Manual do usuário

Elastic Load Balancing

Elastic Load Balancing: Manual do usuário

Table of Contents

O que é Elastic Load Balancing? ... 1
Benefícios do balanceador de carga ... 1
Recursos do Elastic Load Balancing ... 1
Como acessar o Elastic Load Balancing ... 2
Serviços relacionados ... 2
Preços ... 3
Como o Elastic Load Balancing funciona ... 4
Zonas de disponibilidade e nós de balanceador de carga ... 4
Balanceamento de carga entre zonas ... 5
Mudança de zona ... 7
Roteamento de solicitação ... 9
Algoritmo de roteamento ... 9
Conexões HTTP ... 10
Cabeçalhos HTTP ... 11
Limites de cabeçalho HTTP ... 12
Esquema do balanceador de carga ... 12
Conexão MTU ... 13
Conceitos básicos ... 15
Criar um Application Load Balancer ... 15
Criar um Network Load Balancer ... 15
Criar um Gateway Load Balancer ... 16
Criar um Classic Load Balancer ... 16
Segurança ... 17
Proteção de dados ... 18
Criptografia em repouso ... 19
Criptografia em trânsito ... 19
Gerenciamento de identidade e acesso ... 19
Público ... 20
Autenticando com identidades ... 20
Gerenciamento do acesso usando políticas ... 24
Como o Elastic Load Balancing funciona com o IAM ... 27
Permissões de API ... 41
Permissões de API para marcação de recursos ... 44
Perfil vinculado a serviço ... 46

AWS políticas gerenciadas 48
Validação de compatibilidade 51
Resiliência 52
Segurança da infraestrutura 53
Isolamento de rede 53
Controlar o tráfego de rede 54
AWS PrivateLink 55
Criar um endpoint de interface para o Elastic Load Balancing 55
Criar uma política de endpoint da VPC para o Elastic Load Balancing 55
Migrar seu Classic Load Balancer 57
Benefícios da migração 57
Assistente de migração 58
Migração do utilitário de cópia 60
Migração manual 60
..... lxiv

O que é Elastic Load Balancing?

O Elastic Load Balancing distribui automaticamente seu tráfego de entrada entre vários destinos, como instâncias do EC2, contêineres e endereços IP, em uma ou mais zonas de disponibilidade. Ele monitora a integridade dos destinos registrados e roteia o tráfego apenas para os destinos íntegros. O Elastic Load Balancing escalona automaticamente sua capacidade de balanceador de carga em resposta a mudanças ao tráfego de entrada.

Benefícios do balanceador de carga

Um load balancer distribui cargas de trabalho para vários recursos computacionais, como servidores virtuais. Usar um load balancer aumenta a disponibilidade e a tolerância a falhas dos aplicativos.

Adicione e remova recursos computacionais do load balancer conforme mudarem suas necessidades, sem perturbar o fluxo geral de solicitações para os aplicativos.

Configure as verificações de integridade, que monitoram a integridade dos recursos computacionais, para que o load balancer envie solicitações somente para as instâncias íntegras. Também é possível descarregar o trabalho de criptografia e descriptografia no load balancer, para que os recursos computacionais possam se concentrar no trabalho principal.

Recursos do Elastic Load Balancing

O Elastic Load Balancing oferece suporte aos seguintes balanceadores de carga: balanceadores de carga da aplicação, balanceadores de carga da rede, balanceadores de carga do gateway e balanceadores de carga clássicos. Você pode selecionar o tipo de balanceador de carga que melhor se adapte às suas necessidades. Para obter mais informações, consulte Comparações de produtos .

Para obter informações sobre como usar cada balanceador de carga, consulte as documentações a seguir:

- Guia do usuário para Application Load Balancers
- Guia do usuário para Network Load Balancers
- Guia do usuário para Gateway Load Balancers
- Guia do usuário para Classic Load Balancers

Como acessar o Elastic Load Balancing

Você pode criar, acessar e gerenciar seus load balancers usando qualquer uma das interfaces a seguir:

- AWS Management Console: fornece uma interface Web que você pode usar para acessar o Elastic Load Balancing.
- AWS Command Line Interface (AWS CLI CLI): fornece comandos para um amplo conjunto de serviços da AWS, inclusive o Elastic Load Balancing. A AWS CLI é compatível com Windows, macOS e Linux. Para obter mais informações, consulte AWS Command Line Interface.
- AWS SDKs: fornecem APIs específicas da linguagem e cuidam de muitos dos detalhes da conexão, como cálculo de assinaturas, manejo com novas tentativas de solicitação e tratamento de erros. Para obter mais informações, consulte AWS SDKs.
- API de consulta: fornece ações de API de baixo nível que são chamadas usando solicitações HTTPS. Usar a API de consulta é a maneira mais direta de acessar o Elastic Load Balancing. No entanto, a API de consulta requer que o aplicativo lide com detalhes de baixo nível, como gerar o hash para assinar a solicitação e tratamento de erros. Para obter mais informações, consulte as informações a seguir.
  - Application Load Balancers e Network Load Balancers: API versão de 01/12/2015
  - Classic Load Balancers: API versão de 01/06/2012

Serviços relacionados

O Elastic Load Balancing funciona com os serviços a seguir para melhorar a disponibilidade e a escalabilidade das suas aplicações.

- Amazon EC2: servidores virtuais que executam suas aplicações na nuvem. Você pode configurar o load balancer para rotear o tráfego para suas instâncias EC2. Para obter mais informações, consulte o Guia do usuário do Amazon EC2 para instâncias do Linux ou o Guia do Amazon EC2 para instâncias do Windows.
- Amazon EC2 Auto Scaling: garante que você esteja executando o número desejado de instâncias, mesmo se uma instância falhar. O Amazon EC2 Auto Scaling também permite que você aumente ou diminua automaticamente o número de instâncias conforme a demanda nas instâncias mudar. Se você habilitar o Auto Scaling com o Elastic Load Balancing, as instâncias executadas pelo Auto Scaling serão automaticamente registradas no balanceador de carga. Da mesma forma, as instâncias que forem encerradas pelo Auto Scaling terão o registro cancelado automaticamente do

balanceador de carga. Para obter mais informações, consulte o Guia do usuário do Amazon EC2 Auto Scaling.

- AWS Certificate Manager: ao criar um receptor HTTPS, você pode especificar certificados fornecidos pelo ACM. O load balancer usa certificados para encerrar conexões e descriptografar solicitações de clientes.
- Amazon CloudWatch: permite que você monitore o seu balanceador de carga e adote as devidas medidas de acordo com a necessidade. Para obter mais informações, consulte o Guia do usuário do Amazon CloudWatch.
- Amazon ECS: permite que você execute, interrompa e gerencie contêineres do Docker em um cluster de instâncias do EC2. Você pode configurar o load balancer para rotear o tráfego para seus contêineres. Para obter mais informações, consulte o Guia do desenvolvedor do Amazon Elastic Container Service.
- AWS Global Accelerator: melhora a disponibilidade e o desempenho da sua aplicação. Use uma aceleradora para distribuir o tráfego entre vários balanceadores de carga em uma ou mais regiões da AWS. Para obter mais informações, consulte o Guia do desenvolvedor do AWS Global Accelerator.
- Route 53: fornece uma forma confiável e econômica para rotear os visitantes dos sites ao traduzir nomes de domínio em endereços IP numéricos que os computadores usam para estabelecer conexão uns com os outros. Por exemplo, ele traduziria `www.example.com` no endereço IP numérico `192.0.2.1`. A AWS atribui URLs aos seus recursos, como balanceadores de carga. No entanto, você pode querer um URL que seja fácil para seus usuários se lembrarem. Por exemplo, você pode mapear o nome de domínio a um load balancer. Para obter mais informações, consulte o Guia do desenvolvedor do Amazon Route 53.
- AWS WAF: você pode usar o AWS WAF com seu Application Load Balancer para permitir ou bloquear solicitações com base nas regras de uma lista de controle de acesso da Web (ACL da Web). Para obter mais informações, consulte o Guia do desenvolvedor do AWS WAF.

Preços

Com o load balancer, você paga somente pelo que utilizar. Para obter mais informações, consulte Preço do Elastic Load Balancing.

Como o Elastic Load Balancing funciona

Um load balancer aceita o tráfego de entrada de clientes e roteia solicitações para seus destinos registrados (como instâncias do EC2) em uma ou mais Zonas de disponibilidade. O load balancer também monitora a integridade de seus destinos registrados e roteia o tráfego apenas para destinos íntegros. Quando o load balancer detecta um destino não íntegro, ele interrompe o roteamento do tráfego para esse destino. Depois, ele retoma o roteamento do tráfego para esse destino quando detecta que o destino está íntegro novamente.

Você configura seu load balancer para aceitar o tráfego de entrada especificando um ou mais listeners. Um listener é um processo que verifica se há solicitações de conexão. Ele é configurado com um protocolo e um número de porta para as conexões de clientes com o load balancer. Da mesma forma, ele é configurado com um protocolo e um número de porta para conexões do load balancer com os destinos.

O Elastic Load Balancing é compatível com os seguintes tipos de balanceadores de carga:

- Application Load Balancers
- Network Load Balancers
- Balanceadores de carga de gateway
- Classic Load Balancers

Há uma diferença fundamental em como os tipos de balanceadores de carga são configurados. Com Application Load Balancers, Network Load Balancers e Gateway Load Balancers, você registra destinos em grupos de destino e roteia o tráfego para os grupos de destino. Com Classic Load Balancers, você registra as instâncias diretamente no balanceador de carga.

Zonas de disponibilidade e nós de balanceador de carga

Quando você habilita uma zona de disponibilidade para seu balanceador de carga, o Elastic Load Balancing cria um nó de balanceador de carga na zona de disponibilidade. Se você registrar destinos em uma Zona de disponibilidade mas não ativá-la, esses destinos registrados não receberão tráfego. O load balancer é mais eficaz se você garantir que cada zona de disponibilidade habilitada tenha pelo menos um destino registrado.

Recomendamos habilitar várias zonas de disponibilidade para todos os balanceadores de carga. No entanto, com um Application Load Balancer, é necessário que você habilite pelo menos duas

ou mais zonas de disponibilidade. Essa configuração ajuda a garantir que o load balancer possa continuar a rotear o tráfego. Se uma zona de disponibilidade ficar indisponível ou não tiver destinos íntegros, o load balancer poderá continuar a rotear o tráfego para destinos íntegros de outra zona de disponibilidade.

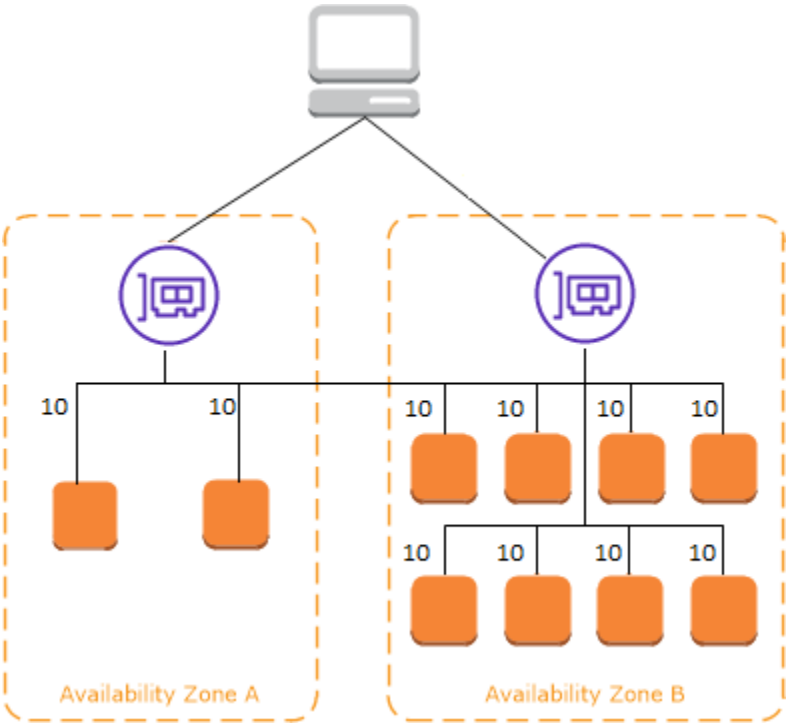
Depois de desabilitar uma zona de disponibilidade, os destinos nessa zona de disponibilidade permanecem registrados com o load balancer. No entanto, mesmo que permaneçam registrados, o load balancer não roteará o tráfego para eles.

Balanceamento de carga entre zonas

Os nós do load balancer distribuem solicitações de clientes para destinos registrados. Quando o balanceamento de carga entre zonas estiver habilitado, cada nó do load balancer distribuirá o tráfego aos destinos registrados em todas as zonas de disponibilidade habilitadas. Quando o balanceamento de carga entre zonas estiver desabilitado, cada nó do load balancer distribuirá o tráfego somente para os destinos registrados na respectiva zona de disponibilidade.

Os diagramas a seguir demonstram o efeito do balanceamento de carga entre zonas com ida e volta como o algoritmo padrão de roteamento. Há duas zonas de disponibilidade habilitadas, com dois destinos na zona de disponibilidade A e oito destinos na zona de disponibilidade B. Os clientes enviam solicitações e o Amazon Route 53 responde a cada solicitação com o endereço IP de um dos nós do balanceador de carga. Com base no algoritmo de roteamento de ida e volta, o tráfego é distribuído de modo que cada nó do balanceador de carga receba 50% do tráfego dos clientes. Cada nó de load balancer distribui a respectiva parcela de tráfego entre os destinos registrados no escopo.

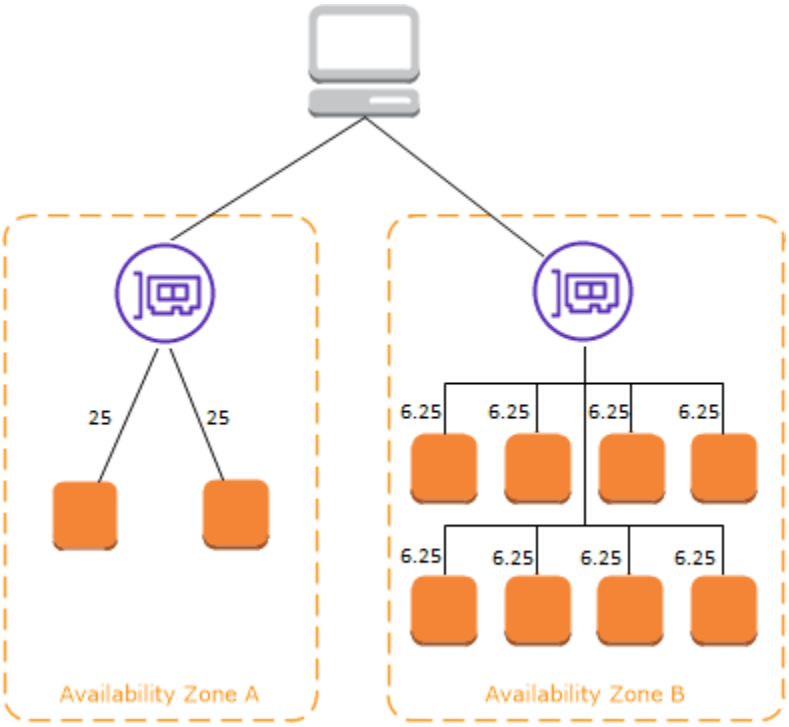
Caso o balanceamento de carga entre zonas esteja habilitado, cada um dos dez destinos recebe 10% do tráfego. Isso ocorre porque cada nó do load balancer pode rotear 50% do tráfego do cliente para todos os dez destinos.



Quando o balanceamento de carga entre zonas está desabilitado:

- Cada um dos dois destinos na zona de disponibilidade A recebe 25% do tráfego.
- Cada um dos oito destinos na zona de disponibilidade B recebe 6,25% do tráfego.

Isso ocorre porque cada nó do load balancer pode rotear 50% do tráfego do cliente apenas para destinos na respectiva zona de disponibilidade.



Com os Application Load Balancers, o balanceamento de carga entre zonas sempre está habilitado por balanceador de carga. É possível desabilitar o balanceamento de carga entre zonas por grupo de destino. Para obter mais informações, consulte Desativar o balanceamento de carga entre zonas no Guia do usuário de Application Load Balancers.

Com Network Load Balancers e Gateway Load Balancers, o balanceamento de carga entre zonas é desabilitado por padrão. Depois de criar o balanceador de carga, você pode habilitar ou desabilitar o balanceamento de carga entre zonas a qualquer momento.

Quando você cria um Classic Load Balancer, o padrão para balanceamento de carga entre zonas depende de como você cria o balanceador de carga. Com a API ou a CLI, o balanceamento de carga entre zonas é desativado por padrão. Com o AWS Management Console, a opção de ativar o balanceamento de carga entre zonas é selecionada por padrão. Depois de criar um Classic Load Balancer, você pode habilitar ou desabilitar o balanceamento de carga entre zonas a qualquer momento. Para obter mais informações, consulte Habilitar o balanceamento de carga entre zonas no Guia do usuário de Classic Load Balancers.

Mudança de zona

A mudança de zona é um recurso do Controlador de Recuperação de Aplicações do Amazon Route 53 (Route 53 ARC). Com a mudança de zona, você pode retirar um recurso do balanceador de carga

de uma zona de disponibilidade prejudicada com uma única ação. Dessa forma, é possível continuar a operar em outras zonas de disponibilidade íntegras em uma Região da AWS.

Quando você inicia uma mudança de zona, o balanceador de carga para de enviar o tráfego do recurso para a zona de disponibilidade afetada. O Route 53 ARC cria a mudança de zona imediatamente. No entanto, a efetivação das conexões existentes e em andamento na zona de disponibilidade afetada pode levar algum tempo, normalmente alguns minutos. Para obter mais informações, consulte Funcionamento da mudança de zona: verificações de integridade e endereços IP de zona no Guia do desenvolvedor do Controlador de Recuperação de Aplicações do Amazon Route 53.

As mudanças de zona só são compatíveis com Application Load Balancers e Network Load Balancers com o balanceamento de carga entre zonas desativado. Caso ative o balanceamento de carga entre zonas, você não poderá iniciar uma mudança de zona. Para obter mais informações, consulte Recursos compatíveis com mudanças de zona no Guia do desenvolvedor do Controlador de Recuperação de Aplicações do Amazon Route 53.

Antes de usar uma mudança de zona, analise o seguinte:

- O balanceamento de carga entre zonas não é compatível com mudanças de zona. Você deve desativar o balanceamento de carga entre zonas para usar esse recurso.
- A mudança de zona não é compatível quando você usa um Application Load Balancer como um endpoint do acelerador no AWS Global Accelerator.
- Você pode iniciar uma mudança de zona para um balanceador de carga específico somente para uma única zona de disponibilidade. Você não pode iniciar uma mudança de zona para várias zonas de disponibilidade.
- AWS remove proativamente os endereços IP do balanceador de carga zonal do DNS quando vários problemas de infraestrutura afetam os serviços. Antes de iniciar uma mudança de zona, sempre verifique a capacidade atual da zona de disponibilidade. Se os balanceadores de carga estiverem com o balanceamento de carga entre zonas desativado e você usar uma mudança de zona para remover o endereço IP de um balanceador de carga de zona, a zona de disponibilidade afetada pela mudança de zona também perderá a capacidade de destino.
- Quando um Application Load Balancer for o destino de um Network Load Balancer, sempre inicie a mudança de zona pelo Network Load Balancer. Se você iniciar uma mudança de zona pelo Application Load Balancer, o Network Load Balancer não reconhecerá a mudança e continuará a enviar tráfego para o Application Load Balancer.

Para obter mais orientações e informações, consulte Práticas recomendadas para mudanças de zona com o Route 53 ARC no Guia do desenvolvedor do Controlador de Recuperação de Aplicações do Amazon Route 53.

Roteamento de solicitação

Antes de um cliente enviar uma solicitação para seu load balancer, ele resolverá o nome de domínio do load balancer usando um servidor Domain Name System (DNS, Sistema de Nomes de Domínios) do servidor. A entrada do DNS é controlada pela Amazon, pois seus load balancers estão no domínio `amazonaws.com`. Os servidores DNS da Amazon retornam um ou mais endereços IP ao cliente. Esses são os endereços IP dos nós do load balancer para o seu load balancer. Com os Network Load Balancers, o Elastic Load Balancing cria uma interface de rede para cada zona de disponibilidade que você habilita e a usa para obter um endereço IP estático. Opcionalmente, você pode associar um endereço IP elástico a cada interface de rede ao criar o Network Load Balancer.

Conforme ocorram mudanças no tráfego para sua aplicação ao longo do tempo, o Elastic Load Balancing dimensionará o balanceador de carga e atualizará a entrada do DNS. A entrada DNS também especifica o time-to-live (TTL) de 60 segundos. Isso ajuda a garantir que os endereços IP possam ser remapeados rapidamente em resposta às alterações de tráfego.

O cliente determina qual endereço IP usar para enviar solicitações para o load balancer. O nó do load balancer que recebe a solicitação seleciona um destino íntegro registrado e envia a solicitação para o destino usando seu endereço IP privado.

Para obter mais informações, consulte Rotear tráfego para um balanceador de carga ELB no Guia do desenvolvedor do Amazon Route 53.

Algoritmo de roteamento

Com Application Load Balancers, o nó do balanceador de carga que recebe a solicitação aplica o seguinte processo:

1. Avalia as regras de listener em ordem de prioridade para determinar qual regra aplicar.
2. Seleciona um destino do grupo de destino para a ação da regra, usando o algoritmo de roteamento configurado para o grupo de destino. O algoritmo de roteamento padrão é o round robin. O roteamento é realizado de forma independente para cada grupo de destino, até mesmo quando um destino é registrado com vários grupos de destino.

Com Network Load Balancers, o nó do balanceador de carga que recebe a conexão aplica o seguinte processo:

1. Seleciona um destino do grupo de destino para a regra padrão usando um algoritmo de hash de fluxo. Ele baseia o algoritmo:
   - No protocolo.
   - No endereço IP de origem e na porta de origem
   - No endereço IP de destino e na porta de destino
   - No número de sequência TCP
2. Cada conexão TCP individual é roteada para um único destino durante a vida útil da conexão. As conexões TCP de um cliente têm diferentes portas de origem e números de sequência e podem ser direcionadas para destinos diferentes.

Com Classic Load Balancers, o nó do balanceador de carga que recebe a solicitação seleciona uma instância registrada da seguinte maneira:

- Usa o algoritmo de roteamento round robin para listeners TCP
- Usa o algoritmo de roteamento de solicitações menos pendentes para listeners HTTP e HTTPS

Conexões HTTP

Os Classic Load Balancers usam conexões pré-abertas, mas os Application Load Balancers não. Tanto os Classic Load Balancers quanto os Application Load Balancers usam multiplexação de conexão. Isso significa que solicitações de vários clientes em várias conexões front-end podem ser roteadas para um determinado destino por meio de uma única conexão back-end. A multiplexação de conexão melhora a latência e reduz a carga em seus aplicativos. Para evitar a multiplexação de conexão, desabilite os cabeçalhos `keep-alive` HTTP definindo o cabeçalho `Connection: close` em suas respostas HTTP.

Os Application Load Balancers e os Classic Load Balancers são compatíveis com HTTP canalizado em conexões de front-end. Eles não são compatíveis com HTTP com pipeline em conexões back-end.

Os Application Load Balancers oferecem suporte aos seguintes métodos de solicitação HTTP: GET, HEAD, POST, PUT, DELETE, OPTIONS e PATCH.

Os Application Load Balancers são compatíveis com os seguintes protocolos em conexões de front-end: HTTP/0.9, HTTP/1.0, HTTP/1.1 e HTTP/2. É possível usar HTTP/2 somente com listeners HTTPS e enviar até 128 solicitações em paralelo usando uma conexão HTTP/2. Os Application Load Balancers também oferecem suporte a atualizações de conexão de HTTP para o. WebSockets No entanto, se houver um upgrade de conexão, as regras e AWS WAF integrações de roteamento de ouvintes do Application Load Balancer não se aplicarão mais.

Os Application Load Balancers usam HTTP/1.1 em conexões de back-end (balanceador de carga para o destino registrado) por padrão. No entanto, você pode usar a versão do protocolo para enviar a solicitação aos destinos usando HTTP/2 ou gRPC. Para obter mais informações, consulte Versões de protocolo. Por padrão, o cabeçalho `keep-alive` é compatível com conexões de back-end. Para solicitações HTTP/1.0 de clientes que não tenham um cabeçalho de host, o load balancer gerará um cabeçalho de host para as solicitações HTTP/1.1 enviadas nas conexões back-end. O cabeçalho do host contém o nome DNS do balanceador de carga.

Os Classic Load Balancers são compatíveis com os seguintes protocolos em conexões de front-end (cliente para balanceador de carga): HTTP/0.9, HTTP/1.0 e HTTP/1.1. Eles usam HTTP/1.1 em conexões de back-end (balanceador de carga para destino registrado). Por padrão, o cabeçalho `keep-alive` é compatível com conexões de back-end. Para solicitações HTTP/1.0 de clientes que não tenham um cabeçalho de host, o load balancer gerará um cabeçalho de host para as solicitações HTTP/1.1 enviadas nas conexões back-end. O cabeçalho do host contém o endereço IP do nó do balanceador de carga.

Cabeçalhos HTTP

Os Application Load Balancers e Classic Load Balancers adicionam automaticamente os cabeçalhos X-Forwarded-For, X-Forwarded-Proto e X-Forwarded-Port à solicitação.

Os Application Load Balancers convertem os nomes de host nos cabeçalhos de host HTTP em letras minúsculas antes de enviá-los aos destinos.

Para conexões front-end que usam HTTP/2, os nomes de cabeçalho estão em minúsculas. Antes de enviar a solicitação ao destino usando HTTP/1.1, os seguintes nomes de cabeçalhos são convertidos para letras maiúsculas e minúsculas: X-Forwarded-For, X-Forwarded-Proto, X-Forwarded-Port, Host, X-Amzn-Trace-Id, Upgrade e Conexão. Todos os outros nomes de cabeçalho estão em minúsculas.

Os Application Load Balancers e Classic Load Balancers diferenciam o cabeçalho de conexão da solicitação de entrada do cliente após enviar um proxy da resposta de volta para o cliente.

Quando os Application Load Balancers e Classic Load Balancers que usam HTTP/1.1 recebem um cabeçalho Expect: 100-Continue, eles respondem imediatamente com HTTP/1.1 100 Continue sem testar o comprimento do cabeçalho do conteúdo. O cabeçalho da solicitação Expect: 100-Continue não é encaminhado para seus destinos.

Ao usar HTTP/2, os Application Load Balancers não são compatíveis com o cabeçalho Expect: 100-Continue das solicitações do cliente. O Application Load Balancer não responderá com HTTP/2 100 Continue nem encaminhará esse cabeçalho para seus destinos.

Limites de cabeçalho HTTP

Os seguintes limites de tamanho para Application Load Balancers são limites inflexíveis e que não podem ser alterados:

- Linha de solicitação: 16 K
- Cabeçalho único: 16 K
- Cabeçalho de resposta inteiro: 32 K
- Cabeçalho da solicitação inteira: 64 K

Esquema do balanceador de carga

Ao criar um load balancer, você deverá optar se deve fazer dele um load balancer interno ou um load balancer voltado para a Internet.

Os nós de um load balancer voltado para a Internet têm endereços IP públicos. O nome DNS de um load balancer voltado para a Internet é resolvível publicamente para os endereços IP públicos dos nós. Portanto, os load balancers voltados para a Internet podem rotear solicitações de clientes pela Internet.

Os nós de um load balancer interno têm somente endereços IP privados. O nome DNS de um load balancer interno é resolvido publicamente para os endereços IP privados dos nós. Portanto, load balancers internos só podem rotear solicitações de clientes com acesso à VPC para o load balancer.

Tanto os load balancers voltados para a Internet quanto os internos roteiam as solicitações para seus destinos usando endereços IP privados. Portanto, seus destinos não precisam de endereços IP públicos para receber solicitações de um load balancer interno ou voltado para a Internet.

Se o seu aplicativo tiver vários níveis, você poderá projetar uma arquitetura que use load balancers internos e load balancers voltados para a Internet. Por exemplo, isso é válido se o aplicativo usa

servidores da web que devem estar conectados à Internet e servidores de aplicativos que estão conectados somente aos servidores da web. Crie um load balancer voltado para a Internet e registre os servidores da web nele. Crie um load balancer interno e registre os servidores de aplicativos nele. Os servidores da web recebem solicitações do load balancer voltado para a Internet e enviam solicitações dos servidores de aplicativos para o load balancer interno. Os servidores de aplicativos recebem solicitações do load balancer interno.

MTU de rede para seu balanceador de carga

A unidade máxima de transmissão (MTU) determina o tamanho, em bytes, do maior pacote que pode ser enviado pela rede. Quanto maior a MTU de uma conexão, mais dados podem ser passados em um único pacote. Os frames de Ethernet consistem no pacote, ou nos próprios dados que você envia, e nas informações de overhead de rede que o cercam. O tráfego enviado por um gateway da Internet tem um MTU de 1500. Isso significa que, se um pacote tiver mais de 1500 bytes, ele será fragmentado para ser enviado usando vários frames ou será descartado se `Don't Fragment` estiver definido no cabeçalho IP.

Não é possível configurar o tamanho da MTU nos nós do balanceador de carga. Os frames jumbo (9.001 MTU) são padrão em todos os nós do balanceador de carga para Application Load Balancers, Network Load Balancers e Classic Load Balancers. Os balanceadores de carga de gateway são compatíveis com 8.500 MTU. Para obter mais informações, consulte Unidade máxima de transmissão (MTU) no Guia do usuário para Gateway Load Balancers.

A MTU do caminho é o tamanho máximo de pacote compatível no caminho entre o host de origem e o host receptor. A Path MTU Discovery (PMTUD – Descoberta de MTU do caminho) é usada para determinar a MTU do caminho entre dois dispositivos. A descoberta de MTU do caminho é especialmente importante se o cliente ou o destino não for compatível com frames jumbo.

Se um host enviar um pacote maior que a MTU do host receptor ou maior que a MTU de um dispositivo no caminho, o host ou dispositivo receptor descartará o pacote e retornará a seguinte mensagem ICMP: `Destination Unreachable: Fragmentation Needed and Don't Fragment was Set (Type 3, Code 4)`. Isso instrui o host de transmissão a dividir a carga útil em vários pacotes menores e retransmiti-los.

Se continuar havendo descarte de pacotes maiores que o tamanho da MTU do cliente ou da interface de destino, é provável que a descoberta de MTU do caminho (PMTUD) não esteja funcionando. Para evitar isso, certifique-se de que a descoberta de MTU do caminho esteja funcionando de ponta a ponta e que você tenha habilitado frames jumbo em seus clientes e destinos.

Para obter mais informações sobre a descoberta de MTU do caminho e a habilitação de frames jumbo, consulte Descoberta de MTU do caminho no Guia do usuário do Amazon EC2.

Como começar a usar o Elastic Load Balancing

O Elastic Load Balancing oferece suporte aos seguintes balanceadores de carga: balanceadores de carga da aplicação, balanceadores de carga da rede, balanceadores de carga do gateway e balanceadores de carga clássicos. Você pode selecionar o tipo de balanceador de carga que melhor se adapte às suas necessidades. Para obter mais informações, consulte Comparações de produtos .

Para demonstrações de configurações comuns do balanceador de carga, consulte Demonstrações do Elastic Load Balancing.

Se você tiver um Classic Load Balancer, poderá migrar para um Application Load Balancer ou um Network Load Balancer. Para obter mais informações, consulte Migrar seu Classic Load Balancer.

Índice

- Criar um Application Load Balancer
- Criar um Network Load Balancer
- Criar um Gateway Load Balancer
- Criar um Classic Load Balancer

Criar um Application Load Balancer

Para criar um Application Load Balancer usando o AWS Management Console, consulte Como começar a usar Network Load Balancers no Guia do usuário para Network Load Balancers.

Para criar um Application Load Balancer usando a AWS CLI, consulte o Criar um Application Load Balancer usando a AWS CLI no Guia do usuário para Application Load Balancers.

Criar um Network Load Balancer

Para criar um Network Load Balancer usando o AWS Management Console, consulte Como começar a usar Network Load Balancers no Guia do usuário para Network Load Balancers.

Para criar um Network Load Balancer usando a AWS CLI, consulte Criar um Network Load Balancer usando a AWS CLI no Guia do usuário para Network Load Balancers.

Criar um Gateway Load Balancer

Para criar um Gateway Load Balancer usando o AWS Management Console, consulte Como começar a usar Gateway Load Balancers no Guia do usuário para Gateway Load Balancers.

Para criar um Gateway Load Balancer usando a AWS CLI, consulte Como começar a usar Gateway Load Balancers com a AWS CLI no Guia do usuário para Gateway Load Balancers.

Criar um Classic Load Balancer

Para criar um Classic Load Balancer usando o AWS Management Console, consulte Criar um Classic Load Balancer no Guia do usuário para Classic Load Balancers.

Segurança no Elastic Load Balancing

A segurança para com a nuvem na AWS é a nossa maior prioridade. Como cliente da AWS, você se beneficiará de um datacenter e de uma arquitetura de rede criados para atender aos requisitos das empresas com as maiores exigências de segurança.

A segurança é uma responsabilidade compartilhada entre a AWS e você. O modelo de responsabilidade compartilhada descreve isso como a segurança da nuvem e a segurança na nuvem:

- Segurança da nuvem: a AWS é responsável pela proteção da infraestrutura que executa produtos da AWS na Nuvem AWS. A AWS também fornece serviços que podem ser usados com segurança. Auditores externos testam e verificam regularmente a eficácia da nossa segurança como parte dos Programas de conformidade da AWS. Para saber mais sobre os programas de conformidade que se aplicam ao Elastic Load Balancing, consulte Serviços da AWS no escopo por programa de conformidade.
- Segurança na nuvem: sua responsabilidade é determinada pelo serviço da AWS que você usa. Você também é responsável por outros fatores, incluindo a confidencialidade de seus dados, os requisitos da sua empresa e as leis e regulamentos aplicáveis.

Esta documentação ajuda a entender como aplicar o modelo de responsabilidade compartilhada ao usar o Elastic Load Balancing. Ela mostra como configurar o Elastic Load Balancing para atender aos seus objetivos de segurança e conformidade. Você também aprende como usar outros serviços da AWS que ajudam você a monitorar e proteger seus recursos do Elastic Load Balancing.

Com um Gateway Load Balancer, você é responsável por escolher e qualificar o software dos fornecedores de equipamento. Você deve confiar no software do equipamento para inspecionar ou modificar o tráfego do balanceador de carga, que opera na camada 3 do modelo Open Systems Interconnection (OSI), a camada de rede. Os fornecedores de dispositivos listados como Parceiros do Elastic Load Balancing integraram e qualificaram seu software de dispositivos com a AWS. Você pode confiar mais no software dos dispositivos dos fornecedores desta lista. No entanto, a AWS não garante a segurança ou a confiabilidade do software desses fornecedores.

Índice

- Proteção de dados no Elastic Load Balancing
- Gerenciamento de identidade e acesso para o Elastic Load Balancing

- Validação de conformidade para o Elastic Load Balancing
- Resiliência no Elastic Load Balancing
- Segurança de infraestrutura no Elastic Load Balancing
- Acessar o Elastic Load Balancing usando um endpoint de interface (AWS PrivateLink)

Proteção de dados no Elastic Load Balancing

O modelo de responsabilidade AWS compartilhada O modelo se aplica à proteção de dados no Elastic Load Balancing. Conforme descrito neste modelo, AWS é responsável por proteger a infraestrutura global que executa todos os Nuvem AWS. Você é responsável por manter o controle sobre seu conteúdo hospedado nessa infraestrutura. Você também é responsável pelas tarefas de configuração e gerenciamento de segurança dos Serviços da AWS que usa. Para ter mais informações sobre a privacidade de dados, consulte as Perguntas frequentes sobre privacidade de dados. Para ter mais informações sobre a proteção de dados na Europa, consulte a postagem do blog AWS Shared Responsibility Model and GDPR no Blog de segurança da AWS .

Para fins de proteção de dados, recomendamos que você proteja Conta da AWS as credenciais e configure usuários individuais com AWS IAM Identity Center ou AWS Identity and Access Management (IAM). Dessa maneira, cada usuário receberá apenas as permissões necessárias para cumprir suas obrigações de trabalho. Recomendamos também que você proteja seus dados das seguintes formas:

- Use uma autenticação multifator (MFA) com cada conta.
- Use SSL/TLS para se comunicar com os recursos. AWS Exigimos TLS 1.2 e recomendamos TLS 1.3.
- Configure a API e o registro de atividades do usuário com AWS CloudTrail.
- Use soluções de AWS criptografia, juntamente com todos os controles de segurança padrão Serviços da AWS.
- Use serviços gerenciados de segurança avançada, como o Amazon Macie, que ajuda a localizar e proteger dados sigilosos armazenados no Amazon S3.
- Se você precisar de módulos criptográficos validados pelo FIPS 140-2 ao acessar AWS por meio de uma interface de linha de comando ou de uma API, use um endpoint FIPS. Para ter mais informações sobre endpoints do FIPS, consulte Federal Information Processing Standard (FIPS) 140-2.

É altamente recomendável que nunca sejam colocadas informações de identificação confidenciais, como endereços de e-mail dos seus clientes, em marcações ou campos de formato livre, como um campo Nome. Isso inclui quando você trabalha com o Elastic Load Balancing ou outro Serviços da AWS usando o console, a API ou AWS os AWS CLI SDKs. Quaisquer dados inseridos em tags ou campos de texto de formato livre usados para nomes podem ser usados para logs de faturamento ou de diagnóstico. Se você fornecer um URL para um servidor externo, recomendamos fortemente que não sejam incluídas informações de credenciais no URL para validar a solicitação a esse servidor.

Criptografia em repouso

Se você habilitar a criptografia no lado do servidor com chaves de criptografia gerenciadas pelo Amazon S3 (SSE-S3) para seu bucket do S3 para logs de acesso do Elastic Load Balancing, o Elastic Load Balancing vai criptografar automaticamente cada arquivo de log de acesso antes de armazená-lo no seu bucket do S3. O Elastic Load Balancing também descriptografa os arquivos de log de acesso quando você os acessa. Cada arquivo de log é criptografado com uma chave exclusiva, que por sua vez é criptografada com uma chave KMS que é rotacionada regularmente.

Criptografia em trânsito

O Elastic Load Balancing simplifica o processo de criação de aplicações Web seguras ao encerrar o tráfego HTTPS e TLS dos clientes no balanceador de carga. O load balancer executa o trabalho de criptografar e descriptografar o tráfego, em vez de exigir que cada instância do EC2 lide com o trabalho de encerramento do TLS. Ao configurar um listener seguro, especifique os pacotes de criptografia e as versões de protocolo compatíveis com seu aplicativo e um certificado de servidor a ser instalado no load balancer. Você pode usar AWS Certificate Manager (ACM) ou AWS Identity and Access Management (IAM) para gerenciar seus certificados de servidor. Application Load Balancers são compatíveis com receptores HTTPS. Network Load Balancers são compatíveis com receptores TLS. Classic Load Balancers são compatíveis com receptores HTTPS e TLS.

Gerenciamento de identidade e acesso para o Elastic Load Balancing

AWS Identity and Access Management (IAM) é uma ferramenta AWS service (Serviço da AWS) que ajuda o administrador a controlar com segurança o acesso aos AWS recursos. Os administradores do IAM controlam quem pode ser autenticado (fazer login) e autorizado (ter permissões) para usar os recursos do Elastic Load Balancing. O IAM é um AWS service (Serviço da AWS) que você pode usar sem custo adicional.

Conteúdo

- Público
- Autenticando com identidades
- Gerenciamento do acesso usando políticas
- Como o Elastic Load Balancing funciona com o IAM
- Permissões de API do Elastic Load Balancing
- Permissões de API do Elastic Load Balancing para marcar recursos durante a criação
- Perfil vinculado a serviço para o Elastic Load Balancing
- AWS políticas gerenciadas para o Elastic Load Balancing

Público

A forma como você usa AWS Identity and Access Management (IAM) difere, dependendo do trabalho que você faz no Elastic Load Balancing.

Usuário do serviço: se você usar o serviço Elastic Load Balancing para fazer o trabalho, o administrador fornecerá as credenciais e as permissões necessárias. Conforme use mais recursos do Elastic Load Balancing para realizar seu trabalho, talvez seja necessário obter permissões adicionais. Entender como o acesso é gerenciado pode ajudar você a solicitar as permissões corretas ao seu administrador.

Administrador de serviço: se você for o responsável pelos recursos do Elastic Load Balancing na empresa, provavelmente terá acesso total ao Elastic Load Balancing. Cabe a você determinar quais funcionalidades e recursos do Elastic Load Balancing os usuários do serviço devem acessar. Assim, você deve enviar solicitações ao administrador do IAM para alterar as permissões dos usuários de seu serviço. Revise as informações nesta página para entender os Introdução ao IAM.

Administrador do IAM: se você for um administrador do IAM, talvez queira saber detalhes sobre como escrever políticas para gerenciar o acesso ao Elastic Load Balancing.

Autenticando com identidades

A autenticação é a forma como você faz login AWS usando suas credenciais de identidade. Você deve estar autenticado (conectado AWS) como o Usuário raiz da conta da AWS, como usuário do IAM ou assumindo uma função do IAM.

Você pode entrar AWS como uma identidade federada usando credenciais fornecidas por meio de uma fonte de identidade. AWS IAM Identity Center Usuários (IAM Identity Center), a autenticação de login único da sua empresa e suas credenciais do Google ou do Facebook são exemplos de identidades federadas. Quando você faz login como uma identidade federada, o administrador já configurou anteriormente a federação de identidades usando perfis do IAM. Ao acessar AWS usando a federação, você está assumindo indiretamente uma função.

Dependendo do tipo de usuário que você é, você pode entrar no AWS Management Console ou no portal de AWS acesso. Para obter mais informações sobre como fazer login em AWS, consulte Como fazer login Conta da AWS no Guia do Início de Sessão da AWS usuário.

Se você acessar AWS programaticamente, AWS fornece um kit de desenvolvimento de software (SDK) e uma interface de linha de comando (CLI) para assinar criptograficamente suas solicitações usando suas credenciais. Se você não usa AWS ferramentas, você mesmo deve assinar as solicitações. Para obter mais informações sobre como usar o método recomendado para assinar solicitações por conta própria, consulte Assinatura de solicitações de AWS API no Guia do usuário do IAM.

Independentemente do método de autenticação usado, também pode ser exigido que você forneça informações adicionais de segurança. Por exemplo, AWS recomenda que você use a autenticação multifator (MFA) para aumentar a segurança da sua conta. Para saber mais, consulte Autenticação multifator no Guia do usuário do AWS IAM Identity Center . Usar a autenticação multifator (MFA) na AWS no Guia do usuário do IAM.

Conta da AWS usuário root

Ao criar uma Conta da AWS, você começa com uma identidade de login que tem acesso completo a todos Serviços da AWS os recursos da conta. Essa identidade é chamada de usuário Conta da AWS raiz e é acessada fazendo login com o endereço de e-mail e a senha que você usou para criar a conta. É altamente recomendável não usar o usuário-raiz para tarefas diárias. Proteja as credenciais do usuário-raiz e use-as para executar as tarefas que somente ele pode executar. Para obter a lista completa das tarefas que exigem login como usuário-raiz, consulte Tarefas que exigem credenciais de usuário-raiz no Guia do usuário do IAM.

Identidade federada

Como prática recomendada, exija que usuários humanos, incluindo usuários que precisam de acesso de administrador, usem a federação com um provedor de identidade para acessar Serviços da AWS usando credenciais temporárias.

Uma identidade federada é um usuário do seu diretório de usuários corporativo, de um provedor de identidade da web AWS Directory Service, do diretório do Identity Center ou de qualquer usuário que acesse usando credenciais fornecidas Serviços da AWS por meio de uma fonte de identidade. Quando as identidades federadas são acessadas Contas da AWS, elas assumem funções, e as funções fornecem credenciais temporárias.

Para o gerenciamento de acesso centralizado, recomendamos usar o . AWS IAM Identity Center Você pode criar usuários e grupos no IAM Identity Center ou pode se conectar e sincronizar com um conjunto de usuários e grupos em sua própria fonte de identidade para uso em todos os seus Contas da AWS aplicativos. Para obter mais informações sobre o Centro de Identidade do IAM, consulte "What is IAM Identity Center?" (O que é o Centro de Identidade do IAM?) noAWS IAM Identity Center Guia do usuário do .

Grupos e usuários do IAM

Um usuário do IAM é uma identidade dentro da sua Conta da AWS que tem permissões específicas para uma única pessoa ou aplicativo. Sempre que possível, recomendamos depender de credenciais temporárias em vez de criar usuários do IAM com credenciais de longo prazo, como senhas e chaves de acesso. No entanto, se você tiver casos de uso específicos que exijam credenciais de longo prazo com usuários do IAM, recomendamos alternar as chaves de acesso. Para obter mais informações, consulte Alterne as chaves de acesso regularmente para casos de uso que exijam credenciais de longo prazo no Guia do usuário do IAM.

Um grupo do IAM é uma identidade que especifica uma coleção de usuários do IAM. Não é possível fazer login como um grupo. É possível usar grupos para especificar permissões para vários usuários de uma vez. Os grupos facilitam o gerenciamento de permissões para grandes conjuntos de usuários. Por exemplo, você pode ter um grupo chamado IAMAdmins e atribuir a esse grupo permissões para administrar atributos do IAM.

Usuários são diferentes de perfis. Um usuário é exclusivamente associado a uma pessoa ou a uma aplicação, mas um perfil pode ser assumido por qualquer pessoa que precisar dele. Os usuários têm credenciais permanentes de longo prazo, mas os perfis fornecem credenciais temporárias. Para saber mais, consulte Quando criar um usuário do IAM (em vez de um perfil) no Guia do usuário do IAM.

Perfis do IAM

Uma função do IAM é uma identidade dentro da sua Conta da AWS que tem permissões específicas. Ele é semelhante a um usuário do IAM, mas não está associado a uma pessoa específica. Você

pode assumir temporariamente uma função do IAM no AWS Management Console trocando de funções. Você pode assumir uma função chamando uma operação de AWS API AWS CLI ou usando uma URL personalizada. Para obter mais informações sobre métodos para o uso de perfis, consulte Usar perfis do IAM no Guia do usuário do IAM.

Perfis do IAM com credenciais temporárias são úteis nas seguintes situações:

- Acesso de usuário federado: para atribuir permissões a identidades federadas, você pode criar um perfil e definir permissões para ele. Quando uma identidade federada é autenticada, essa identidade é associada ao perfil e recebe as permissões definidas pelo mesmo. Para obter mais informações sobre perfis para federação, consulte Criar um perfil para um provedor de identidades de terceiros no Guia do usuário do IAM. Se você usar o IAM Identity Center, configure um conjunto de permissões. Para controlar o que suas identidades podem acessar após a autenticação, o IAM Identity Center correlaciona o conjunto de permissões a um perfil no IAM. Para obter informações sobre conjuntos de permissões, consulte Conjuntos de permissões noAWS IAM Identity Center Guia do usuário do .
- Permissões temporárias para usuários do IAM: um usuário ou um perfil do IAM pode assumir um perfil do IAM para obter temporariamente permissões diferentes para uma tarefa específica.
- Acesso entre contas: é possível usar um perfil do IAM para permitir que alguém (uma entidade principal confiável) em outra conta acesse recursos em sua conta. Os perfis são a principal forma de conceder acesso entre contas. No entanto, com alguns Serviços da AWS, você pode anexar uma política diretamente a um recurso (em vez de usar uma função como proxy). Para saber a diferença entre perfis e políticas baseadas em atributo para acesso entre contas, consulte Como os perfis do IAM diferem das políticas baseadas em atributo no Guia do usuário do IAM.
- Acesso entre serviços — Alguns Serviços da AWS usam recursos em outros Serviços da AWS. Por exemplo, quando você faz uma chamada em um serviço, é comum que esse serviço execute aplicações no Amazon EC2 ou armazene objetos no Amazon S3. Um serviço pode fazer isso usando as permissões da entidade principal de chamada, usando um perfil de serviço ou uma função vinculada ao serviço.
  - Sessões de acesso direto (FAS) — Quando você usa um usuário ou uma função do IAM para realizar ações AWS, você é considerado principal. Ao usar alguns serviços, você pode executar uma ação que inicia outra ação em um serviço diferente. O FAS usa as permissões do diretor chamando um AWS service (Serviço da AWS), combinadas com a solicitação AWS service (Serviço da AWS) para fazer solicitações aos serviços posteriores. As solicitações do FAS são feitas somente quando um serviço recebe uma solicitação que requer interações com outros Serviços da AWS ou com recursos para ser concluída. Nesse caso, você precisa ter permissões

para executar ambas as ações. Para obter detalhes da política ao fazer solicitações de FAS, consulte Encaminhar sessões de acesso.

- Perfil de serviço: um perfil de serviço é um perfil do IAMhttps://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles.html que um serviço assume para realizar ações em seu nome. Um administrador do IAM pode criar, modificar e excluir um perfil de serviço do IAM. Para obter mais informações, consulte Criar um perfil para delegar permissões a um AWS service (Serviço da AWS) no Guia do usuário do IAM.
- Função vinculada ao serviço — Uma função vinculada ao serviço é um tipo de função de serviço vinculada a um. AWS service (Serviço da AWS) O serviço pode assumir o perfil de executar uma ação em seu nome. As funções vinculadas ao serviço aparecem em você Conta da AWS e são de propriedade do serviço. Um administrador do IAM pode visualizar, mas não pode editar as permissões para perfis vinculados ao serviço.

- Aplicativos em execução no Amazon EC2 — Você pode usar uma função do IAM para gerenciar credenciais temporárias para aplicativos que estão sendo executados em uma instância do EC2 e fazendo AWS CLI solicitações de API. AWS É preferível fazer isso a armazenar chaves de acesso na instância do EC2. Para atribuir uma AWS função a uma instância do EC2 e disponibilizá-la para todos os seus aplicativos, você cria um perfil de instância anexado à instância. Um perfil de instância contém a perfil e permite que os programas em execução na instância do EC2 obtenham credenciais temporárias. Para mais informações, consulte Usar um perfil do IAM para conceder permissões a aplicações em execução nas instâncias do Amazon EC2 no Guia do usuário do IAM.

Para saber se deseja usar as funções do IAM, consulte Quando criar um perfil do IAM (em vez de um usuário) no Guia do usuário do IAM.

Gerenciamento do acesso usando políticas

Você controla o acesso AWS criando políticas e anexando-as a AWS identidades ou recursos. Uma política é um objeto AWS que, quando associada a uma identidade ou recurso, define suas permissões. AWS avalia essas políticas quando um principal (usuário, usuário raiz ou sessão de função) faz uma solicitação. As permissões nas políticas determinam se a solicitação será permitida ou negada. A maioria das políticas é armazenada AWS como documentos JSON. Para obter mais informações sobre a estrutura e o conteúdo de documentos de políticas JSON, consulte Visão geral das políticas JSON no Guia do usuário do IAM.

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual entidade principal pode executar ações em quais recursos e em que condições.

Por padrão, usuários e funções não têm permissões. Para conceder aos usuários permissão para executar ações nos recursos de que eles precisam, um administrador do IAM pode criar políticas do IAM. O administrador pode então adicionar as políticas do IAM a perfis, e os usuários podem assumir os perfis.

As políticas do IAM definem permissões para uma ação, independentemente do método usado para executar a operação. Por exemplo, suponha que você tenha uma política que permite a ação `iam:GetRole`. Um usuário com essa política pode obter informações de função da AWS Management Console AWS CLI, da ou da AWS API.

Políticas baseadas em identidade

As políticas baseadas em identidade são documentos de políticas de permissões JSON que você pode anexar a uma identidade, como usuário do IAM, grupo de usuários ou perfil do IAM. Essas políticas controlam quais ações os usuários e funções podem realizar, em quais recursos e em que condições. Para saber como criar uma política baseada em identidade, consulte Criar políticas do IAM no Guia do usuário do IAM.

As políticas baseadas em identidade podem ser categorizadas ainda mais como políticas em linha ou políticas gerenciadas. As políticas em linha são anexadas diretamente a um único usuário, grupo ou perfil. As políticas gerenciadas são políticas autônomas que você pode associar a vários usuários, grupos e funções em seu Conta da AWS. As políticas AWS gerenciadas incluem políticas gerenciadas e políticas gerenciadas pelo cliente. Para saber como escolher entre uma política gerenciada ou uma política em linha, consulte Escolher entre políticas gerenciadas e políticas em linha no Guia do usuário do IAM.

Políticas baseadas em recursos

Políticas baseadas em recursos são documentos de políticas JSON que você anexa a um recurso. São exemplos de políticas baseadas em recursos as políticas de confiança de perfil do IAM e as políticas de bucket do Amazon S3. Em serviços compatíveis com políticas baseadas em recursos, os administradores de serviço podem usá-las para controlar o acesso a um recurso específico. Para o atributo ao qual a política está anexada, a política define quais ações uma entidade principal especificada pode executar nesse atributo e em que condições. Você deve especificar uma entidade principal em uma política baseada em recursos. Os diretores podem incluir contas, usuários, funções, usuários federados ou. Serviços da AWS

Políticas baseadas em atributos são políticas em linha que estão localizadas nesse serviço. Você não pode usar políticas AWS gerenciadas do IAM em uma política baseada em recursos.

Listas de controle de acesso (ACLs)

As listas de controle de acesso (ACLs) controlam quais entidades principais (membros, usuários ou funções da conta) têm permissões para acessar um recurso. As ACLs são semelhantes às políticas baseadas em recursos, embora não usem o formato de documento de política JSON.

O Amazon S3 e o Amazon VPC são exemplos de serviços que oferecem suporte a ACLs. AWS WAF Para saber mais sobre ACLs, consulte Visão geral da lista de controle de acesso (ACL) no Guia do desenvolvedor do Amazon Simple Storage Service.

Outros tipos de política

AWS oferece suporte a tipos de políticas adicionais menos comuns. Esses tipos de política podem definir o máximo de permissões concedidas a você pelos tipos de política mais comuns.

- Limites de permissões: um limite de permissões é um recurso avançado no qual você define o máximo de permissões que uma política baseada em identidade pode conceder a uma entidade do IAM (perfil ou usuário do IAM). É possível definir um limite de permissões para uma entidade. As permissões resultantes são a interseção das políticas baseadas em identidade de uma entidade e dos seus limites de permissões. As políticas baseadas em atributo que especificam o usuário ou o perfil no campo `Principal` não são limitadas pelo limite de permissões. Uma negação explícita em qualquer uma dessas políticas substitui a permissão. Para obter mais informações sobre limites de permissões, consulte Limites de permissões para identidades do IAM no Guia do usuário do IAM.
- Políticas de controle de serviço (SCPs) — SCPs são políticas JSON que especificam as permissões máximas para uma organização ou unidade organizacional (OU) em. AWS Organizations AWS Organizations é um serviço para agrupar e gerenciar centralmente vários Contas da AWS que sua empresa possui. Se você habilitar todos os atributos em uma organização, poderá aplicar políticas de controle de serviço (SCPs) a qualquer uma ou a todas as contas. O SCP limita as permissões para entidades nas contas dos membros, incluindo cada uma Usuário raiz da conta da AWS. Para obter mais informações sobre o Organizações e SCPs, consulte Como os SCPs funcionam no Guia do usuário do AWS Organizations .
- Políticas de sessão: são políticas avançadas que você transmite como um parâmetro quando cria de forma programática uma sessão temporária para um perfil ou um usuário federado. As permissões da sessão resultante são a interseção das políticas baseadas em identidade do usuário ou do perfil e das políticas de sessão. As permissões também podem ser provenientes de uma política baseada em atributo. Uma negação explícita em qualquer uma dessas políticas

substitui a permissão. Para obter mais informações, consulte Políticas de sessão no Guia do usuário do IAM.

Vários tipos de política

Quando vários tipos de política são aplicáveis a uma solicitação, é mais complicado compreender as permissões resultantes. Para saber como AWS determinar se uma solicitação deve ser permitida quando vários tipos de políticas estão envolvidos, consulte Lógica de avaliação de políticas no Guia do usuário do IAM.

Como o Elastic Load Balancing funciona com o IAM

Antes de usar o IAM para gerenciar o acesso ao Elastic Load Balancing, conheça quais recursos do IAM estão disponíveis para uso com o Elastic Load Balancing.

Recursos do IAM que você pode usar com o Elastic Load Balancing

Atributo do IAM	Compatibilidade com o Elastic Load Balancing
Políticas baseadas em identidade	Sim
Políticas baseadas em recursos	Não
Ações de políticas	Sim
atributos de políticas	Sim
Chaves de condição de política (específicas do serviço)	Sim
ACLs	Não
ABAC (etiquetas em políticas)	Sim
Credenciais temporárias	Sim
Permissões de entidade principal	Sim
Perfis de serviço	Não
Funções vinculadas ao serviço	Sim

Políticas baseadas em identidade para o Elastic Load Balancing

É compatível com políticas baseadas em identidade	Sim

As políticas baseadas em identidade são documentos de políticas de permissões JSON que você pode anexar a uma identidade, como usuário, grupo de usuários ou perfil do IAM. Essas políticas controlam quais ações os usuários e funções podem realizar, em quais recursos e em que condições. Para saber como criar uma política baseada em identidade, consulte Criar políticas do IAM no Guia do usuário do IAM.

Com as políticas baseadas em identidade do IAM, é possível especificar ações ou atributos permitidos ou negados, bem como as condições sob as quais as ações são permitidas ou negadas. Não é possível especificar a entidade principal em uma política baseada em identidade porque ela se aplica ao usuário ou função à qual ela está anexado. Para saber mais sobre todos os elementos que podem ser usados em uma política JSON, consulte Referência de elementos da política JSON do IAM no Guia do Usuário do IAM.

Políticas baseadas em recursos no Elastic Load Balancing

Oferece suporte a políticas baseadas em recursos	Não

Políticas baseadas em recursos são documentos de políticas JSON que você anexa a um recurso. São exemplos de políticas baseadas em recursos as políticas de confiança de perfil do IAM e as políticas de bucket do Amazon S3. Em serviços compatíveis com políticas baseadas em recursos, os administradores de serviço podem usá-las para controlar o acesso a um recurso específico. Para o atributo ao qual a política está anexada, a política define quais ações uma entidade principal especificada pode executar nesse atributo e em que condições. Você deve especificar uma entidade principal em uma política baseada em recursos. Os diretores podem incluir contas, usuários, funções, usuários federados ou. Serviços da AWS

Para permitir o acesso entre contas, você pode especificar uma conta inteira ou as entidades do IAM em outra conta como a entidade principal em uma política baseada em atributo. Adicionar uma entidade principal entre contas à política baseada em atributo é apenas metade da tarefa de estabelecimento da relação de confiança. Quando o principal e o recurso são diferentes Contas da

AWS, um administrador do IAM na conta confiável também deve conceder permissão à entidade principal (usuário ou função) para acessar o recurso. Eles concedem permissão ao anexar uma política baseada em identidade para a entidade. No entanto, se uma política baseada em atributo conceder acesso a uma entidade principal na mesma conta, nenhuma política baseada em identidade adicional será necessária. Para obter mais informações, consulte Como os perfis do IAM diferem de políticas baseadas em recursos no Guia do usuário do IAM.

Ações de política para o Elastic Load Balancing

Oferece suporte a ações de políticas	Sim

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual entidade principal pode executar ações em quais recursos, e em que condições.

O elemento `Action` de uma política JSON descreve as ações que você pode usar para permitir ou negar acesso em uma política. As ações de política geralmente têm o mesmo nome da operação de AWS API associada. Existem algumas exceções, como ações somente de permissão, que não têm uma operação de API correspondente. Há também algumas operações que exigem várias ações em uma política. Essas ações adicionais são chamadas de ações dependentes.

Incluem ações em uma política para conceder permissões para executar a operação associada.

Para ver uma lista das ações do Elastic Load Balancing, consulte Ações definidas pelo Elastic Load Balancing na Referência de autorização do serviço.

As ações de política no Elastic Load Balancing usam o seguinte prefixo antes da ação:

```
elasticloadbalancing
```

Para especificar várias ações em uma única instrução, separe-as com vírgulas.

```
"Action": [
    "elasticloadbalancing:action1",
    "elasticloadbalancing:action2"
]
```

Você também pode especificar várias ações usando caracteres-curinga (*). Por exemplo, para especificar todas as ações que começam com a palavra `Describe`, inclua a seguinte ação:

```
"Action": "elasticloadbalancing:Describe*"
```

Para ver a lista completa de todas as ações de API para o Elastic Load Balancing, consulte a documentação a seguir:

- Application Load Balancers, Network Load Balancers e Gateway Load Balancers: referência de API versão de 01/12/2015
- Classic Load Balancers: referência de API versão de 01/06/2012

Para obter mais informações sobre as permissões exigidas por cada ação do Elastic Load Balancing, consulte Permissões de API do Elastic Load Balancing.

Recursos de política para o Elastic Load Balancing

Oferece suporte a atributos de políticas	Sim

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual entidade principal pode executar ações em quais recursos, e em que condições.

O elemento `Resource` de política JSON especifica o objeto ou os objetos aos quais a ação se aplica. As instruções devem incluir um elemento `Resource` ou um elemento `NotResource`. Como prática recomendada, especifique um recurso usando Nome do recurso da Amazon (ARN). Isso pode ser feito para ações que oferecem suporte a um tipo de atributo específico, conhecido como permissões em nível de atributo.

Para ações não compatíveis com permissões no nível de recurso, como operações de listagem, use um curinga (*) para indicar que a instrução se aplica a todos os recursos.

```
"Resource": "*"
```

Algumas ações de API do Elastic Load Balancing são compatíveis com vários recursos. Para especificar vários recursos em uma única instrução, separe os ARNs com vírgulas.

```
"Resource": [
    "resource1",
    "resource2"
]
```

Para obter uma lista dos tipos de recursos do Elastic Load Balancing e seus ARNs, consulte Recursos definidos pelo Elastic Load Balancing na Referência de autorização do serviço. Para saber com quais ações é possível especificar o ARN de cada recurso, consulte Ações definidas pelo Elastic Load Balancing.

Chaves de condição de política para o Elastic Load Balancing

Compatível com chaves de condição de política específicas do serviço	Sim

Os administradores podem usar políticas AWS JSON para especificar quem tem acesso ao quê. Ou seja, qual principal pode executar ações em quais recursos, e em que condições.

O elemento `Condition` (ou `Condition` bloco de) permite que você especifique condições nas quais uma instrução está em vigor. O elemento `Condition` é opcional. É possível criar expressões condicionais que usam agentes de condição, como “igual a” ou “menor que”, para fazer a condição da política corresponder aos valores na solicitação.

Se você especificar vários elementos `Condition` em uma instrução ou várias chaves em um único `Condition` elemento, a AWS os avaliará usando uma operação lógica `AND`. Se você especificar vários valores para uma única chave de condição, AWS avalia a condição usando uma `OR` operação lógica. Todas as condições devem ser atendidas para que as permissões da instrução sejam concedidas.

Você também pode usar variáveis de espaço reservado ao especificar as condições. Por exemplo, é possível conceder a um usuário do IAM permissão para acessar um atributo somente se ele estiver marcado com seu nome de usuário do IAM. Para obter mais informações, consulte Elementos de política do IAM: variáveis e tags no Guia do usuário do IAM.

AWS suporta chaves de condição globais e chaves de condição específicas do serviço. Para ver todas as chaves de condição AWS globais, consulte as chaves de contexto de condição AWS global no Guia do usuário do IAM.

Para ver uma lista de chaves de condição do Elastic Load Balancing, consulte Chaves de condição do Elastic Load Balancing na Referência de autorização do serviço. Para saber com quais ações e recursos é possível usar uma chave de condição, consulte Ações definidas pelo Elastic Load Balancing.

Chave da condição **elasticloadbalancing:ResourceTag**

A chave de condição`elasticloadbalancing:ResourceTag/`*`key`* é específica do Elastic Load Balancing. As ações a seguir oferecem suporte a essa chave de condição:

Versão da API de 01/12/2015

- `AddTags`
- `CreateListener`
- `CreateLoadBalancer`
- `DeleteLoadBalancer`
- `DeleteTargetGroup`
- `DeregisterTargets`
- `ModifyLoadBalancerAttributes`
- `ModifyTargetGroup`
- `ModifyTargetGroupAttributes`
- `RegisterTargets`
- `RemoveTags`
- `SetIpAddressType`
- `SetSecurityGroups`
- `SetSubnets`

API versão de 01/06/2012

- `AddTags`
- `ApplySecurityGroupsToLoadBalancer`
- `AttachLoadBalancersToSubnets`
- `ConfigureHealthCheck`
- `CreateAppCookieStickinessPolicy`
- `CreateLBCookieStickinessPolicy`
- `CreateLoadBalancer`
- `CreateLoadBalancerListeners`

- `CreateLoadBalancerPolicy`
- `DeleteLoadBalancer`
- `DeleteLoadBalancerListeners`
- `DeleteLoadBalancerPolicy`
- `DeregisterInstancesFromLoadBalancer`
- `DetachLoadBalancersFromSubnets`
- `DisableAvailabilityZonesForLoadBalancer`
- `EnableAvailabilityZonesForLoadBalancer`
- `ModifyLoadBalancerAttributes`
- `RegisterInstancesWithLoadBalancer`
- `RemoveTags`
- `SetLoadBalancerListenerSSLCertificate`
- `SetLoadBalancerPoliciesForBackendServer`
- `SetLoadBalancerPoliciesOfListener`

Chave da condição **`elasticloadbalancing:ListenerProtocol`**

A chave de `elasticloadbalancing:ListenerProtocol` condição pode ser usada para condições que definem os tipos de ouvintes que podem ser criados e usados. As ações a seguir oferecem suporte a essa chave de condição:

Versão da API de 01/12/2015

- `CreateListener`
- `ModifyListener`

API versão de 01/06/2012

- `CreateLoadBalancer`
- `CreateLoadBalancerListeners`

A política está disponível para balanceadores de carga de aplicativos, balanceadores de carga de rede e balanceadores de carga clássicos. Veja a seguir um exemplo de política que só permite que os usuários selecionem um dos protocolos especificados para seu ouvinte.

Protocolos compatíveis:

- HTTPS
- HTTP
- TCP
- SSL
- TLS
- UDP
- TCP_UDP

```
"Version": "2015-12-01",
    "Statement": {"Effect": "Allow",
        "Action": [
            "elasticloadbalancing:CreateListener",
            "elasticloadbalancing:ModifyListener"
        ],
        "Resource": "*",
        "Condition": {
            "StringEquals":{
                    "elasticloadbalancing:ListenerProtocol": [
                    "HTTPS",
                    "TLS"
                    ]
            },
        }
```

Chave da condição **`elasticloadbalancing:SecurityPolicy`**

A chave de `elasticloadbalancing:SecurityPolicy` condição pode ser usada para condições que definem e impõem políticas de segurança específicas nos balanceadores de carga. As ações a seguir oferecem suporte a essa chave de condição:

Versão da API de 01/12/2015

- `CreateListener`
- `ModifyListener`

API versão de 01/06/2012

- `CreateLoadBalancerPolicy`
- `SetLoadBalancerPoliciesOfListener`

A política está disponível para balanceadores de carga de aplicativos, balanceadores de carga de rede e balanceadores de carga clássicos. Veja a seguir um exemplo de política que só permite que os usuários selecionem uma das políticas de segurança especificadas para o balanceador de carga.

```
"Resource": [
"Version": "2015-12-01",
    "Statement": {"Effect": "Allow",
        "Action": [
            "elasticloadbalancing:CreateListener",
            "elasticloadbalancing:ModifyListener"
        ],
        "Resource": "*",
        "Condition": {
            "ForAnyValue:StringEquals":{
                "elasticloadbalancing:SecurityPolicy": [
                    "ELBSecurityPolicy-TLS13-1-2-2021-06",
                    "ELBSecurityPolicy-TLS13-1-2-Res-2021-06",
                    "ELBSEcurityPolicy-TLS13-1-1-2021-06"
                ]
            },
        }
]
```

Chave da condição **`elasticloadbalancing:Scheme`**

A chave de `elasticloadbalancing:Scheme` condição pode ser usada para condições que definem qual esquema pode ser selecionado durante a criação do balanceador de carga. As ações a seguir oferecem suporte a essa chave de condição:

Versão da API de 01/12/2015

- `CreateLoadBalancer`

API versão de 01/06/2012

- `CreateLoadBalancer`

A política está disponível para balanceadores de carga de aplicativos, balanceadores de carga de rede e balanceadores de carga clássicos. Veja a seguir um exemplo de política que permite que os usuários selecionem apenas um dos esquemas especificados para o balanceador de carga.

```
"Version": "2015-12-01",
    "Statement": {"Effect": "Allow",
        "Action":"elasticloadbalancing:CreateLoadBalancer",
        "Resource": "*",
        "Condition": {
            "StringEquals":{
                "elasticloadbalancing:Scheme": "internal"
            },
        }
```

Chave da condição **`elasticloadbalancing:Subnet`**

> **Important**
>
> O Elastic Load Balancing aceita todas as capitalizações de IDs de sub-rede. No entanto, certifique-se de usar os operadores de condições adequados que não diferenciam maiúsculas e minúsculas, por exemplo`StringEqualsIgnoreCase`.

A chave de `elasticloadbalancing:Subnet` condição pode ser usada para condições que definem quais sub-redes podem ser criadas e anexadas aos balanceadores de carga. As ações a seguir oferecem suporte a essa chave de condição:

Versão da API de 01/12/2015

- `CreateLoadBalancer`
- `SetSubnets`

API versão de 01/06/2012

- `CreateLoadBalancer`
- `AttachLoadBalancerToSubnets`

A política está disponível para balanceadores de carga de aplicativos, balanceadores de carga de rede, balanceadores de carga de gateway e balanceadores de carga clássicos. Veja a seguir um

exemplo de política que só permite que os usuários selecionem uma das sub-redes especificadas para o balanceador de carga.

```
"Version": "2015-12-01",
    "Statement": {"Effect": "Allow",
        "Action": [
            "elasticloadbalancing:CreateLoadBalancer",
            "elasticloadbalancing:SetSubnets"
        ],
        "Resource": "*",
        "Condition": {
            "ForAnyValue:StringEqualsIgnoreCase":{
                "elasticloadbalancing:Subnet": [
                    "subnet-01234567890abcdef",
                    "subnet-01234567890abcdeg "
                ]
            },
        }
```

Chave da condição **`elasticloadbalancing:SecurityGroup`**

> **Important**
>
> O Elastic Load Balancing aceita todas as capitalizações de IDs. SecurityGroup No entanto, certifique-se de usar os operadores de condições adequados que não diferenciam maiúsculas e minúsculas, por exemplo`StringEqualsIgnoreCase`.

A chave de `elasticloadbalancing:SecurityGroup` condição pode ser usada para condições que definem quais grupos de segurança podem ser aplicados aos balanceadores de carga. As ações a seguir oferecem suporte a essa chave de condição:

Versão da API de 01/12/2015

- `CreateLoadBalancer`
- `SetSecurityGroups`

API versão de 01/06/2012

- `CreateLoadBalancer`

- `ApplySecurityGroupsToLoadBalancer`

A política está disponível para balanceadores de carga de aplicativos, balanceadores de carga de rede e balanceadores de carga clássicos. Veja a seguir um exemplo de política que só permite que os usuários selecionem um dos grupos de segurança especificados para o balanceador de carga.

```
"Version": "2015-12-01",
    "Statement": {"Effect": "Allow",
        "Action": [
            "elasticloadbalancing:CreateLoadBalancer",
            "elasticloadbalancing:SetSecurityGroup"
        ],
        "Resource": "*",
        "Condition": {
            "ForAnyValue:StringEqualsIgnoreCase":{
                "elasticloadbalancing:SecurityGroup": [
                    "sg-51530134",
                    "sg-51530144",
                    "sg-51530139"
                ]
            },
        }
```

ACLs no Elastic Load Balancing

Oferece suporte a ACLs	Não

As listas de controle de acesso (ACLs) controlam quais entidades principais (membros, usuários ou funções da conta) têm permissões para acessar um recurso. As ACLs são semelhantes às políticas baseadas em recursos, embora não usem o formato de documento de política JSON.

ABAC com o Elastic Load Balancing

Oferece suporte a ABAC (tags em políticas)	Sim

O controle de acesso por atributo (ABAC) é uma estratégia de autorização que define permissões com base em atributos. Em AWS, esses atributos são chamados de tags. Você pode anexar tags

a entidades do IAM (usuários ou funções) e a vários AWS recursos. A marcação de entidades e atributos é a primeira etapa do ABAC. Em seguida, você cria políticas de ABAC para permitir operações quando a tag da entidade principal corresponder à tag do atributo que ela está tentando acessar.

O ABAC é útil em ambientes que estão crescendo rapidamente e ajuda em situações em que o gerenciamento de políticas se torna um problema.

Para controlar o acesso baseado em tags, forneça informações sobre as tags no elemento de condição de uma política usando as chaves de condição `aws:ResourceTag/key-name`, `aws:RequestTag/key-name` ou `aws:TagKeys`.

Se um serviço oferecer suporte às três chaves de condição para cada tipo de recurso, o valor será Sim para o serviço. Se um serviço oferecer suporte às três chaves de condição somente para alguns tipos de recursos, o valor será Parcial.

Para obter mais informações sobre o ABAC, consulte O que é ABAC? no Guia do usuário do IAM. Para visualizar um tutorial com etapas para configurar o ABAC, consulte Usar controle de acesso baseado em atributos (ABAC) no Guia do usuário do IAM.

Como usar credenciais temporárias com o Elastic Load Balancing

Oferece suporte a credenciais temporárias	Sim

Alguns Serviços da AWS não funcionam quando você faz login usando credenciais temporárias. Para obter informações adicionais, incluindo quais Serviços da AWS funcionam com credenciais temporárias, consulte Serviços da AWS “Trabalhe com o IAM” no Guia do usuário do IAM.

Você está usando credenciais temporárias se fizer login AWS Management Console usando qualquer método, exceto um nome de usuário e senha. Por exemplo, quando você acessa AWS usando o link de login único (SSO) da sua empresa, esse processo cria automaticamente credenciais temporárias. Você também cria automaticamente credenciais temporárias quando faz login no console como usuário e, em seguida, alterna perfis. Para obter mais informações sobre como alternar perfis, consulte Alternar para um perfil (console) no Guia do usuário do IAM.

Você pode criar manualmente credenciais temporárias usando a AWS API AWS CLI ou. Em seguida, você pode usar essas credenciais temporárias para acessar AWS. AWS recomenda que você gere

credenciais temporárias dinamicamente em vez de usar chaves de acesso de longo prazo. Para mais informações, consulte Credenciais de segurança temporárias no IAM.

Permissões de entidade principal entre serviços para o Elastic Load Balancing

Suporte para o recurso Encaminhamento de sessões de acesso (FAS)	Sim

Quando você usa um usuário ou uma função do IAM para realizar ações AWS, você é considerado um principal. Ao usar alguns serviços, você pode executar uma ação que inicia outra ação em um serviço diferente. O FAS usa as permissões do diretor chamando um AWS service (Serviço da AWS), combinadas com a solicitação AWS service (Serviço da AWS) para fazer solicitações aos serviços posteriores. As solicitações do FAS são feitas somente quando um serviço recebe uma solicitação que requer interações com outros Serviços da AWS ou com recursos para ser concluída. Nesse caso, você precisa ter permissões para executar ambas as ações. Para obter detalhes da política ao fazer solicitações de FAS, consulte Encaminhar sessões de acesso.

Perfis de serviço para o Elastic Load Balancing

Oferece suporte a perfis de serviço	Não

O perfil de serviço é um perfil do IAMhttps://docs.aws.amazon.com/IAM/latest/UserGuide/id_roles.html que um serviço assume para realizar ações em seu nome. Um administrador do IAM pode criar, modificar e excluir um perfil de serviço do IAM. Para obter mais informações, consulte Criar um perfil para delegar permissões a um AWS service (Serviço da AWS) no Guia do usuário do IAM.

Perfis vinculados a serviço para o Elastic Load Balancing

Oferece suporte a perfis vinculados ao serviço	Sim

Uma função vinculada ao serviço é um tipo de função de serviço vinculada a um. AWS service (Serviço da AWS) O serviço pode assumir o perfil de executar uma ação em seu nome. As funções vinculadas ao serviço aparecem em você Conta da AWS e são de propriedade do serviço. Um

administrador do IAM pode visualizar, mas não pode editar as permissões para perfis vinculados ao serviço.

Para obter detalhes sobre como criar ou gerenciar perfis vinculadas a serviço do Elastic Load Balancing, consulte Perfil vinculado a serviço para o Elastic Load Balancing.

Permissões de API do Elastic Load Balancing

Você deve conceder permissões para que os usuários chamem as ações de API do Elastic Load Balancing de que precisam. Além disso, para algumas ações do Elastic Load Balancing, você deve conceder permissões aos usuários para chamar ações específicas da API do Amazon EC2.

Permissões necessárias para a API de 01/12/2015

Ao chamar as seguintes ações da API de 01/12/2015, você deve conceder permissões aos usuários para chamar as ações especificadas.

`CreateLoadBalancer`

- `elasticloadbalancing:CreateLoadBalancer`
- `ec2:DescribeAccountAttributes`
- `ec2:DescribeAddresses`
- `ec2:DescribeInternetGateways`
- `ec2:DescribeSecurityGroups`
- `ec2:DescribeSubnets`
- `ec2:DescribeVpcs`
- `iam:CreateServiceLinkedRole`

`CreateTargetGroup`

- `elasticloadbalancing:CreateTargetGroup`
- `ec2:DescribeInternetGateways`
- `ec2:DescribeVpcs`

`RegisterTargets`

- `elasticloadbalancing:RegisterTargets`
- `ec2:DescribeInstances`

- ec2:DescribeInternetGateways
- ec2:DescribeSubnets
- ec2:DescribeVpcs

SetIpAddressType

- elasticloadbalancing:SetIpAddressType
- ec2:DescribeSubnets

SetSubnets

- elasticloadbalancing:SetSubnets
- ec2:DescribeSubnets

Permissões necessárias para a API de 01/06/2012

Ao chamar as seguintes ações da API de 01/06/2012, você deve conceder permissões aos usuários para chamar as ações especificadas.

ApplySecurityGroupsToLoadBalancer

- elasticloadbalancing:ApplySecurityGroupsToLoadBalancer
- ec2:DescribeAccountAttributes
- ec2:DescribeSecurityGroups

AttachLoadBalancerToSubnets

- elasticloadbalancing:AttachLoadBalancerToSubnets
- ec2:DescribeSubnets

CreateLoadBalancer

- elasticloadbalancing:CreateLoadBalancer
- ec2:CreateSecurityGroup
- ec2:DescribeAccountAttributes
- ec2:DescribeInternetGateways
- ec2:DescribeSecurityGroups
- ec2:DescribeSubnets
- ec2:DescribeVpcs

- iam:CreateServiceLinkedRole

DeregisterInstancesFromLoadBalancer

- elasticloadbalancing:DeregisterInstancesFromLoadBalancer
- ec2:DescribeClassicLinkInstances
- ec2:DescribeInstances

DescribeInstanceHealth

- elasticloadbalancing:DescribeInstanceHealth
- ec2:DescribeClassicLinkInstances
- ec2:DescribeInstances

DescribeLoadBalancers

- elasticloadbalancing:DescribeLoadBalancers
- ec2:DescribeSecurityGroups

DisableAvailabilityZonesForLoadBalancer

- elasticloadbalancing:DisableAvailabilityZonesForLoadBalancer
- ec2:DescribeAccountAttributes
- ec2:DescribeInternetGateways
- ec2:DescribeVpcs

EnableAvailabilityZonesForLoadBalancer

- elasticloadbalancing:EnableAvailabilityZonesForLoadBalancer
- ec2:DescribeAccountAttributes
- ec2:DescribeInternetGateways
- ec2:DescribeSubnets
- ec2:DescribeVpcs

RegisterInstancesWithLoadBalancer

- elasticloadbalancing:RegisterInstancesWithLoadBalancer
- ec2:DescribeAccountAttributes
- ec2:DescribeClassicLinkInstances
- ec2:DescribeInstances

- `ec2:DescribeVpcClassicLink`

Permissões de API do Elastic Load Balancing para marcar recursos durante a criação

Para que os usuários marquem recursos durante a criação, eles devem ter permissões para usar a ação que cria o recurso, como `elasticloadbalancing:CreateLoadBalancer` ou `elasticloadbalancing:CreateTargetGroup`. Se as tags forem especificadas na ação resource-creating, será necessário ter autorização adicional na ação `elasticloadbalancing:AddTags` para verificar se os usuários têm permissões para aplicar tags aos recursos que estão sendo criados. Portanto, os usuários também precisam ter permissões para usar a ação `elasticloadbalancing:AddTags`.

Na definição de política do IAM para a ação `elasticloadbalancing:AddTags`, é possível usar o elemento `Condition` com a chave de condição `elasticloadbalancing:CreateAction` para conceder permissões de marcação à ação que cria o recurso.

O exemplo a seguir demonstra uma política que permite que os usuários criem grupos de destino e apliquem qualquer tag a eles durante a criação. Os usuários não têm permissão para marcar recursos existentes (não podem chamar a ação `elasticloadbalancing:AddTags` diretamente).

```
{
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
         "elasticloadbalancing:CreateTargetGroup"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
         "elasticloadbalancing:AddTags"
      ],
      "Resource": "*",
      "Condition": {
         "StringEquals": {
             "elasticloadbalancing:CreateAction" : "CreateTargetGroup"
```

```
          }
       }
    }
  ]
}
```

De modo semelhante, a política a seguir permite que os usuários criem um balanceador de carga e apliquem tags durante a criação. Os usuários não têm permissão para marcar recursos existentes (não podem chamar a ação `elasticloadbalancing:AddTags` diretamente).

```
{
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
         "elasticloadbalancing:CreateLoadBalancer"
      ],
      "Resource": "*"
    },
    {
      "Effect": "Allow",
      "Action": [
         "elasticloadbalancing:AddTags"
      ],
      "Resource": "*",
      "Condition": {
         "StringEquals": {
             "elasticloadbalancing:CreateAction" : "CreateLoadBalancer"
          }
       }
    }
  ]
}
```

A ação `elasticloadbalancing:AddTags` será avaliada somente se as tags forem aplicadas durante a ação resource-creating. Portanto, um usuário que tiver permissões para criar um recurso (pressupondo-se que não existam condições de marcação) não precisa de permissão para usar a ação `elasticloadbalancing:AddTags` se nenhuma tag for especificada na solicitação. Contudo,

se o usuário tentar criar um recurso com tags, haverá falha na solicitação se o usuário não tiver permissão para usar a ação `elasticloadbalancing:AddTags`.

Perfil vinculado a serviço para o Elastic Load Balancing

O Elastic Load Balancing usa um perfil vinculado a serviço para as permissões necessárias para chamar outros serviços da AWS em seu nome. Para obter mais informações, consulte Usar funções vinculadas a serviço no Guia do usuário do IAM.

Permissões concedidas pela função vinculada ao serviço

O Elastic Load Balancing usa a função vinculada ao serviço nomeada AWSServiceRoleForElasticLoadBalancingpara chamar as seguintes ações em seu nome:

- `ec2:AssignIpv6Addresses`
- `ec2:AssignPrivateIpAddresses`
- `ec2:AssociateAddress`
- `ec2:AttachNetworkInterface`
- `ec2:AuthorizeSecurityGroupIngress`
- `ec2:CreateNetworkInterface`
- `ec2:CreateSecurityGroup`
- `ec2:DeleteNetworkInterface`
- `ec2:DescribeAccountAttributes`
- `ec2:DescribeAddresses`
- `ec2:DescribeClassicLinkInstances`
- `ec2:DescribeCoipPools`
- `ec2:DescribeInstances`
- `ec2:DescribeInternetGateways`
- `ec2:DescribeNetworkInterfaces`
- `ec2:DescribeSecurityGroups`
- `ec2:DescribeSubnets`
- `ec2:DescribeVpcClassicLink`
- `ec2:DescribeVpcPeeringConnections`

- `ec2:DescribeVpcs`
- `ec2:DetachNetworkInterface`
- `ec2:DisassociateAddress`
- `ec2:GetCoipPoolUsage`
- `ec2:ModifyNetworkInterfaceAttribute`
- `ec2:ReleaseAddress`
- `ec2:UnassignIpv6Addresses`
- `logs:CreateLogDelivery`
- `logs:DeleteLogDelivery`
- `logs:GetLogDelivery`
- `logs:ListLogDeliveries`
- `logs:UpdateLogDelivery`
- `outposts:GetOutpostInstanceTypes`

AWSServiceRoleForElasticLoadBalancingconfia no `elasticloadbalancing.amazonaws.com` serviço para assumir a função.

Criar a função vinculada ao serviço

Você não precisa criar a AWSServiceRoleForElasticLoadBalancingfunção manualmente. O Elastic Load Balancing cria esse perfil quando você cria um balanceador de carga ou um grupo de destino.

Para o Elastic Load Balancing criar um perfil vinculado a serviço em seu nome, você deve ter as permissões necessárias. Para obter mais informações, consulte Permissões de perfil vinculado a serviços no Guia do usuário do IAM.

Se você criou um load balancer antes de 11 de janeiro de 2018, o Elastic Load Balancing AWSServiceRoleForElasticLoadBalancingcriou em AWS sua conta. Para obter mais informações, consulte Uma nova função apareceu em minha AWS conta no Guia do usuário do IAM.

Editar a função vinculada ao serviço

Você pode editar a descrição do AWSServiceRoleForElasticLoadBalancinguso do IAM. Para obter mais informações, consulte Editar uma função vinculada a serviço no Manual do usuário do IAM.

Excluir a função vinculada ao serviço

Se você não precisar mais usar o Elastic Load Balancing, recomendamos que você exclua. AWSServiceRoleForElasticLoadBalancing

Você pode excluir essa função vinculada ao serviço somente depois de excluir todos os balanceadores de carga da sua conta. AWS Isso garante que você não remova por engano a permissão para acessar os load balancers. Para obter mais informações, consulte Excluir um Application Load Balancer, Excluir um Network Load Balancer e Excluir um Classic Load Balancer.

Você pode usar o console, a CLI ou a API do IAM para excluir funções vinculadas ao serviço. Para obter mais informações, consulte Excluir uma função vinculada ao serviço no Guia do usuário do IAM.

Depois de excluir AWSServiceRoleForElasticLoadBalancing, o Elastic Load Balancing cria a função novamente se você criar um balanceador de carga.

AWS políticas gerenciadas para o Elastic Load Balancing

Uma política AWS gerenciada é uma política autônoma criada e administrada por AWS. AWS as políticas gerenciadas são projetadas para fornecer permissões para muitos casos de uso comuns, para que você possa começar a atribuir permissões a usuários, grupos e funções.

Lembre-se de que as políticas AWS gerenciadas podem não conceder permissões de privilégio mínimo para seus casos de uso específicos porque elas estão disponíveis para uso de todos os AWS clientes. Recomendamos que você reduza ainda mais as permissões definindo políticas gerenciadas pelo cliente da específicas para seus casos de uso.

Você não pode alterar as permissões definidas nas políticas AWS gerenciadas. Se AWS atualizar as permissões definidas em uma política AWS gerenciada, a atualização afetará todas as identidades principais (usuários, grupos e funções) às quais a política está anexada. AWS é mais provável que atualize uma política AWS gerenciada quando uma nova AWS service (Serviço da AWS) é lançada ou novas operações de API são disponibilizadas para serviços existentes.

Para mais informações, consulte Políticas gerenciadas pela AWS no Manual do usuário do IAM.

AWS política gerenciada: AWSElasticLoadBalancingClassicServiceRolePolicy

Essa política inclui todas as permissões que o Elastic Load Balancing (Classic Load Balancer) exige para chamar AWS outros serviços em seu nome. Os perfis vinculados a serviço são predefinidos. Com os perfis predefinidos, você não precisa adicionar manualmente as permissões necessárias

para o Elastic Load Balancing concluir as ações em seu nome. Você não pode anexar, desanexar, modificar ou excluir essa política.

Para ver as permissões dessa política, consulte AWSElasticLoadBalancingClassicServiceRolePolicyna Referência de política AWS gerenciada.

AWS política gerenciada: AWSElasticLoadBalancingServiceRolePolicy

Essa política inclui todas as permissões que o Elastic Load Balancing requer para chamar outros serviços da AWS em seu nome. Os perfis vinculados a serviço são predefinidos. Com os perfis predefinidos, você não precisa adicionar manualmente as permissões necessárias para o Elastic Load Balancing concluir as ações em seu nome. Você não pode anexar, desanexar, modificar ou excluir essa política.

Para ver as permissões dessa política, consulte AWSElasticLoadBalancingServiceRolePolicyna Referência de política AWS gerenciada.

AWS política gerenciada: ElasticLoadBalancingFullAccess

Essa política dá acesso total ao serviço Elastic Load Balancing e acesso limitado a outros serviços por meio do AWS Management Console.

Para ver as permissões dessa política, consulte ElasticLoadBalancingFullAccessna Referência de política AWS gerenciada.

AWS política gerenciada: ElasticLoadBalancingReadOnly

Essa política fornece acesso somente leitura ao Elastic Load Balancing e a serviços dependentes.

Para ver as permissões dessa política, consulte ElasticLoadBalancingReadOnlyna Referência de política AWS gerenciada.

Atualizações do Elastic Load Balancing nas políticas AWS gerenciadas

Veja detalhes sobre as atualizações das políticas AWS gerenciadas do Elastic Load Balancing desde que esse serviço começou a rastrear essas mudanças.

Alteração	Descrição	Data
AWS política gerenciada: ElasticLoadBalancingFullAcc	O Elastic Load Balancing adicionou uma nova ação para conceder permissões para usar a mudança de zona. Essa ação foi adicionad	28 de novembro de 2022

Alteração	Descrição	Data
ess: atualização em uma política existente.	a à política de acesso total do Elastic Load Balancing. Ela está associado às operações de API `arc-zonal-shift:*` .	
AWS política gerenciada: ElasticLoadBalancingReadOnly: atualização em uma política existente.	O Elastic Load Balancing adicionou uma nova ação para conceder permissões para usar a mudança de zona. Essa ação foi adicionad a à política de acesso somente leitura do Elastic Load Balancing. Ela está associado às operações de API `arc-zonal-shift:GetManagedResource` , `arc-zonal-shift:ListManagedResources` e `arc-zonal-shift:ListZonalShifts` .	28 de novembro de 2022
AWS política gerenciada: AWSElasticLoadBalancingServiceRolePolicy: atualização em uma política existente.	O Elastic Load Balancing adicionou uma nova ação para conceder permissões para usar conexões de emparelhamento. Essa ação foi adicionada à política de perfil vinculado a serviço para o ambiente de gerenciamento do Elastic Load Balancing. Ela está associado à operação de API `ec2:DescribeVpcPeeringConnections` .	11 de outubro de 2021
AWS política gerenciada: ElasticLoadBalancingFullAccess: atualização em uma política existente.	O Elastic Load Balancing adicionou uma nova ação para conceder permissões para usar conexões de emparelhamento. Essa ação foi adicionada à política de acesso total do Elastic Load Balancing. Ela está associado à operação de API `ec2:DescribeVpcPeeringConnections` .	11 de outubro de 2021
AWS política gerenciada: AWSElasticLoadBalancingClassicServiceRolePolicy: atualização em uma política existente.	O Elastic Load Balancing adicionou uma política de perfil vinculado a serviço (para o ambiente de gerenciamento) para o Classic Load Balancer. Essa atualização é para a versão 2 (padrão).	7 de outubro de 2019

Alteração	Descrição	Data
AWS política gerenciada: ElasticLoadBalancingReadOnly	Fornece acesso somente leitura ao Elastic Load Balancing e a serviços dependentes. Essa é a versão 1 (padrão).	20 de setembro de 2018
O Elastic Load Balancing começou a rastrear as alterações.	O Elastic Load Balancing começou a monitorar as mudanças em suas políticas AWS gerenciadas.	23 de julho de 2021

Validação de conformidade para o Elastic Load Balancing

Para saber se um AWS service (Serviço da AWS) está no escopo de programas de conformidade específicos, consulte Serviços da AWS no escopo por programa de conformidade e selecione o programa de conformidade em que você está interessado. Para obter informações gerais, consulteAWS Programas de conformidade.

É possível fazer download de relatórios de auditoria de terceiros usando o AWS Artifact. Para obter mais informações, consulte Downloading Reports inAWS Artifact.

Sua responsabilidade de conformidade ao usar o Serviços da AWS é determinada pela confidencialidade dos seus dados, pelos objetivos de conformidade da sua empresa e pelos regulamentos e leis aplicáveis. A AWS fornece os seguintes atributos para ajudar com a conformidade:

- Guias de referência rápida de conformidade e segurança - estes guias de implantação discutem considerações sobre arquitetura e fornecem as etapas para a implantação de ambientes de linha de base focados em segurança e conformidade na AWS.
- Arquitetura para segurança e conformidade com HIPAA no Amazon Web Services: esse whitepaper descreve como as empresas podem usar a AWS para criar aplicações adequadas aos padrões HIPAA.

> **Note**
>
> Nem todos os Serviços da AWS estão qualificados pela HIPAA. Para obter mais informações, consulte a Referência dos serviços qualificados pela HIPAA.

- atributos de conformidade da AWS: essa coleção de manuais e guias pode ser aplicada a seu setor e local.
- Guias de conformidade do cliente da AWS: entenda o modelo de responsabilidade compartilhada sob a ótica da conformidade. Os guias resumem as práticas recomendadas para proteção de Serviços da AWS e mapeiam as diretrizes para controles de segurança em várias estruturas (incluindo o Instituto Nacional de Padrões e Tecnologia (NIST), o Conselho de Padrões de Segurança do Setor de Cartões de Pagamento (PCI) e a Organização Internacional de Padronização (ISO)).
- Avaliar atributos com regras no AWS Config Guia do desenvolvedor: o serviço AWS Config avalia como as configurações de atributos estão em conformidade com práticas internas, diretrizes do setor e regulamentos.
- AWS Security Hub: este AWS service (Serviço da AWS) fornece uma visão abrangente do seu estado de segurança na AWS. O Security Hub usa controles de segurança para avaliar os atributos da AWS e verificar a conformidade com os padrões e as práticas recomendadas do setor de segurança. Para obter uma lista dos serviços e controles aceitos, consulte a Referência de controles do Security Hub.
- AWS Audit Manager – Esse AWS service (Serviço da AWS) ajuda a auditar continuamente seu uso da AWS para simplificar a forma como você gerencia os riscos e a conformidade com regulamentos e padrões do setor.

Resiliência no Elastic Load Balancing

A infraestrutura global da AWS é criada com base em regiões e zonas de disponibilidade da AWS. As regiões fornecem várias zonas de disponibilidade separadas e isoladas fisicamente, que são conectadas com baixa latência, alta throughput e redes altamente redundantes. Com as zonas de disponibilidade, é possível projetar e operar aplicações e bancos de dados que automaticamente executam o failover entre as zonas sem interrupção. As zonas de disponibilidade são mais altamente disponíveis, tolerantes a falhas e escaláveis que uma ou várias infraestruturas de data center tradicionais.

Para obter mais informações sobre regiões e zonas de disponibilidade da AWS, consulte Infraestrutura global da AWS.

Além da infraestrutura global da AWS, o Elastic Load Balancing fornece os seguintes recursos para viabilizar a resiliência dos seus dados:

- Distribui o tráfego de entrada entre várias instâncias em uma única zona de disponibilidade ou em várias zonas de disponibilidade.
- É possível usar o AWS Global Accelerator com seus Application Load Balancers para distribuir o tráfego de entrada entre vários balanceadores de carga em uma ou mais regiões da AWS. Para obter mais informações, consulte o Guia do desenvolvedor do AWS Global Accelerator.
- O Amazon ECS permite que você execute, interrompa e gerencie contêineres do Docker em um cluster de instâncias do EC2. É possível configurar o serviço do Amazon ECS para usar um balanceador de carga a fim de distribuir o tráfego de entrada entre os serviços em um cluster. Para obter mais informações, consulte o Guia do desenvolvedor do Amazon Elastic Container Service.

Segurança de infraestrutura no Elastic Load Balancing

Por ser um serviço gerenciado, o Elastic Load Balancing é protegido pela segurança da rede global da AWS. Para obter informações sobre serviços de segurança da AWS e como a AWS protege a infraestrutura, consulte Segurança na nuvem da AWS. Para projetar seu ambiente da AWS usando as práticas recomendadas de segurança de infraestrutura, consulte Proteção de infraestrutura em Pilar de segurança: AWS Well-Architected Framework.

Você usa chamadas de API publicadas pela AWS para acessar o Elastic Load Balancing por meio da rede. Os clientes precisam oferecer suporte para:

- Transport Layer Security (TLS). Exigimos TLS 1.2 e recomendamos TLS 1.3.
- Conjuntos de criptografia com sigilo de encaminhamento perfeito (perfect forward secrecy, ou PFS) como DHE (Ephemeral Diffie-Hellman, ou Efêmero Diffie-Hellman) ou ECDHE (Ephemeral Elliptic Curve Diffie-Hellman, ou Curva elíptica efêmera Diffie-Hellman). A maioria dos sistemas modernos, como Java 7 e versões posteriores, comporta esses modos.

Além disso, as solicitações devem ser assinadas utilizando um ID da chave de acesso e uma chave de acesso secreta associada a uma entidade principal do IAM. Ou você pode usar o AWS Security Token Service (AWS STS) para gerar credenciais de segurança temporárias para assinar solicitações.

Isolamento de rede

Uma nuvem virtual privada (VPC) é uma rede virtual na área isolada logicamente na Nuvem AWS. Uma sub-rede é um intervalo de endereços IP em uma VPC. Ao criar um load balancer, é possível especificar uma ou mais sub-redes para os nós do load balancer. Você pode implantar instâncias do

EC2 nas sub-redes da sua VPC e registrá-las no load balancer. Para obter mais informações sobre VPC e sub-redes, consulte o Guia do usuário da Amazon VPC.

Quando você cria um load balancer em uma VPC, ele pode ser voltado para a Internet ou interno. Um load balancer interno só pode rotear solicitações de clientes com acesso à VPC para o load balancer.

O load balancer envia solicitações para seus destinos registrados usando endereços IP privados. Portanto, seus destinos não precisam de endereços IP públicos para receber solicitações de um load balancer.

Para chamar a API do Elastic Load Balancing diretamente da sua VPC usando endereços IP privados, use o AWS PrivateLink. Para ter mais informações, consulte Acessar o Elastic Load Balancing usando um endpoint de interface (AWS PrivateLink).

Controlar o tráfego de rede

Considere as opções a seguir para proteger o tráfego de rede ao usar um load balancer:

- Use receptores protegidos para oferecer suporte à comunicação criptografada entre clientes e seus balanceadores de carga. Application Load Balancers são compatíveis com receptores HTTPS. Network Load Balancers são compatíveis com receptores TLS. Classic Load Balancers são compatíveis com receptores HTTPS e TLS. É possível escolher entre políticas de segurança predefinidas para o load balancer a fim de especificar os pacotes de criptografia e as versões de protocolo compatíveis com seu aplicativo. Você pode usar o AWS Certificate Manager (ACM) ou o AWS Identity and Access Management (IAM) para gerenciar os certificados de servidor instalados no balanceador de carga. É possível usar o protocolo SNI (Server Name Indication) para atender vários sites seguros usando um único listener seguro. O SNI é habilitado automaticamente para o load balancer ao associar mais de um certificado de servidor a um listener seguro.
- Configure os grupos de segurança para que seus Application Load Balancers e Classic Load Balancers aceitem tráfego somente de clientes específicos. Esses grupos de segurança devem permitir tráfego de entrada de clientes nas portas do listener e tráfego de saída para os clientes.
- Configure os grupos de segurança para que suas instâncias do Amazon EC2 aceitem tráfego somente do balanceador de carga. Esses grupos de segurança devem permitir tráfego de entrada do load balancer nas portas do listener e nas portas da verificação de integridade.
- Configure seu Application Load Balancer para autenticar usuários com segurança por meio de um provedor de identidade ou usando identidades corporativas. Para obter mais informações, consulte Como autenticar usuários usando um Application Load Balancer.

- Use o AWS WAF com seus Application Load Balancers para permitir ou bloquear solicitações com base nas regras de uma lista de controle de acesso da Web (ACL da Web).

Acessar o Elastic Load Balancing usando um endpoint de interface (AWS PrivateLink)

Você pode estabelecer uma conexão privada entre a nuvem privada virtual (VPC) e a API do Elastic Load Balancing criando um endpoint da VPC de interface. É possível usar essa conexão para chamar a API do Elastic Load Balancing em sua VPC sem precisar conectar um gateway da Internet, instância NAT ou conexão VPN à sua VPC. O endpoint fornece conectividade confiável e escalável à API do Elastic Load Balancing, versões 01/12/2015 e 01/06/2012, que você usa para criar e gerenciar seus balanceadores de carga.

Os endpoints da VPC de interface são habilitados pelo AWS PrivateLink, um recurso que permite a comunicação entre suas aplicações e os Serviços da AWS usando endereços IP privados. Para obter mais informações, consulte AWS PrivateLink.

Limite

O AWS PrivateLink não é compatível com Network Load Balancers com mais de 50 receptores.

Criar um endpoint de interface para o Elastic Load Balancing

Criar um endpoint para o Elastic Load Balancing usando o seguinte nome de serviço:

```
com.amazonaws.region.elasticloadbalancing
```

Para mais informações, consulte Criar um endpoint de interface no Guia do usuário do AWS PrivateLink.

Criar uma política de endpoint da VPC para o Elastic Load Balancing

Você pode anexar uma política ao seu endpoint da VPC para controlar o aceso à API do Elastic Load Balancing. A política especifica:

- O principal que pode executar ações.
- As ações que podem ser executadas.
- O recurso no qual as ações podem ser executadas.

O exemplo a seguir mostra uma política de VPC endpoint que nega a todos permissão para criar um load balancer pelo endpoint. O exemplo de política também concede a todos permissão para executar todas as outras ações.

```
{
   "Statement": [
        {
            "Action": "*",
            "Effect": "Allow",
            "Resource": "*",
            "Principal": "*"
        },
        {
            "Action": "elasticloadbalancing:CreateLoadBalancer",
            "Effect": "Deny",
            "Resource": "*",
            "Principal": "*"
        }
    ]
}
```

Para obter mais informações, consulte Control access to services using endpoint policies (Controlar o acesso a serviços usando políticas de endpoint) no Guia do AWS PrivateLink.

Migrar seu Classic Load Balancer

O Elastic Load Balancing é compatível com os seguintes tipos de balanceadores de carga: Application Load Balancers, Network Load Balancers, Gateway Load Balancers e Classic Load Balancers. Para obter informações sobre os diferentes recursos de cada tipo de balanceador de carga, consulte o Comparativo de produtos do Elastic Load Balancing.

Você também pode optar por migrar um Classic Load Balancer existente em uma VPC para um Application Load Balancer ou um Network Load Balancer.

Benefícios da migração de um Classic Load Balancer

Cada tipo de balanceador de carga tem seus próprios recursos, funções e configurações exclusivos. Analise os benefícios de cada balanceador de carga para ajudar a decidir qual é o melhor para você.

Application Load Balancer

Usar um Application Load Balancer em vez de um Classic Load Balancer tem os seguintes benefícios:

Support for:

- Condições do caminho, condições do host e condições do cabeçalho HTTP.
- Redirecionando solicitações de uma URL para outra e roteando solicitações para vários aplicativos em uma única instância do EC2.
- Retornando respostas HTTP personalizadas.
- Registrando alvos por endereço IP e registrando funções do Lambda como alvos. Incluindo destinos fora da VPC para o balanceador de carga.
- Autenticar usuários por meio de identidades corporativas ou sociais.
- Aplicativos em contêineres do Amazon Elastic Container Service (Amazon ECS).
- Monitoramento independente da integridade de cada serviço.

Os registros de acesso contêm informações adicionais e são armazenados em um formato compactado.

Melhor desempenho geral do balanceador de carga.

Network Load Balancer

Usar um Network Load Balancer em vez de um Classic Load Balancer tem os seguintes benefícios:

Support for:

- Endereços IP estáticos, que permitem atribuir um endereço IP elástico por sub-rede habilitada para o balanceador de carga.
- Registro de destinos por endereço IP, incluindo destinos fora da VPC para o balanceador de carga.
- Roteamento de solicitações para vários aplicativos em uma única instância do EC2.
- Aplicativos em contêineres do Amazon Elastic Container Service (Amazon ECS).
- Monitoramento independente da integridade de cada serviço.

Capacidade de processar cargas de trabalho voláteis e de alterar a escala para milhões de solicitações por segundo.

Migre usando o assistente de migração

O assistente de migração usa a configuração do Classic Load Balancer para criar um Application Load Balancer ou Network Load Balancer equivalente. Isso reduz o tempo e o esforço necessários para migrar um Classic Load Balancer em comparação com outros métodos.

> **Note**
>
> O assistente cria um novo balanceador de carga. O assistente não converte o Classic Load Balancer existente em um Application Load Balancer ou Network Load Balancer. Você deve redirecionar manualmente o tráfego para o balanceador de carga recém-criado.

Limitações

- O nome do novo balanceador de carga não pode ser o mesmo de um balanceador de carga existente do mesmo tipo, na mesma região.
- Se o Classic Load Balancer tiver alguma tag contendo o `aws:` prefixo em sua chave, essas tags não serão migradas.

Ao migrar para um Application Load Balancer

- Se o Classic Load Balancer tiver somente uma sub-rede, você deverá especificar uma segunda sub-rede.
- Se o Classic Load Balancer tiver ouvintes HTTP/HTTPS que usam verificações de integridade TCP, o protocolo de verificação de integridade será atualizado para HTTP e o caminho será definido como “/”.
- Se o Classic Load Balancer tiver ouvintes HTTPS usando uma política de segurança personalizada ou não suportada, o assistente de migração usará a política de segurança padrão para o novo tipo de balanceador de carga.

Ao migrar para um Network Load Balancer

- Os seguintes tipos de instância não serão registrados no novo grupo de destino: C1, CC1, CC2, CG1, CG2, CR1, CS1, G1, G2, HI1, HS1, M1, M2, M3, T1
- Algumas configurações de verificação de integridade do Classic Load Balancer podem não ser transferíveis para o novo grupo-alvo. Esses casos serão indicados como uma alteração na seção de resumo do assistente de migração.
- Se o Classic Load Balancer tiver ouvintes SSL, o assistente de migração cria um ouvinte TLS usando o certificado e a política de segurança do ouvinte SSL.

Processo do assistente de migração

Para migrar um Classic Load Balancer usando o assistente de migração

1. Abra o console do Amazon EC2 em https://console.aws.amazon.com/ec2/.
2. No painel de navegação, em Load Balancing (Balanceamento de carga), escolha Load balancers (Balanceadores de carga).
3. Selecione o Classic Load Balancer que você deseja migrar.
4. Na seção Detalhes dos balanceadores de carga, escolha Iniciar assistente de migração.
5. Escolha Migrar para o Application Load Balancer ou Migrar para o Network Load Balancer para abrir o assistente de migração.
6. Em Nome do novo balanceador de carga, em Nome do balanceador de carga, insira um nome para seu novo balanceador de carga.

7. Em Nomear novo grupo-alvo e revisar alvos, em Nome do grupo-alvo, insira um nome para seu novo grupo-alvo.
8. (Opcional) Em Targets, você pode revisar as instâncias de destino que serão registradas no novo grupo-alvo.
9. (Opcional) Em Revisar tags, você pode revisar as tags que serão aplicadas ao seu novo balanceador de carga
10. Em Summary for Application Load Balancer ou Summary for Network Load Balancer, revise e verifique as opções de configuração atribuídas pelo assistente de migração.
11. Depois de ficar satisfeito com o resumo da configuração, escolha Create Application Load Balancer ou Create Network Load Balancer para iniciar a migração.

Migre usando o utilitário de cópia do balanceador de carga

Os utilitários de cópia do balanceador de carga estão disponíveis no repositório do Elastic Load Balancing Tools, na página. AWS GitHub

Recursos

- Ferramentas do Elastic Load Balancing
- Utilitário de cópia do Classic Load Balancer para o Application Load Balancer
- Utilitário de cópia do Classic Load Balancer para Network Load Balancer

Migre seu balanceador de carga manualmente

As informações a seguir fornecem instruções gerais para criar manualmente um novo Application Load Balancer ou Network Load Balancer com base em um Classic Load Balancer existente em uma VPC. Você pode migrar usando o AWS Management Console AWS CLI, o ou um AWS SDK. Para ter mais informações, consulte Como começar a usar o Elastic Load Balancing.

Depois de concluir o processo de migração, você poderá aproveitar os recursos do seu novo load balancer.

Processo de migração manual

Etapa 1: criar um novo balanceador de carga

Crie um balanceador de carga com uma configuração equivalente ao Classic Load Balancer para migrar.

1. Crie um novo balanceador de carga com o mesmo esquema (voltado para a Internet ou interno), sub-redes e grupos de segurança do Classic Load Balancer.
2. Crie um grupo de destino para o seu balanceador de carga com as mesmas configurações de verificação de integridade presentes no seu Classic Load Balancer.
3. Execute um destes procedimentos:
   - Se o Classic Load Balancer estiver anexado a um grupo do Auto Scaling, anexe o grupo de destino ao grupo do Auto Scaling. Isso também registrará as instâncias do Auto Scaling com o grupo de destino.
   - Registre suas instâncias EC2 com o seu grupo de destino.
4. Crie um ou mais listeners, cada um com uma regra padrão que encaminha solicitações para o grupo de destino. Se você criar um receptor HTTPS, poderá especificar o mesmo certificado que foi especificado para o seu Classic Load Balancer. Recomendamos que você use a política de segurança padrão.
5. Se o seu Classic Load Balancer tiver tags, verifique e adicione as tags relevantes ao seu novo balanceador de carga.

Etapa 2: redirecionar gradualmente o tráfego para seu novo balanceador de carga

Depois que suas instâncias forem registradas com o novo balanceador de carga, você poderá iniciar o processo de redirecionamento do tráfego do balanceador de carga antigo para o novo. Isso permite que você teste seu novo balanceador de carga enquanto minimiza os riscos à disponibilidade da sua aplicação.

Para redirecionar o tráfego gradualmente para seu novo load balancer

1. Cole o nome DNS do seu novo load balancer no campo de endereço de um navegador da Web conectado à Internet. Se tudo estiver funcionando, o navegador exibirá a página padrão da sua aplicação.
2. Crie um novo registro de DNS que associe seu nome de domínio ao seu novo load balancer. Se o serviço de DNS for compatível com o recurso de ponderação, especifique um peso de 1 no novo registro de DNS e um peso de 9 no registro de DNS existente para seu balanceador de carga antigo. Isso direcionará 10% do tráfego para o novo balanceador de carga e 90% do tráfego para o balanceador de carga antigo.

3. Monitore seu novo load balancer para verificar se ele está recebendo o tráfego e solicitações de roteamento para suas instâncias.

> **Important**
>
> O time-to-live (TTL) no registro DNS é de 60 segundos. Isso significa que qualquer servidor DNS que resolver o nome de domínio manterá as informações do registro em cache por 60 segundos, enquanto as alterações são propagadas. Portanto, esses servidores DNS ainda poderão rotear o tráfego para o balanceador de carga antigo por até 60 segundos após você concluir a etapa anterior. Durante a propagação, o tráfego pode ser direcionado para o load balancer.

4. Continue para atualizar a ponderação dos seus registros DNS até que todo o tráfego seja direcionado para o novo load balancer. Após concluir, você poderá excluir o registro DNS do seu balanceador de carga antigo.

Etapa 3: atualizar políticas, scripts e código

Se você tiver migrado o Classic Load Balancer para um Application Load Balancer ou Network Load Balancer, não esqueça de fazer o seguinte:

- Atualize as políticas do IAM que usam a versão de API de 01/06/2012 para usar a versão de 01/12/2015.
- Atualize processos que usam CloudWatch métricas no `AWS/ELB` namespace para usar métricas do namespace `AWS/ApplicationELB` or`AWS/NetworkELB`.
- Atualize scripts que usam aws elb AWS CLI comandos para usar aws elbv2 AWS CLI comandos.
- Atualize AWS CloudFormation modelos que usam o `AWS::ElasticLoadBalancing::LoadBalancer` recurso para usar os `AWS::ElasticLoadBalancingV2` recursos.
- Atualize o código que usa a versão de API de 01/06/2012 do Elastic Load Balancing para usar a versão de 01/12/2015.

Recursos

- elbv2 na Referência de comandos da AWS CLI
- Referência de API do Elastic Load Balancing versão de 01/12/2015

- Gerenciamento de identidade e acesso para o Elastic Load Balancing
- Métricas do Application Load Balancer no Guia do usuário para Application Load Balancers
- Métricas do Network Load Balancer no Guia do usuário para Network Load Balancers
- AWS::ElasticLoadBalancingV2::LoadBalancer no AWS CloudFormation Guia do usuário

Etapa 4: excluir o balanceador de carga antigo

Você pode excluir o antigo Classic Load Balancer depois de:

- Ter redirecionado todo o tráfego do balanceador de carga antigo para o novo.
- Todas as solicitações existentes que foram roteadas para o balanceador de carga antigo tiverem sido concluídas.

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.