4. Open a new terminal.
- **Use the command line to source the available profiles:**

```
$ source /etc/profile && source $HOME/.bashrc
```

Troubleshooting issues in clusters with AWS Batch integration

This section provides possible troubleshooting tips for clusters with AWS Batch scheduler integration, specifically with head node issues, compute issues, job failures, and timeout errors.

Topics

Head node issues

You can troubleshoot head node setup issues in the same way as a Slurm cluster (except for Slurm specific logs). For more information about these issues, see Head node.

Compute issues

AWS Batch manages the scaling and compute aspects of your services. If you encounter compute related issues, see the AWS Batch troubleshooting documentation for help.

Job failures

If a job fails, you can run the `awsbout` command to retrieve the job output. You can also run the `awsbstat` command to obtain a link to the job logs stored by Amazon CloudWatch.

Connect timeout on endpoint URL error

If multi-node parallel jobs fail with error: `Connect timeout on endpoint URL`:

- In the `awsbout` output log, check that the job is multi-node parallel from the output: `Detected 3/3 compute nodes. Waiting for all compute nodes to start.`
- Verify whether the compute nodes subnet is public.

Multi-node parallel jobs don't support the use of public subnets when using AWS Batch in AWS ParallelCluster. Use a private subnet for your compute nodes and jobs. For more information, see Compute environment considerations in the *AWS Batch User Guide*. To configure a private subnet for your compute nodes, see AWS ParallelCluster with AWS Batch scheduler.

Troubleshooting multi-user integration with Active Directory

This section is relevant to clusters integrated with an Active Directory.

If the Active Directory integration feature isn't working as expected the SSSD logs can provide useful diagnostic information. These logs are located in `/var/log/sssd` on cluster nodes. By default, they're also stored in a cluster's Amazon CloudWatch log group.

Topics

- How to move from LDAPS to LDAP
- How to disable LDAPS server certificate verification
- How to log in with an SSH key rather than password
- How to reset a user password and expired passwords
- How to verify the joined domain
- How to troubleshoot issues with certificates
- How to verify that the integration with Active Directory is working
- How to troubleshoot logging in to compute nodes
- Known issues with SimCenter StarCCM+ jobs in a multi-user environment
- Known issues with username resolution
- How to resolve home directory create issues

Active Directory specific troubleshooting

This section is relevant to troubleshooting specific to an Active Directory type.

Simple AD

- The `DomainReadOnlyUser` value must match the Simple AD directory base search for users:

  `cn=ReadOnlyUser,cn=Users,dc=corp,dc=example,dc=com`

  Note `cn` for `Users`.
- Default admin user is `Administrator`.
- `Ldapsearch` requires NetBIOS name before the username.

  `Ldapsearch` syntax must be as follows:

```
$ ldapsearch -x -D "corp\\Administrator" -w "Password" -H ldap://192.0.2.103 \
   -b "cn=Users,dc=corp,dc=example,dc=com"
```

AWS Managed Microsoft AD

- The `DomainReadOnlyUser` value must match the AWS Managed Microsoft AD directory base search for users:

`cn=ReadOnlyUser,ou=Users,ou=CORP,dc=corp,dc=example,dc=com`

- Default admin user is `Admin`.
- `Ldapsearch` syntax must be as follows:

```
$ ldapsearch -x -D "Admin" -w "Password" -H ldap://192.0.2.103 \
   -b "ou=Users,ou=CORP,dc=corp,dc=example,dc=com"
```

Enable debug mode

Debug logs from SSSD can be useful to troubleshoot issues. To enable debug mode, you must update the cluster with the following changes made to the cluster configuration:

```
DirectoryService:
  AdditionalSssdConfigs:
    debug_level: "0x1ff"
```

How to move from LDAPS to LDAP

Moving from LDAPS (LDAP with TLS/SSL) to LDAP is discouraged because LDAP alone doesn't provide any encryption. Nevertheless, it can be useful for testing purposes and troubleshooting.

You can restore the cluster to its previous configuration by updating the cluster with the previous configuration definition.

To move from LDAPS to LDAP, you must update the cluster with the following changes in the cluster configuration:

```
DirectoryService:
  LdapTlsReqCert: never
  AdditionalSssdConfigs:
    ldap_auth_disable_tls_never_use_in_production: True
```

How to disable LDAPS server certificate verification

It can be useful to temporarily disable LDAPS server certificate verification on the head node, for testing or troubleshooting purposes.

You can restore the cluster to its previous configuration by updating the cluster with the previous configuration definition.

To disable the LDAPS server certificate verification, you must update the cluster with the following changes in the cluster configuration:

```
DirectoryService:
  LdapTlsReqCert: never
```

How to log in with an SSH key rather than password

The SSH key is created in `/home/$user/.ssh/id_rsa` after the first time that you log in with a password. To log in with the SSH key, you must log in with your password, copy the SSH key locally, and then use it to SSH password-less as usual:

```
$ ssh -i $LOCAL_PATH_TO_SSH_KEY $username@$head_node_ip
```

How to reset a user password and expired passwords

If a user loses access to a cluster, their AWS Managed Microsoft AD password might have expired.

To reset the password, run the following command with a user and role having write permission on the directory:

```
$ aws ds reset-user-password \
  --directory-id "d-abcdef01234567890" \
  --user-name "USER_NAME" \
  --new-password "NEW_PASSWORD" \
  --region "region-id"
```

If you reset the password for the `DirectoryService` / `DomainReadOnlyUser`:

1. Be sure to update the `DirectoryService` / `PasswordSecretArn` secret with the new password.
2. Update the cluster for the new secret value:
   a. Stop the compute fleet with the `pcluster update-compute-fleet` command.
   b. Run the following command from within the cluster head node.

```
$ sudo /opt/parallelcluster/scripts/directory_service/
update_directory_service_password.sh
```

After the password reset and cluster update, the user's cluster access should be restored.

For more information, see Reset a user password in the *AWS Directory Service Administration Guide*.

How to verify the joined domain

The following command must run from an instance that's joined to the domain, not the head node.

```
$ realm list corp.example.com \
  type: kerberos \
  realm-name: CORP.EXAMPLE.COM \
  domain-name: corp.example.com \
  configured: kerberos-member \
  server-software: active-directory \
  client-software: sssd \
  required-package: oddjob \
  required-package: oddjob-mkhomedir \
  required-package: sssd \
  required-package: adcli \
  required-package: samba-common-tools \
  login-formats: %U \
  login-policy: allow-realm-logins
```

How to troubleshoot issues with certificates

When LDAPS communication isn't working, it can be due to errors in the TLS communication, which in turn can be due to issues with certificates.

Notes about certificates:

- The certificate specified in cluster config `LdapTlsCaCert` must be a bundle of PEM certificates containing the certificates for the whole certificate of authority (CA) chain that issued certificates for the domain controllers.
- A bundle of PEM certificates is a file made of the concatenation of PEM certificates.
- A certificate in PEM format (typically used in Linux) is equivalent to a certificate in base64 DER format (typically exported by Windows).

- If the certificate for domain controllers is issued by a subordinate CA, then the certificate bundle must contain the certificate of both the subordinate and root CA.

Troubleshooting verification steps:

The following verification steps assume that the commands are run from within the cluster head node and that the domain controller is reachable at *SERVER*:*PORT*.

To troubleshoot an issue that's related to certificates, follow these verification steps:

Verification steps:

1. **Check the connection to the Active Directory domain controllers:**

   Verify that you can connect to a domain controller. If this step succeeds, then the SSL connection to the domain controller succeeds and the certificate is verified. Your issue isn't related to certificates.

   If this step fails, go ahead with next verification.

   ```
   $ openssl s_client -connect SERVER:PORT -CAfile PATH_TO_CA_BUNDLE_CERTIFICATE
   ```

2. **Check the certificate verification:**

   Verify that the local CA certificate bundle can validate the certificate provided by the domain controller. If this step succeeds, then your issue isn't related to certificates, but to other networking issues.

   If this step fails, go ahead with next verification.

   ```
   $ openssl verify -verbose -
   CAfile PATH_TO_CA_BUNDLE_CERTIFICATE PATH_TO_A_SERVER_CERTIFICATE
   ```

3. **Check the certificate provided by the Active Directory domain controllers:**

   Verify that the content of the certificate provided by the domain controllers is as expected. If this step succeeds, you probably have issues with the CA certificate used to verify controllers, go to the next troubleshooting step.

   If this step fails, you must correct the certificate issued for the domain controllers and re-execute the troubleshooting steps.

```
$ openssl s_client -connect SERVER:PORT -showcerts
```

4. **Check the content of a certificate:**

   Verify that the content of the certificate that's provided by the domain controllers is as expected. If this step succeeds, you probably have issues with the CA certificate used to verify controller's, go to the next troubleshooting step.

   If this step fails, you must correct the certificate issued for the domain controllers and rerun the troubleshooting steps.

   ```
   $ openssl s_client -connect SERVER:PORT -showcerts
   ```

5. **Check the content of the local CA certificate bundle:**

   Verify that the content of the local CA certificate bundle used to validate domain controllers certificate is as expected. If this step succeeds, you probably have issues with the certificate that are provided by the domain controllers.

   If this step fails, you must correct CA certificate bundle issued for the domain controllers and rerun the troubleshooting steps.

   ```
   $ openssl x509 -in PATH_TO_A_CERTIFICATE -text
   ```

How to verify that the integration with Active Directory is working

If the following two checks succeed, the integration with the Active Directory is working.

Checks:

1. **You can discover users defined in the directory:**

   From within the cluster head node, as an `ec2-user`:

   ```
   $ getent passwd $ANY_AD_USER
   ```

2. **You can SSH into the head node providing the user password:**

```
$ ssh $ANY_AD_USER@$HEAD_NODE_IP
```

If check one fails, we expect check two to fail also.

Additional troubleshooting checks:

- Verify that the user exists in the directory.
- Enable debug logging.
- Consider temporarily disabling encryption by moving from LDAPS to LDAP to rule out LDAPS issues.

How to troubleshoot logging in to compute nodes

This section is relevant to logging in to compute nodes in clusters integrated with Active Directory.

With AWS ParallelCluster, password logins to cluster compute nodes are disabled by design.

All users must use their own SSH key to log in to compute nodes.

Users can retrieve their SSH key in the head node after first authentication (for example login), if `GenerateSshKeysForUsers` is enabled in the cluster configuration.

When users authenticate on the head node for the first time, they can retrieve SSH keys that are automatically generated for them as directory users. Home directories for the user are also created. This can also happen the first time a sudo-user switches to a user in the head node.

If a user hasn't logged into the head node, SSH keys aren't generated and the user won't be able to log in to compute nodes.

Known issues with SimCenter StarCCM+ jobs in a multi-user environment

This section is relevant to jobs launched in a multi-user environment by Simcenter StarCCM+ computational fluid dynamics software from Siemens.

If you run StarCCM+ v16 jobs configured to use the embedded IntelMPI, by default the MPI processes are bootstrapped using SSH.

Due to a known Slurm bug that causes username resolution to be wrong, jobs might fail with an error like `error setting up the bootstrap proxies`. This bug only impacts AWS ParallelCluster versions 3.1.1 and 3.1.2.

To prevent this from occurring, force IntelMPI to use Slurm as MPI bootstrap method. Export the environment variable `I_MPI_HYDRA_BOOTSTRAP=slurm` into the job script that launches StarCCM+, as described in the IntelMPI official documentation.

Known issues with username resolution

This section is relevant to retrieving usernames within jobs.

Due to a known bug in Slurm, the username retrieved within a job process might be nobody if you run a job without `srun`. This bug only impacts AWS ParallelCluster versions 3.1.1 and 3.1.2.

For example, if you run the command `sbatch --wrap 'srun id'` as a directory user, the correct username is returned. However, if you run the `sbatch --wrap 'id'` as a directory user, nobody might be returned as the username.

You can use the following workarounds.

1. Launch your job with `'srun'` instead of `'sbatch'`, if possible.
2. Enable SSSD enumeration by setting the AdditionalSssdConfigs in cluster configuration as follows.

```
AdditionalSssdConfigs:
  enumerate: true
```

How to resolve home directory create issues

This section is relevant to home directory creation issues.

If you see errors like the one shown in the following example, a home directory wasn't created for you when you first logged in to the head node. Or, a home directory wasn't created for you when you first switched from a sudoer to an Active Directory user in the head node.

```
$ ssh AD_USER@$HEAD_NODE_IP
/opt/parallelcluster/scripts/generate_ssh_key.sh failed: exit code 1

       __|  __|_  )
```

```
       _|  (     /   Amazon Linux 2 AMI
      ___|\___|___|

https://aws.amazon.com/amazon-linux-2/
Could not chdir to home directory /home/PclusterUser85: No such file or directory
```

The home directory create failure can be caused by the `oddjob` and `oddjob-mkhomedir` packages installed in the cluster head node.

Without a home directory and SSH key, the user can't submit jobs or SSH into the cluster nodes.

If you need the `oddjob` packages in your system, verify that the `oddjobd` service is running and refresh the PAM config files to make sure that the home directory is created. To do this, run the commands in the head node as shown in the following example.

```
sudo systemctl start oddjobd
sudo authconfig --enablemkhomedir --updateall
```

If you don't need the `oddjob` packages in your system, uninstall them and refresh the PAM config files to make sure that the home directory is created. To do this, run the commands in the head node as shown in the following example.

```
sudo yum remove -y oddjob oddjob-mkhomedir
sudo authconfig --enablemkhomedir --updateall
```

Troubleshooting custom AMI issues

This section provides possible troubleshooting tips for custom AMI issues.

When you use a custom AMI, you can see the following warnings:

```
"validationMessages": [
  {
    "level": "WARNING",
    "type": "CustomAmiTagValidator",
    "message": "The custom AMI may not have been created by pcluster. You can ignore
 this warning if the AMI is shared or copied from another pcluster AMI. If the
 AMI is indeed not created by pcluster, cluster creation will fail. If the cluster
 creation fails, please go to https://docs.aws.amazon.com/parallelcluster/latest/ug/
troubleshooting.html#troubleshooting-stack-creation-failures for troubleshooting."
  },
```

```
  {
   "level": "WARNING",
   "type": "AmiOsCompatibleValidator",
   "message": "Could not check node AMI ami-0000012345 OS and cluster OS alinux2
 compatibility, please make sure they are compatible before cluster creation and update
 operations."
  }
]
```

If you're sure that the correct AMI is being used, you can ignore these warnings.

If you don't want to see these warnings in the future, tag the custom AMI with the following tags, where *my-os* is one of `alinux2`, `alinux2023`, ubuntu2404, ubuntu2204, ubuntu2004, `rhel8`, or `rhel9` and *"3.13.2"* is the `pcluster` version in use:

```
$ aws ec2 create-tags \
  --resources ami-yourcustomAmi \
  --tags Key="parallelcluster:version",Value="3.13.2"
 Key="parallelcluster:os",Value="my-os"
```

Troubleshooting a cluster update timeout when `cfn-hup` isn't running

The `cfn-hup` helper is a daemon that detects changes in resource metadata and runs user-specified actions when a change is detected. This is how you make configuration updates on your running Amazon EC2 instances through the `UpdateStack` API action.

Currently the `cfn-hup` daemon is launched by the `supervisord`. But after launch, the `cfn-hup` process is detached from `supervisord` control. If the `cfn-hup` demon is killed by an external actor, it's not restarted automatically. If `cfn-hup` isn't running, during a cluster update, the CloudFormation stack starts the update process as expected but the update procedure isn't activated on the head node and the stack eventually goes into timeout. From the cluster logs `/var/log/chef-client`, you can see that the update recipe is never invoked.

Check and restart `cfn-hup` in case of failures

1. On the head node, check if `cfn-hup` is running:

```
$ ps aux | grep cfn-hup
```

2. Check `cfn-hup` log `/var/log/cfn-hup.log` and `/var/log/supervisord.log` on the head node.
3. If `cfn-hup` isn't running, try restarting it by running:

```
$ sudo /opt/parallelcluster/pyenv/versions/cookbook_virtualenv/bin/supervisorctl
 start cfn-hup
```

Network troubleshooting

This section provides a troubleshooting tip for when you come across network issues, specifically when dealing with a cluster in a single public subnet issue.

Cluster in a single public subnet issues

Check the `cloud-init-output.log` from one of the compute nodes. If you find something like the following that indicates the node is stuck in Slurm initialization, it is most likely due to a missing DynamoDB VPC endpoint. Add the DynamoDB endpoint. For more information see AWS ParallelCluster in a single subnet with no internet access.

```
ruby_block[retrieve compute node info] action run[2022-03-11T17:47:11+00:00] INFO:
 Processing ruby_block[retrieve compute node info] action run (aws-parallelcluster-
slurm::init line 31)
```

Cluster update failed on onNodeUpdated custom action

When a `HeadNode` / `CustomActions` / `OnNodeUpdated` script fails, the update fails and the script is not run at rollback time. It's your responsibility to manually perform the cleanups needed after the rollback is completed. For example, if the `OnNodeUpdated` script changes the status of a field in a configuration file (for example, from `true` to `false`) and then fails, you need to manually restore that field value to the pre-update state (for example, `false` to `true`). For more information, see Custom bootstrap actions.

Seeing errors with custom Slurm configuration

Starting in AWS ParallelCluster version 3.6.0, you can no longer target single `prolog` or `epilog` scripts by including them in a custom Slurm configuration. In AWS ParallelCluster version 3.6.0 and

later versions, you must locate custom `prolog` and `epilog` scripts in the respective `Prolog` and `Epilog` folders. These folders are configured by default to point to:

- `Prolog` points to `/opt/slurm/etc/scripts/prolog.d/`.
- `Epilog` points to `/opt/slurm/etc/scripts/epilog.d/`.

We recommend that you keep the `90_plcuster_health_check_manager` prolog script and the `90_pcluster_noop` epilog script in place.

Slurm runs the scripts in reverse alphabetical order. Both the `Prolog` and `Epilog` folder must contain at least one file. For more information, see Slurm`prolog and epilog` and Slurm configuration customization.

Cluster alarms

Cluster health monitoring is essential for ensuring optimal performance. AWS ParallelCluster enables you to monitor multiple CloudWatch based alarms for the cluster's head node.

This section provides detail for each type of Head node cluster alarms including its naming conventions, specific conditions that trigger alarms, and suggested troubleshooting steps.

The naming convention for cluster alarms is `CLUSTER_NAME-COMPONENT-METRIC`, e.g. `mycluster-HeadNode-Cpu`.

- `CLUSTER_NAME-HeadNode`: signals the overall status of the head node. It is red if at least one of the alarms below is.
- `CLUSTER_NAME-HeadNode-Health`: red if there is at least one Amazon EC2 Health Check failure. In case of alarm, we suggest to have a look at Troubleshoot instances with failed status checks.
- `CLUSTER_NAME-HeadNode-Cpu`: red if CPU utilization is greater than 90%. In case of alarm, check the processes that are consuming the CPU the most with `ps -aux --sort=-%cpu | head -n 10`.
- `CLUSTER_NAME-HeadNode-Mem`: red if memory utilization is greater than 90%. In case of alarm, check the processes that are consuming the memory the most with `ps -aux --sort=-%mem | head -n 10`.

- `CLUSTER_NAME-HeadNode-Disk`: red if the occupied disk space is greater than 90% on path /. In case of alarm, check the folders consuming the majority of the space with `du -h --max-depth=2 / 2> /dev/null | sort -hr`.

Resolving OS configuration changes that cause errors or failures

When making OS configuration changes to AWS ParallelCluster nodes, various issues can arise that may cause cluster creation, update, or operation failures. This section provides guidance on identifying and resolving common OS configuration-related issues.

Common OS configuration issues

Locale configuration issues

One of the most common OS configuration issues is related to locale settings. If you see errors like:

```
cannot change locale (en_US.utf-8) because it has an invalid name
```

This typically occurs when:

- A `yum` installation process was unsuccessful and left locale settings in an inconsistent state
- A user terminated an installation process prematurely
- Locale packages are missing or corrupted

How to diagnose

1. Check if you can switch to the pcluster-admin user:

   ```
   $ su - pcluster-admin
   ```

   If you see an error like `cannot change locale...no such file or directory`, this confirms the issue.
2. Check available locales:

   ```
   $ localedef --list
   ```

If this returns an empty list or doesn't contain the default locale, your locale configuration is broken.

3. Check the last `yum` command:

```
$ yum history
$ yum history info #ID
```

If the last ID doesn't have `Return-Code: Success`, the post-install scripts might not have run successfully.

How to resolve

Rebuild the locale by reinstalling the language packs:

```
$ sudo yum reinstall glibc-all-langpacks
```

After the rebuild, verify the issue is fixed by running:

```
$ su - pcluster-admin
```

If no error or warning appears, the issue has been resolved.

OS package conflicts

When installing custom packages or modifying system packages, conflicts can arise that prevent proper cluster operation.

How to diagnose

1. Check the chef-client log for package-related errors:

```
$ less /var/log/chef-client.log
```

2. Look for package dependency conflicts in the cfn-init log:

```
$ less /var/log/cfn-init.log
```

How to resolve

1. If a specific package is causing issues, try reinstalling it:

```
$ sudo yum reinstall package-name
```

2. For dependency conflicts, you may need to remove conflicting packages:

```
$ sudo yum remove conflicting-package
```

3. If the issue persists, consider creating a custom AMI with your required packages pre-installed using the `pcluster build-image` command. For more information, see AWS ParallelCluster AMI customization.

System configuration file modifications

Modifying critical system configuration files can cause cluster failures, especially if these files are managed by AWS ParallelCluster.

How to diagnose

1. Check for errors in the chef-client log that mention specific configuration files:

```
$ grep -i "config" /var/log/chef-client.log
```

2. Look for permission or syntax errors in configuration files:

```
$ less /var/log/cfn-init.log
```

How to resolve

1. Restore modified configuration files to their original state:

```
$ sudo cp /etc/file.conf.bak /etc/file.conf
```

2. If you need to make persistent changes to system configuration files, use custom bootstrap actions instead of directly modifying files:

```
HeadNode:
```

```
  CustomActions:
    OnNodeConfigured:
      Script: s3://bucket-name/config-script.sh
```

For more information, see Custom bootstrap actions.

3. For configuration changes that must be made directly to system files, consider creating a custom AMI. For more information, see AWS ParallelCluster AMI customization.

Kernel updates and compatibility issues

Kernel updates can cause compatibility issues with certain AWS services, particularly with Amazon FSx for Lustre.

How to diagnose

1. Check if kernel updates have been applied:

```
$ uname -r
```

2. Look for Amazon FSx mount failures in the logs:

```
$ grep -i "fsx" /var/log/chef-client.log
```

How to resolve

1. For Ubuntu 22.04, avoid updating to the latest kernel as there is no Amazon FSx client for that kernel. For more information, see Operating system considerations.
2. If you've already updated the kernel and are experiencing issues, consider downgrading to a compatible kernel version:

```
$ sudo apt install linux-image-previous-version
```

3. For persistent kernel customizations, create a custom AMI with the specific kernel version you need. For more information, see AWS ParallelCluster AMI customization.

Best practices for OS configuration changes

To minimize issues when making OS configuration changes:

1. **Use Custom Bootstrap Actions**: Instead of directly modifying system files, use `OnNodeStart` or `OnNodeConfigured` scripts to make changes in a controlled manner. For more information, see Custom bootstrap actions.
2. **Create Custom AMIs**: For significant OS modifications, create a custom AMI using `pcluster build-image` rather than making changes to running instances. For more information, see AWS ParallelCluster AMI customization.
3. **Test Changes First**: Before applying changes to a production cluster, test them on a small test cluster to ensure compatibility.
4. **Document Changes**: Keep track of all OS configuration changes made to facilitate troubleshooting.
5. **Backup Configuration Files**: Before modifying any system configuration file, create a backup:

```
$ sudo cp /etc/file.conf /etc/file.conf.bak
```

6. **Check Logs After Changes**: After making OS configuration changes, check the logs for any errors:

```
$ less /var/log/cfn-init.log
$ less /var/log/chef-client.log
```

By following these guidelines, you can minimize the risk of OS configuration changes causing cluster failures and more effectively troubleshoot any issues that do arise.

AWS ParallelCluster support policy

AWS ParallelCluster supports multiple releases at the same time. Every AWS ParallelCluster release has a scheduled End of Support Life (EOSL) date. After the EOSL date, no further support or maintenance is provided for that release.

AWS ParallelCluster uses a `major.minor.patch` version scheme. New features, performance improvements, security updates, and bug fixes are included in new minor version releases for the latest major version release. Minor versions are backward compatible within a major version. For critical issues, AWS provides fixes through patch releases, but only for the latest minor versions of releases that have not reached EOSL. If you want to use the updates from a new version release, you need to upgrade to the new minor or patch version.

AWS ParallelCluster versions	End of supported life (EOSL) date
3.0.*x*	3/31/2023
3.1.*x*	8/31/2023
3.2.*x*	1/31/2024
3.3.*x*	5/31/2024
3.4.*x*	6/28/2024
3.5.*x*	8/31/2024
3.6.*x*	11/30/2024
3.7.*x*	2/28/2025
3.8.*x*	6/30/2025
3.9.*x*	09/05/2025
3.10.*x*	12/27/2025
3.11.*x*	03/25/2026

AWS ParallelCluster versions	End of supported life (EOSL) date
3.12.*x*	06/30/2026
3.13.*x*	09/30/2026

Security in AWS ParallelCluster

Cloud security at AWS is the highest priority. As an AWS customer, you benefit from a data center and network architecture that is built to meet the requirements of the most security sensitive organizations.

Security is a shared responsibility between AWS and you. The shared responsibility model describes this as security *of* the cloud and security *in* the cloud:

- **Security of the cloud** – AWS is responsible for protecting the infrastructure that runs AWS services in the AWS Cloud. AWS also provides you with services that you can use securely. Third-party auditors regularly test and verify the effectiveness of our security as part of the AWS Compliance Programs. To learn about the compliance programs that apply to AWS ParallelCluster, see AWS Services in Scope by Compliance Program.
- **Security in the cloud** – Your responsibility is determined by the specific AWS service or services that you use. You are also responsible for several other related factors including the sensitivity of your data, your company's requirements, and applicable laws and regulations.

This documentation describes how you should apply the shared responsibility model when using AWS ParallelCluster. The following topics show you how to configure AWS ParallelCluster to meet your security and compliance objectives. You also learn how to use AWS ParallelCluster in a way that helps you to monitor and secure your AWS resources.

Topics

Security information for services used by AWS ParallelCluster

- Security in Amazon API Gateway
- Security in AWS Batch
- Security in AWS CloudFormation
- Security in Amazon CloudWatch
- Security in AWS CodeBuild
- Security in Amazon DynamoDB
- Security in Amazon ECR
- Security in Amazon ECS
- Security in Amazon EFS
- Security in FSx for Lustre
- Security in AWS Identity and Access Management (IAM)
- Security in EC2 Image Builder
- Security in AWS Lambda
- Security in Amazon Route 53
- Security in Amazon SNS
- Security in Amazon SQS (For AWS ParallelCluster version 2.x.)
- Security in Amazon S3
- Security in Amazon VPC

Data protection in AWS ParallelCluster

The AWS shared responsibility model applies to data protection in AWS ParallelCluster. As described in this model, AWS is responsible for protecting the global infrastructure that runs all of the AWS Cloud. You are responsible for maintaining control over your content that is hosted on this infrastructure. You are also responsible for the security configuration and management tasks for the AWS services that you use. For more information about data privacy, see the Data Privacy FAQ. For information about data protection in Europe, see the AWS Shared Responsibility Model and GDPR blog post on the *AWS Security Blog*.

For data protection purposes, we recommend that you protect AWS account credentials and set up individual users with AWS IAM Identity Center or AWS Identity and Access Management (IAM). That way, each user is given only the permissions necessary to fulfill their job duties. We also recommend that you secure your data in the following ways:

- Use multi-factor authentication (MFA) with each account.
- Use SSL/TLS to communicate with AWS resources. We require TLS 1.2 and recommend TLS 1.3.
- Set up API and user activity logging with AWS CloudTrail. For information about using CloudTrail trails to capture AWS activities, see Working with CloudTrail trails in the *AWS CloudTrail User Guide*.
- Use AWS encryption solutions, along with all default security controls within AWS services.
- Use advanced managed security services such as Amazon Macie, which assists in discovering and securing sensitive data that is stored in Amazon S3.
- If you require FIPS 140-3 validated cryptographic modules when accessing AWS through a command line interface or an API, use a FIPS endpoint. For more information about the available FIPS endpoints, see Federal Information Processing Standard (FIPS) 140-3.

We strongly recommend that you never put confidential or sensitive information, such as your customers' email addresses, into tags or free-form text fields such as a **Name** field. This includes when you work with AWS ParallelCluster or other AWS services using the console, API, AWS CLI, or AWS SDKs. Any data that you enter into tags or free-form text fields used for names may be used for billing or diagnostic logs. If you provide a URL to an external server, we strongly recommend that you do not include credentials information in the URL to validate your request to that server.

Data encryption

A key feature of any secure service is that information is encrypted when it is not being actively used.

Encryption at rest

AWS ParallelCluster does not itself store any customer data other than the credentials it needs to interact with the AWS services on the user's behalf.

For data on the nodes in the cluster, data can be encrypted at rest.

For Amazon EBS volumes, encryption is configured using the `EbsSettings`/`Encrypted` and `EbsSettings`/`KmsKeyId` settings in the `EbsSettings` section. For more information, see Amazon EBS encryption in the Amazon EC2 User Guide.

For Amazon EFS volumes, encryption is configured using the `EfsSettings`/`Encrypted` and `EfsSettings`/`KmsKeyId` settings in the `EfsSettings` section. For more information, see How encryption at rest works in the *Amazon Elastic File System User Guide.*

For FSx for Lustre file systems, encryption of data at rest is automatically enabled when creating an Amazon FSx file system. For more information, see Encrypting data at rest in the *Amazon FSx for Lustre User Guide*.

For instance types with NVMe volumes, the data on NVMe instance store volumes is encrypted using an XTS-AES-256 cipher implemented on a hardware module on the instance. The encryption keys are generated using the hardware module and are unique to each NVMe instance storage device. All encryption keys are destroyed when the instance is stopped or terminated and cannot be recovered. You cannot disable this encryption and you cannot provide your own encryption key. For more information, see Encryption at rest in the *Amazon EC2 User Guide*.

If you use AWS ParallelCluster to invoke an AWS service that transmits customer data to your local computer for storage, then refer to the Security and Compliance chapter in that service's User Guide for information on how that data is stored, protected, and encrypted.

Encryption in transit

By default, all data transmitted from the client computer running AWS ParallelCluster and AWS service endpoints is encrypted by sending everything through a HTTPS/TLS connection. Traffic between the nodes in the cluster can be automatically encrypted, depending on the instance types selected. For more information, see Encryption in transit in the *Amazon EC2 User Guide*.

See also

- Data protection in Amazon EC2
- Data protection in EC2 Image Builder
- Data protection in AWS CloudFormation
- Data protection in Amazon EFS
- Data protection in Amazon S3
- Data protection in FSx for Lustre

Identity and Access Management for AWS ParallelCluster

AWS ParallelCluster uses roles to access your AWS resources and their services. The instance and user policies that AWS ParallelCluster uses to grant permissions are documented at AWS Identity and Access Management permissions in AWS ParallelCluster.

The only major difference is how you authenticate when using a standard user and long-term credentials. Although an user requires a password to access an AWS service's console, that same user requires an access key pair to perform the same operations using AWS ParallelCluster. All other short-term credentials are used in the same way they are used with the console.

The credentials used by AWS ParallelCluster are stored in plaintext files and are ***not*** encrypted.

- The `$HOME/.aws/credentials` file stores long-term credentials required to access your AWS resources. These include your access key ID and secret access key.
- Short-term credentials, such as those for roles that you assume, or that are for AWS IAM Identity Center services, are also stored in the `$HOME/.aws/cli/cache` and `$HOME/.aws/sso/cache` folders, respectively.

Mitigation of Risk

- We strongly recommend that you configure your file system permissions on the `$HOME/.aws` folder and its child folders and files to restrict access to only authorized users.
- Use roles with temporary credentials wherever possible to reduce the opportunity for damage if the credentials are compromised. Use long-term credentials only to request and refresh short-term role credentials.

Compliance validation for AWS ParallelCluster

Third-party auditors assess the security and compliance of AWS services as part of multiple AWS compliance programs. Using AWS ParallelCluster to access a service does not alter that service's compliance.

For a list of AWS services in scope of specific compliance programs, see AWS services in scope by compliance program. For general information, see AWS compliance programs.

You can download third-party audit reports using the AWS Artifact. For more information, see Downloading reports in AWS Artifact.

Your compliance responsibility when using AWS ParallelCluster is determined by the sensitivity of your data, your company's compliance objectives, and applicable laws and regulations. AWS provides the following resources to help with compliance:

- Security and compliance quick start guides – These deployment guides discuss architectural considerations and provide steps for deploying security- and compliance-focused baseline environments on AWS.
- Architecting for HIPAA security and Compliance on Amazon Web Services AWS Whitepaper – This whitepaper describes how companies can use AWS to create HIPAA-compliant applications.
- AWS compliance resources – This collection of workbooks and guides might apply to your industry and location.
- Evaluating resources with rules in the *AWS Config Developer Guide* – The AWS Config service assesses how well your resource configurations comply with internal practices, industry guidelines, and regulations.
- AWS Security Hub – This AWS service provides a comprehensive view of your security state within AWS that helps you check your compliance with security industry standards and best practices.

Enforcing a Minimum Version of TLS 1.2

To add increased security when communicating with AWS services, you should configure your AWS ParallelCluster to use TLS 1.2 or later. When you use AWS ParallelCluster, Python is used to set the TLS version.

To ensure AWS ParallelCluster uses no TLS version earlier than TLS 1.2, you might need to recompile OpenSSL to enforce this minimum and then recompile Python to use the newly built OpenSSL.

Determine Your Currently Supported Protocols

First, create a self-signed certificate to use for the test server and the Python SDK using OpenSSL.

```
$ openssl req -subj '/CN=localhost' -x509 -newkey rsa:4096 -nodes -keyout key.pem -out
 cert.pem -days 365
```

Then spin up a test server using OpenSSL.

```
$ openssl s_server -key key.pem -cert cert.pem -www
```

In a new terminal window, create a virtual environment and install the Python SDK.

```
$ python3 -m venv test-env
```

```
source test-env/bin/activate
pip install botocore
```

Create a new Python script named `check.py` that uses the SDK's underlying HTTP library.

```
$ import urllib3
URL = 'https://localhost:4433/'

http = urllib3.PoolManager(
ca_certs='cert.pem',
cert_reqs='CERT_REQUIRED',
)
r = http.request('GET', URL)
print(r.data.decode('utf-8'))
```

Run your new script.

```
$ python check.py
```

This displays details about the connection made. Search for "Protocol : " in the output. If the output is "TLSv1.2" or later, the SDK defaults to TLS v1.2 or later. If it's an earlier version, you need to recompile OpenSSL and recompile Python.

However, even if your installation of Python defaults to TLS v1.2 or later, it's still possible for Python to renegotiate to a version earlier than TLS v1.2 if the server doesn't support TLS v1.2 or later. To check that Python doesn't automatically renegotiate to earlier versions, restart the test server with the following.

```
$ openssl s_server -key key.pem -cert cert.pem -no_tls1_3 -no_tls1_2 -www
```

If you're using an earlier version of OpenSSL, you might not have the `-no_tls_3` flag available. If this is the case, remove the flag because the version of OpenSSL you're using doesn't support TLS v1.3. Then rerun the Python script.

```
$ python check.py
```

If your installation of Python correctly doesn't renegotiate for versions earlier than TLS 1.2, you should receive an SSL error.

```
$ urllib3.exceptions.MaxRetryError: HTTPSConnectionPool(host='localhost',
 port=4433): Max retries exceeded with url: / (Caused by SSLError(SSLError(1, '[SSL:
 UNSUPPORTED_PROTOCOL] unsupported protocol (_ssl.c:1108)')))
```

If you're able to make a connection, you need to recompile OpenSSL and Python to disable negotiation of protocols earlier than TLS v1.2.

Compile OpenSSL and Python

To ensure that AWS ParallelCluster doesn't negotiate for anything earlier than TLS 1.2, you need to recompile OpenSSL and Python. To do this, copy the following content to create a script and run it.

```
#!/usr/bin/env bash
set -e

OPENSSL_VERSION="1.1.1d"
OPENSSL_PREFIX="/opt/openssl-with-min-tls1_2"
PYTHON_VERSION="3.8.1"
PYTHON_PREFIX="/opt/python-with-min-tls1_2"


curl -O "https://www.openssl.org/source/openssl-$OPENSSL_VERSION.tar.gz"
tar -xzf "openssl-$OPENSSL_VERSION.tar.gz"
cd openssl-$OPENSSL_VERSION
./config --prefix=$OPENSSL_PREFIX no-ssl3 no-tls1 no-tls1_1 no-shared
make > /dev/null
sudo make install_sw > /dev/null


cd /tmp
curl -O "https://www.python.org/ftp/python/$PYTHON_VERSION/Python-$PYTHON_VERSION.tgz"
tar -xzf "Python-$PYTHON_VERSION.tgz"
cd Python-$PYTHON_VERSION
./configure --prefix=$PYTHON_PREFIX --with-openssl=$OPENSSL_PREFIX --disable-shared > /
dev/null
make > /dev/null
sudo make install > /dev/null
```

This compiles a version of Python that has a statically linked OpenSSL that doesn't automatically negotiate anything earlier than TLS 1.2. This also installs OpenSSL in the `/opt/openssl-with-`

`min-tls1_2` directory and installs Python in the `/opt/python-with-min-tls1_2` directory. After you run this script, confirm installation of the new version of Python.

```
$ /opt/python-with-min-tls1_2/bin/python3 --version
```

This should print out the following.

```
Python 3.8.1
```

To confirm this new version of Python doesn't negotiate a version earlier than TLS 1.2, rerun the steps from Determine Your Currently Supported Protocols using the newly installed Python version (that is, `/opt/python-with-min-tls1_2/bin/python3`).

Configuring security groups for restricted environments

By default, AWS ParallelCluster creates and configures security groups that allow all traffic between cluster nodes. In highly restricted environments, you might need to limit network access to only the ports required for cluster operation. This section describes how to configure custom security groups with restricted access for your AWS ParallelCluster deployment.

Security groups overview

AWS ParallelCluster uses security groups to control network traffic between the head node, compute nodes, and login nodes (if configured). By default, when AWS ParallelCluster creates a cluster, it creates security groups that allow all traffic between nodes within the cluster. In environments with strict security requirements, you can provide custom security groups that limit traffic to only the necessary ports.

Security groups can be configured in the following sections of your cluster configuration:

- `HeadNode / Networking` - Controls access to and from the head node
- `Scheduling / SlurmQueues / Networking` - Controls access to and from compute nodes
- `LoginNodes` - Controls access to and from login nodes (if configured)

For each of these sections, you can specify:

- `SecurityGroups` - Replaces the default security groups that AWS ParallelCluster would create

- `AdditionalSecurityGroups` - Adds security groups in addition to the default ones created by AWS ParallelCluster

Required ports for cluster operation

When configuring custom security groups, you must ensure that the following ports are open between the appropriate nodes:

Required ports for head node

Port	Protocol	Direction	Purpose
22	TCP	Inbound	SSH access to the head node (from allowed IP ranges)
6817-6819	TCP	Inbound	Slurm controller ports (from compute and login nodes)
6817-6819	TCP	Outbound	Slurm controller ports (to compute and login nodes)
8443	TCP	Inbound	NICE DCV (if enabled, from allowed IP ranges)
111, 2049	TCP/UDP	Inbound	NFS (from compute and login nodes, if using NFS for shared storage)
443	TCP	Outbound	HTTPS access to AWS services (if not using VPC endpoints)

Required ports for compute nodes

Port	Protocol	Direction	Purpose
22	TCP	Inbound	SSH access (from head node and login nodes)
6818	TCP	Inbound	Slurm daemon port (from head node)

Port	Protocol	Direction	Purpose
6817-6819	TCP	Outbound	Slurm controller ports (to head node)
111, 2049	TCP/UDP	Outbound	NFS (to head node, if using NFS for shared storage)
443	TCP	Outbound	HTTPS access to AWS services (if not using VPC endpoints)

If you're using EFA (Elastic Fabric Adapter), you must also allow all traffic between compute nodes that have EFA enabled:

- All TCP and UDP traffic between compute nodes with EFA
- All traffic on the EFA device between compute nodes with EFA

> **Note**
>
> If you're using shared storage systems like FSx for Lustre, Amazon EFS, or other storage solutions, you'll need to ensure that the appropriate ports are open for those services as well.

Creating custom security groups

To create custom security groups for your AWS ParallelCluster deployment, follow these steps:

1. Create security groups for the head node, compute nodes, and login nodes (if applicable) using the AWS Management Console, AWS CLI, or AWS CloudFormation.
2. Configure the security group rules to allow only the necessary traffic as outlined in the previous section.
3. Reference these security groups in your cluster configuration file.

Here's an example of how to create security groups using the AWS CLI:

```
# Create security group for head node
aws ec2 create-security-group \
  --group-name pcluster-head-node-sg \
  --description "Security group for ParallelCluster head node" \
  --vpc-id vpc-12345678

# Create security group for compute nodes
aws ec2 create-security-group \
  --group-name pcluster-compute-node-sg \
  --description "Security group for ParallelCluster compute nodes" \
  --vpc-id vpc-12345678

# Add rules to allow necessary traffic between head and compute nodes
# (Add specific rules based on the required ports listed above)
```

Configuring security groups in the cluster configuration

Once you've created your custom security groups, you can reference them in your cluster configuration file:

```
# Example cluster configuration with custom security groups
HeadNode:
  ...
  Networking:
    SubnetId: subnet-12345678
    SecurityGroups:
      - sg-headnode12345  # Custom security group for head node
    # Or use AdditionalSecurityGroups if you want to keep the default security groups
    # AdditionalSecurityGroups:
    #   - sg-additional12345
  ...

Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - Name: queue1
      ...
      Networking:
        SubnetIds:
          - subnet-12345678
        SecurityGroups:
          - sg-computenode12345  # Custom security group for compute nodes
```

```
        # Or use AdditionalSecurityGroups if you want to keep the default security
 groups
        # AdditionalSecurityGroups:
        #   - sg-additional12345
      ...

# If using login nodes
LoginNodes:
  Pools:
    - Name: login-pool
      ...
      Networking:
        SubnetIds:
          - subnet-12345678
        SecurityGroups:
          - sg-loginnode12345  # Custom security group for login nodes
        # Or use AdditionalSecurityGroups if you want to keep the default security
 groups
        # AdditionalSecurityGroups:
        #   - sg-additional12345
      ...
```

When using `SecurityGroups`, AWS ParallelCluster will use only the security groups you specify, replacing the default ones. When using `AdditionalSecurityGroups`, AWS ParallelCluster will use both the default security groups it creates and the additional ones you specify.

> **Warning**
>
> If you enable Elastic Fabric Adapter (EFA) for your compute instances, make sure that your EFA-enabled instances are members of a security group that allows all inbound and outbound traffic to itself. This is required for EFA to function properly.

Using VPC endpoints in restricted environments

In highly restricted environments, you might want to deploy AWS ParallelCluster in a subnet without internet access. In this case, you'll need to configure VPC endpoints to allow the cluster to communicate with AWS services. For detailed instructions, see AWS ParallelCluster in a single subnet with no internet access.

When using VPC endpoints, ensure that your security groups allow traffic to and from the VPC endpoints. You can do this by adding the security groups associated with the VPC endpoints to the `AdditionalSecurityGroups` configuration for your head node and compute nodes.

```
HeadNode:
  ...
  Networking:
    SubnetId: subnet-1234567890abcdef0
    AdditionalSecurityGroups:
      - sg-abcdef01234567890  # Security group that enables communication with VPC
 endpoints
  ...

Scheduling:
  Scheduler: slurm
  SlurmQueues:
    - ...
      Networking:
        SubnetIds:
          - subnet-1234567890abcdef0
        AdditionalSecurityGroups:
          - sg-1abcdef01234567890  # Security group that enables communication with VPC
 endpoints
```

Best practices for security group configuration

When configuring security groups for AWS ParallelCluster in restricted environments, consider the following best practices:

- **Principle of least privilege**: Only open the ports that are necessary for cluster operation.
- **Use security group references**: When possible, use security group references (allowing traffic from another security group) rather than CIDR blocks to limit traffic between cluster components.
- **Restrict SSH access**: Limit SSH access to the head node to only the IP ranges that need it using the `HeadNode / Ssh / AllowedIps` configuration.
- **Restrict DCV access**: If using NICE DCV, limit access to only the IP ranges that need it using the `HeadNode / Dcv / AllowedIps` configuration.
- **Test thoroughly**: After configuring custom security groups, thoroughly test all cluster functionality to ensure that all required communication paths are working.

- **Document your configuration**: Maintain documentation of your security group configuration, including which ports are open and why they are needed.

Troubleshooting security group issues

If you encounter issues after configuring custom security groups, consider the following troubleshooting steps:

- **Check cluster logs**: Review the cluster logs in CloudWatch Logs for any connection errors.
- **Verify security group rules**: Ensure that all required ports are open between the appropriate nodes.
- **Test connectivity**: Use tools like `telnet` or `nc` to test connectivity between nodes on specific ports.
- **Temporarily expand rules**: If you're having trouble identifying which ports are needed, temporarily allow all traffic between cluster nodes and then gradually restrict it as you identify the required ports.
- **Check VPC endpoint configuration**: If you're using VPC endpoints, ensure that they are properly configured and that the security groups allow traffic to and from them.

If you continue to experience issues, you can revert to using the default security groups created by AWS ParallelCluster by removing the `SecurityGroups` configuration from your cluster configuration file.

Supported AWS Regions for AWS ParallelCluster

AWS ParallelCluster version 3 is available in the following AWS Regions:

Region Name	Region	First AWS ParallelCluster version supporting the region
US East (Ohio)	us-east-2	3.0.0
US East (N. Virginia)	us-east-1	3.0.0
US West (N. California)	us-west-1	3.0.0
US West (Oregon)	us-west-2	3.0.0
Africa (Cape Town)	af-south-1	3.0.0
Asia Pacific (Hong Kong)	ap-east-1	3.0.0
Asia Pacific (Mumbai)	ap-south-1	3.0.0
Asia Pacific (Seoul)	ap-northeast-2	3.0.0
Asia Pacific (Singapore)	ap-southeast-1	3.0.0
Asia Pacific (Sydney)	ap-southeast-2	3.0.0
Asia Pacific (Jakarta)	ap-southeast-3	3.10.0
Asia Pacific (Malaysia)	ap-southeast-5	3.13.0
Asia Pacific (Thailand)	ap-southeast-7	3.13.0
Asia Pacific (Tokyo)	ap-northeast-1	3.0.0
Canada (Central)	ca-central-1	3.0.0
China (Beijing)	cn-north-1	3.0.0
China (Ningxia)	cn-northwest-1	3.0.0

Region Name	Region	First AWS ParallelCluster version supporting the region
Europe (Frankfurt)	eu-central-1	3.0.0
Europe (Ireland)	eu-west-1	3.0.0
Europe (London)	eu-west-2	3.0.0
Europe (Milan)	eu-south-1	3.0.0
Europe (Paris)	eu-west-3	3.0.0
Europe (Stockholm)	eu-north-1	3.0.0
Middle East (Bahrain)	me-south-1	3.0.0
South America (São Paulo)	sa-east-1	3.0.0
AWS GovCloud (US-East)	us-gov-east-1	3.0.0
AWS GovCloud (US-West)	us-gov-west-1	3.0.0
Israel (Tel Aviv)	il-central-1	3.8.0

Release notes and document history

The following tables describe the major updates and new features for the *AWS ParallelCluster User Guide*. We also update the documentation frequently to address the feedback that you send us.

AWS ParallelCluster

Change	Description	Date
AWS ParallelCluster version 3.13.2 released	To upgrade, enter the following: `sudo pip install --upgrade aws-parallelcluster .` Bug fixes: • Fix a bug which may cause `update-cluster` and `update-compute-fleet` to fail when compute resources reference an expired Capacity Reservation that is no longer accessible via EC2 APIs. • Fix `build-image` failure on Rocky 9 that occurs when the parent image does not ship the latest kernel version. See https://github.com/aws/aws-parallelcluster/issues/6874.	June 24, 2025
AWS ParallelCluster version 3.13.1 released	To upgrade, enter the following: `sudo pip install --upgrade aws-parallelcluster .` Changes:	June 04, 2025

Change	Description	Date

	• Upgrade Slurm to version 24.05.8.
	• Upgrade EFA installer to 1.41.0 (from 1.38.1).
	  • Efa-driver: efa-2.15.0-1
	  • Efa-config: efa-config-1.18-1
	  • Efa-profile: efa-profile-1.7-1
	  • Libfabric-aws: libfabric-aws-2.1.0-1
	  • Rdma-core: rdma-core-57.0-1
	  • Open MPI: openmpi40-aws-4.1.7-2 and openmpi50-aws-5.0.6
	• Upgrade amazon-efs-utils to version 2.3.1 (from v2.1.0) for non-Amazon Linux AMI's.
	• Support DCV in us-isob-east-1 and us-iso-east-1.
	• Support FSX for Lustre and Ontap in us-isob-east-1 and us-iso-east-1.
	• Ensure kernel consistency throughout ParallelCluster image build by pinning at the beginning and unpinning at completion.	

Change	Description	Date
	Bug fixes: • Fix a bug in the installat ion of ARM Performance Library that was causing the build image fail in isolated environments. • Fix a bug that was preventing the script 'update_directory_service_p assword.sh' from updating the AD password.	

Change	Description	Date
AWS ParallelCluster version 3.13.0 released	To upgrade, enter the following: `sudo pip install --upgrade aws-parallelcluster .` Deprecations: • This is the last ParallelCluster release supporting Ubuntu 20.04 as Ubuntu 20.04 will be in End-Of-Standard-Support in May 2025. Enhancements: • Add support for Ubuntu 24.04. • Add support for ap-southeast-7 region. • Disable unused services cups and wpa_supplicant from Official ParallelCluster AMIs to improve security. Changes: • Upgrade Slurm to version 24.05.7. • Upgrade NVIDIA driver to version 570.86.15 (from 550.127.08) for all OSs except AL2.	April 01, 2025

Change	Description	Date

	• Upgrade CUDA Toolkit to version 12.8.0 (from 12.4.1) for all OSs except AL2.
	• Upgrade Python to 3.12.8 for all OSs except AL2 (from 3.9.20).
	• On Ubuntu 22.04, install the Nvidia driver with the same compiler version used to compile the kernel.
	• Upgrade `aws-cfn-bootstrap` to version 2.0-33.
	• Upgrade EFA installer to 1.38.0 (from 1.36.0).
	  • Efa-driver: efa-2.13.0-1
	  • Efa-config: efa-config-1.17-1
	  • Efa-profile: efa-profile-1.7-1
	  • `Libfabric-aws` : `libfabric-aws-1.22.0-1`
	  • Rdma-core: rdma-core-54.0-1
	  • Open MPI: `openmpi40-aws-4.1.7-1` and `openmpi50-aws-5.0.5`
	• Upgrade amazon-efs-utils to version 2.1.0.	

Change	Description	Date

	• Remove third-party cookbook: apt-7.5.22 and pyenv-4.2.3.
	• Upgrade third-party cookbook dependencies:
	  • line-4.5.21 (from line-4.5.13)
	  • nfs-5.1.5 (from nfs-5.1.2)
	  • openssh-2.11.14 (from openssh-2.11.12)
	  • yum-7.4.20 (from yum-7.4.13)
	  • yum-epel-5.0.8 (from yum-epel-5.0.2)
	• Upgrade Pmix to 5.0.6 (from 5.0.3).
	• Upgrade ARM PL to version 24.10 (from 23.10).
	• Upgrade Python to version 3.12.8 (from 3.9.17) in Lambda layer and installer.
	• Upgrade NodeJS to version 20.18.3 (from 18.20.3) in Lambda layer and installer.
	• Remove generation of DSA keys for login nodes as DSA, which became unsupported in OpenSSH 9.7+.
	• Set instance ID and instance type information in Slurm	

Change	Description	Date
	upon compute nodes launch. • Install NVIDIA drivers without the option 'no-cc-version-check', which is now deprecated in the NVIDIA installer. • Add validator to enforce up to 10- login node pools. • Update the default root volume size to 45 GB. Bug fixes: • Remove usage of cfn-init for compute node bootstrapping to reduce node scale up time. • Fix an issue causing compute node bootstrap failure when a proxy is used. • On Ubuntu 22.04, install the Nvidia driver with the same compiler version used to compile the kernel to prevent installation failures.- Fix the execution of overriding aws-paral lelcluster-node package only on the head node during update. • Fix an issue where container ized jobs executed through	

Change	Description	Date
	Pyxis/Enroot in a multi-user environment (integrated with Active Directory) would fail. • Fix usage of authselect causing node bootstrap failures on Rocky 9.5+ when directory service is used.	

Change	Description	Date
AWS ParallelCluster version 3.12.0 released	To upgrade, enter the following: `sudo pip install --upgrade aws-parallelcluster .` Enhancements: • Add new build image configuration section `Build/Installation` to turn on/off NVIDIA software and Lustre client installations. By default, NVIDIA software, although included in official ParallelCluster AMIs, is not installed by `build-image` . By default, Lustre client is installed. • The CLI commands `export-cluster-logs` and `export-image-logs` can now by default export the logs to the default ParallelCluster bucket or to the CustomS3Bucket if specified in the config. • Extend Amazon DCV support to Ubuntu2204 on ARM instances. Changes:	December 19, 2024

Change	Description	Date

	• Upgrade NVIDIA driver to version 550.127.08 (from 550.90.07). This addresses a known issue from NVIDIA. For more information, see Known Issues in the *NVIDIA Data Center Documentation*.
	• Upgrade Amazon DCV to version `2024.0-18131` .
	  • server: `2024.0-18131-1`
	  • xdcv: `2024.0.631-1`
	  • gl: `2024.0.1078-1`
	  • web_viewer: `2024.0-18131-1`
	• Upgrade EFA installer to `1.36.0`.
	  • Efa-driver: `efa-2.13.0-1`
	  • Efa-config: `efa-config-1.17-1`
	  • Efa-profile: `efa-profile-1.7-1`
	  • Libfabric-aws: `libfabric-aws-1.22.0-1`
	  • Rdma-core: `rdma-core-54.0-1`
	  • Open MPI: `openmpi40-aws-4.1.7-1` and	

Change	Description	Date
	`openmpi50-aws-5.0.5` • Auto-restart slurmctld on failure. • Upgrade mysql-community-client to version 8.0.39. • Remove support for Python 3.7 and 3.8, which are end of life. Bug fixes: • Fix an issue where changes in sequence of custom actions scripts were not detected during cluster updates. • Add missing permissions for the AWS ParallelCluster API to create the service linked roles for Elastic Load Balancing and Auto Scaling, that are required to deploy login nodes. • Fix an issue in the way we get the region when manage volumes so that it can correctly handle local zone. • Fix an issue where adding EFS filesystems with `AccessPointIds` during an update would fail.	

Change	Description	Date
	• Fix an issue where when using PCAPI, cluster update could fail when updating a parameter that is not type `String` (for example, `MaxCount`). • When mounting an external OpenZFS, it is no longer required to set the outbound rules for ports 111, 2049, 20001, 20002, 20003.	

Change	Description	Date
AWS ParallelCluster version 3.11.1 released	Features: • Pyxis is now disabled by default, so it must be manually enabled as documented in the product documentation. • Upgrade Python runtime to version 3.12 in ParallelCluster Lambda Layer. • Remove version pinning for setuptools to version prior to 70.0.0. • Upgrade libjwt to version 1.17.0. • Full Changelog Bug fixes • Fix an issue in the way we configure the Pyxis Slurm plugin in ParallelCluster that can lead to job submission failures. • Fix an issue that was causing failing deployment in configurations with login nodes by add missing permissions required by login nodes in the public template of policies. https://github.com/aws/a	October 21, 2024

Change	Description	Date
	ws-parallelcluster/issues/6483	

Change	Description	Date
AWS ParallelCluster version 3.11.0 released	Enhancements • Add support for custom actions on login nodes. • Allow DCV connection to login nodes. • Add support for ap-southeast-3 region. • Add security groups to login node network load balancer. • Add `AllowedIps` configuration for login nodes. • Add new configuration `SharedStorage/EfsSettings/AccessPointId` to specify an optional EFS access point for a mount • Allow up to 10 login node pools. • Install enroot and pyxis in official pcluster AMIs Changes • [*BREAKING*] The `loginNodes` field returned by the API `DescribeCluster` and the CLI command `describe-cluster`	September 26, 2024

Change	Description	Date

	has been changed from a dictionary to an array to support multiple pools of login nodes. This change breaks backward compatibility, making these operations incompatible with clusters deployed with older versions.
	• Upgrade Slurm to 23.11.10 (from 23.11.7).
	• Upgrade Pmix to 5.0.3 (from 5.0.2).
	• Upgrade EFA installer to `1.34.0`.
	  • Efa-driver: `efa-2.10.0-1`
	  • Efa-config: `efa-config-1.17-1`
	  • Efa-profile: `efa-profile-1.7-1`
	  • Libfabric-aws: `libfabric-aws-1.22.0-1`
	  • Rdma-core: `rdma-core-52.0-1`
	  • Open MPI: `openmpi40-aws-4.1.6-3` and `openmpi50-aws-5.0.3-11`
	• Upgrade NVIDIA driver to version 550.90.07 (from 535.183.01).	

Change	Description	Date
	• Upgrade CUDA Toolkit to version 12.4.1 (from 12.2.2). • Upgrade Python to 3.9.20 (from 3.9.19). • Upgrade Intel MPI Library to 2021.13.1.769 (from 2021.12.1.8). Bug fixes • Fix validator `EfaPlacementGroupValidator` so that it does not suggest to configure a Placement Group when Capacity Blocks are used. • Fix occasional cluster creation failures by ensuring that FSx for Lustre file systems are created after security group rules. • Fix cluster deletion failure when placement group is enabled. • Fix issue with login nodes being marked unhealthy when restricting SSH access. • Fix `retrieve_supported_regions` so that it can get the correct S3 url.	

Change	Description	Date
	• Fix `describe_images` to use pagination. • Fix `No route tables found` bug when specifying default VPC subnet to LoginNodes/Networking/SubnetIds.	
AWS ParallelCluster version 3.10.1 released	Bug fixes • Fix image build failure in China regions.	July 8, 2024

Change	Description	Date
AWS ParallelCluster version 3.10.0 released	Enhancements: • Add new configuration section `Scheduling/SlurmSettings/ExternalSlurmdbd` to connect the cluster to an external Slurmdbd. • Allow build-image to be run in an isolated network. • Add support for Amazon Linux 2023. • Add support for `price-capacity-optimized` as an `AllocationStrategy`. • Add validator to prevent the use of Placement Groups with Capacity Blocks. Changes: • CentOS 7 is no longer supported. • Upgrade Cinc Client to version to 18.4.12 from 18.2.7. • Upgrade munge to version 0.5.16 (from 0.5.15). • Upgrade Pmix to 5.0.2 (from 4.2.9).	June 27, 2024

Change	Description	Date

	• Upgrade third-party cookbook dependencies:
	  • apt-7.5.22 (from apt-7.5.14)
	  • openssh-2.11.12 (from openssh-2.11.3)
	• Remove third-party cookbook: selinux-6.1.12.
	• Upgrade EFA installer to `1.32.0`.
	  • Efa-driver: `efa-2.8.0-1`
	  • Efa-config: `efa-config-1.16-1`
	  • Efa-profile: `efa-profile-1.7-1`
	  • Libfabric-aws: `libfabric-aws-1.21.0-1`
	  • Rdma-core: `rdma-core-50.0-1`
	  • Open MPI: `openmpi40-aws-4.1.6-3` and `openmpi50-aws-5.0.2-12`
	• Upgrade NVIDIA driver to version 535.183.01 (from 535.154.05).
	• Upgrade Python to 3.9.19 (from 3.9.17).	

Change	Description	Date
	• Upgrade Intel MPI Library to 2021.12.1.8 (from 2021.9.0.43482). Bug fixes: • Fix Data Repository Associations configuration to make `AutoExportPolicy` and `AutoImportPolicy` optional. • Fixed an issue during cluster deletion that now completes compute fleet cleanup when instances are either in shutting-down or terminated state. This is to avoid cluster deletion failures for instance types with longer termination cycles. • Allow cloudwatch dashboard to be enabled and alarms to be disabled in the `Monitoring` section of the cluster config. • Allow ParallelCluster Custom Resource to suppress validators using `PclusterCluster/SuppressValidators` . • Removing `/etc/profile.d/pcluster.sh`	

Change	Description	Date
	so that it's not executed at every user login and `cfn_bootstrap_virtualenv` is not added in PATH environment variable. • Fix ParallelCluster API spec by replacing field `failureReason` with `failures` in `DescribeCluster` response. • Fix ParallelCluster API spec by adding the CloudFormation stack status that were missing: `IMPORT_*`, `REVIEW_IN_PROGRESS` and `UPDATE_FAILED` . • Fix an issue that prevented cluster updates from including EFS filesystems with encryption in transit. • Fix an issue that prevented slurmctld and slurmdbd services from restarting on head node reboot when EFS is used for shared internal data. • On Ubuntu systems, remove default logrotate configuration for cloud-init log files that clashed with the configuration coming from ParallelCluster.	

Change	Description	Date
	• Fix image build failure with RHEL 8.10 or newer.	
AWS ParallelCluster version 3.9.3 released	To upgrade, type `sudo pip install --upgrade aws-parallelcluster` Features: • Added support for `FSx Lustre` as a shared storage type in `us-iso-east-1 .` Bug fixes: • Remove `cloud_dns` from the `Slurmctld Parameters` in the Slurm config to avoid Slurm fanout issues. This isn't required, since we set the IP addresses on instance launch.	June 19, 2024
AWS ParallelCluster version 3.9.2 released	Features: • Upgrade Slurm to 23.11.7 (from 23.11.4). • For more details, see the `CHANGELOG 3.9.2` on GitHub.	May 28, 2024

Change	Description	Date
AWS ParallelCluster version 3.9.1 released	To upgrade, enter the following: `sudo pip install --upgrade aws-parallelcluster` Bug fixes • Remove recursive deletion of shared storage mountdir when unmounting filesystems as part of update-cluster operation.	April 11, 2024

Change	Description	Date
AWS ParallelCluster version 3.9.0 released	To upgrade, enter the following: `sudo pip install --upgrade aws-parallelcluster` Enhancements: • Add the configuration parameter `DeploymentSettings/DefaultUserHome` to allow users to move the default user's home directory to `/local/home` instead of `/home` (default). • Permit to update `MinCount`, `MaxCount`, `Queue` and `ComputeResource` configuration parameters without the need to stop the compute fleet. It's now possible to update them by setting `Scheduling/SlurmSettings/QueueUpdateStrategy` to TERMINATE. AWS ParallelCluster will terminate only the nodes removed during a resize of the cluster capacity performed through a cluster update. • Permit to update the external shared storage of type Efs, FsxLustre,	March 5, 2024

Change	Description	Date
	FsxOntap, FsxOpenZfs and FileCache without replacing the compute and login fleet. • Add support for RHEL9. • Add support for Rocky Linux 9 as `CustomAmi` created through `build-image` process. No public official AWS ParallelCluster Rocky9 Linux AMI is made available at this time. • Remove `CommunicationParameters` from the Custom Slurm Settings deny list. • Add `DeploymentSettings/DisableSudoAccessForDefaultUser` parameter to disable sudo access of default user in supported OSes. • Changes to FSx for Lustre file systems created by ParallelCluster: Change the Lustre server version to `2.15`. • Add possibility to choose between Open and Closed Source Nvidia Drivers when building an AMI, through the `['cluster`	

Change	Description	Date
	`']['nvidia']['kernel_open']` cookbook node attribute. • * Add a clustermgtd config option `ec2_instance_missing_max_count` to allow a configurable amount of retries for eventual Amazon EC2 describe instances consistency with run instances. Changes • Upgrade Slurm to 23.11.4 (from 23.02.7). • Upgrade NVIDIA driver to version 535.154.05. • Add support for Python 3.11, 3.12 in pcluster CLI and aws-parallelcluster-batch-cli. • Build network interfaces using network card index from `NetworkCardIndex` list of Amazon EC2 DescribeInstances response, instead of looping over `MaximumNetworkCards` range. • Fail cluster creation when using instance types P3, G3, P2 and G2 because their GPU architecture	

Change	Description	Date

	is not compatible with Open Source Nvidia Drivers (OpenRM) introduced as part of 3.8.0 release.
	• Upgrade third-party cookbook dependencies: nfs-5.1.2 (from nfs-5.0.0)
	• Upgrade EFA installer to `1.30.0.`
	  • Efa-driver: `efa-2.6.0-1`
	  • Efa-config: `efa-config-1.15-1`
	  • Efa-profile: `efa-profile-1.6-1`
	  • Libfabric-aws: `libfabric-aws-1.19.0`
	  • Rdma-core: `rdma-core-46.0-1`
	  • Open MPI: `openmpi40-aws-4.1.6-2` and `openmpi50-aws-5.0.0-11`
	• Upgrade NICE DCV to version `2023.1-16388.`
	  • server: `2023.1.16388-1`
	  • xdcv: `2023.1.565-1`
	  • gl: `2023.1.1047-1`
	  • web_viewer: `2023.1.16388-1`	

Change	Description	Date
	Bug fixes • Fix issue making job fail when submitted as active directory user from login nodes. The issue was caused by an incomplet e configuration of the integration with the external Active Directory on the head node. • Refactor IAM policies defined in CloudForm ation template parallelc lutser-policies.yaml to prevent ParallelCluster API deployment failure caused by policies exceeding IAM limits. • Fix issue making login nodes fail to bootstrap when the head node takes more time than expected in writing keys. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster-ui package on GitHub.	

Change	Description	Date
AWS ParallelCluster version 3.8.0 released	AWS ParallelCluster version 3.8.0 released. Enhancements: • Add support for Amazon EC2 Capacity Blocks for ML. • Add support for Rocky Linux 8 as `CustomAmi` created through `build-image` process. No public official AWS ParallelCluster Rocky8 Linux AMI is made available at this time. • Add `Scheduling/ScalingStrategy` parameter to control the cluster scaling strategy to use when launching Amazon EC2 instances for Slurm compute nodes. Possible values are `all-or-nothing` , `greedy-all-or-nothing` , `best-effort` , with `all-or-nothing` being the default. • Add `HeadNode/SharedStorageType` parameter to use EFS storage instead of NFS exports from the head node root volume for intra-cluster shared file system resources: ParallelCluster,	December 19, 2023

Change	Description	Date
	Intel, Slurm, and /home data. This enhancement reduces the load on the head node networking. • Allow for mounting /home as an EFS or FSx external shared storage via the SharedStorage section of the config file. • Add new parameter SlurmSettings/MungeKeySecretArn to permit to use an external user-defined MUNGE key from AWS Secrets Manager. • Add Monitoring/Alarms/Enabled parameter to toggle Amazon CloudWatch Alarms for the cluster. • Add head node alarms to monitor Amazon EC2 health checks, CPU utilization and the overall status of the head node, and add them to the CloudWatch Dashboard created with the cluster. • Add support for Data Repository Associations when using PERSISTENT_2 as DeploymentType for a managed FSx for Lustre.	

Change	Description	Date
	• Add `Scheduling/SlurmSettings/Database/DatabaseName` parameter to allow users to specify a custom name for the database on the database server to be used for Slurm accounting. • Make `InstanceType` an optional configuration parameter when configuring `CapacityReservationTarget/CapacityReservationId` in the compute resource. • Add possibility to specify a prefix for IAM roles and policies created by AWS ParallelCluster API. • Add possibility to specify a permissions boundary to be applied for IAM roles and policies created by AWS ParallelCluster API. Changes • Upgrade Slurm to 23.02.7 (from 23.02.6). • Upgrade NVIDIA driver to version 535.129.03. • Upgrade CUDA Toolkit to version 12.2.2.	

Change	Description	Date

	• Use Open Source NVIDIA GPU drivers (OpenRM) as NVIDIA kernel module for Linux instead of NVIDIA closed source module.
	• Remove support of `all_or_nothing_batch` configuration parameter in the Slurm resume program, in favor of the new `Scheduling/ScalingStrategy` cluster configuration.
	• Changed cluster alarms naming convention to '[cluster-name]-[component-name]-[metric]'.
	• Change default EBS volume types in ADC regions from gp2 to gp3, for both the root and additional volumes.
	• The optional permissions boundary for the AWS ParallelCluster API is now applied to every IAM role created by the API infrastructure.
	  • Upgrade EFA installer to `1.29.1`.
	  • Efa-driver: `efa-2.6.0-1`	

Change	Description	Date
	• Efa-config: `efa-config-1.15-1` • Efa-profile: `efa-profile-1.5-1` • Libfabric-aws: `libfabric-aws-1.19.0-1` • Rdma-core: `rdma-core-46.0-1` • Open MPI: `openmpi40-aws-4.1.6-1` • Upgrade GDRCopy to version 2.4 in all supported OSes, except for Centos 7 where version 2.3.1 is used. • Upgrade `aws-cfn-bootstrap` to version 2.0-28. • Add support for Python 3.10 in aws-parallelcluster-batch-cli. Bug fixes • Fix inconsistent scaling configuration after cluster update rollback when modifying the list of instance types declared in the Compute Resources. • Fix users SSH keys generation when switching users without root privilege	

Change	Description	Date
	in clusters integrated with an external LDAP server through cluster configuration files. • Fix disabling Slurm power save mode when setting `ScaledownIdletime = -1`. • Fix hard-coded path to Slurm installation dir in `update_slurm_database_password.sh` script for Slurm Accounting.	
AWS ParallelCluster version 3.7.2 released	AWS ParallelCluster version 3.7.2 released. Changes: • Upgrade Slurm to 23.02.6.	October 25, 2023

Change	Description	Date

AWS ParallelCluster version 3.7.1 released	AWS ParallelCluster version 3.7.1 released.
	
	Changes:
	
	• Upgrade Slurm to 23.02.5 (from 23.02.4).
	  • Upgrade Pmix to 4.2.6 (from 3.2.3).
	  • Upgrade libjwt to 1.15.3 (from 1.12.0).
	• Upgrade EFA installer to `1.26.1`, fixing RDMA writedata issue in P5.
	  • Efa-driver: `efa-2.5.0-1 .`
	  • Efa-config: `efa-config-1.15-1 .`
	  • Efa-profile: `efa-profile-1.5-1 .`
	  • Libfabric-aws: `libfabric-aws-1.18.2-1 .`
	  • ERdma-core: `rdma-core-46.0-1 .`
	  • Open MPI: `openmpi40-aws-4.1.5-4 .`
		September 22, 2023

Change	Description	Date
AWS ParallelCluster version 3.7.0 released	AWS ParallelCluster version 3.7.0 released. Enhancements: • Support configuration of static and dynamic node priorities in compute resources by using a AWS ParallelCluster configuration YAML file. • Add support for Ubuntu 22. RSA keys are not supported by default. • Add the queue configuration setting `JobExclusiveAllocation` to allocate nodes in a partition exclusively to a single job at any given time. • Allow Override `aws-parallelcluster-node` package at cluster create and cluster update time. For the head node, this applies for cluster update. Useful for development purposes only. • Avoid NFS server start on compute nodes. • Add support for log-in nodes. • Allow memory-based scheduling when multiple	August 30, 2023

Change	Description	Date
	instance types are specified for a Slurm Compute Resource. • Add support to mount existing Amazon File Cache as shared storage. Changes: • Assign Slurm dynamic nodes a priority (weight) of 1000 by default. By doing this, Slurm can prioritize idle static nodes over idle dynamic nodes. • Make `aws-parallelcluster-node` daemons only handle AWS ParallelCluster managed Slurm partitions. • Increase `EFS-utils` watchdog poll interval to 10 seconds. This change applies when `EncryptionInTransit` is set to `true`, which is the only condition that causes the watchdog to run. • Upgrade the EFA installer to `1.25.1`. • Efa-driver: `efa-2.5.0-1` (from `efa-2.1.1g`)	

Change	Description	Date
	• Efa-config: `efa-config-1.15-1` (from `efa-config-1.13-1`) • Efa-profile: `efa-profile-1.5-1` (no change) • Libfabric-aws: `libfabric-aws-1.18.1-0` (from `libfabric-aws-1.17.1-1`) • Rdma-core: `rdma-core-46.0-1` (from `rdma-core-43.0-1`) • Open MPI: `openmpi40-aws-4.1.5-4` (from `openmpi40-aws-4.1.5-1`) • Upgrade Slurm to version 23.02.4. • Change the default value of Imds/ImdsSupport from v1.0 to v2.0. • Deprecate Ubuntu 18. • Update the default root volume size to 40 GB to account for limits on Centos 7. • Restrict permission on file /tmp/wait_condition_handle.txt within the head node so that only root can read it.	

Change	Description	Date

	• Create a Slurm partition -nodelist mapping JSON file to be used by the node package daemons to recognize PC-manage d Slurm partitions and nodelists.
	• Upgrade NVIDIA driver to version 535.54.03.
	• Upgrade CUDA library to version 12.2.0.
	• Upgrade NVIDIA Fabric manager to nvidia-fa bricmanager-535.
	• Upgrade ARM PL to version 23.04.1 for Ubuntu 22.04 only.
	• Upgrade NICE DCV to version `2023.0-15487` .
	  • Server: `2023.0.15 487-1`
	  • xdcv: `2023.0.551-1`
	  • gl: `2023.0.1039-1`
	  • web_viewer: `2023.0.15 487-1`
	
	Bug fixes:
	• Add validation to the `ScaledownIdletime` value, to prevent setting a value lower than -1.	

Change	Description	Date
	• Fix cluster create failure with Ubuntu Deep Learning AMI on GPU instances with DCV enabled. • Fix issue causing dangling IAM policies to be created when creating ParallelCluster CloudFormation custom resource provider with CustomLambdaRole. • Fix an issue that was causing misalignment of compute nodes DNS name on instances with multiple network interfaces, when using `SlurmSettings/Dns/UseEc2Hostnames` equals to `True` For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	
Documentation only release	AWS ParallelCluster version 3 specific user guide published. Documentation only release: • AWS ParallelCluster version 3 has its own separate user guide.	July 17, 2023

Change	Description	Date
AWS ParallelCluster version 3.6.1 released	AWS ParallelCluster version 3.6.1 released. Changes: • Avoid duplication of nodes seen by `clustermgtd` if compute nodes are added to multiple Slurm partitions. Bug fixes: • Remove hard coding of root volume device name (`/dev/sda1` and `/dev/xvda`) and retrieve it from the AMIs used during `create-cluster`. • Fix cluster create failure when using CloudFormation custom resource with `ElasticIp` set to `True`. • Fix cluster create and update failures when using a AWS CloudFormation custom resource with large configuration files. • Fix an issue that prevented `ptrace` protection from being disabled on Ubuntu and that didn't permit Cross Memory Attach (CMA) in libfabric.	July 5, 2023

Change	Description	Date
	• Fix fast insufficient capacity fail-over logic when using multiple instance types and no instances are returned. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	

Change	Description	Date
AWS ParallelCluster version 3.6.0 released	AWS ParallelCluster version 3.6.0 released. Documentation: • Add documentation for the AWS ParallelCluster Python library API. Enhancements: • Add support for RHEL8. • Add an AWS CloudFormation custom resource for creating and managing clusters with CloudFormation. • Add support for customizing the cluster Slurm configuration in the AWS ParallelCluster configuration YAML file. • Build Slurm with support for LUA. • Increase the limit on the maximum number of queues per cluster from 10 to 50. Each queue can have up to 50 compute resources. Each cluster can have up to 50 compute resources. • Add support for specifying a sequence of multiple custom action scripts	May 22, 2023

Change	Description	Date
	for an event configure d in `OnNodeStart` , `OnNodeConfigured` , and `OnNodeUpdated` parameters. • Add new configuration section `HealthChecks` / `Gpu`, for applying GPU health checks on a compute node before a job is run. • Add support for `Tags` in the `SlurmQueues` and `SlurmQueues` / `ComputeResources` configuration. • Add support for DetailedMonitoring in the `Monitoring` configura tion. • Add `mem_used_ percent` and `disk_used _percent` metrics for head node memory and root volume disk utilizati on tracking in the AWS ParallelCluster CloudWatc h dashboard, and set up alarms for monitoring these metrics. • Add log rotation support for AWS ParallelCluster managed logs.	

Change	Description	Date
	• Track common compute node errors and dynamic node longest idle time in the CloudWatch Dashboard. • Enforce the DCV Authentic ator Server to use at least `TLS-1.2` protocol when creating the SSL Socket. • Install the NVIDIA Data Center GPU Manager (DCGM) package on all supported operating systems except `aarch64` `centos7` and `alinux2`. • Load the kernel module nvidia-uvm by default to provide Unified Virtual Memory (UVM) functiona lity to the CUDA driver. • Install the NVIDIA Persisten ce Daemon as a system service. Changes: • Upgrade Slurm to version `23.02.2` (from version `22.05.8`). • Upgrade munge to version `0.5.15` (from version `0.5.14`). • Set the Slurm `TreeWidth` to 30.	

Change	Description	Date

	• Set the Slurm `prolog` and `epilog` configurations to target directory `/opt/slurm/etc/scripts/prolog.d/` and `/opt/slurm/etc/scripts/epilog.d/` respectively.
	• Set Slurm `BatchStartTimeout` to 3 minutes maximum for running `Prolog` scripts during compute node registration.
	• Increase the default `RetentionInDays` of CloudWatch logs from 14 to 180 days.
	• Upgrade the EFA installer to `1.22.1`.
	  • Dkms: `2.8.3-2`
	  • Efa-driver: `efa-2.1.1g` (no change)
	  • Efa-config: `efa-config-1.13-1` (no change)
	  • Efa-profile: `efa-profile-1.5-1` (no change)
	  • Libfabric-aws: `libfabric-aws-1.17.1-1` (from `libfabric-aws-1.17.0-1`)
	  • Rdma-core: `rdma-core-43.0-1` (no change)	

Change	Description	Date
	• Open MPI: `openmpi40-aws-4.1.5-1` (no change) • Upgrade the Lustre client version to `2.12` on Amazon Linux 2. Lustre client `2.12` has been installed on Ubuntu 20.04, 18.04, and CentOS >= 7.7. • Upgrade the Lustre client version to `2.10.8` on CentOS 7.6. • Upgrade the NVIDIA driver to version `470.182.03` (from version `470.141.03`). • Upgrade the NVIDIA Fabric Manager to version `470.182.03` (from version `470.141.03`). • Upgrade the NVIDIA CUDA Toolkit to version `11.8.0` (from version `11.7.1`). • Upgrade the NVIDIA CUDA sample to version `11.8.0`. • Upgrade the Intel MPI Library to Version 2021 Update 9 (from Version 2021 Update 6). For more information, see Intel® MPI Library 2021 Update 9. • Upgrade NICE DCV to version `2023.0-15022`	

Change	Description	Date
	(from version `2022.2-14521`). • server: `2023.0.15022-1` (from version `2022.2-14521-1`). • xdcv: `2023.0.547-1` (from version `2022.2.519-1`). • gl: `2023.0.1027-1` (from version `2022.2.1012-1`). • web_viewer: `2023.0.15022-1` (from version `2022.2.14521-1`). • Upgrade `aws-cfn-bootstrap` to version `2.0-24`. • Upgrade image used by the CodeBuild environment when building container images for AWS Batch clusters: • `aws/codebuild/amazonlinux2-x86_64-standard:4.0` (from `aws/codebuild/amazonlinux2-x86_64-standard:3.0`). • `aws/codebuild/amazonlinux2-aarch64-standard:2.0` (from `aws/codebuild/amaz`	

Change	Description	Date
	onlinux2-aarch64-standard:1.0). Bug fixes: • Fix Amazon EFS and Amazon FSx network security group validators to avoid reporting false errors. • Fix missing tagging of resources created by Image Builder during the build-image operation. • Fix update policy for MaxCount to always perform numerical comparisons on the MaxCount property. • Fix IP alignment on compute node instances with multiple network cards. • Fix replacement of StoragePass in the slurm_parallelcluster_slurmdbd.conf when a queue parameter update is performed and the Slurm accounting configurations are not updated. • Fix issue that causes dangling security groups to be created when creating a	

Change	Description	Date
	cluster with an existing EFS file system. • Fix issue causing the `cfn-hup` daemon to fail when it gets restarted. • Consider dynamic nodes with `INVALID_REG` flag as bootstrap failures for Slurm protected mode. Static nodes failing Slurm registrat ion are already treated as bootstrap failures after the `node_replacement_t imeout` . For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	

Change	Description	Date

AWS ParallelCluster version 3.5.1 released	AWS ParallelCluster version 3.5.1 released.
	
	Enhancements:
	
	• Add a stand-alone `pcluster` CLI installer executable.
	
	Changes:
	
	• Upgrade EFA installer to `1.22.0`.
	  • Efa-driver: `efa-2.1.1g` (from `efa-2.1.1-1`)
	  • Efa-config: `efa-config-1.13-1` (from efa-config-1.12-1)
	  • Efa-profile: `efa-profile-1.5-1` (no change)
	  • Libfabric-aws: `libfabric-aws-1.17.0-1` (from `libfabric-aws-1.16.1amzn3.0-1`)
	  • Rdma-core: `rdma-core-43.0-1` (no change)
	  • Open MPI: `openmpi40-aws-4.1.5-1` (from `openmpi40-aws-4.1.4-3`)
	
	Upgrade NICE DCV to version `2022.2-14521` .
		March 29, 2023

Change	Description	Date
	• server: `2022.2.14521-1` • xdcv: `2022.2.519-1` • gl: `2022.2.1012-1` • web_viewer: `2022.2.14521-1` Bug fixes: • Fix potential node launch failures caused by pattern matching between `MountDir` and `/etc/exports` when removing shared Amazon EBS volumes as part of a cluster update. • Fix to prevent `compute_console_output` log file truncation at every `clustermgtd` iteration. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	

Change	Description	Date
AWS ParallelCluster version 3.5.0 released	AWS ParallelCluster version 3.5.0 released. Enhancements: • Access and manage clusters with the AWS ParallelCluster UI. • Add versioned AWS ParallelCluster policies in a CloudFormation template that you can reference in your workloads. • Add a AWS ParallelCluster Python library that you can use with your own code. • Add logging of compute node console output to Amazon CloudWatch on compute node bootstrap failure. • Add failures field containing failure code and reason to `describe-cluster` output when cluster creation fails. • Add validators to prevent malicious string injection while calling the subprocess module. • Fail cluster creation if cluster status changes to `PROTECTED` while provisioning static nodes.	February 20, 2023

Change	Description	Date

	Changes:
	• Upgrade to Slurm version `22.05.8` (from version `22.05.7`)
	• Upgrade EFA installer to `1.21.0`.
	  • Efa-driver: `efa-2.1.1-1` (from `efa-2.1`)
	  • Efa-config: `efa-config-1.12-1` (from efa-config-1.11-1)
	  • Efa-profile: `efa-profile-1.5-1` (no change)
	  • Libfabric-aws: `libfabric-aws-1.16.1amzn3.0-1` (from `libfabric-aws-1.16.1`)
	  • Rdma-core: `rdma-core-43.0-1` (from `rdma-core-43.0-2`)
	  • Open MPI: `openmpi40-aws-4.1.4-3` (no change)
	• Make Slurm controller logs more verbose and enable additional logging for the Slurm power save plugin.
	
	Bug fixes:	

Change	Description	Date
	• Fix cluster database creation by verifying that the cluster name is not longer than 40 characters when Slurm accounting is enabled. • Fix an issue in `clustermgtd` that caused compute nodes, rebooted through Slurm, to be replaced if the Amazon EC2 instance status checks fail. • Fix an issue that prevented compute nodes, with capacity reservations shared by other accounts, from launching because of an incorrect IAM policy on the head node. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, aws-parallelcluster-node, and aws-parallelcluster-ui packages on GitHub.	

Change	Description	Date
AWS ParallelCluster version 3.4.1 released	AWS ParallelCluster version 3.4.1 released. Bug fixes: • Fix a Slurm scheduler issue that could cause the incorrect application of updates to its internal registry of compute nodes. As a result if this issue, EC2 instances could become unavailable or could be backed by an incorrect instance type. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	January 13, 2023

Change	Description	Date
AWS ParallelCluster version 3.4.0 released	AWS ParallelCluster version 3.4.0 released. Enhancements: • Add support for launching nodes across multiple availability zones to increase capacity availabil ity. • Add support for specifying multiple subnets for each queue to increase capacity availability. • Add new configuration parameter in Iam / ResourcePrefix to specify a prefix for path and name of IAM resources created by AWS ParallelCluster. • Add new configuration section Deploymen tSettings / LambdaFun ctionsVpcConfig for specifying the Vpc config used by AWS ParallelCluster Lambda functions. • Add the ability to specify a custom script to run in the head node during a cluster update. The script can be specified with HeadNode / CustomActions	December 22, 2022

Change	Description	Date

	/ OnNodeUpdated when using Slurm as scheduler.
	
	Changes:
	
	• Remove creation of Amazon EFS mount targets for existing file systems.
	• Mount EFS file systems using `amazon-efs-utils`. EFS files systems can be mounted using in-transit encryption and an IAM authorized user.
	• Install stunnel 5.67 on CentOS7 and Ubuntu to support EFS in-transit encryption.
	• Upgrade EFA installer to `1.20.0` (from `1.18.0`).
	  • Efa-driver: `efa-2.1` (from `efa-1.16.0-1`)
	  • Efa-config: `efa-config-1.11-1` (no change)
	  • Efa-profile: `efa-profile-1.5-1` (no change)
	  • Libfabric-aws: `libfabric-aws-1.16.1` (from `libfabric-aws-1.16.0~amzn4.0-1`)	

Change	Description	Date
	• Rdma-core: `rdma-core-43.0-2` from (`rdma-core-41.0-2`) • Open MPI: `openmpi40-aws-4.1.4-3` from (`openmpi40-aws-4.1.4-2`) • Upgrade Slurm to version `22.05.7` (from `22.05.5`). • Upgrade Python to `3.9.16` and `3.7.16`. (from `3.9.15` and `3.7.13`). • With Slurm `22.05.7`, dynamic nodes in `IDLE+CLOUD+COMPLETING+POWER_DOWN+NOT_RESPONDING` status aren't considered unhealthy . For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	

Change	Description	Date
AWS ParallelCluster version 3.3.1 released	AWS ParallelCluster version 3.3.1 released. Changes: • Official AWS ParallelCluster product AMIs are now available after Amazon EC2 deprecation at two years. • Increase memory size of the AWS ParallelCluster API Lambda to 2048 in order to reduce cold start penalties and avoid timeouts. Bug fixes: • Prevent replacement of managed FSx for Lustre file systems and loss of data on cluster updates that include changes to the compute fleet subnet ID. • SharedStorage `DeletionPolicy` applies to cluster update actions. For details of the changes, see the `CHANGELOG` file for the aws-parallelcluster package on GitHub.	December 2, 2022

Change	Description	Date
AWS ParallelCluster documentation only hpc6id note	AWS ParallelCluster documentation-only update • AWS ParallelCluster doesn't support the hpc6id instance type for the HeadNode / InstanceType setting.	December 2, 2022

Change	Description	Date

AWS ParallelCluster version 3.1.5 released	AWS ParallelCluster version 3.1.5 released.
	
	Enhancements:
	
	• Fix Slurm issue that prevents idle nodes termination.
	• Upgrade EFA installer to 1.18.0
	  • Efa-driver: `efa-1.16.0-1`
	  • Efa-config: `efa-config-1.11-1` (from `efa-config-1.9-1`)
	  • Efa-profile: `efa-profile-1.5-1` (no change)
	  • Libfabric-aws: `libfabric-aws-1.16.0~amzn4.0-1` (from `libfabric-1.13.2`).
	  • Rdma-core: `rdma-core-41.0-2` (from `rdma-core-37.0`)
	  • Open MPI: `openmpi40-aws-4.1.4-2` (from `openmpi40-aws-4.1.1-2`)
	
	Changes:
	
	• Add `lambda:ListTags` and `lambda:Un`
		November 16, 2022

Change	Description	Date
	`tagResource` to the `ParallelClusterUserRole` used by the AWS ParallelCluster API stack for a cluster update. • Upgrade Intel MPI Library to Version 2021 Update 6 (from Version 2021 Update 4). For more information, see Intel® MPI Library 2021 Update 6. • Upgrade NVIDIA driver to version 470.141.03 (from 470.103.01). • Upgrade NVIDIA Fabric Manager to version 470.141.03 (from 470.103.01). For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	

Change	Description	Date
AWS ParallelCluster version 3.3.0 released	AWS ParallelCluster version 3.3.0 released. Enhancements: • Add support for multiple instance allocation configuration for a compute resource when using Slurm as a scheduler . For more information, see Multiple instance type allocation with Slurm. • Add support for adding and removing SharedStorage with a cluster update, using an updated configuration. For more information, see Shared storage. • Add new configuration parameter `DeletionPolicy` for Efs and FsxLustre shared storage settings to support storage retention. • Add support for Slurm accounting with new configuration parameter Scheduling / SlurmSettings / Database. For more information, see Slurm accounting with AWS ParallelCluster.	November 2, 2022

Change	Description	Date
	• Add support for On-Demand Capacity Reservations (ODCR) and capacity reservation resource groups. For more information, see Launch instances with On-Demand Capacity Reservations (ODCR). • Add new configuration parameter to specify the IMDS version to support in a cluster or build image infrastructure in the cluster, Imds / ImdsSupport, and build, Imds / ImdsSupport, configurations. • Add support for Networking / PlacementGroup in the SlurmQueues / ComputeResources section. • Add support for instances with multiple network interfaces that are limited to only one ENI per device. • Improve validation of networking for external Amazon EFS file systems by checking the CIDR block in the attached security group. • Add validator to check if configured instance types support placement groups.	

Change	Description	Date
	• Configure NFS threads to be min(256, max(8, num_cores * 4)) to ensure better stability and performance. • Move NFS installation at build time to reduce configuration time. • Enable server-side encryption for the EcrImageBuilder SNS topic that's created when deploying AWS ParallelCluster API and is used to notify on docker image build events. Changes: • Change the behavior of SlurmQueues / Networking / PlacementGroup / `Enabled`. It now creates a unique managed placement group for each compute resource instead of a single managed placement group for all compute resources. • Add support for SlurmQueues / Networking / PlacementGroup / Name as the preferred naming method.	

Change	Description	Date
	• Move head node tags from Launch Template to instance definition to avoid head node replacement on tags updates. • Disable multithreading through script executed by `cloud-init` and not through `CpuOptions` set in the Launch Template. • Upgrade Python to version 3.9 and NodeJS to version 16 in the API infrastructure, API Docker container, and cluster Lambda resources. • Remove support for Python 3.6 in `aws-parallelcluster-batch-cli` . • Upgrade Slurm to version `22.05.5` (from `21.08.8-2`). • Upgrade NVIDIA driver to version `470.141.03` (from `470.129.06`). • Upgrade NVIDIA Fabric Manager to version `470.141.03` (from `470.129.06`). • Upgrade NVIDIA CUDA Toolkit to version 11.7.1 (`from 11.4.4`).	

Change	Description	Date

	• Upgrade Python used in AWS ParallelCluster virtualenvs from `3.7.13` to `3.9.15`.
	• Upgrade EFA installer to version 1.18.0.
	  • Efa-driver: `efa-1.16.0-1` (no change)
	  • Efa-config: `efa-config-1.11-1` (from `efa-config-1.10-1`)
	  • Efa-profile: `efa-profile-1.5-1` (no change)
	  • Libfabric-aws: `libfabric-aws-1.16.0~amzn4.0-1` (from `libfabric-aws-1.16.0~amzn2.0-1`).
	  • Rdma-core: `rdma-core-41.0-2` (from `rdma-core-37.0`)
	  • Open MPI: `openmpi40-aws-4.1.4-2` (from `openmpi40-aws-4.1.1-2`)
	• Upgrade NICE DCV to version `2022.1-13300` (from `2022.0-12760`).
	• Enable suppression of the `SingleSubnetValidator` for `Queues`.	

Change	Description	Date
	• Do not replace `DRAIN` nodes when nodes are in `COMPLETING` state as Epilog may be still running. Bug fixes: • Fix validation of filters parameter in the AWS ParallelCluster `ListClusterLogStreams` command to fail when incorrect filters are passed. • Fix validation of parameter SharedStorage / EfsSettings to fail validation when `FileSystemId` is specified along with other SharedStorage / EfsSettings parameters. Previously, `FileSystemId` wasn't included. • Fix cluster update when changing the order of SharedStorage together with other changes in the configuration. • Fix `UpdateParallelClusterLambdaRole` in the AWS ParallelCluster API to upload logs to CloudWatch. • Fix Cinc not using the local CA certificates bundle	

Change	Description	Date
	when installing packages before any cookbooks are executed. • Fix a hang in upgrading ubuntu with `pcluster build-image` when `Build:UpdateOsPackages:Enabled:true` is set. • Fix parsing of YAML cluster configuration by failing on duplicate keys. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	
AWS ParallelCluster documentation only API reference added.	AWS ParallelCluster documentation-only update • Added the version 3 AWS ParallelCluster API reference to the documentation.	October 27, 2022

Change	Description	Date
AWS ParallelCluster version 3.2.1 released	AWS ParallelCluster version 3.2.1 released. Enhancements: • Improve the logic to associate the host routing tables to the different network cards to better support Amazon EC2 instances with several NICs. Changes: • Upgrade NVIDIA driver to version 470.141.03. • Upgrade NVIDIA Fabric Manager to version 470.141.03. • Disable `cron` job tasks `man-db` and `mlocate`, which may have a negative impact on node performan ce. • Upgrade Intel MPI Library to 2021.6.0.602. • Upgrade Python from 3.7.10 to 3.7.13 in response to this security risk. Bug fixes: • Avoid failing on `DescribeCluster` when	October 3, 2022

Change	Description	Date
	cluster configuration is not available. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	

Change	Description	Date

AWS ParallelCluster version 3.2.0 released	AWS ParallelCluster version 3.2.0 released.
	
	Enhancements:
	
	• Add support for memory-based scheduling in Slurm.
	  • Configure compute nodes real memory in the Slurm cluster configuration.
	  • Add new configuration parameter Scheduling / SlurmSettings / EnableMemoryBasedScheduling to enable memory-based scheduling in Slurm.
	  • Add new configuration parameter Scheduling / SlurmQueues / ComputeResources / SchedulableMemory to override the default value of the memory seen by the scheduler on compute nodes.
	• Improve flexibility on cluster configuration updates to avoid the stop and start of the entire cluster whenever possible. Add new configuration parameter Scheduling / SlurmSettings /
		July 27, 2022

Change	Description	Date
	QueueUpdateStrategy to set the preferred strategy to use when compute nodes need a configuration update and replacement. • Improve fail-over mechanism over available compute resources when hitting insufficient capacity issues with Amazon EC2 instances. Disable compute nodes by a configurable amount of time when a node launch fails due to insufficient capacity. • Add support to mount existing FSx for ONTAP and FSx for OpenZFS file systems. • Add support to mount multiple instances of existing Amazon Elastic File Systems, FSx for Lustre, FSx for ONTAP, and FSx for OpenZFS file systems. • Add support for FSx for Lustre Persistent_2 deployment type when creating a new file system. • Prompt user to enable EFA for supported instance types when using the pcluster configure wizard.	

Change	Description	Date

	• Add support for rebooting compute nodes using Slurm.
	• Improve handling of Slurm power states to also account for manual powering down of nodes.
	• Install NVIDIA GDRCopy 2.3 into the product AMIs to enable low-latency GPU memory copy.
	
	Changes:
	
	• Upgrade EFA installer to version 1.17.2.
	  • EFA driver: `efa-1.16.0-1`
	  • EFA configuration: `efa-config-1.10-1`
	  • EFA profile: `efa-profile-1.5-1`
	  • Libfabric: `libfabric-aws-1.16.0~amzn2.0-1`
	  • RDMA core: `rdma-core-41.0-2`
	  • Open MPI: `openmpi40-aws-4.1.4-2`
	• Upgrade NICE DCV to version 2022.0-12760.
	• Upgrade NVIDIA driver to version 470.129.06.	

Change	Description	Date

	• Upgrade NVIDIA Fabric Manager to version 470.129.06.
	• Change default EBS volume types from gp2 to gp3 in both the root and additiona l volumes.
	• Changes to FSx for Lustre file systems created by AWS ParallelCluster:
	  • Change the default deployment type to `Scratch_2` .
	  • Change the Lustre server version to `2.12`.
	• Doesn't require Placement Group / Enabled to be set to `true` when passing an existing `Placement Group / Id`.
	• Doesn't allow setting `PlacementGroup / Id` when `PlacementGroup / Enabled` is explicitly set to `false`.
	• Add `parallelc luster:cluster-name` tag to all resources created by AWS ParallelC luster.
	• Add `lambda:Li stTags` and `lambda:Un tagResource` to	

Change	Description	Date
	`ParallelClusterUserRole` used by the AWS ParallelCluster API stack for cluster update. • Restrict IPv6 access to `IMDS` to root and cluster admin users only, when configuration parameter `HeadNode` / `Imds` / `Secured` is enabled. • With a custom AMI, use the AMI root volume size instead of the ParallelCluster default of 35 GiB. The value can be changed in cluster configuration file. • Automatic disabling of the compute fleet when the configuration parameter `Scheduling` / `SlurmQueues` / `ComputeResources` / `SpotPrice` is lower than the minimum required Spot request fulfillment price. • Show `requested_value` and `current_value` values in the change set when adding or removing a section during an update. • Disable `aws-ubuntu-eni-helper` service, available in Deep Learning AMIs, to avoid conflicts with `configure_nw_inter`	

Change	Description	Date
	face.sh when configuring instances with multiple network cards. • Remove support for Python 3.6. • Set MTU to 9001 for all the network interfaces when configuring instances with multiple network cards. • Remove the trailing dot when configuring the compute node FQDN. • Manage static nodes in POWERING_DOWN . • Doesn't replace dynamic node in POWER_DOWN as jobs may be still running. • Restart clustermgtd and slurmctld daemons at cluster update time only when Scheduling parameters are updated in the cluster configuration. • Update slurmctld and slurmd systemd service files. • Restrict IPv6 access to IMDS to root and cluster admin users only, when configuration parameter HeadNode / Imds / Secured is enabled. • Set Slurm configuration AuthInfo=cred_expi	

Change	Description	Date

	`re=70` to reduce the time requeued jobs must wait before starting again when nodes are not available.
	• Upgrade third-party cookbook dependencies:
	  • apt-7.4.2 (from apt-7.4.0)
	  • line-4.5.2 (from line-4.0.1)
	  • openssh-2.10.3 (from openssh-2.9.1)
	  • pyenv-3.5.1 (from pyenv-3.4.2)
	  • selinux-6.0.4 (from selinux-3.1.1)
	  • yum-7.4.0 (from yum-6.1.1)
	  • yum-epel-4.5.0 (from yum-epel-4.1.2)
	
	Bug fixes:
	• Fix the default behavior to skip the AWS ParallelCluster validation and test steps when building a custom AMI.
	• Fix file handle leak in `computemgtd` .
	• Fix race condition that was sporadically causing launched instances to be immediately terminate	

Change	Description	Date
	d because they were not yet available in the EC2 DescribeInstances response. • Fix support for the `DisableSimultaneousMultithreading` parameter on instance types with Arm processors. • Fix AWS ParallelCluster API stack update failure when upgrading from a previous version. Add resource pattern used for the `ListImagePipelineImages` Action in the `EcrImageDeletionLambdaRole` . • Fix AWS ParallelCluster API adding missing permissions needed to import or export from Amazon S3 when creating an FSx for Lustre file system. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	

Change	Description	Date
AWS ParallelCluster documentation-only updates this year to date	AWS ParallelCluster documentation-only updates. New sections: • Best practices: budget alerts V3 • Best practices: moving a cluster to a new AWS ParallelCluster minor or patch version V3 • Working with Amazon S3 V3 • Working with Spot Instances V3 • Slurm cluster protected mode V3 • AWS ParallelCluster resources and tagging V3 • Amazon CloudWatch dashboard V3 • Integration with Amazon CloudWatch Logs V3 • Elastic Fabric Adapter V3 • AWS ParallelCluster AMI customization V3 • Launch instances with On-Demand Capacity Reservations (ODCR) V3 • AMI patching and Amazon EC2 instance replacement V3	July 6, 2022

Change	Description	Date
	• How AWS ParallelCluster works V3 • Configuring shared storage encryption with an AWS KMS key V3 • Running jobs in a multiple queue mode cluster V3 • Using the AWS ParallelCluster API V3 Section updates: • Best practices: network performance V3: Added best practices for using Elastic Fabric Adaptor. • AWS Identity and Access Management permissions in AWS ParallelCluster V3: Various updates and added Additional AWS ParallelCluster pcluster user policy when using Amazon FSx for Lustre. • AWS ParallelCluster troubleshooting V3: Various updates.	

Change	Description	Date
AWS ParallelCluster version 3.1.4 released	AWS ParallelCluster version 3.1.4 released. Enhancements: • Add validation for Directory Service / PasswordSecretArn to fail if the secret doesn't exist. Add support for enabling JWT authentication Slurm. Changes: • Upgrade Slurm to version 21.08.8-2. • Build Slurm with JWT support. • Doesn't require Placement Group / Enabled to be set to `true` when passing an existing `Placement Group / Id`. • Add `lambda:TagResource` to `ParallelClusterUserRole` used by ParallelCluster API stack for cluster creation and image creation. Bug fixes:	May 16, 2022

Change	Description	Date
	• Fix the ability to export a cluster's logs when using the `export-cluster-logs` command with the `--filters` option. • Fix AWS Batch Docker entry point to use `/home` shared directory to coordinate Multi-node-Parallel job execution. • Reset node address when setting Slurm unhealthy static node to down to avoid treating static node failed with insufficient capacity as a bootstrap failure node. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	

Change	Description	Date
AWS ParallelCluster version 3.1.3 released	AWS ParallelCluster version 3.1.3 released. Enhancements: • Execute SSH key creation alongside with the creation of HOME directory, for example, during SSH login, when switching to another user and when executing a command as another user. • Add support for both FQDN and LDAP Distinguished Names in the configuration parameter DirectoryService / DomainName. The new validator now checks both the syntaxes. • New `update_directory_service_password.sh` script deployed on the head node supports the manual update of the Active Directory password in the SSSD configuration. The password is retrieved by the AWS Secrets Manager as from the cluster configuration. • Add support to deploy API infrastructure in environments without a default VPC.	April 20, 2022

Change	Description	Date
	Changes: • Disable deeper C-States in x86_64 official AMIs and AMIs created through `build-image` command, to guarantee high performance and low latency. • OS package updates and security fixes. • Change Amazon Linux 2 base images to use AMIs with Kernel 5.10. Bug fixes: • Fix build-image stack in `DELETE_FAILED` after image built successful, due to new EC2 Image Builder policies. • Fix the configuration parameter DirectoryService / DomainAddr conversion to ldap_uri SSSD property when it contains multiples domain addresses. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, and aws-parallelcluster-cookbook packages on GitHub.	

Change	Description	Date
AWS ParallelCluster version 3.1.2 released	AWS ParallelCluster version 3.1.2 released. Changes: • Upgrade Slurm to version `21.08.6` (from `21.08.5`). Bug fixes: • Fix the update of `/etc/hosts` file on compute nodes when a cluster is deployed in subnets without internet access. • Fix compute nodes bootstrap to wait for ephemeral drives initialization before joining the cluster. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster package on GitHub.	March 2, 2022

Change	Description	Date
AWS ParallelCluster version 3.1.1 released	AWS ParallelCluster version 3.1.1 released. • Add support for multiple user cluster environments by integrating with Active Directory (AD) domains managed through AWS Directory Service. • Add support for UseEc2Hostnames in the cluster configuration file. When set to true, use Amazon EC2 default hostnames (e.g. ip-1-2-3-4) for compute nodes. • Add support for cluster creation in subnets with no internet access. • Add support for multiple compute instance types per queue. • Add support for GPU scheduling with Slurm on ARM instances with NVIDIA cards. • Add abbreviated flags for `cluster-name` (`-n`), `region` (`-r`), `image-id` (`-i`) and `cluster-configuration` / `image-configuration` (`-c`) to the AWS ParallelCluster CLI.	February 10, 2022

Change	Description	Date
	• Add support for `NEW_CHANGED_DELETED` option for FSx for Lustre AutoImportPolicy parameter. • Add `parallelcluster:compute-resource-name` tag to EC2 `LaunchTemplates` resources used by compute nodes. • Improve security groups created within the cluster to allow inbound connections from custom security groups when `SecurityGroups` parameters are specified for some head node and/or queues. • Install NVIDIA drivers and CUDA library for ARM. Changes: • Upgrade Slurm to version `21.08.5` (from `20.11.8`). • Upgrade Slurm plugin to version `21.08` (from `20.11`). • Upgrade NICE DCV to version `2021.3-11591` (from `2021.1-10851`).	

Change	Description	Date
	• Upgrade NVIDIA driver to version `470.103.01` (from `470.57.02`). • Upgrade NVIDIA Fabric manager to version `470.103.01` (from `470.57.02`). • Upgrade CUDA to version `11.4.4` (from `11.4.0`). • Intel MPI updated to Version 2021 Update 4 (updated from Version 2019 Update 8). For more information, see Intel® MPI Library 2021 Update 4. • Upgrade PMIx to version `3.2.3` (from `3.1.5`). • Remove dumping of failed compute nodes to `/home/logs/compute` . Compute nodes log files are available in CloudWatch and in Amazon EC2 console logs. • Enable potential to suppress `SlurmQueues` and `ComputeResources` length validators. • Disable package update at instance launch time on Amazon Linux 2. • Disable Amazon EC2 ImageBuilder enhanced	

Change	Description	Date
	image metadata when building AWS ParallelC luster custom images. • Explicitly set `cloud-init` datasource to be EC2. This saves boot time for Ubuntu and CentOS platforms. • Use compute resource name rather than instance type in compute fleet launch template name. • Redirect stderr and stdout to CLI log file to prevent unwanted text in the pcluster CLI output. • Move the configure/ install recipes to separate cookbooks that are called from the main one. Existing entrypoints are maintained and backwards compatible. • Download dependencies of Intel HPC platform during AMI build time to avoid contacting internet during cluster creation time. • Do not strip `-` from compute resource name when configuring Slurm nodes. • Do not configure GPUs in Slurm when NVIDIA driver is not installed.	

Change	Description	Date
	• Fix `ecs:ListContainerInstances` permission in `BatchUserRole`. • Fix exporting of cluster logs when there is no prefix specified, previously exported to a `None` prefix. • Fix rollback not being performed in case of cluster update failure. • Fix `ecs:ListContainerInstances` permission in `BatchUserRole`. • Fix `RootVolume` schema for the `HeadNode` by raising an error if an unsupported `KmsKeyId` is specified. • Fix Amazon FSx missing metrics to be displayed in CloudWatch Dashboard. • Fix `EfaSecurityGroupValidator`. Previously, it had potential to produce false failures when custom security groups were provided and EFA was enabled. For details of the changes, see the `CHANGELOG` files for the	

Change	Description	Date
	aws-parallelcluster, aws-parallelcluster-cookbook, and aws-parallelcluster-node packages on GitHub.	
AWS ParallelCluster version 3.0.3 released	AWS ParallelCluster version 3.0.3 released. • Disable `log4j-cve-2021-44228-hotpatch` agent (`Log4jHotPatch`) on Amazon Linux 2 to avoid potential performance degradation. For more information, see Amazon Linux Hotpatch Announcement for Apache Log4j. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster and aws-parallelcluster-cookbook packages on GitHub.	January 17, 2022

Change	Description	Date
AWS ParallelCluster version 3.0.2 released	AWS ParallelCluster version 3.0.2 released. Upgrade Elastic Fabric Adapter installer to 1.14.1 • EFA config: `efa-config-1.9-1` (from `efa-config-1.9`) • EFA profile: `efa-profile-1.5-1` (from `efa-profile-1.5`) • EFA Kernel module: `efa-1.14.2` (from `efa-1.13.0`) • RDMA core: `rdma-core-37.0` (from `rdma-core-35`) • Libfabric: `libfabric-1.13.2` (from `libfabric-1.13.0`) • Open MPI: `openmpi40-aws-4.1.1-2` (no change) GPUDirect RDMA is always enabled if supported by the instance type. The GdrSupport configuration option has no effect. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-paral	November 5, 2021

Change	Description	Date
	lelcluster-cookbook and aws-parallelcluster-node packages on GitHub.	

Change	Description	Date
AWS ParallelCluster version 3.0.1 released	AWS ParallelCluster version 3.0.1 released. Cluster configuration migration tool • Customers can now migrate their cluster configurations from the AWS ParallelCluster version 2 format to the YAML-based AWS ParallelCluster version 3 format. For more information, see pcluster3-config-converter. Head node can be stopped • After stopping the compute fleet, the head node can be stopped and later restarted using the Amazon EC2 console or the stop-instances AWS CLI command. Default AWS Region read from `~/.aws/config` file • For the pcluster command, if the AWS Region is not specified in the configuration file, in the environment, or on the command line, the default AWS Region specified in the `region`	October 27, 2021

Change	Description	Date
	setting in the `[default]` section of the `~/.aws/config` file is used. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster, aws-parallelcluster-cookbook and aws-parallelcluster-node packages on GitHub.	

Change	Description	Date
AWS ParallelCluster version 3.0.0 released	AWS ParallelCluster version 3.0.0 released. Support for cluster management via Amazon API Gateway • Customers can now manage and deploy clusters through HTTP endpoints with Amazon API Gateway. This opens up new possibili ties for scripted or event-dri ven workflows. The AWS ParallelCluster command line interface (CLI) has also been redesigned for compatibi lity with this API and includes a new JSON output option. This new functionality makes it possible for customers to implement similar building block capabilities using the CLI as well. Improved custom AMI creation • Customers now have access to a more robust process for creating and managing custom AMIs using EC2	September 10, 2021

Change	Description	Date
	Image Builder. Custom AMIs can now be managed through a separate AWS ParallelCluster configuration file, and can be created using the pcluster build-image command in the AWS ParallelCluster command line interface. For details of the changes, see the CHANGELOG files for the aws-parallelcluster, aws-parallelcluster-cookbook and aws-parallelcluster-node packages on GitHub.	

PCUI

Change	Description	Date
PCUI version 2025.04.0 released	PCUI version 2025.04.0 released Breaking changes: • Remove default value for the PC version. Now the user must specify the PC version to use. Features: • Add new stack parameter 'AdditionalPolicies	April 16, 2025

Change	Description	Date
	`PCAPI '` to add custom permissions for the ParallelCluster API Lambda role, in addition to the default ones. Bug fixes: • Fix PCUI deployment in private subnets by making the PCUI template use and return the correct URLs. • Fix an issue that prevents the loading of 200+ jobs in the Job status Tab. (See https://github.com/aws/aws-parallelcluster-ui/issues/376). Security: • Upgrade Python from 3.9 to 3.12. • Upgrade cross-spawn from 7.0.3 to 7.0.6 to address vulnerability CVE-2024-21538. • Upgrade requests from 2.31.0 to 2.32.0 to address CVE-2024-35195. • Upgrade urllib3 from 1.26.18 to 1.26.19 to address CVE-2024-37891.	

Change	Description	Date
	• Upgrade cryptography from 42.0.4 to 44.0.1 to address CWE-1395. • Upgrade certifi from 2023.7.22 to 2024.7.4 to address CVE-2024-39689. • Upgrade jinja2 from 3.1.3 to 3.1.6 to address CVE-2024-56201 and CVE-2024-56326. • Upgrade serverless_wsgi.py to version 3.0.5. • Upgrade Werkzeug from 2.3.8 to version 3.0.6 to address CVE-2024-34069, CVE-2024-49766 and CVE-2024-49767. • Upgrade Axios from 1.6.7 to version 1.8.2 to address CVE-2024-39338. • Upgrade Next.js from 14.1.1 to version 14.2.25 to address CVE-2024-51479, CVE-2024-46982 and CVE-2025-29927. • Upgrade idna from 3.4 to version 3.7 to address CVE-2024-3651. • Upgrade nanoid from 3.3.7 to version 3.3.8 to address CVE-2024-55565.	

Change	Description	Date
	• Upgrade python-jose from 3.3.0 to version 3.4.0 to address CVE-2022-29217.	
PCUI version 2024.11.0 released	PCUI version 2024.11.0 released Bug fixes: • Explicitly set the policy for an ECR private repository to prevent policy removal on a stack update impacting a Lambda function. The policy includes the permissions required by the Lambda function to fetch the code.	November 22, 2024

Change	Description	Date
PCUI version 2024.10.0 released	PCUI version 2024.10.0 released Changes: • Add support for AWS ParallelCluster 3.11.1. • Add support for On-Demand Capacity Reservations and Capacity Block in the wizard. • Add g6, m7 and p5 families to the list of supported instance types in the wizard. • Add new stack optional parameters to configure custom domain for both PCUI and Cognito. Bug fixes: • Fixes a bug that was breaking the setup of custom domain. Security: • Upgrade Flask-CORS from 3.0.10 to 4.0.2 to address vulnerability CVE-2024-6221. • Upgrade lint-staged from 13.0.3 to 15.2.5 to address	October 22, 2024

Change	Description	Date
	vulnerability CVE-2024-4068. • Full Changelog	
PCUI version 2024.05.0 released	PCUI version 2024.05.0 released. Bug Fixes: • Fixed a bug in the frontend blocking the UI when the user opens the Job Status panel. • Full Changelog	May 14, 2024
PCUI version 2024.04.0 released	PCUI version 2024.04.0 released. Features: • Added support for AWS ParallelCluster version 3.9.1 • Full Changelog	April 17, 2024

Change	Description	Date
PCUI version 2024.03.0 released	PCUI version 2024.03.0 released. Features: • Added support for AWS ParallelCluster version 3.9.0 • Added support for Ubuntu 22.04 and Red Hat Enterprise Linux 9 • Deprecated Ubuntu 18.04 Bug fixes • Fixed issue causing some clusters to not appear when using many clusters For details of the changes, see the `CHANGELOG` files for the `aws-parallelcluster-ui` package on GitHub.	March 12, 2024

Change	Description	Date
PCUI version 2024.02.0 released	PCUI version 2024.02.0 released Changes: • Updated the Lambda runtime environment to Python v3.9 For details of the changes, see the `CHANGELOG` files for the `aws-parallelcluster-ui` package on GitHub.	February 8, 2024

Change	Description	Date
PCUI version 2023.12.0 released	PCUI version 2023.12.0 released. Features: • Added support for PCUI deployment with private networking. • Added possibility to optionally apply a Permissions Boundary to every IAM role created by the PCUI and PCAPI infrastructures • Added possibility to optionally apply a prefix to every IAM role and policy created by the PCUI and PCAPI infrastructure. • Added support for ParallelCluster version 3.8.0, without feature parity in the wizard. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster-ui package on GitHub.	December 21, 2023

Change	Description	Date
PCUI version 2023.10.0 released	PCUI version 2023.10.0 released. Features: • Added support for ParallelCluster 3.7.2 with feature parity in the wizard limited to FSx File Cache and memory based scheduling compatibility with multiple instance types. Bug fixes: • Fixed issue causing UI errors when PCUI does not have permissions to interact with Cost Explorer. Improvements • Improved security by reducing the access token TTL from 10 minutes to 5 minutes. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster-ui package on GitHub.	October 20, 2023

Change	Description	Date
PCUI version 2023.06.0 released	PCUI version 2023.06.0 released. Changes: • Upgraded the default AWS ParallelCluster API version to 3.6.0. Bug fixes: • Fixed broken deployment for AWS GovCloud (US-West) Region. • Split panel now correctly loads cluster details after creation has started. Notes: • The Cost Monitoring feature is not available in AWS GovCloud (US) Regions. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster-ui package on GitHub.	June 7, 2023

Change	Description	Date
PCUI version 2023.05.0 released	PCUI version 2023.05.0 released. Enhancements: • Starting with AWS ParallelCluster version 3.6.0, add support for RHEL 8. • Add cluster cost monitoring. • Starting with AWS ParallelCluster version 3.6.0, increase queue and compute resource quotas. Changes: • Improved the cluster creation wizard user interface. • Increased the speed of PCUI deployment. • Improved the interface for adding a new user. • Queues are in the head node subnet by default. Bug fixes: • Switch to the correct region after cluster creation completes.	May 16, 2023

Change	Description	Date
	• Fix the loading indicator display in the "Edit cluster" feature. • Fix cluster creation when the EBS SnapshotId property is removed. For details of the changes, see the `CHANGELOG` files for the aws-parallelcluster-ui package on GitHub.	

Change	Description	Date
PCUI version 2023.04.0 released	PCUI version 2023.04.0 released. Enhancements: • Cluster create wizard re-design. • Cluster logs page re-design. • Add custom name setting for shared storage. • Add multiple storage selection when adding storage to a cluster. • Add `DeletionPolicy` support for Amazon EFS and FSx for Lustre. • Add `ImdsSupport` setting in cluster configuration. • Add support for C7 instance types. • Added tutorial Reverting to a previous AWS Systems Manager document version. Changes: • Cluster configuration YAML up to 1MB in size. • User isn't logged out due to an authorization with Boto3 IAM temporary credentials.	April 17, 2023

Change	Description	Date
	• Disabled multi-threading options when an HPC instance is selected. • Removed disable rollback on cluster create page. • User is prevented from using the PCUI until the required information is provided. • Up to 10 queues can be added. • The `SSM-SessionManagerRunShell` document is not overwritten during PCUI installation. Bug fixes: • Fix broken reset password link. • Fix broken `delete stack` caused by `EcrPrivateRepository` not being empty • Fixed initialization issue of the Generate SSH Keys check-box in Multiple user management properties section. • Fixed crash caused be a job with undefined properties. • Fixed SCRATCH FSx settings.	

Change	Description	Date
	• Fixed Start and Stop instances button, still enabled after being clicked once. For details of the changes, see the CHANGELOG files for the aws-parallelcluster-ui package on GitHub.	

Terraform

Change	Description	Date
Terraform Provider for AWS ParallelCluster 1.1.0 released	Bug fixes: • Fixed an issue that was causing terraform-apply failure when ParallelCluster API 3.11.x is used to deploy clusters with login nodes.	December 6, 2024
Terraform Module for AWS ParallelCluster 1.1.0 released	Changes: • Use AWS ParallelCluster Terraform Provider 1.x in all module examples. • Use ParallelCluster API 3.11.1 in all examples with stack name ParallelClusterAPI. • Deploy login nodes in all module examples.	December 6, 2024

Change	Description	Date
Terraform Provider for AWS ParallelCluster 1.0.0 released	Features: • Full changelog	June 26, 2024
Terraform Module for AWS ParallelCluster 1.0.0 released	Features: • Full changelog	June 26, 2024