AWS Livro branco

AWS Outposts Considerações sobre design e arquitetura de alta disponibilidade

AWS Outposts Considerações sobre design e arquitetura de alta disponibilidade: AWS Livro branco

Table of Contents

AWS Outposts Considerações sobre design e arquitetura de alta disponibilidade

Data de publicação: 12 de agosto de 2021 (Histórico do documento)

Este whitepaper discute considerações de arquitetura e práticas recomendadas que os gerentes de TI e arquitetos de sistemas podem aplicar para criar ambientes de aplicativos locais altamente disponíveis. AWS Outposts

Você é Well-Architected?

O AWS Well-Architected Framework ajuda você a entender os prós e os contras das decisões que você toma ao criar sistemas na nuvem. Os seis pilares do Framework permitem que você aprenda as práticas recomendadas arquitetônicas para projetar e operar sistemas confiáveis, seguros, eficientes, economicamente viáveis e sustentáveis. Usando o AWS Well-Architected Tool, disponível gratuitamente no AWS Management Console, você pode analisar suas workloads em relação a essas práticas recomendadas respondendo a um conjunto de perguntas para cada pilar.

Para obter orientações especializadas e práticas recomendadas adicionais para a arquitetura de sua nuvem (implantações de arquitetura de referência, diagramas e whitepapers), consulte o Centro de Arquitetura da AWS.

Introdução

Este paper é destinado a gerentes de TI e arquitetos de sistemas que desejam implantar, migrar e operar aplicativos usando a plataforma de AWS nuvem e executá-los localmente com AWS Outposts rack, o formato de rack de 42U da. AWS Outposts

Ele apresenta os padrões de arquitetura, os antipadrões e as práticas recomendadas para criar sistemas altamente disponíveis que incluem AWS Outposts rack. Você aprenderá como gerenciar sua capacidade de AWS Outposts rack e usar serviços de rede e instalações de data center para configurar soluções de infraestrutura de AWS Outposts rack altamente disponíveis.

AWS Outposts O rack é um serviço totalmente gerenciado que fornece um pool lógico de recursos de computação, armazenamento e rede em nuvem. Com os racks do Outposts, os clientes podem

usar serviços gerenciados da AWS compatíveis em seus ambientes on-premises, incluindo: Amazon Elastic Compute Cloud (Amazon EC2), Amazon Elastic Block Store (Amazon EBS), Amazon S3 no Outposts, Amazon Elastic Kubernetes Service (Amazon EKS), Amazon Elastic Container Service (Amazon ECS), Amazon Relational Database Service (Amazon RDS) e outros serviços da AWS no Outposts. Os serviços no Outposts são fornecidos no mesmo AWS Nitro System usado nas Regiões da AWS.

Ao aproveitar o AWS Outposts rack, você pode criar, gerenciar e escalar aplicativos locais altamente disponíveis usando serviços e ferramentas de AWS nuvem familiares. AWS Outposts O rack é ideal para cargas de trabalho que exigem acesso de baixa latência a sistemas locais, processamento de dados local, residência de dados e migração de aplicativos com interdependências do sistema local.

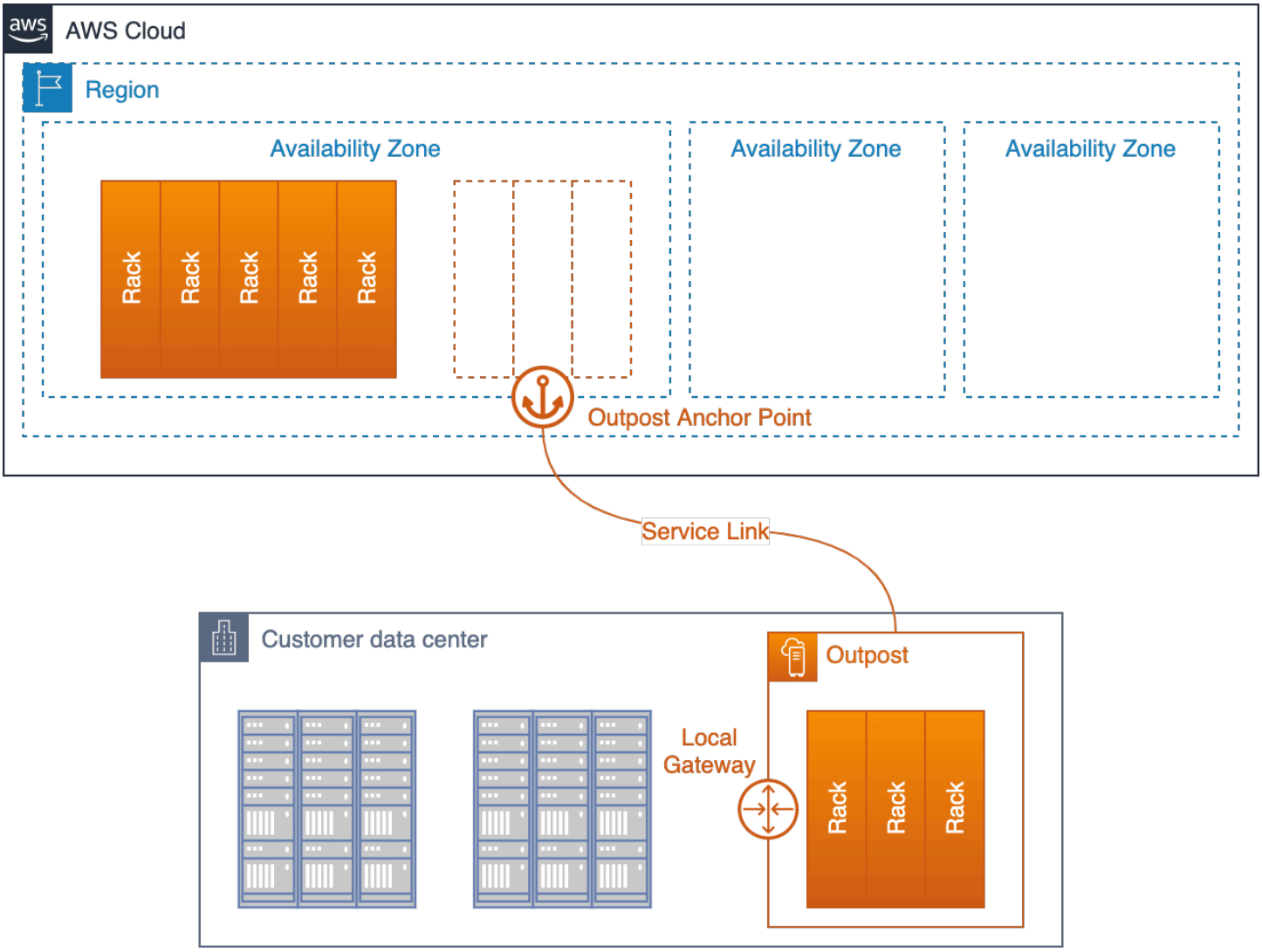
Estendendo a AWS infraestrutura e os serviços para locais locais

O AWS Outposts serviço fornece AWS infraestrutura e serviços para locais locais em mais de 50 países e territórios, oferecendo aos clientes a capacidade de implantar a mesma AWS infraestrutura, AWS serviços, APIs e ferramentas em praticamente qualquer data center, espaço de co-localização ou instalação local para uma experiência híbrida verdadeiramente consistente. Para entender como projetar com o Outposts, você deve entender os diferentes níveis que compõem a nuvem. AWS

Uma Região da AWS é uma área geográfica do mundo. Cada um Região da AWS é uma coleção de data centers agrupados logicamente em zonas de disponibilidade (AZs). Regiões da AWS forneça várias (pelo menos duas) zonas de disponibilidade fisicamente separadas e isoladas, conectadas com baixa latência, alta taxa de transferência e conectividade de rede redundante. Cada AZ consiste em um ou mais datacenters físicos.

Um Posto Avançado lógico (doravante denominado Posto Avançado) é uma implantação de um ou mais AWS Outposts racks conectados fisicamente gerenciados como uma única entidade. Um Outpost fornece um pool de capacidade de AWS computação e armazenamento em um de seus sites como uma extensão privada de um AZ em um. Região da AWS

Talvez o melhor modelo conceitual AWS Outposts seja pensar em desconectar um ou mais racks de um data center em uma AZ de um. Região da AWS Você transfere os racks do datacenter da AZ para o seu datacenter. Em seguida, você conecta os racks aos pontos de ancoragem no datacenter da AZ com um cabo (muito) longo para que os racks continuem funcionando como parte da Região da AWS. Você também os conecta à sua rede local para fornecer conectividade de baixa latência entre suas redes on-premises e as cargas de trabalho em execução nesses racks.



Um Outpost implantado em um datacenter de clientes e conectado de volta à sua AZ âncora e região pai

O Outpost funciona como uma extensão do AZ, onde está ancorado. AWS opera, monitora e gerencia a AWS Outposts infraestrutura como parte do Região da AWS. Em vez de um cabo físico muito longo, um Outpost se conecta de volta à sua região principal por meio de um conjunto de túneis VPN criptografados chamados de link de serviço.

O link de serviço termina em um conjunto de pontos de ancoragem em uma zona de disponibilidade (AZ) na região principal do Outpost.

Você escolhe onde seu conteúdo é armazenado. Você pode replicar e fazer backup do seu conteúdo no Região da AWS ou em outros locais. Seu conteúdo não será movido ou copiado para fora dos locais escolhidos sem o seu consentimento, exceto conforme o necessário para cumprir a lei ou uma

exigência de um órgão governamental. Para obter mais informações, consulte Perguntas frequentes sobre a privacidade de dados da AWS.

As workloads que você implanta nesses racks são executadas localmente. E, embora a capacidade de computação e armazenamento disponível nesses racks seja finita e não possa acomodar a execução dos serviços em escala de nuvem de um Região da AWS, os recursos implantados no rack (suas instâncias e seu armazenamento local) recebem os benefícios de serem executados localmente enquanto o plano de gerenciamento continua operando no. Região da AWS

Para implantar workloads em um Outpost, você adiciona sub-redes aos seus ambientes de Virtual Private Cloud (VPC) e especifica um Outpost como o local das sub-redes. Em seguida, você seleciona a sub-rede desejada ao implantar AWS recursos compatíveis por meio de AWS Management Console CLI, APIs, CDK ou ferramentas de infraestrutura como código (IaC). As instâncias nas sub-redes do Outpost se comunicam com outras instâncias no Outpost ou na região por meio da rede VPC.

O Outpost Service Link transporta tanto o tráfego de gerenciamento do Outpost quanto o tráfego VPC do cliente (tráfego VPC entre as sub-redes no Outpost e as sub-redes na região).

Termos importantes:

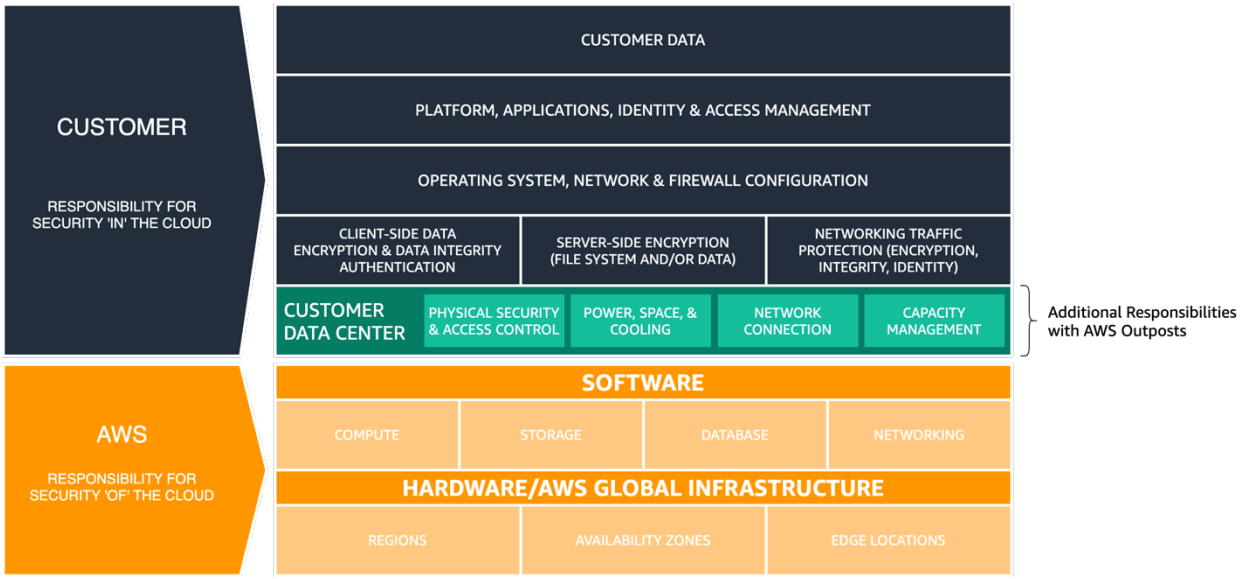
- AWS Outposts— é um serviço totalmente gerenciado que oferece a mesma AWS infraestrutura, AWS serviços, APIs e ferramentas para praticamente qualquer data center, espaço de co-localização ou instalação local para uma experiência híbrida verdadeiramente consistente.
- Outpost — é uma implantação de um ou mais AWS Outposts racks fisicamente conectados que é gerenciada como uma única entidade lógica e um pool de AWS computação, armazenamento e rede implantado no local do cliente.
- Região principal — a Região da AWS que fornece gerenciamento, serviços de plano de controle e AWS serviços regionais para a implantação de um Posto Avançado.
- Zona de disponibilidade âncora (âncora AZ): a zona de disponibilidade na região principal que hospeda os pontos de ancoragem de um Outpost. Um Outpost funciona como uma extensão de sua Zona de Disponibilidade âncora.
- Pontos de ancoragem: terminais no AZ âncora que recebem as conexões de Outposts implantados remotamente.
- Link de serviço: um conjunto de túneis VPN criptografados que conectam um Outpost à sua zona de disponibilidade âncora em sua região principal.

- Gateway local (LGW): um roteador virtual de interconexão lógica que permite a comunicação entre seu Outpost e sua rede on-premises.

Entendendo o modelo de responsabilidade compartilhada atualizado

Ao implantar a AWS Outposts infraestrutura em seus data centers ou instalações de co-localização, você assume responsabilidades adicionais no modelo de Responsabilidade AWS Compartilhada. Por exemplo, na região, a AWS fornece diversas fontes de energia, rede central redundante e conectividade resiliente de rede de longa distância (WAN) para garantir que os serviços estejam disponíveis no caso de falhas de um ou mais componentes.

Com o Outposts, você é responsável por fornecer energia resiliente e conectividade de rede aos racks do Outpost para atender aos seus requisitos de disponibilidade para workloads executados nos Outposts.



AWS Modelo de responsabilidade compartilhada atualizado para AWS Outposts

Com AWS Outposts, você é responsável pela segurança física e pelos controles de acesso do ambiente do data center. Você deve fornecer energia, espaço e resfriamento suficientes para manter o Outpost operacional e as conexões de rede para conectar o Outpost à Região.

Como a capacidade do Outpost é finita e determinada pelo tamanho e pelo número de AWS instalações de racks em seu local, você deve decidir quanta capacidade de EC2, EBS e S3 no Outposts você precisa para executar suas cargas de trabalho iniciais, acomodar o crescimento futuro e fornecer capacidade extra para mitigar falhas no servidor e eventos de manutenção.

AWS é responsável pela disponibilidade da infraestrutura do Outposts, incluindo as fontes de alimentação, servidores e equipamentos de rede dentro dos AWS Outposts racks. AWS também gerencia o hipervisor de virtualização, os sistemas de armazenamento e os AWS serviços executados no Outposts.

Uma prateleira de alimentação central em cada rack do Outposts é convertida de AC para CC e fornece energia aos servidores no rack por meio de uma arquitetura de barramento. Com a arquitetura de barramento, metade das fontes de alimentação no rack podem falhar e todos os servidores continuarão funcionando sem interrupções.

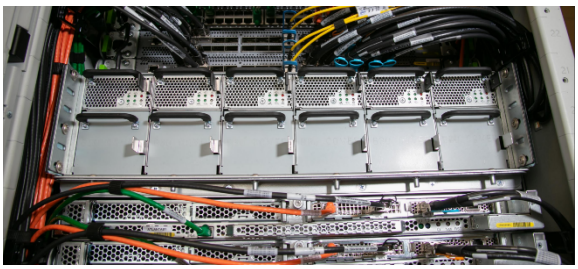


Figura 3 - Fontes de alimentação AWS Outposts de CA para CC e distribuição de energia por barramento

Os comutadores de rede e o cabeamento dentro e entre os racks do Outposts também são totalmente redundantes. Um patch panel de fibra fornece conectividade entre um rack Outpost e a rede local e serve como ponto de demarcação entre o ambiente de data center gerenciado pelo cliente e o ambiente gerenciado. AWS Outposts

Assim como na região, AWS é responsável pelos serviços de nuvem oferecidos nos Outposts e assume responsabilidades adicionais à medida que você seleciona e implementa serviços gerenciados de alto nível, como o Amazon RDS on Outposts. Você deve revisar o Modelo de Responsabilidade compartilhada da AWS e as páginas de Perguntas Frequentes (FAQ) de serviços individuais ao considerar e selecionar os serviços a serem implantados no Outposts. Esses recursos fornecem detalhes adicionais sobre a divisão de responsabilidades entre você AWS e.

Pensando em termos de modos de falha

Ao projetar um aplicativo ou sistema altamente disponível, você deve considerar quais componentes podem falhar, qual impacto as falhas de componentes terão no sistema e quais mecanismos você pode implementar para mitigar ou eliminar o impacto das falhas de componentes. Seu aplicativo é executado em um único servidor, em um único rack ou em um único datacenter? O que acontecerá quando um servidor, rack ou datacenter sofrer uma falha temporária ou permanente? O que acontece quando há uma falha em um subsistema crítico, como rede ou no próprio aplicativo? Esses são modos de falha.

Você deve considerar os modos de falha nesta seção ao planejar seus Outposts e implantações de aplicativos. As seções a seguir analisarão como mitigar esses modos de falha para fornecer um maior nível de alta disponibilidade para seu ambiente de aplicativos.

Modo de falha 1: Rede

A implantação de um Outpost depende de uma conexão resiliente com sua região principal para gerenciamento e monitoramento. As interrupções na rede podem ser causadas por uma variedade de falhas, como erros do operador, falhas no equipamento e interrupções no provedor de serviços. Um Outpost, que pode ser composto por um ou mais racks conectados entre si no local, é considerado desconectado quando não consegue se comunicar com a Região por meio do Link de Serviço.

Caminhos de rede redundantes podem ajudar a reduzir o risco de eventos de desconexão. Você deve mapear as dependências do aplicativo e o tráfego de rede para entender o impacto que os eventos de desconexão terão nas operações da carga de trabalho. Planeje redundância de rede suficiente para atender aos requisitos de disponibilidade de seus aplicativos.

Durante um evento de desconexão, as instâncias executadas em um Outpost continuam em execução e podem ser acessadas a partir de redes locais por meio do Gateway local do Outpost (LGW). As cargas de trabalho e os serviços locais podem ser prejudicados ou falhar se dependerem de serviços na região. Solicitações mutantes (como iniciar ou interromper instâncias no Posto Avançado), operações do plano de controle e telemetria de serviço (por exemplo, CloudWatch métricas) falharão enquanto o Posto Avançado estiver desconectado da Região.

Modo de falha 2: Instâncias

As instâncias do EC2 podem ficar danificadas ou falhar se o servidor em que estão sendo executadas tiver um problema ou se a instância apresentar uma falha no sistema operacional ou no aplicativo. A forma como os aplicativos lidam com esses tipos de falhas depende da arquitetura do aplicativo. Os aplicativos monolíticos geralmente usam recursos do aplicativo ou do sistema para recuperação, enquanto as arquiteturas modulares orientadas a serviços ou de microsserviços geralmente substituem os componentes com falha para manter a disponibilidade do serviço.

Você pode substituir instâncias com falha por novas instâncias usando mecanismos automatizados, como grupos do EC2 Auto Scaling. A recuperação automática de instâncias pode reiniciar instâncias que falham devido a falhas no servidor, desde que haja capacidade disponível suficiente nos servidores restantes.

Modo de falha 3: Computação

Os servidores podem falhar ou ficar danificados e podem precisar ser retirados de operação (temporária ou permanentemente) por vários motivos, como falhas de componentes e operações de manutenção programadas. A forma como os serviços no rack Outposts lidam com falhas e deficiências do servidor varia e pode depender de como os clientes configuram as opções de alta disponibilidade.

Você deve solicitar capacidade computacional suficiente para suportar um modelo de disponibilidade `N+M`, onde `N` é a capacidade necessária e `M` a capacidade ociosa provisionada para acomodar falhas no servidor.

As substituições de hardware para servidores com falha são fornecidas como parte do serviço de AWS Outposts rack totalmente gerenciado. AWS monitora ativamente a integridade de todos os servidores e dispositivos de rede em uma implantação do Outpost. Se houver necessidade de realizar manutenção física, a AWS agendará um horário para visitar seu site para substituir os componentes com defeito. O provisionamento de capacidade não utilizada permite que você mantenha suas cargas de trabalho em execução enquanto os servidores com falha são retirados de serviço e substituídos.

Modo de falha 4: racks ou datacenters

As falhas nos racks podem ocorrer devido à perda total de energia dos racks ou devido a falhas ambientais, como perda de resfriamento ou danos físicos ao datacenter causados por uma

inundação ou terremoto. Deficiências nas arquiteturas de distribuição de energia do datacenter ou erros durante a manutenção padrão da energia do datacenter podem resultar na perda de energia de um ou mais racks ou até mesmo de todo o datacenter.

Esses cenários podem ser mitigados com a implantação de infraestrutura em vários andares ou locais de datacenter que sejam independentes uns dos outros dentro do mesmo campus ou área metropolitana.

Adotar essa abordagem com o AWS Outposts rack exigirá uma análise cuidadosa de como os aplicativos são arquitetados e distribuídos para serem executados em vários Outposts lógicos separados para manter a disponibilidade dos aplicativos.

Modo de falha 5: zona ou região de AWS disponibilidade

Cada Outpost está ancorado em uma Zona de Disponibilidade (AZ) específica dentro de uma Região da AWS. Falhas na AZ âncora ou na região-mãe podem causar a perda do gerenciamento e da mutabilidade do Outpost e podem interromper a comunicação de rede entre o Outpost e a Região.

Semelhantes às falhas de rede, as falhas na AZ ou na região podem fazer o Outpost ser desconectado da região. As instâncias executadas em um Outpost continuam em execução e são acessíveis a partir de redes locais por meio do Outpost Local Gateway (LGW) e podem ser prejudicadas ou falhar se dependerem de serviços na Região, conforme descrito anteriormente.

Para mitigar o impacto das falhas de AWS AZ e região, você pode implantar vários Outposts, cada um ancorado em uma AZ ou região diferente. Em seguida, você pode projetar sua carga de trabalho para operar em um modelo distribuído de implantação Multi-Outpost usando muitos dos mecanismos e padrões de arquitetura semelhantes que você usa para projetar e implantar atualmente. AWS

Criação de aplicativos de alta disponibilidade e soluções de infraestrutura com AWS Outposts rack

Com o AWS Outposts rack, você pode criar, gerenciar e escalar aplicativos locais altamente disponíveis usando serviços e ferramentas de AWS nuvem familiares. É importante entender que as arquiteturas e abordagens de HA na nuvem geralmente são diferentes das arquiteturas tradicionais de HA on-premises que você pode estar executando em seu datacenter atualmente.

Com as implantações tradicionais de aplicativos de HA on-premises, os aplicativos são implantados em máquinas virtuais (VMs). Sistemas e infraestrutura de TI complexos são implantados e mantidos para garantir que essas máquinas virtuais estejam funcionando e íntegras. As VMs geralmente têm identidades específicas e cada VM pode desempenhar um papel fundamental na arquitetura total do aplicativo.

As funções arquitetônicas estão fortemente vinculadas às identidades da VM. Os arquitetos de sistemas utilizam os recursos de infraestrutura de TI para fornecer ambientes de execução de VM altamente disponíveis que fornecem a cada VM acesso confiável à capacidade computacional, volumes de armazenamento e serviços de rede. Se uma VM falhar, os processos de recuperação automatizados ou manuais são executados para restaurar a VM com falha para um estado íntegro, geralmente em outra infraestrutura ou em outro datacenter.

As arquiteturas de HA em nuvem adotam uma abordagem diferente. AWS os serviços em nuvem fornecem recursos confiáveis de computação, armazenamento e rede. Os componentes do aplicativo são implantados em instâncias do EC2, contêineres, funções sem servidores ou outros serviços gerenciados.

Uma instância é uma instanciação de um componente do aplicativo, talvez uma das muitas que desempenham essa função. Os componentes do aplicativo estão fracamente acoplados entre si e ao papel que desempenham na arquitetura total do aplicativo. A identidade individual de uma instância geralmente não é importante. Instâncias adicionais podem ser criadas ou destruídas para aumentar ou diminuir a escala em resposta à demanda. Instâncias com falha ou instâncias não íntegras são simplesmente substituídas por novas instâncias íntegras.

AWS Outposts O rack é um serviço totalmente gerenciado que estende AWS computação, armazenamento, rede, banco de dados e outros serviços de nuvem para locais locais para uma experiência híbrida verdadeiramente consistente. Você não deve pensar no serviço de rack do Outposts como um substituto imediato para sistemas de infraestrutura de TI com mecanismos

tradicionais de HA on-premises. Tentar usar AWS serviços e Outposts para dar suporte a uma arquitetura tradicional de HA local é um antipadrão.

As cargas de trabalho executadas no AWS Outposts rack usam mecanismos de HA na nuvem, como o Amazon EC2 Auto Scaling (para escalar horizontalmente para atender às demandas da carga de trabalho), verificações de integridade do EC2 (para detectar e remover instâncias não íntegras) e Application Load Balancers (para redirecionar o tráfego de carga de trabalho de entrada para instâncias escaladas ou substituídas). Ao migrar aplicativos para a nuvem, seja para um AWS Outposts rack Região da AWS ou para um rack, você deve atualizar sua arquitetura de aplicativos de HA para começar a aproveitar os serviços de nuvem gerenciados e os mecanismos de HA na nuvem.

As seções a seguir apresentam padrões de arquitetura, antipadrões e práticas recomendadas para implantar o rack do AWS Outposts em seus ambientes on-premises para executar workloads com requisitos de alta disponibilidade. Essas seções apresentam padrões e práticas; no entanto, elas não fornecem detalhes de configuração e implementação. Você deve ler e se familiarizar com as perguntas frequentes do do rack do AWS Outposts e o guia do usuário e as perguntas frequentes e a documentação dos serviços executados no rack do Outposts ao preparar seu ambiente para o rack dp Outposts e seus aplicativos para migração para serviços da AWS .

Tópicos

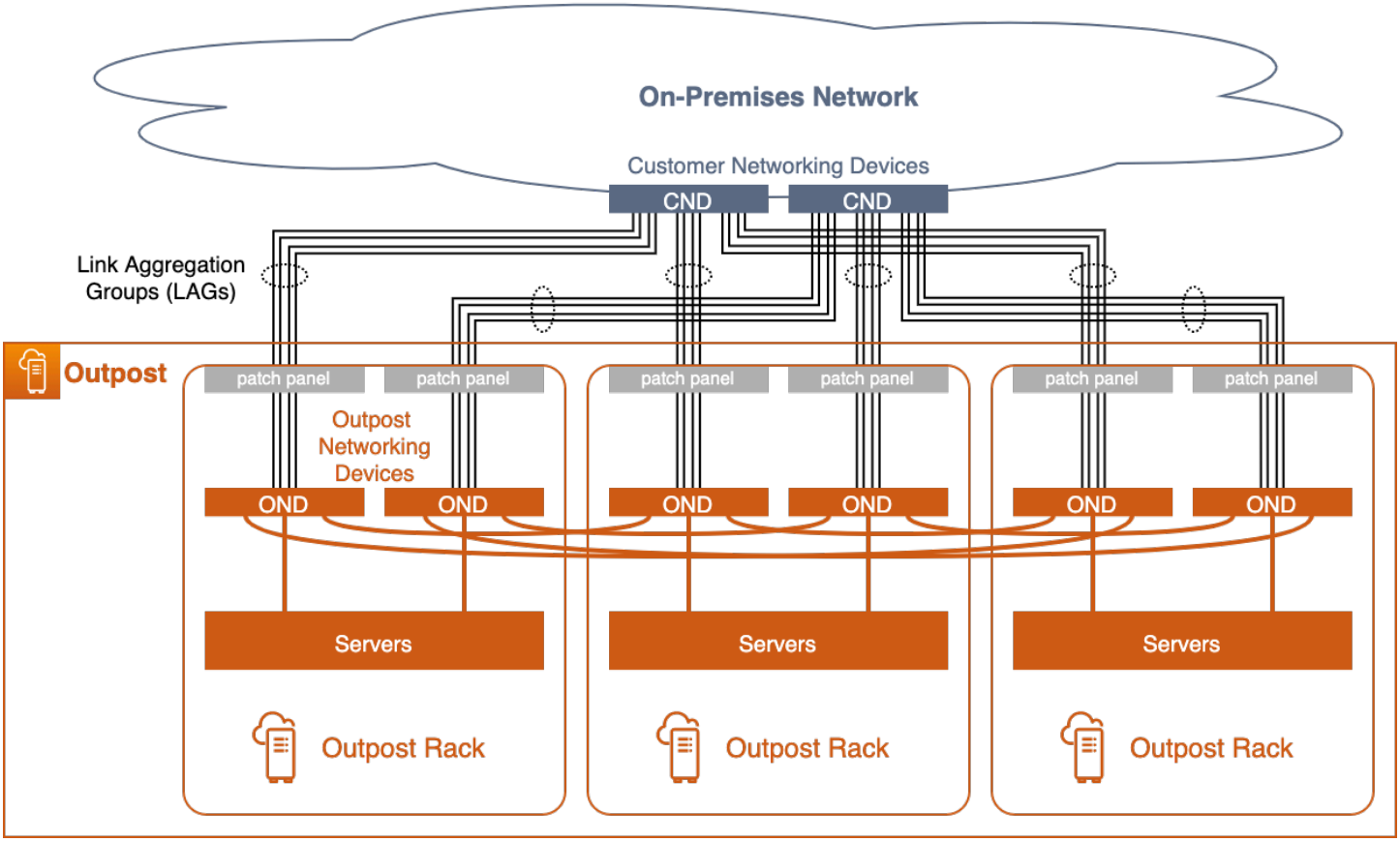
Redes

A implantação de um Outpost depende de uma conexão resiliente com seu AZ âncora para que as operações de gerenciamento, monitoramento e serviço funcionem adequadamente. Você deve provisionar sua rede local para fornecer conexões de rede redundantes para cada rack Outpost e conectividade confiável de volta aos pontos de ancoragem na nuvem. AWS Considere também os caminhos de rede entre as workloads de aplicativos em execução no Outpost e os outros sistemas on-premises e na nuvem com os quais elas se comunicam. Como você direcionará esse tráfego em sua rede?

Tópicos

Anexo de rede

Cada AWS Outposts rack é configurado com top-of-rack switches redundantes chamados Outpost Networking Devices (ONDs). Os servidores de computação e armazenamento em cada rack se conectam aos dois ONDs. Você deve conectar cada OND a um switch separado chamado Customer Networking Device (CND) em seu datacenter para fornecer diversos caminhos físicos e lógicos para cada rack do Outpost. Os ONDs se conectam aos seus CNDs com uma ou mais conexões físicas usando cabos de fibra óptica e transceptores ópticos. As conexões físicas são configuradas em links lógicos de grupo de agregação de links (LAG).



Outpost com vários racks com conexões de rede redundantes

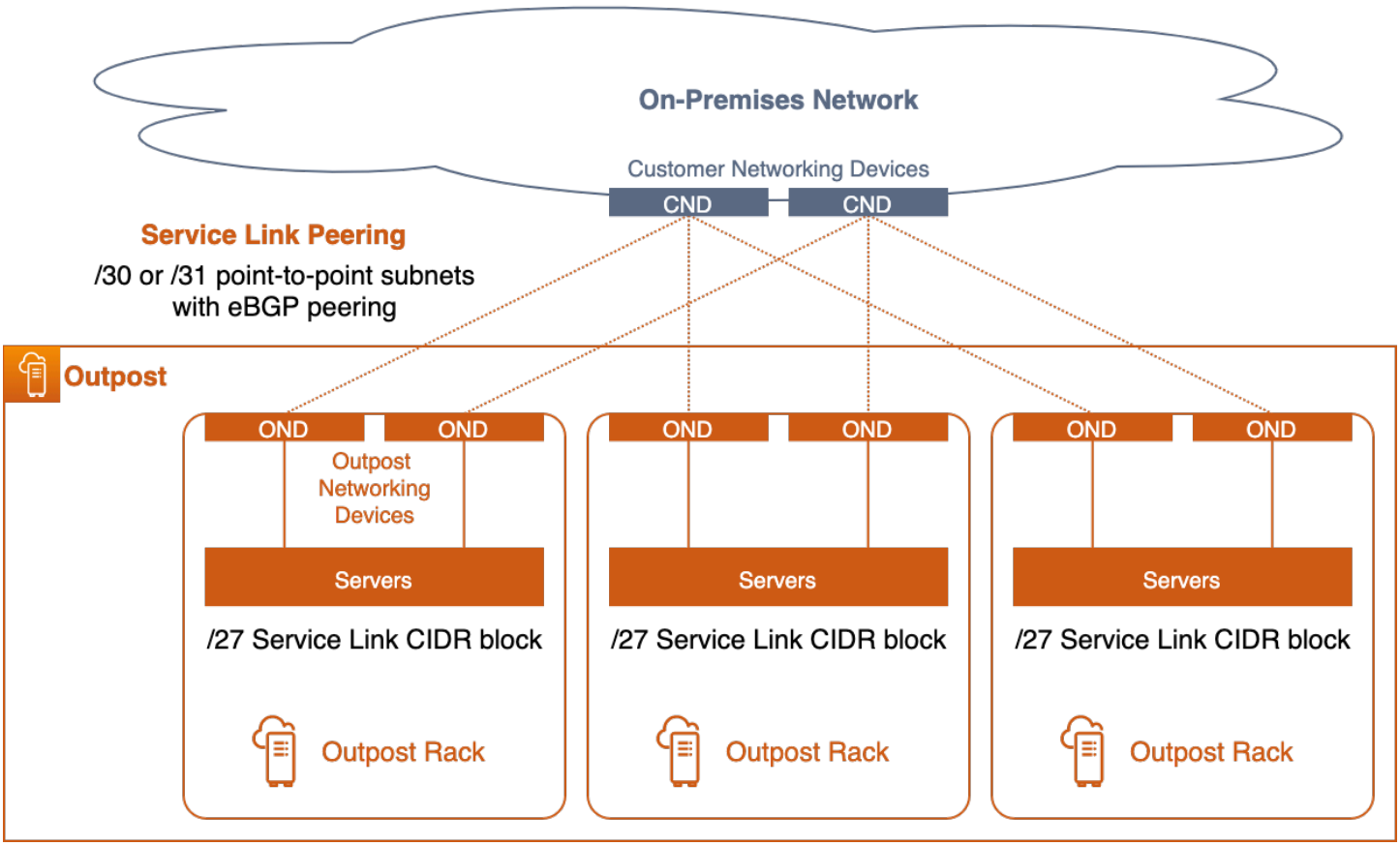
Os links OND para CND são sempre configurados em um LAG, mesmo que a conexão física seja um único cabo de fibra óptica. A configuração dos links como grupos de LAG permite aumentar a largura

de banda do link adicionando conexões físicas ao grupo lógico. Os links LAG são configurados como troncos Ethernet IEEE 802.1q para permitir a rede segregada entre o Outpost e a rede on-premises.

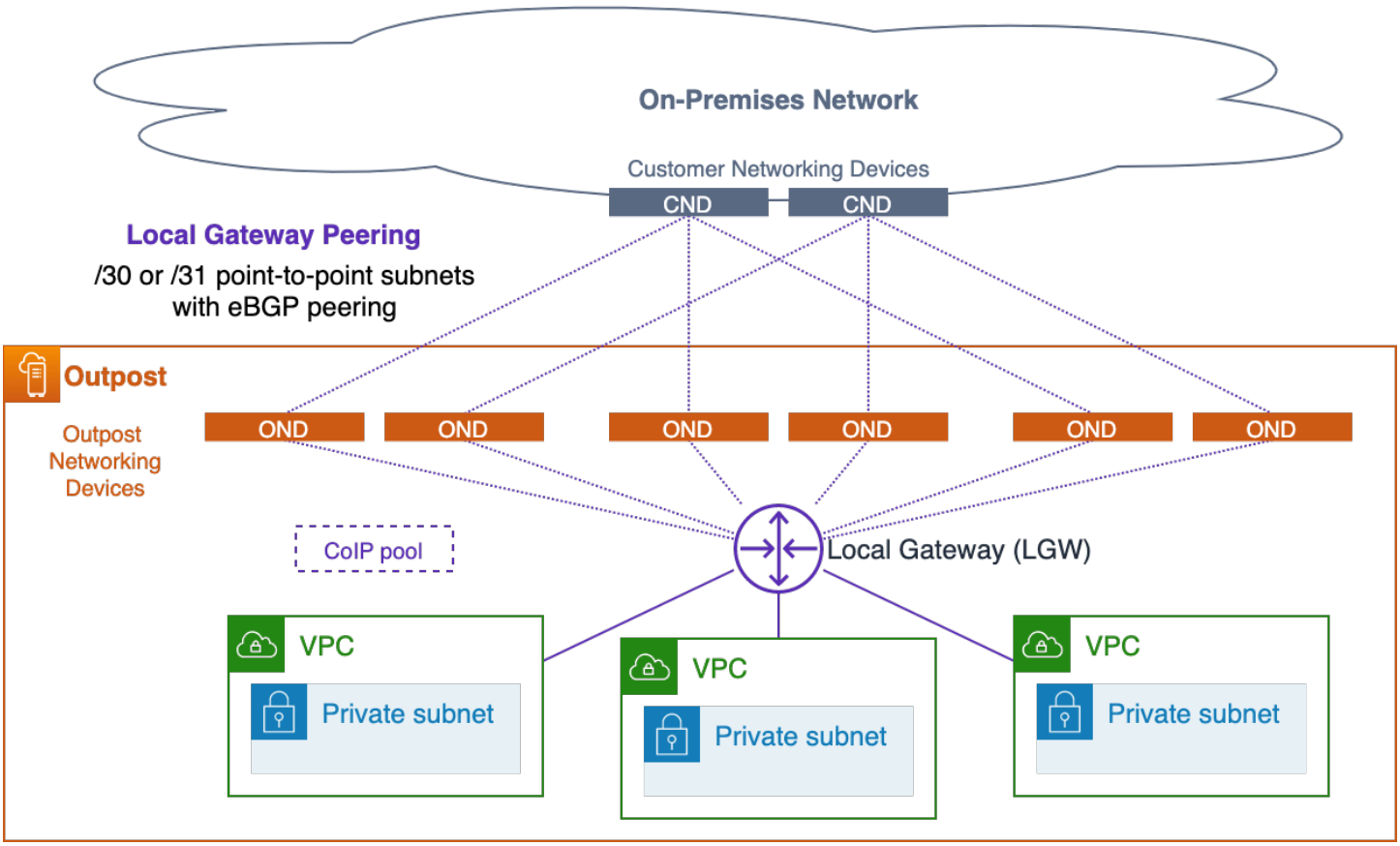
Cada Outpost tem pelo menos duas redes logicamente segregadas que precisam se comunicar com ou através da rede do cliente:

- Rede de link de serviço: aloca endereços IP do link de serviço para os servidores do Outpost e facilita a comunicação com a rede on-premises para permitir que os servidores se conectem novamente aos pontos de ancoragem do Outpost na região.
- Rede de gateway local: permite a comunicação entre as sub-redes VPC no Outpost e a rede on-premises por meio do Gateway local do Outpost (LGW).

Essas redes segregadas se conectam à rede local por meio de um conjunto de conexões point-to-point IP nos links LAG. Cada link OND para CND LAG é configurado com IDs de VLAN, sub-redes IP point-to-point (/30 ou /31) e emparelhamento eBGP para cada rede segregada (Service Link e LGW). Você deve considerar os links LAG, com suas point-to-point VLANs e sub-redes, como conexões de camada 2 segmentadas e roteadas de camada 3. As conexões IP roteadas fornecem caminhos lógicos redundantes que facilitam a comunicação entre as redes segregadas no Outpost e a rede on-premises.



Emparelhamento de links de serviço



Emparelhamento do gateway local

Você deve encerrar os links LAG de camada 2 (e suas VLANs) nos switches CND conectados diretamente e configurar as interfaces IP e o emparelhamento BGP nos switches CND. Você não deve conectar as VLANs LAG entre os switches do datacenter. Para obter mais informações, consulte Conectividade da camada de rede no Manual do usuário do AWS Outposts .

Dentro de um Outpost lógico de vários racks, os ONDs são interconectados de forma redundante para fornecer conectividade de rede altamente disponível entre os racks e as cargas de trabalho executadas nos servidores. AWS é responsável pela disponibilidade da rede dentro do Outpost.

Práticas recomendadas para anexo de rede altamente disponível

- Conecte cada Outpost Networking Device (OND) em um rack do Outpost a um Customer Networking Device (CND) separado no datacenter.
- Encerre os links de camada 2, as VLANs, as sub-redes IP de camada 3 e o emparelhamento BGP nos switches Customer Networking Device (CND) anexado diretamente. Não conecte as VLANs OND às CND entre as CNDs ou na rede on-premises.

- Adicione links aos grupos de agregação de links (LAGs) para aumentar a largura de banda disponível entre o Outpost e o datacenter. Não confie na largura de banda agregada dos diversos caminhos por meio dos dois ONDs.
- Use os diversos caminhos através dos ONDs redundantes para fornecer conectividade resiliente entre as redes do Outpost e a rede on-premises.
- Para obter a redundância ideal e permitir a manutenção do OND sem interrupções, recomendamos que os clientes configurem os anúncios e as políticas do BGP da seguinte forma:
  - O equipamento de rede do cliente deve receber anúncios de BGP do Outpost sem alterar os atributos do BGP e permitir o balanceamento de vários caminhos/carga do BGP para obter fluxos de tráfego de entrada ideais (do cliente para o Outpost). O acréscimo do caminho AS é usado para prefixos BGP do Outpost para afastar o tráfego de um OND/Uplink específico, caso seja necessária manutenção. A rede do cliente deve preferir rotas do Outpost com caminho AS de comprimento 1 em vez de rotas com caminho AS de comprimento 4, ou seja, reagir ao acréscimo do caminho AS.
  - A rede de clientes deve anunciar prefixos BGP iguais com os mesmos atributos para todos os ONDs no Outpost. Por padrão, a carga da rede do Outpost equilibra o tráfego de saída (para o cliente) entre todos os uplinks. As políticas de roteamento são usadas no lado do Outpost para afastar o tráfego de um OND específico, caso seja necessária manutenção. Prefixos BGP iguais do lado do cliente em todos os ONDs são necessários para realizar essa mudança de tráfego e realizar a manutenção sem interrupções. Quando a manutenção é necessária na rede do cliente, recomendamos usar o acréscimo do caminho AS para afastar temporariamente o tráfego de um determinado uplink ou dispositivo.

Conectividade de âncora

Um link de serviço do Outpost se conecta a âncoras públicas ou privadas (não ambas) em uma zona de disponibilidade específica (AZ) na região principal do Outpost. Os servidores Outpost iniciam conexões VPN de saída do link de serviço a partir de seus endereços IP do link de serviço até os pontos de ancoragem na AZ âncora. Essas conexões usam a porta UDP e TCP 443. AWS é responsável pela disponibilidade dos pontos de ancoragem na Região.

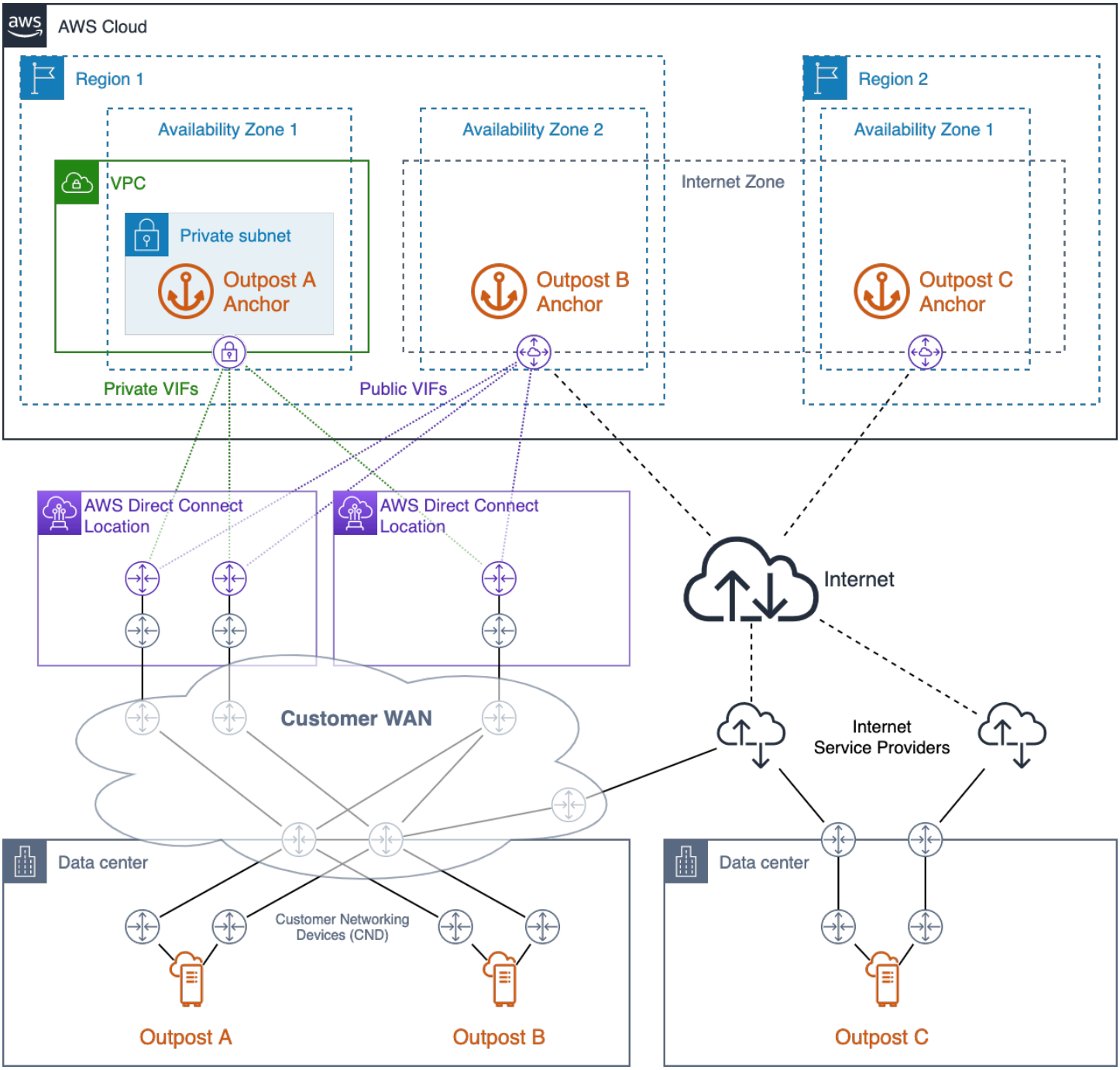
Você deve garantir que os endereços IP do Outpost Service Link possam se conectar por meio de sua rede aos pontos de ancoragem na AZ âncora. Os endereços IP do link de serviço não precisam se comunicar com outros hosts em sua rede on-premises.

Os pontos de ancoragem públicos residem nos intervalos de IP públicos da região (nos blocos CIDR do serviço EC2) e podem ser acessados pela Internet ou interfaces virtuais públicas (VIFs) da AWS Direct Connect (DX). O uso de pontos de ancoragem públicos permite uma seleção de caminhos mais flexível, pois o tráfego do link de serviço pode ser roteado por qualquer caminho disponível que possa alcançar com êxito os pontos de ancoragem na Internet pública.

Os pontos de ancoragem privados permitem que você use seus intervalos de endereços IP para conectividade de âncora. Os pontos de ancoragem privados são criados em uma sub-rede privada dentro de uma VPC dedicada usando endereços IP atribuídos pelo cliente. A VPC é criada no proprietário do recurso Outpost e você é responsável por garantir Conta da AWS que a VPC esteja disponível e configurada corretamente (não a exclua!). Os pontos de ancoragem privados devem ser acessados usando VIFs privadas do Direct Connect.

Você deve provisionar caminhos de rede redundantes entre o Outpost e os pontos de ancoragem na região, com conexões terminando em dispositivos separados em mais de um local. O roteamento dinâmico deve ser configurado para redirecionar automaticamente o tráfego para caminhos alternativos quando as conexões ou os dispositivos de rede falharem. Você deve provisionar capacidade de rede suficiente para garantir que a falha de um caminho de WAN não sobrecarregue os caminhos restantes.

O diagrama a seguir mostra três Outposts com caminhos de rede redundantes para o uso de suas AZs âncoras, AWS Direct Connect além de conectividade pública com a Internet. O Outpost A e o Outpost B estão ancorados em diferentes zonas de disponibilidade na mesma região. O Outpost A se conecta a pontos de ancoragem privados no AZ 1 da região 1. O Outpost B se conecta a pontos de ancoragem públicos na AZ 2 da região 1. O Outpost C se conecta a âncoras públicas no AZ 1 da região 2.



Conectividade âncora altamente disponível AWS Direct Connect e acesso público à Internet

O Outpost A tem três caminhos de rede redundantes para alcançar seu ponto de ancoragem privado. Dois caminhos estão disponíveis por meio de circuitos redundantes do Direct Connect em um único local do Direct Connect. O terceiro caminho está disponível por meio de um circuito do Direct Connect em um segundo local do Direct Connect. Esse design mantém o tráfego do link de serviço

do Outpost A em redes privadas e fornece redundância de caminhos que permite a falha de qualquer um dos circuitos do Direct Connect ou a falha de um local inteiro do Direct Connect.

O Outpost B tem quatro caminhos de rede redundantes para alcançar seu ponto de ancoragem público. Três caminhos estão disponíveis por meio de VIFs públicas provisionadas nos circuitos e locais do Direct Connect usados pelo Outpost A. O quarto caminho está disponível por meio da WAN do cliente e da Internet pública. O tráfego do link de serviço do Outpost B pode ser roteado por qualquer caminho disponível que possa alcançar com sucesso os pontos de ancoragem na Internet pública. O uso dos caminhos do Direct Connect pode fornecer latência mais consistente e maior disponibilidade de largura de banda, enquanto o caminho público da Internet pode ser usado para cenários de recuperação de desastres (DR) ou aumento de largura de banda.

O Outpost C tem dois caminhos de rede redundantes para alcançar seu ponto de ancoragem público. O Outpost C é implantado em um datacenter diferente do Outpost A e B. O datacenter do Outpost C não tem circuitos dedicados conectados à WAN do cliente. Em vez disso, o datacenter tem conexões de Internet redundantes fornecidas por dois provedores de serviços de Internet (ISPs) diferentes. O tráfego do link de serviço do Outpost C pode ser roteado por qualquer uma das redes do ISP para alcançar os pontos de ancoragem na Internet pública. Esse design permite flexibilidade para rotear o tráfego do link de serviço por qualquer conexão pública de Internet disponível. No entanto, o end-to-end caminho depende de redes públicas de terceiros, nas quais a disponibilidade da largura de banda e a latência da rede flutuam.

O caminho de rede entre um Outpost e seus pontos de ancoragem do link de serviço deve atender às seguintes especificações de largura de banda e latência:

- 500 Mbps - 1 Gbps de largura de banda disponível por rack do Outpost (por exemplo, 3 racks: largura de banda disponível de 1,5 a 3 Gbps)
- Latência de menos de 300 milissegundos (ida e volta)

Práticas recomendadas para conectividade de âncora altamente disponível:

- Provisione caminhos de rede redundantes entre cada Outpost e seus pontos de ancoragem na região.
- Use os caminhos do Direct Connect (DX) para controlar a latência e a disponibilidade da largura de banda.

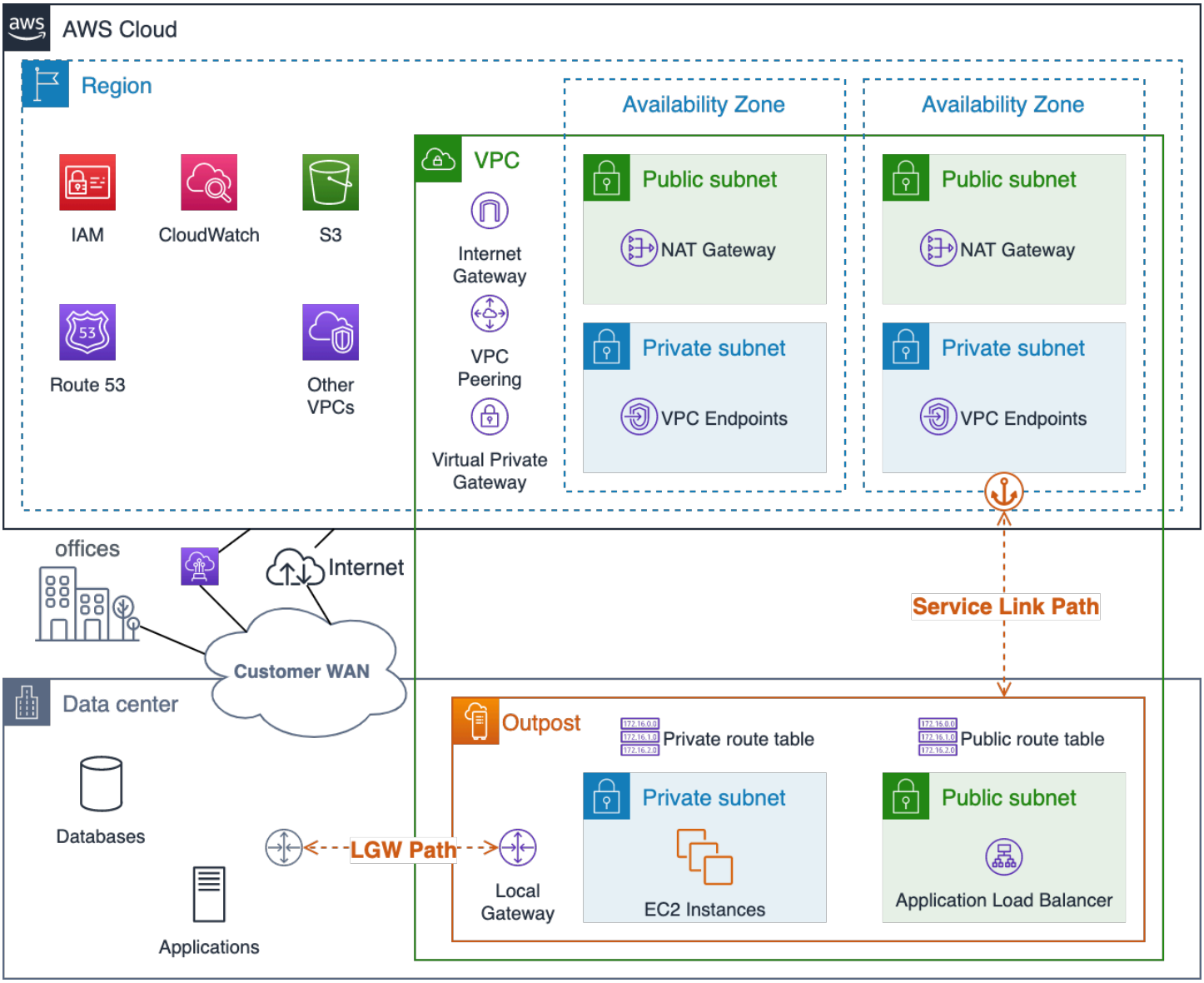
- Verifique se as portas TCP e UDP 443 estão abertas (de saída) dos blocos CIDR do link de serviço do Outpost para os intervalos de endereços IP do EC2 na região principal. Verifique se as portas estão abertas em todos os caminhos de rede.
- Verifique se cada caminho atende aos requisitos de disponibilidade de largura de banda e latência.
- Use o roteamento dinâmico para automatizar o redirecionamento de tráfego em caso de falhas na rede.
- Teste o roteamento do tráfego do link de serviço em cada caminho de rede planejado para garantir que o caminho funcione conforme o esperado.

Roteamento de aplicativos/workloads

Há dois caminhos fora do Outpost para workloads de aplicativos:

- O caminho do link de serviço
- O caminho do gateway local (LGW)

Você configura as tabelas de rotas de sub-rede do Outpost para controlar qual caminho seguir para alcançar as redes de destino. As rotas apontadas para o LGW direcionarão o tráfego para fora do gateway local e para a rede on-premises. As rotas apontadas para destinos na região, como gateways de Internet, gateways NAT, gateways privados virtuais e conexões de emparelhamento de VPC, direcionarão o tráfego pelo link de serviço para alcançar esses destinos.



Visualização dos caminhos de rede do link de serviço do Outpost e do LGW

Ao planejar o roteamento de aplicativos, você deve ter cuidado para considerar tanto a operação normal quanto a disponibilidade limitada do roteamento e do serviço durante falhas na rede. O caminho do link de serviço não está disponível quando um Outpost é desconectado da região.

Você deve provisionar caminhos diversos e configurar o roteamento dinâmico entre o LGW do Outpost e seus aplicativos, sistemas e usuários on-premises críticos. Caminhos de rede redundantes permitem que a rede direcione o tráfego para contornar falhas e garantir que os recursos locais possam se comunicar com as workloads em execução no Outpost durante falhas parciais na rede.

As configurações de rota do VPC do Outpost são estáticas. Você configura tabelas de roteamento de sub-rede por meio de AWS Management Console CLI, APIs e outras ferramentas de Infraestrutura como Código (IaC); no entanto, você não poderá modificar as tabelas de roteamento de sub-rede durante um evento de desconexão. Você precisará restabelecer a conectividade entre o Outpost e a região para atualizar as tabelas de rotas. Use as mesmas rotas para operações normais que você planeja usar durante eventos de desconexão.

Os recursos no Outpost podem acessar a Internet por meio do link de serviço e de um gateway da Internet (IGW) na região ou pelo caminho do gateway local (LGW). O roteamento do tráfego da Internet pelo caminho do LGW e pela rede on-premises permite que você use os pontos de entrada/saída da Internet on-premises existentes e que podem fornecer menor latência, maiores MTUs e taxas de saída de dados da AWS reduzidas em comparação com o uso do caminho do link de serviço para um IGW na região.

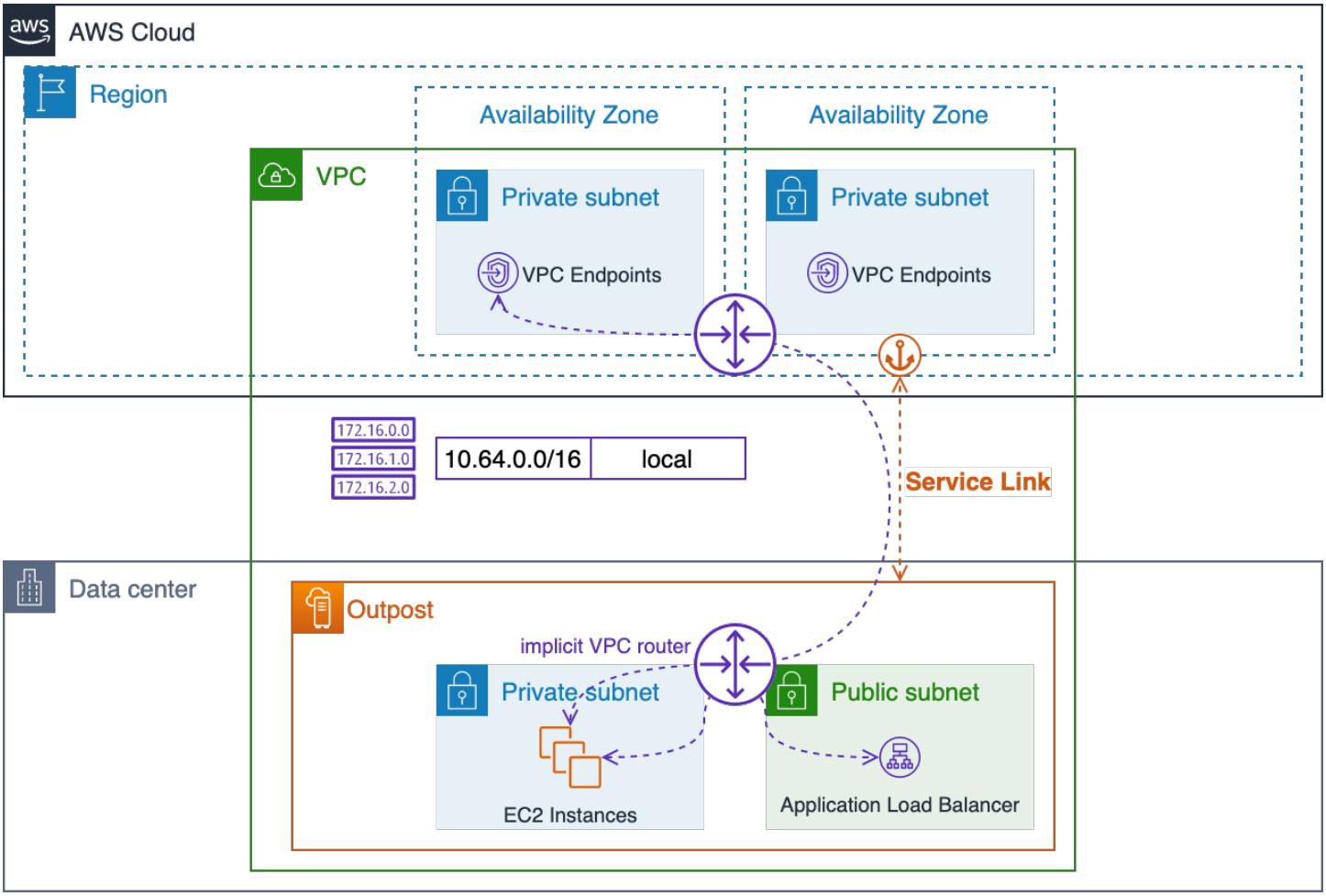
Se seu aplicativo precisar ser executado on-premises e precisar ser acessível pela Internet pública, você deverá rotear o tráfego do aplicativo por meio de suas conexões de Internet on-premises para o LGW para alcançar os recursos no Outpost.

Embora você possa configurar sub-redes em um Outpost como sub-redes públicas na região, essa pode ser uma prática indesejável para a maioria dos casos de uso. O tráfego de entrada da Internet entrará pelo Região da AWS e será roteado pelo Service Link para os recursos em execução no Outpost.

O tráfego de resposta, por sua vez, será roteado pelo Service Link e retornará por meio das conexões Região da AWS de internet do. Esse padrão de tráfego pode aumentar a latência e incorrer em cobranças de saída de dados à medida que o tráfego sai da região em direção ao Outpost e quando o tráfego volta pela região e sai para a Internet. Se seu aplicativo puder ser executado na região, a região será o melhor lugar para executá-lo.

O tráfego entre os recursos da VPC (na mesma VPC) sempre seguirá a rota CIDR da VPC local e será roteado entre sub-redes pelos roteadores VPC implícitos.

Por exemplo, o tráfego entre uma instância do EC2 em execução no Outpost e um endpoint da VPC na região sempre será roteado pelo link de serviço.



Roteamento de VPC local por meio de roteadores implícitos

Práticas recomendadas para roteamento de aplicativos/workload:

- Use o caminho do gateway local (LGW) em vez do caminho do link de serviço sempre que possível.
- Direcione o tráfego da Internet pelo caminho do LGW.
- Configure as tabelas de roteamento de sub-rede do Outpost com um conjunto padrão de rotas: elas serão usadas tanto para operações normais quanto durante eventos de desconexão.
- Provisione caminhos de rede redundantes entre o LGW do Outpost e os recursos essenciais de aplicativos on-premises. Use o roteamento dinâmico para automatizar o redirecionamento de tráfego em caso de falhas na rede on-premises.

Computação

Embora a capacidade de entrada do Amazon EC2 Regiões da AWS seja aparentemente infinita, a capacidade nos Outposts é finita. Você é responsável por planejar e gerenciar a capacidade computacional de suas implantações do Outposts.

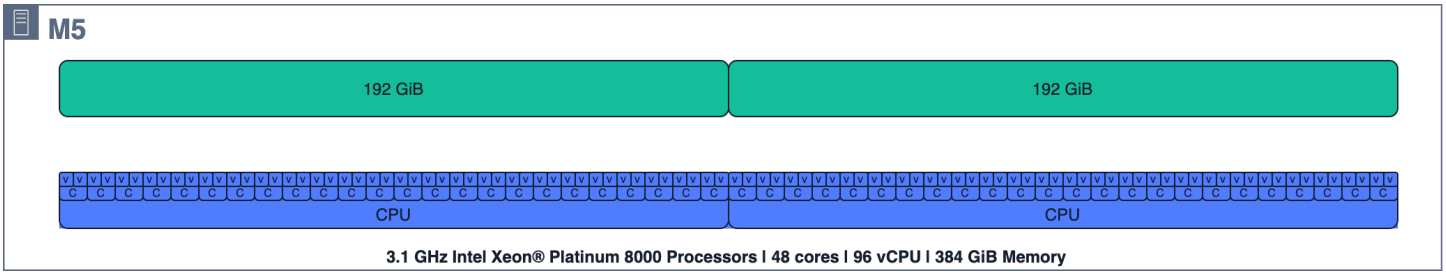
Tópicos

Planejamento de capacidade

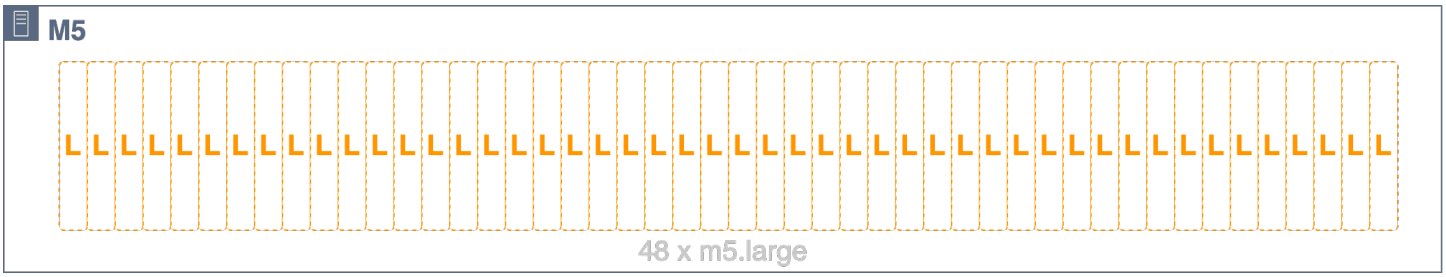
Embora a capacidade de entrada do Amazon EC2 Regiões da AWS seja aparentemente infinita, a capacidade nos Outposts é finita — limitada pelo volume total de capacidade computacional solicitada. Você é responsável por planejar e gerenciar a capacidade computacional de suas implantações do Outposts. Você deve solicitar capacidade computacional suficiente para suportar um modelo de disponibilidade N+M, em que N é o número necessário de servidores e M é o número de servidores não utilizados provisionados para acomodar falhas no servidor. N+1 e N+2 são os níveis de disponibilidade mais comuns.

Cada servidor (`C5`,, `M5R5`, etc.) oferece suporte a uma única família de instâncias do EC2. Antes de iniciar instâncias em servidores computacionais do EC2, você deve fornecer layouts de ranhura que especifiquem os tamanhos de instância do EC2 que você deseja que cada servidor forneça. AWS configura cada servidor com o layout de ranhura solicitado.

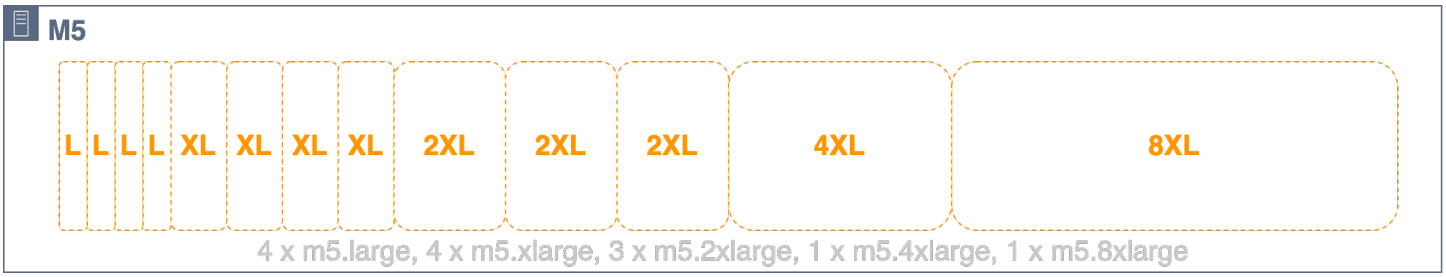
Os servidores podem ser distribuídos de forma homogênea, onde todos os slots têm o mesmo tamanho de instância (por exemplo, 48 `m5.large` slots) ou de forma heterogênea com uma mistura de tipos de instâncias (por exemplo, 4, 3 `m5.large` `m5.xlarge` `m5.2xlargem5.4xlarge`, 1 e 1`m5.8xlarge`) — veja as próximas três figuras para obter visualizações dessas configurações de slots.



*m5.24xlarge*recursos computacionais do servidor



*m5.24xlarge*servidor uniformemente encaixado em 48 slots *m5.large*



*m5.24xlarge*servidor dividido de forma heterogênea em 4*m5.large*, 4, 3 *m5.xlargem5.2xlarge*, 1 e 1 *m5.4xlarge* slots *m5.8xlarge*

Não é necessário encaixar em slots a capacidade total do servidor. Os slots podem ser adicionados a um servidor que tenha capacidade não alocada disponível. Você modifica um layout de slots abrindo um tíquete de suporte. O Enterprise Support pode exigir que você desligue ou reinicie determinadas instâncias para concluir uma solicitação de redistribuição se o novo layout de slot não puder ser aplicado enquanto determinados slots estiverem ocupados por instâncias em execução.

Todos os servidores contribuem com seus slots provisionados para os pools de capacidade do EC2 no Outpost, e todos os slots de um determinado tipo e tamanho de instância são gerenciados como um único pool de capacidade do EC2. Por exemplo, o servidor anterior com slots heterogêneos com slots`m5.large`,, `m5.xlarge` `m5.2xlargem5.4xlarge`, e contribuiria com esses `m5.8xlarge` slots para cinco pools de capacidade do EC2 — um pool para cada tipo e tamanho de instância.

É importante considerar o slot do servidor e os pools de capacidade do EC2 ao planejar a capacidade ociosa para a disponibilidade do servidor N+M. AWS detecta quando um servidor falha ou está degradado e agenda uma visita ao local para substituir o servidor com falha. Você deve projetar seus pools de capacidade do EC2 para tolerar a falha de pelo menos um servidor de cada família de instâncias (N+1) em um Outpost. Com esse nível mínimo de disponibilidade do servidor, quando um servidor falha ou precisa ser retirado de serviço, você pode reiniciar instâncias com falha ou degradadas nos slots não utilizados dos servidores restantes da mesma família.

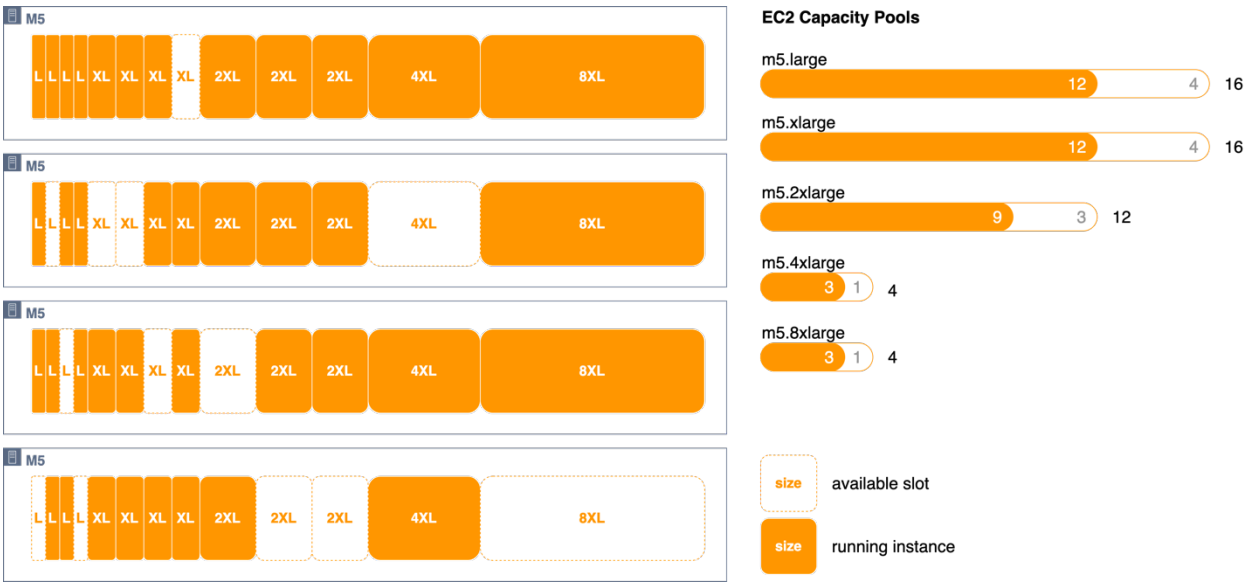
Planejar a disponibilidade de N+M é simples quando você tem servidores com slots homogêneos ou grupos de servidores com slots heterogêneos com layouts de slot idênticos. Basta calcular o número de servidores (N) necessários para executar todas as suas workloads e, em seguida, adicionar (M) servidores para atender aos requisitos de disponibilidade do servidor durante eventos de falha e manutenção.

As seguintes configurações de ranhura não podem ser usadas devido aos limites do NUMA:

- 3 `m5.8xlarge`
- 1 `m5.16xlarge` e 1 `m5.8xlarge`

Consulte sua Conta da AWS equipe para validar sua configuração planejada de ranhura em AWS Outposts rack.

Na figura a seguir, quatro `m5.24xlarge` servidores estão distribuídos de forma heterogênea com um layout de ranhura idêntico. Os quatro servidores criam cinco pools de capacidade do EC2. Cada pool está sendo executado com utilização máxima (75%) para manter a disponibilidade de N+1 para as instâncias em execução nesses quatro servidores. Se algum servidor falhar, haverá espaço suficiente para reiniciar as instâncias com falha nos servidores restantes.



Visualização de slots de servidor EC2, instâncias em execução e pools de slots

Para layouts de slot mais complexos, nos quais os servidores não têm slots idênticos, você precisará calcular a disponibilidade de N+M para cada pool de capacidade do EC2. Você pode usar a fórmula a seguir para calcular quantos servidores (que contribuem com slots para um determinado pool

de capacidade do EC2) podem falhar e ainda permitir que os servidores restantes carreguem as instâncias em execução:

$$M = \left\lceil \frac{poolSlots_{available}}{serverSlots_{max}} \right\rceil$$

Em que:

- $poolSlots_{available}$ é o número de slots disponíveis em um determinado pool de capacidade do EC2 (número total de slots no pool menos o número de instâncias em execução)
- $serverSlots_{max}$ é o número máximo de slots contribuídos por qualquer servidor para um determinado pool de capacidade do EC2
- M é o número de servidores que podem falhar e ainda permitir que os servidores restantes carreguem as instâncias em execução

Exemplo: um posto avançado tem três servidores que contribuem com slots para um pool `m5.2xlarge` de capacidade. O primeiro contribui com 4 slots, o segundo contribui com 3 slots e o terceiro servidor com 2 slots. O pool de `m5.2xlarge` instâncias no Outpost tem uma capacidade total de 9 slots (4 + 3 + 2). O Outpost tem 4 `m5.2xlarge` instâncias em execução. Quantos servidores podem falhar e ainda permitir que os servidores restantes carreguem as instâncias em execução?

$$poolSlots_{available} = total\ capacity - running\ instances = 9 - 4 = 5$$

$$serverSlots_{max} = \max([4, 3, 2]) = 4$$

$$M = \left\lceil \frac{poolSlots_{available}}{serverSlots_{max}} \right\rceil = \left[\frac{5}{4}\right] = [1.25] = 1$$

Resposta: você pode perder qualquer um dos servidores e ainda carregar as instâncias em execução nos servidores restantes.

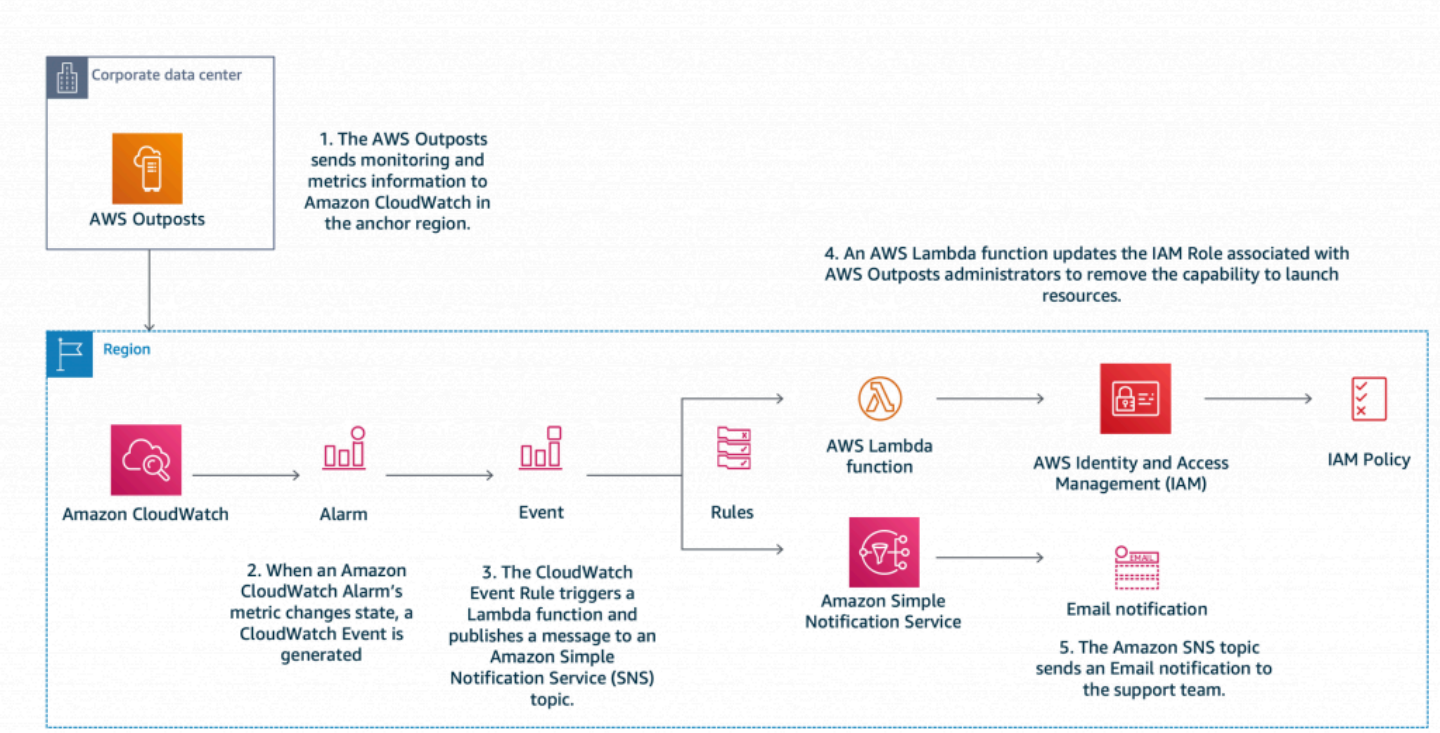
Práticas recomendadas para o planejamento da capacidade computacional:

- Dimensione sua capacidade computacional para fornecer redundância N+M para cada pool de capacidade do EC2 em um Outpost.
  - Implante servidores N+M para servidores homogêneos ou idênticos com slots heterogêneos.
  - Calcule a disponibilidade de N+M para cada pool de capacidade do EC2 e garanta que cada pool atenda aos seus requisitos de disponibilidade.

Gerenciamento de capacidade

Você pode monitorar a utilização do pool de instâncias do Outpost EC2 nas métricas da Amazon e AWS Management Console por meio delas. CloudWatch Entre em contato com o Enterprise Support para recuperar ou alterar os layouts de slot de seus Outposts.

Você usa os mesmos mecanismos de recuperação automática de instâncias e EC2 Auto Scaling para recuperar ou substituir instâncias afetadas por falhas no servidor e eventos de manutenção. Você deve monitorar e gerenciar a capacidade do seu Outpost para garantir que a capacidade disponível suficiente esteja sempre disponível para acomodar falhas no servidor. A postagem Gerenciando sua AWS Outposts capacidade usando a Amazon CloudWatch e o AWS Lambda blog fornece um tutorial prático que mostra como combinar AWS CloudWatch e gerenciar sua capacidade do Outpost AWS Lambda para manter a disponibilidade da instância.



Gerenciando a AWS Outposts capacidade com a Amazon CloudWatch e AWS Lambda

Práticas recomendadas para o gerenciamento da capacidade computacional:

- Configure suas instâncias do EC2 em grupos do Auto Scaling ou use a recuperação automática de instâncias para reiniciar instâncias com falha.
- Automatize o monitoramento da capacidade de suas implantações do Outpost e configure notificações e (opcionalmente) respostas automatizadas para alarmes de capacidade.

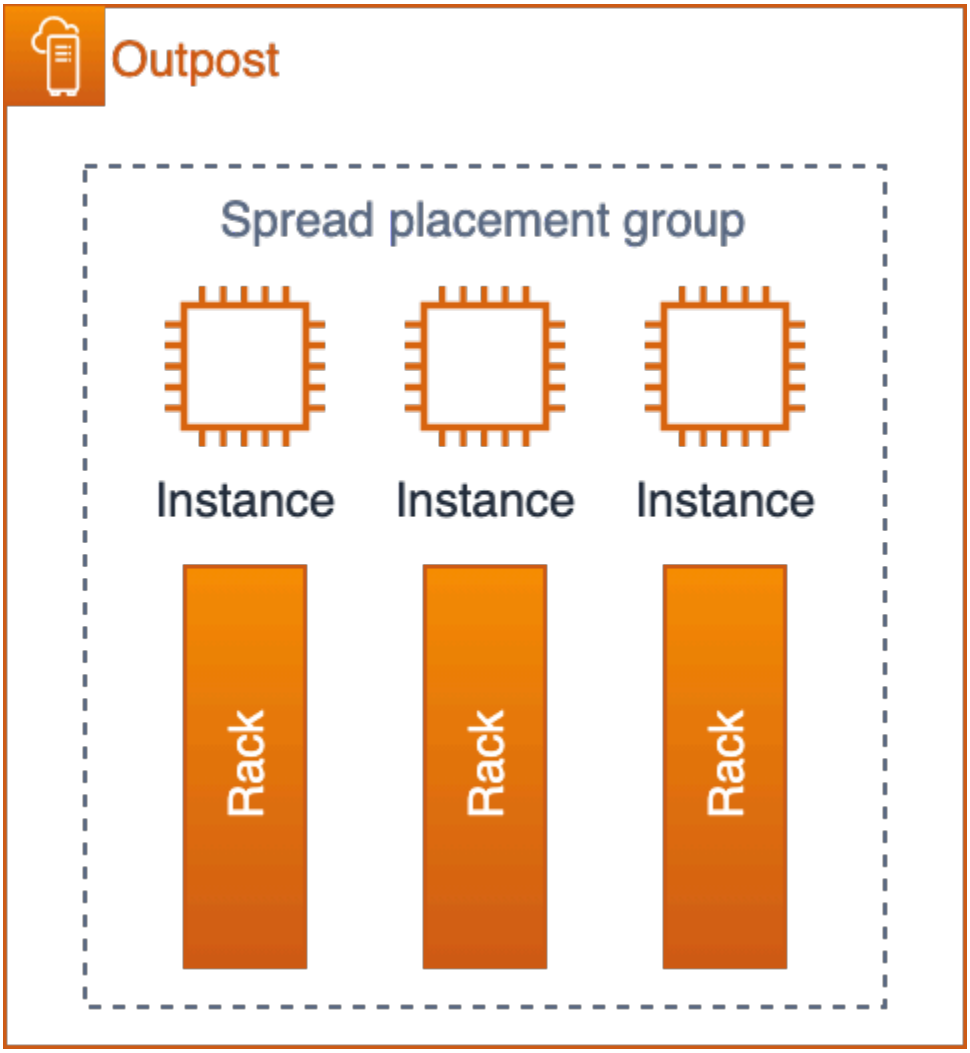
Posicionamento da instância

Outposts têm um número finito de servidores computacionais. Se seu aplicativo implantar várias instâncias relacionadas no Outposts, sem configuração adicional, as instâncias poderão ser implantadas no mesmo servidor ou em servidores no mesmo rack. Atualmente, existem três mecanismos que você pode usar para distribuir instâncias a fim de reduzir o risco de executar instâncias relacionadas na mesma infraestrutura:

Implantação de vários Outposts: semelhante a uma estratégia Multi-AZ na região, você pode implantar Outposts em datacenters separados e implantar recursos de aplicativos no Outposts específico. Isso permite que você execute instâncias no Outpost desejado (um conjunto lógico de racks). Uma estratégia de vários Outposts pode ser empregada para proteção contra modos de falha de rack e datacenter e, se os Outposts estiverem ancorados em AZs ou regiões separadas, também poderão fornecer proteção contra modos de falha de AZ ou região. Para obter mais informações sobre arquiteturas de vários Outposts, consulte os Modos de falha maiores.

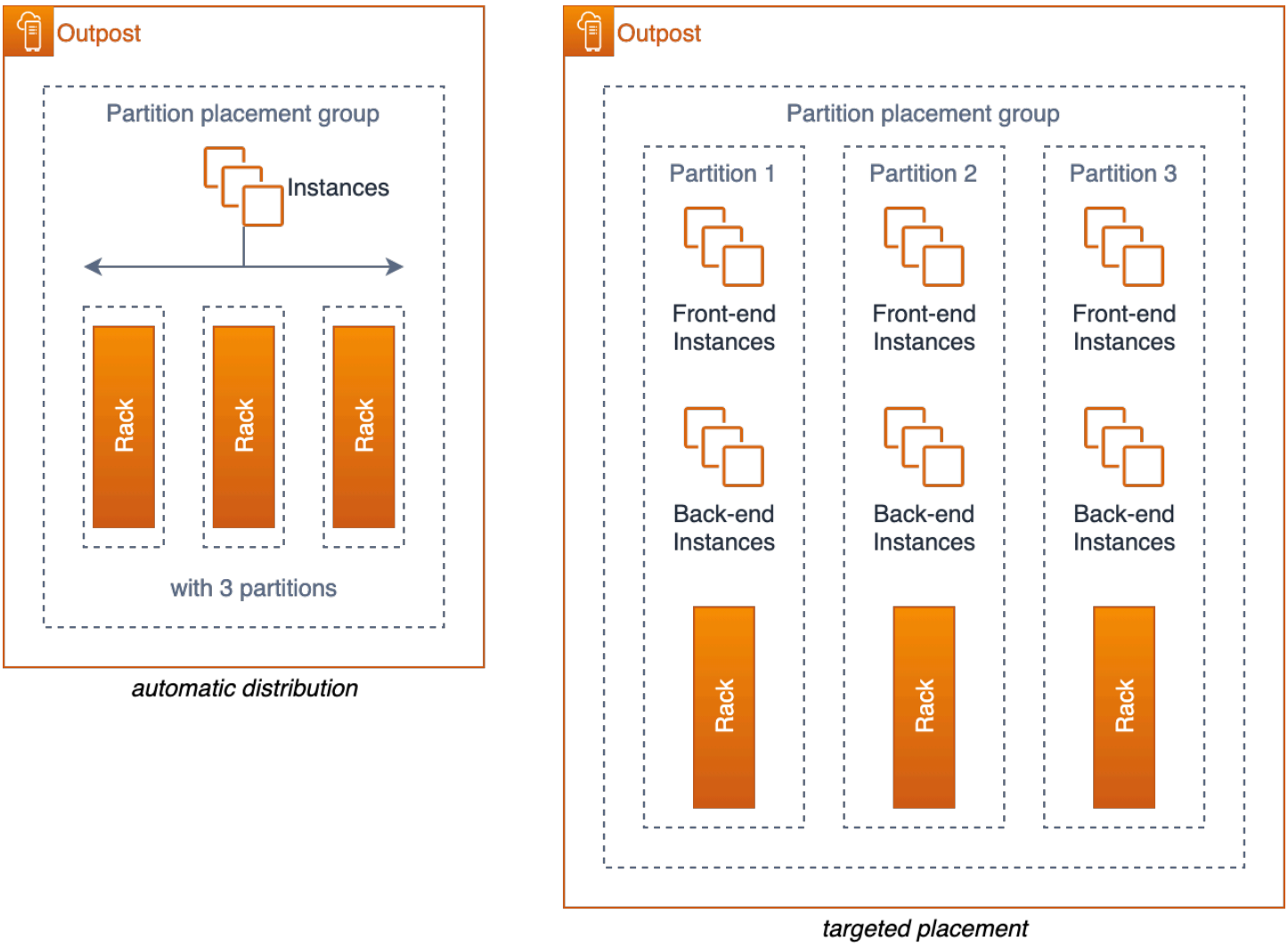
Grupos com posicionamento do Amazon EC2 no Outposts (posicionamento de instância com vários racks de um único OutPost) permitem que você use as estratégias de cluster, spread e partição para influenciar o posicionamento. As estratégias de posicionamento distribuído e em partições permitem que você distribua instâncias entre racks em um Outpost com vários racks.

Um grupo com posicionamento distribuído fornece uma maneira simples de distribuir instâncias únicas em racks para reduzir o potencial de falhas correlacionadas. Você só pode implantar no grupo o mesmo número de instâncias que tiver de racks em seu Outpost.



Grupo com posicionamento distribuído EC2 em um Outpost com três racks

Você também pode distribuir instâncias em vários racks com grupos com posicionamento em partições. Use a distribuição automática para distribuir instâncias entre partições no grupo ou implantar instâncias em partições de destino selecionadas. A implantação de instâncias nas partições de destino permite que você implante recursos selecionados no mesmo rack enquanto distribui outros recursos entre os racks. Por exemplo, se você tiver um Outpost lógico com três racks, criar um grupo com posicionamento em partições com três partições permite distribuir recursos entre os racks.



Grupos com posicionamento em partições EC2 em um Outpost com três racks

Slots criativos de servidor: se você tiver um Outpost de rack único ou se o serviço que você está usando no Outposts não oferecer suporte a grupos com posicionamento, talvez você possa usar o slot criativo para garantir que suas instâncias não sejam implantadas no mesmo servidor físico. Se as instâncias relacionadas tiverem o mesmo tamanho de instância do EC2, você poderá alocar seus servidores para limitar o número de slots desse tamanho configurado em cada servidor, distribuindo os slots entre os servidores. O slot do servidor limitará o número de instâncias (desse tamanho) que podem ser executadas em um único servidor.

Como exemplo, considere o layout de slot mostrado anteriormente na Figura 13. Se seu aplicativo precisasse implantar três `m5.4xlarge` instâncias no Outpost configurado com esse layout de slot, o EC2 colocaria cada instância em um servidor separado e não haveria possibilidade de que essas

instâncias pudessem ser executadas no mesmo servidor — desde que a configuração de slots não mudasse para abrir slots adicionais nos servidores. `m5.4xlarge`

Práticas recomendadas para posicionamento de instâncias de computação:

- Use grupos com posicionamento do Amazon EC2 no Outposts para controlar o posicionamento de instâncias em racks dentro de um único Outpost.
- Em vez de pedir um Outpost com um único rack médio ou grande do Outpost, divida a capacidade em dois racks pequenos ou médios para permitir que você aproveite a capacidade dos grupos com posicionamento do EC2 de distribuir instâncias entre racks.

Armazenamento

O serviço de AWS Outposts rack fornece três tipos de armazenamento:

- Armazenamento de instância em tipos de instância do EC2 com suporte
- Volumes gp2 do Amazon Elastic Block Store (EBS) para armazenamento persistente em bloco
- Amazon Simple Storage Service on Outposts (S3 on Outposts) para armazenamento local de objetos

O armazenamento de instâncias é fornecido em servidores compatíveis (`C5d`, `M5d`, `R5d`, `G4dn` e `I3en`). Assim como na região, os dados em um armazenamento de instâncias persistem somente durante a vida útil (em execução) da instância.

Os volumes do EBS do Outposts e o armazenamento de objeto do S3 on Outposts são fornecidos como parte dos serviços gerenciados do rack do AWS Outposts . Os clientes são responsáveis pelo gerenciamento da capacidade dos pools de armazenamento do Outpost. Os clientes especificam seus requisitos de armazenamento para armazenamento EBS e S3 ao solicitar um Outpost. AWS configura o Outpost com o número de servidores de armazenamento necessários para fornecer a capacidade de armazenamento solicitada. AWS é responsável pela disponibilidade do EBS e do S3 nos serviços de armazenamento da Outposts. Servidores de armazenamento suficientes são provisionados para fornecer serviços de armazenamento de alta disponibilidade para o Outpost. A perda de um único servidor de armazenamento não deve interromper os serviços nem resultar em perda de dados.

Você pode usar as CloudWatch métricas AWS Management Console e para monitorar o Outpost EBS e o S3 na utilização da capacidade do Outposts.

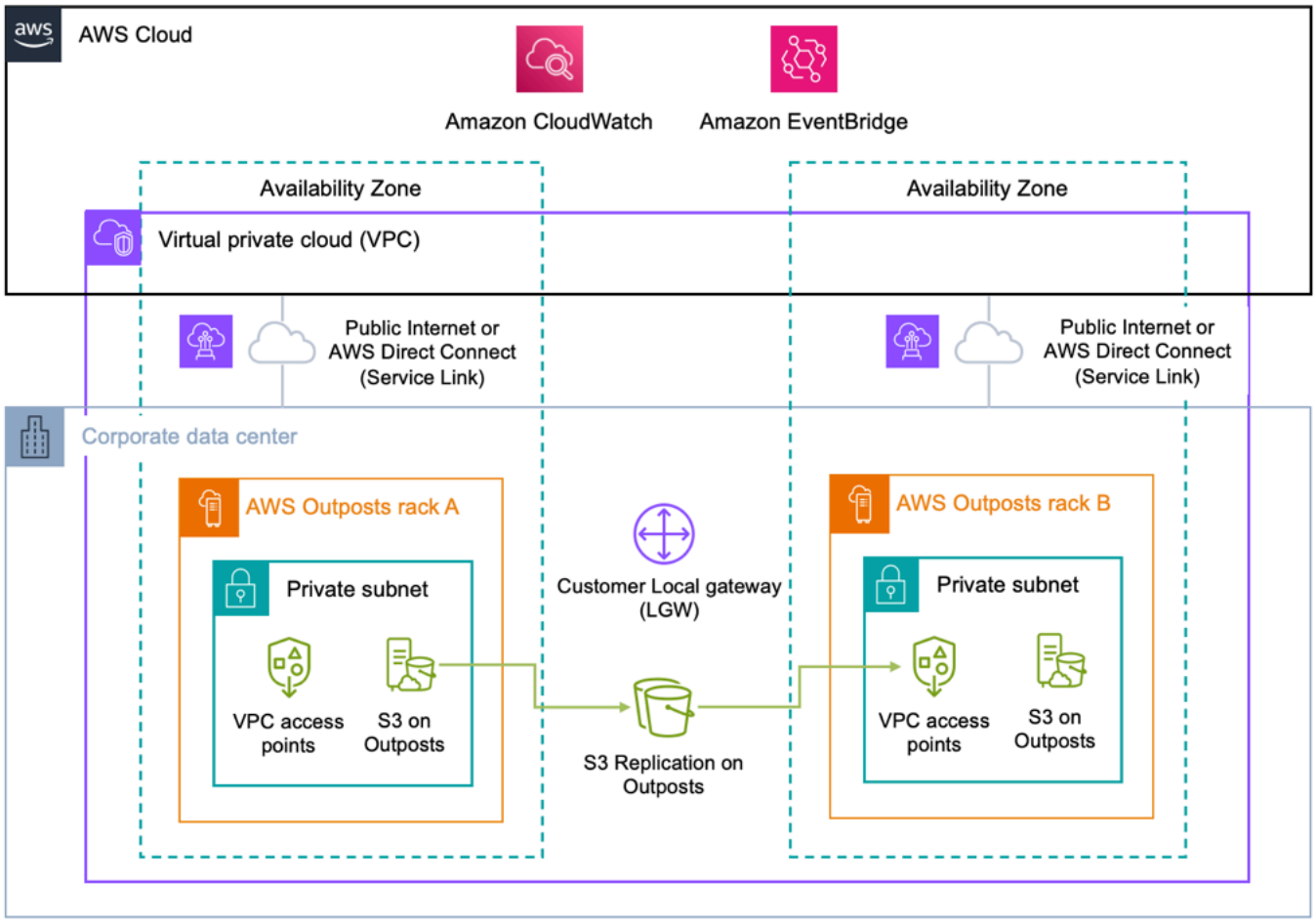
Proteção de dados

Para volumes EBS: o AWS Outposts rack suporta instantâneos de volume do EBS para fornecer um mecanismo de proteção de dados simples e seguro para proteger seus dados de armazenamento em bloco. Os snapshots são backups point-in-time incrementais dos seus volumes do EBS. Por padrão, os snapshots de volumes do Amazon EBS no seu Outpost são armazenados no Amazon S3 na região. Se seus Outposts tiverem sido configurados com a capacidade do S3 on Outposts, você pode usar o EBS Local Snapshots on Outposts para armazenar snapshots localmente no seu Outpost usando o armazenamento do S3 on Outposts.

Para buckets do S3 on Outposts (casos de uso de residência de dados):

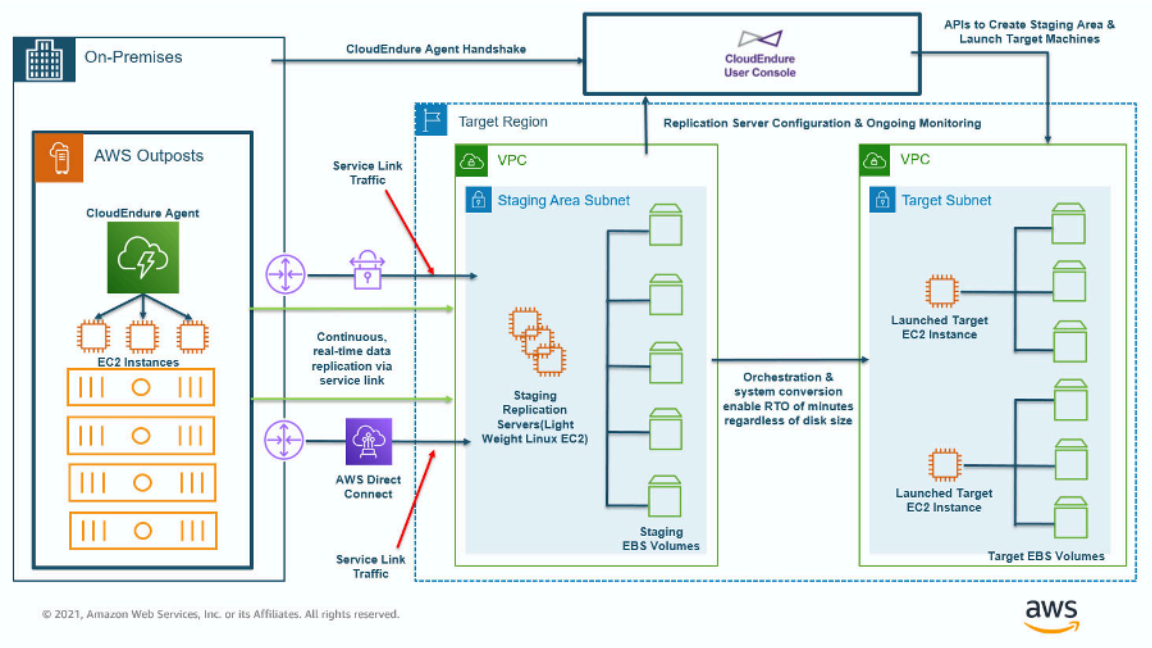
- Você pode usar o versionamento do S3 no Outposts para salvar todas as alterações e o histórico dos objetos. Quando habilitado, o versionamento do S3 salva várias cópias distintas de um objeto no mesmo bucket. O versionamento do S3 pode ser usado para preservar, recuperar e restaurar todas as versões de cada objeto armazenado em buckets do Outposts. O versionamento do S3 ajuda você a se recuperar de ações não intencionais de usuários e de falhas da aplicação.
- Você pode usar a Replicação do S3 no Outposts para criar e configurar regras de replicação para replicar automaticamente seus objetos do S3 para outro Outpost ou para outro bucket no mesmo Outpost. Durante a replicação, os objetos do S3 on Outposts são enviados pelo gateway local (LGW) do cliente, e os objetos não voltam para a Região da AWS. A replicação do S3 no Outposts fornece uma maneira fácil e flexível de replicar dados automaticamente dentro de um perímetro de dados específico para atender aos requisitos de redundância e conformidade de dados.

  A replicação do S3 no Outposts também fornece métricas e notificações detalhadas para monitorar o status da replicação do seu objeto. Você pode monitorar o progresso da replicação rastreando bytes pendentes, operações pendentes e latência de replicação entre seus buckets Outposts de origem e destino usando a Amazon. CloudWatch Você também pode configurar EventBridge as regras da Amazon para receber eventos de falha de replicação para diagnosticar e corrigir problemas de configuração rapidamente.



Para buckets do S3 on Outposts (casos de uso sem residência de dados) Regiões da AWS para: Você pode usar para AWS DataSyncautomatizar as transferências de dados do S3 on Outposts entre seu Outpost e a região. DataSync permite que você escolha o que transferir, quando transferir e quanta largura de banda usar. O backup de seus buckets do S3 on Outposts on-premises em buckets do S3 na Região da AWS permite que você aproveite 99,999999999% (onze 9) de durabilidade de dados e níveis adicionais de armazenamento (Standard, acesso pouco frequente e Glacier) para otimização de custos disponível com o serviço regional do S3.

Replicação de instâncias: você pode usar CloudEndurepara replicar instâncias individuais de sistemas locais para um Posto Avançado, de um Posto Avançado para a Região, da Região para um Posto Avançado ou de um Posto Avançado para outro. A postagem do CloudEndure blog Architecting for DR on AWS Outposts with descreve cada um desses cenários e como projetar uma solução com o. CloudEndure



Recuperação de desastres (DR) de um Outpost para a região

Usar o AWS Outposts rack como CloudEndure destino (destino de replicação) requer armazenamento S3 on Outposts.

Práticas recomendadas para proteção de dados:

- Use snapshots do EBS para criar point-in-time backups de volumes de armazenamento em bloco no Amazon S3 na região ou no S3 no Outposts.
- Use o versionamento de objetos do S3 on Outposts para manter várias versões e o histórico de seus objetos.
- Use a replicação do S3 no Outposts para replicar automaticamente seus dados de objeto para outro Outpost.
- Para casos de uso sem residência de dados, use AWS DataSync para fazer backup de objetos armazenados no S3 no Outpost para o Amazon S3 na região.
- Use CloudEndure para replicar instâncias entre sistemas locais, Outposts lógicos e a região.

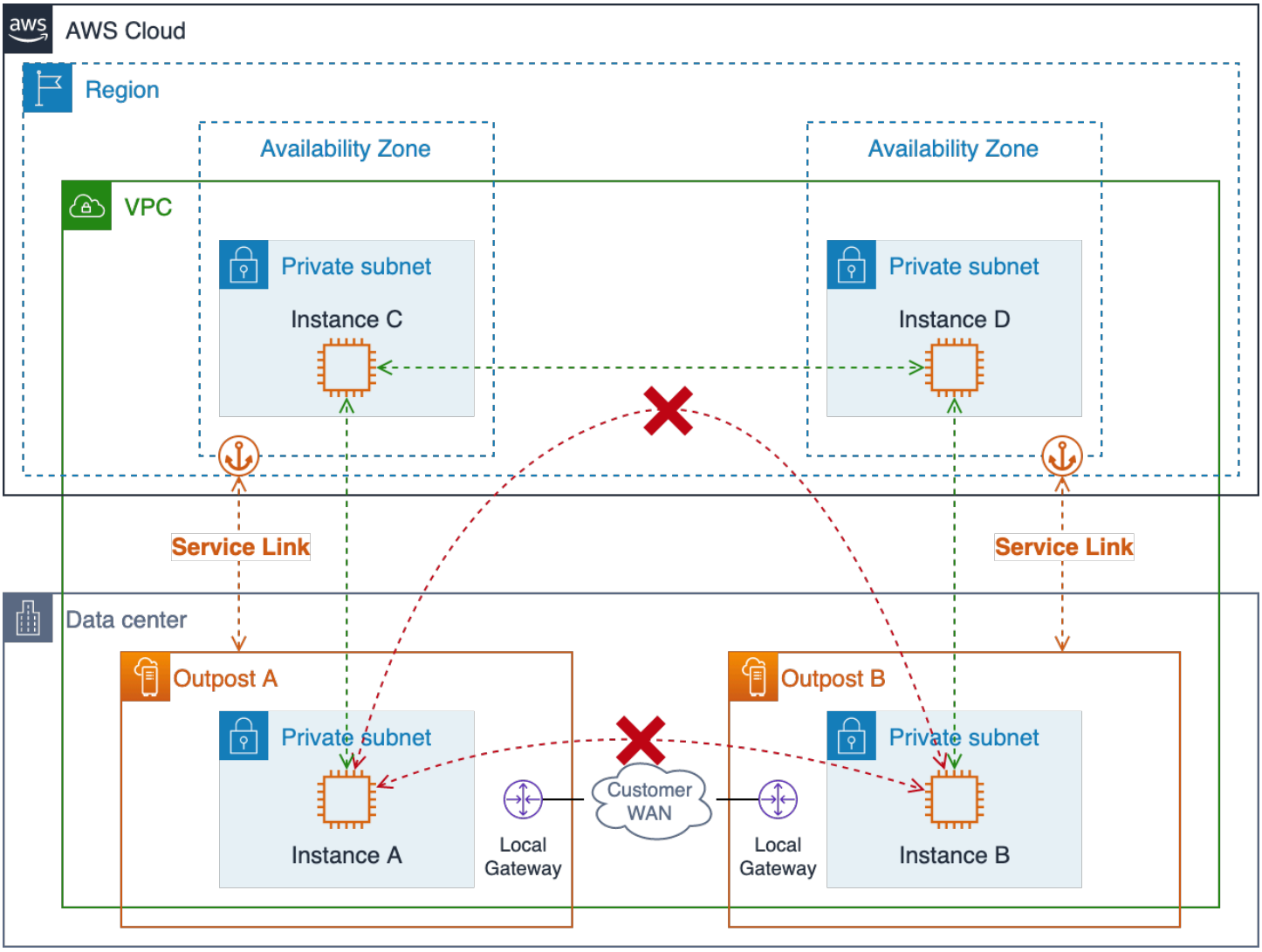
Modos de falha maiores

Para projetar arquiteturas de HA para mitigar modos de falha maiores, como falhas de rack, datacenter, zona de disponibilidade (AZ) ou região, você deve implantar vários Outposts com

capacidade de infraestrutura suficiente em datacenters separados com energia independente e conectividade WAN. Você ancora os Outposts em diferentes zonas de disponibilidade (AZs) dentro de uma Região da AWS ou em várias regiões. Você também deve provisionar site-to-site conectividade resiliente e suficiente entre os locais para oferecer suporte à replicação de dados síncrona ou assíncrona e ao redirecionamento do tráfego da carga de trabalho. Dependendo da arquitetura do seu aplicativo, você pode usar o DNS do Amazon Route 53 disponível globalmente e os serviços do Elastic Load Balancing disponíveis regionalmente para direcionar o tráfego para o local desejado e automatizar o redirecionamento de tráfego para locais sobreviventes no caso de falhas em grande escala.

Há limitações de rede que você deve conhecer ao projetar e implantar workloads de aplicativos em vários Outposts. Os recursos em dois Outposts separados não podem se comunicar uns com os outros ao transitar o tráfego pela região. Os recursos em dois Outposts separados implantados na mesma VPC não podem se comunicar entre si na rede do cliente. Recursos em dois Outposts separados implantados em VPCs diferentes podem se comunicar pela rede do cliente.

As duas figuras a seguir ilustram os caminhos de rede bloqueados e bem-sucedidos.

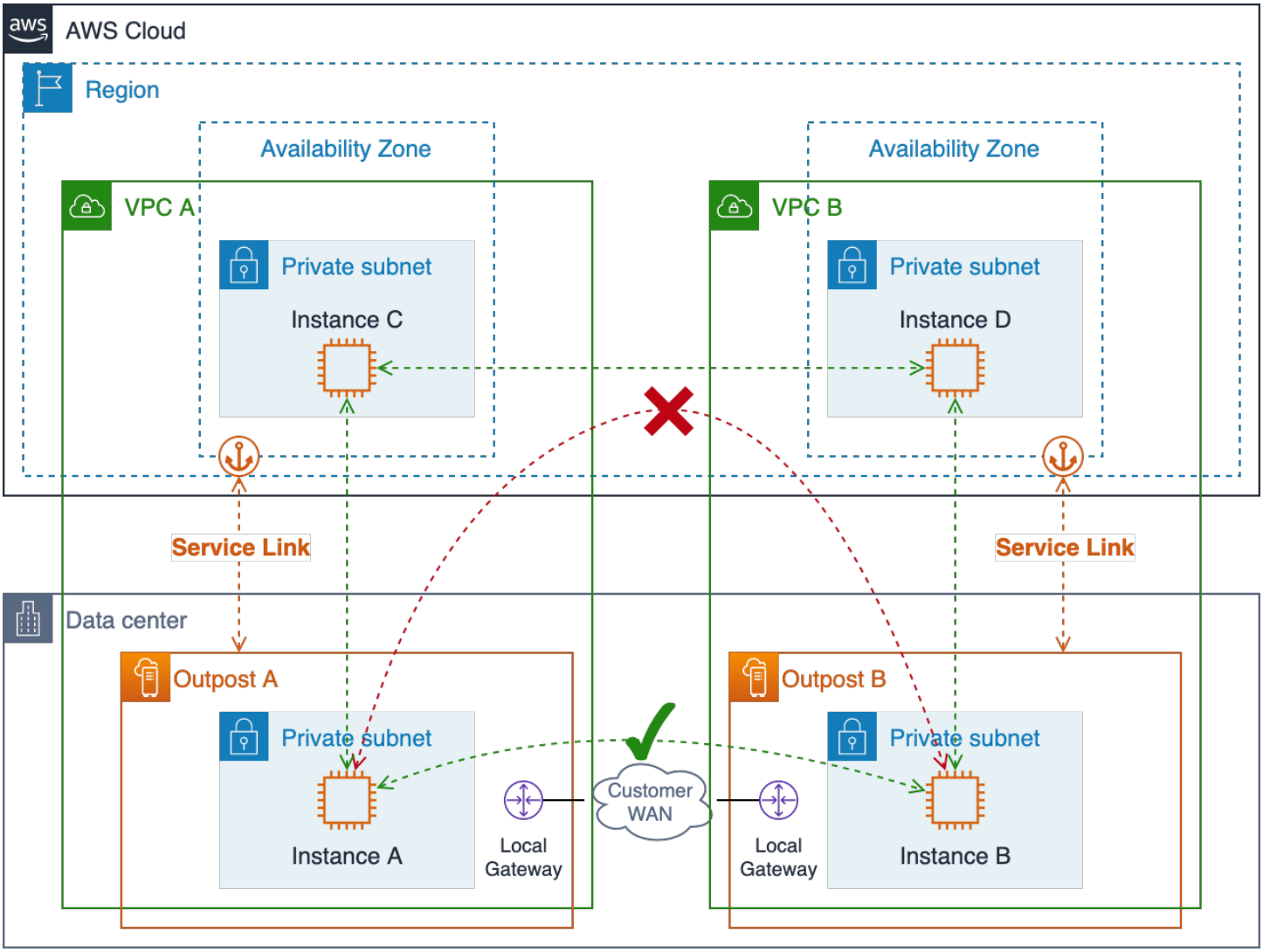


Caminhos de rede de vários Outposts de VPC única

O tráfego de Outpost para Outpost que transita pela região está bloqueado, pois esse é um antipadrão. Esse tráfego incorreria em cobranças de saída em ambas as direções e provavelmente teria uma latência muito maior do que simplesmente rotear o tráfego pela WAN do cliente.

Recursos em vários Outposts na mesma VPC não podem se comunicar entre si. O tráfego entre o Outpost na mesma VPC sempre seguirá a rota CIDR local da VPC pela região onde será bloqueado.

Você deve usar VPCs separadas para implantar recursos em vários Outposts para permitir que você direcione o tráfego de Outpost para Outpost em suas redes locais e WAN.



Caminhos de rede de várias VPCs e vários Outposts

Práticas recomendadas para proteção contra modos de falha maiores:

- Implante vários Outposts ancorados em várias AZs e regiões.
- Use VPCs separadas para cada Outpost em uma implantação de vários Outposts.

Conclusão

Com o AWS Outposts rack, você pode criar, gerenciar e escalar aplicativos locais altamente disponíveis usando AWS ferramentas e serviços familiares, como Amazon EC2, Amazon EBS, Amazon S3 on Outposts, Amazon ECS, Amazon EKS e Amazon RDS. As cargas de trabalho podem ser executadas localmente, atender clientes, acessar aplicativos e sistemas em suas redes locais e acessar o conjunto completo de serviços no. Região da AWS O rack Outposts é ideal para cargas de trabalho que precisam de acesso de baixa latência a sistemas on-premises, processamento de dados local, residência de dados e migração de aplicações com interdependências do sistema local.

Ao fornecer uma implantação do Outpost com energia, espaço e resfriamento adequados e conexões resilientes ao Região da AWS, você pode criar serviços de data center único altamente disponíveis. E, para obter níveis mais altos de disponibilidade e resiliência, você pode implantar vários Outposts e distribuir seus aplicativos entre limites lógicos e geográficos.

O rack Outposts elimina o trabalho pesado indiferenciado de criar pools locais de computação, armazenamento e rede de aplicativos e permite que você estenda o alcance da infraestrutura AWS global aos seus data centers e instalações de co-localização. Agora você pode concentrar seu tempo e energia na modernização de seus aplicativos, na simplificação de suas implantações de aplicativos e no aumento do impacto comercial de seus serviços de TI.

Colaboradores

Os colaboradores deste documento incluem:

- Mallory Gershenfeld, S3 em Outposts, Amazon Web Services
- Chris Lunsford, arquiteto sênior de soluções especializado AWS Outposts, Amazon Web Services
- Rohan Mathews, arquiteto líder AWS Outposts, Amazon Web Services

Histórico do documento

Para ser notificado sobre atualizações desse whitepaper, inscreva-se no feed RSS.

Alteração	Descrição	Data
Atualização secundária	Foram adicionadas orientaçõ es adicionais de ranhura no planejamento de capacidade.	9 de fevereiro de 2024
Atualização secundária	Atualizado para refletir os lançamentos de recursos desde a publicação inicial.	19 de julho de 2023
Atualização secundária	Práticas recomendadas atualizadas para conexão de rede altamente disponível.	29 de junho de 2023
Publicação inicial	Whitepaper publicado pela primeira vez.	12 de agosto de 2021

> **Note**
>
> Para assinar as atualizações de RSS, é preciso ter um plug-in de RSS habilitado para o navegador usado.

Avisos

Os clientes são responsáveis por fazer a própria avaliação independente das informações contidas neste documento. Este documento: (a) é apenas para fins informativos, (b) representa ofertas e práticas atuais de AWS produtos, que estão sujeitas a alterações sem aviso prévio, e (c) não cria nenhum compromisso ou garantia de AWS suas afiliadas, fornecedores ou licenciadores. AWS os produtos ou serviços são fornecidos “como estão” sem garantias, representações ou condições de qualquer tipo, expressas ou implícitas. As responsabilidades e obrigações de AWS seus clientes são controladas por AWS contratos, e este documento não faz parte nem modifica nenhum contrato entre AWS e seus clientes.

AWS Glossário

Para obter a AWS terminologia mais recente, consulte o AWS glossário na Glossário da AWS Referência.

As traduções são geradas por tradução automática. Em caso de conflito entre o conteúdo da tradução e da versão original em inglês, a versão em inglês prevalecerá.