aws

Guia do usuário

Elastic Load Balancing

Elastic Load Balancing: Guia do usuário

Table of Contents

O que é Elastic Load Balancing?

O Elastic Load Balancing distribui automaticamente seu tráfego de entrada em vários destinos, como EC2 instâncias, 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 a vários tipos de balanceadores de carga. Você pode selecionar o tipo de balanceador de carga que melhor se adapte às suas necessidades. Para obter mais informações, consulte Balancing.

Para obter mais informações sobre os balanceadores de carga da geração atual, consulte a documentação a seguir:

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

Os Classic Load Balancers são a geração anterior de balanceadores de carga do Elastic Load Balancing. Recomendamos que você migre para um balanceador de carga da geração atual. Para obter mais informações, consulte Migrar seu Classic Load Balancer.

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 Interface de linha de comando (AWS CLI) — Fornece comandos para um amplo conjunto de AWS serviços, incluindo o Elastic Load Balancing. O AWS CLI é compatível com Windows, macOS e Linux. Para obter mais informações, consulte AWS Command Line Interface.
- AWS SDKs— forneça informações específicas para o idioma APIs e cuide de muitos detalhes da conexão, como calcular assinaturas, lidar 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:
  - Application Load Balancers, Network Load Balancers e Gateway Load Balancers: versão de 01/12/2015 da API
  - 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 seus aplicativos na nuvem. Você pode configurar seu balanceador de carga para direcionar o tráfego para suas EC2 instâncias. Para obter mais informações, consulte o Guia EC2 do usuário da Amazon.
- 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 à medida que a demanda por suas instâncias muda. 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 monitorar seu balanceador de carga e agir conforme necessário. Para obter mais informações, consulte o Guia CloudWatch do usuário da Amazon.
- Amazon ECS — Permite que você execute, interrompa e gerencie contêineres Docker em um cluster de EC2 instâncias. Você pode configurar o load balancer para rotear o tráfego para seus contêineres. Para obter mais informações, consulte o Guia do desenvolvedor do serviço Elastic Container da Amazon.
- AWS Global Accelerator: melhora a disponibilidade e o desempenho da sua aplicação. Use um acelerador para distribuir o tráfego entre vários balanceadores de carga em uma ou mais AWS regiões. 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, isso se `www.example.com` traduziria no endereço `192.0.2.1` IP numérico. 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 AWS WAF com seu Application Load Balancer para permitir ou bloquear solicitações com base nas regras de uma lista de controle de acesso à web (web ACL). 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 balanceador de carga aceita tráfego de entrada de clientes e encaminha solicitações para seus destinos registrados (como EC2 instâncias) 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.

Conteúdo

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.

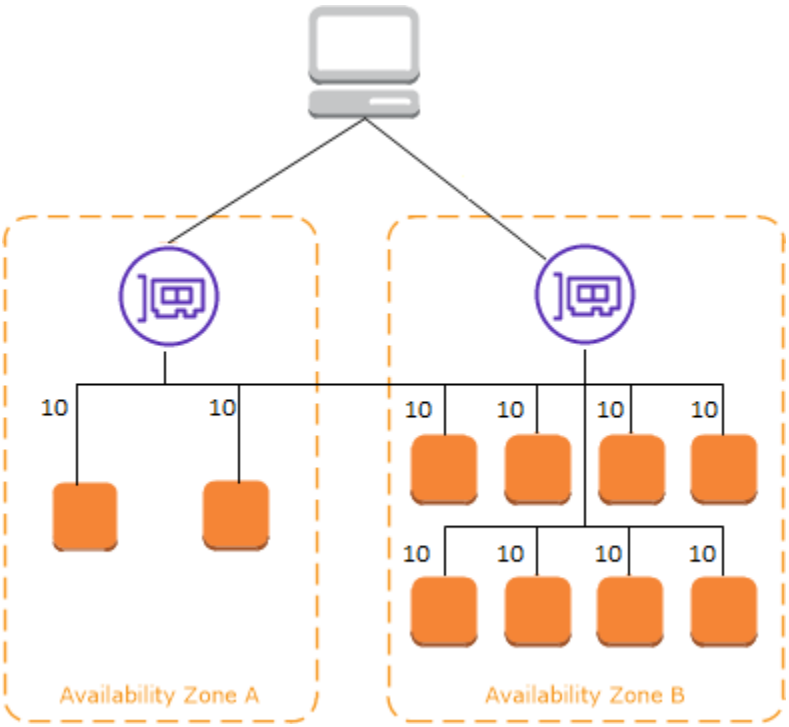
Depois de desabilitar uma zona de disponibilidade, os destinos nessa zona de disponibilidade permanecem registrados com o load balancer. No entanto, mesmo que permaneçam registrados, o load balancer não roteará o tráfego para eles.

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.

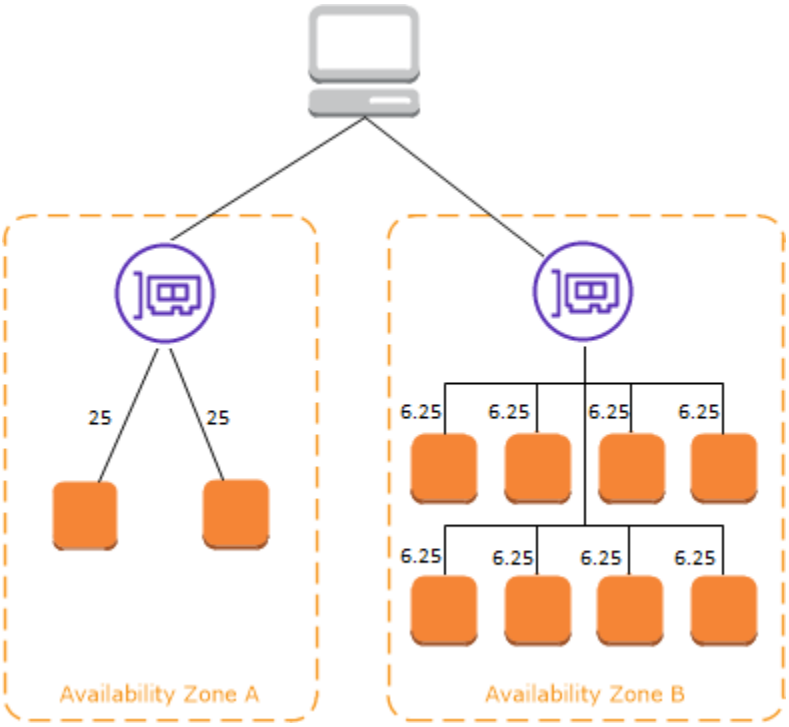
Caso o balanceamento de carga entre zonas esteja habilitado, cada um dos dez destinos recebe 10% do tráfego. Isso ocorre porque cada nó do load balancer pode rotear 50% do tráfego do cliente para todos os dez destinos.



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. Para obter mais informações, consulte Cross-zone load balancing no Guia do usuário de Network Load Balancers.

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 Amazon Application Recovery Controller (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 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 How a zonal shift works: health checks and zonal IP addresses no Guia do desenvolvedor do Amazon Application Recovery Controller (ARC).

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

- A mudança de zona é compatível ao usar um Network Load Balancer com o balanceamento de carga entre zonas ativado ou desativado.
- 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.

Para obter mais orientações e informações, consulte as melhores práticas para mudanças zonais no ARC no Guia do desenvolvedor do Amazon Application Recovery Controller (ARC).

Roteamento de solicitações

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 os balanceadores de carga do Gateway, o nó do balanceador de carga que recebe a conexão usa um algoritmo de hash de fluxo de 5 tuplas para selecionar um dispositivo de destino. Depois que um fluxo é estabelecido, todos os pacotes do mesmo fluxo são roteados consistentemente para o mesmo dispositivo de destino. O balanceador de carga e os dispositivos de destino trocam tráfego usando o protocolo GENEVE na porta 6081.

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 backend. 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 backend.

Os Application Load Balancers são compatíveis com os 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 backend (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 backend. 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 backend. 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 backend (balanceador de carga para destino registrado). Por padrão, o cabeçalho `keep-alive` é compatível com conexões de backend. 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 backend. 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.

Tipos de endereço IP

O tipo de endereço IP que você especifica para o balanceador de carga determina como os clientes podem se comunicar com o balanceador de carga.

- IPv4 somente — Os clientes se comunicam usando IPv4 endereços públicos e privados. As sub-redes que você seleciona para seu balanceador de carga devem ter IPv4 intervalos de endereços.
- Dualstack — Os clientes se comunicam usando endereços públicos e privados IPv4 . IPv6 As sub-redes que você seleciona para seu balanceador de carga devem ter IPv4 intervalos de endereços. IPv6
- Dualstack sem público IPv4 — Os clientes se comunicam usando endereços públicos e privados e IPv6 endereços privados. IPv4 As sub-redes que você seleciona para seu balanceador de carga devem ter IPv4 intervalos de endereços. IPv6 Essa opção não é compatível com o esquema do balanceador de carga `internal`.

A tabela a seguir descreve os endereços IP compatíveis com cada tipo de balanceador de carga.

Tipo de load balancer	IPv4 somente	Pilha dupla	Dualstack sem público IPv4
Application Load Balancer	Yes (Sim)	Yes (Sim)	Yes (Sim)
Network Load Balancer	Yes (Sim)	Yes (Sim)	Não
Gateway Load Balancer	Yes (Sim)	Yes (Sim)	Não

Tipo de load balancer	IPv4 somente	Pilha dupla	Dualstack sem público IPv4
Classic Load Balancer	Sim	Não	Nº

O tipo de endereço IP especificado para o grupo de destino determina como o balanceador de carga pode se comunicar com os destinos.

- IPv4 somente — O balanceador de carga se comunica usando endereços privados IPv4 . Você deve registrar alvos com IPv4 endereços com um IPv4 grupo-alvo.
- IPv6 somente — O balanceador de carga se comunica usando IPv6 endereços. Você deve registrar alvos com IPv6 endereços com um IPv6 grupo-alvo. O grupo de destino deve ser usado com um balanceador de carga dualstack.

A tabela a seguir descreve os tipos de endereço IP compatíveis com cada protocolo de grupo de destino.

Protocolo do grupo de destino	IPv4 somente	IPv6 somente	
HTTP e HTTPS	Yes (Sim)	Yes (Sim)	
TCP	Yes (Sim)	Yes (Sim)	
TLS	Yes (Sim)	Yes (Sim)	
UDP e TCP_UDP	Yes (Sim)	Yes (Sim)	
GENEVE	-	-	

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 quadros de Ethernet consistem no pacote, ou nos dados em si 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 o Path MTU Discovery e a habilitação de jumbo frames, consulte Path MTU Discovery no Guia do usuário da Amazon EC2 .

Como começar a usar o Elastic Load Balancing

O Elastic Load Balancing oferece suporte a vários tipos de balanceadores de carga. Você pode selecionar o tipo de balanceador de carga que melhor se adapte às suas necessidades. Para obter mais informações, consulte Balancing.

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.

Conteúdo

Criar um Application Load Balancer

Para criar um Application Load Balancer usando o AWS Management Console, consulte Introdução aos Application Load Balancers no Guia do usuário de Application Load Balancers.

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

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

Criar um Gateway Load Balancer

Para criar um balanceador de carga de gateway usando o AWS Management Console, consulte Introdução aos balanceadores de carga de gateway no Guia do usuário de balanceadores de carga de gateway.

Para criar um Gateway Load Balancer usando o AWS CLI, consulte Introdução aos Gateway Load Balancers usando o AWS CLI no Guia do usuário para Gateway Load Balancers.

Segurança no Elastic Load Balancing

A segurança na nuvem AWS é a maior prioridade. Como AWS cliente, você se beneficia de um data center e de uma arquitetura de rede criados para atender aos requisitos das organizações mais sensíveis à segurança.

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

- Segurança da nuvem — AWS é responsável por proteger a infraestrutura que executa AWS os serviços na AWS nuvem. AWS também fornece serviços que você pode usar com segurança. Auditores terceirizados testam e verificam regularmente a eficácia de nossa segurança como parte dos programas de AWS conformidade dos programas de de . Para saber mais sobre os programas de conformidade que se aplicam ao Elastic Load Balancing, consulte AWS serviços em escopo por programa de conformidade AWS .
- Segurança na nuvem — Sua responsabilidade é determinada pelo AWS serviço 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 normas 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 a usar outros AWS serviços que ajudam 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 Elastic Load Balancing Partners integraram e qualificaram seu software de dispositivos com. AWS Você pode confiar mais no software dos dispositivos dos fornecedores desta lista. No entanto, AWS não garante a segurança ou a confiabilidade do software desses fornecedores.

Conteúdo

- 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 o conteúdo hospedado nessa infraestrutura. Você também é responsável pelas tarefas de configuração e gerenciamento de segurança dos Serviços da AWS que usa. Para obter mais informações sobre a privacidade de dados, consulte as Data Privacy FAQ. Para obter mais informações sobre a proteção de dados na Europa, consulte a postagem do blog AWS Shared Responsibility Model and RGPD 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. Para obter informações sobre o uso de CloudTrail trilhas para capturar AWS atividades, consulte Como trabalhar com CloudTrail trilhas no Guia AWS CloudTrail do usuário.
- 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 140-3 ao acessar AWS por meio de uma interface de linha de comando ou de uma API, use um endpoint FIPS. Para obter

mais informações sobre os endpoints FIPS disponíveis, consulte Federal Information Processing Standard (FIPS) 140-3.

É altamente recomendável que nunca sejam colocadas informações confidenciais ou sigilosas, como endereços de e-mail de clientes, em tags 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 CLI AWS SDKs Quaisquer dados inseridos em tags ou em 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 é criptografada com uma chave do KMS alternada 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 balanceador de carga executa o trabalho de criptografar e descriptografar o tráfego, em vez de exigir que cada EC2 instância realize 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

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 a solicitar as permissões corretas ao 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. Envie as solicitações ao administrador do IAM para alterar as permissões dos usuários de serviço. Revise as informações nesta página para compreender os conceitos básicos do 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.

Autenticação 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 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 designar solicitações por conta própria, consulte Versão 4 do AWS Signature para solicitações de API no Guia do usuário do IAM.

Independente do método de autenticação usado, também pode ser necessário fornecer 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 e Usar a autenticação multifator da AWS no IAM 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 puder 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, é recomendável 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 O que é o Centro de Identidade do IAM? no Guia do Usuário do AWS IAM Identity Center .

Usuários e grupos 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, é recomendável contar com 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, é recomendável alternar as chaves de acesso. Para obter mais informações, consulte Alternar 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 IAMAdminse conceder a esse grupo permissões para administrar recursos 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 Casos de uso para usuários do IAM 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. Para assumir temporariamente uma função do IAM no AWS Management Console, você pode alternar de um usuário para uma função do IAM (console). 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 usar perfis, consulte Métodos para assumir um perfil 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, é possível 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 por ele. Para ter mais informações sobre perfis para federação, consulte Criar um perfil para um provedor de identidade de terceiros (federação) no Guia do usuário do IAM. Se usar o Centro de Identidade do IAM, configure um conjunto de permissões. Para controlar o que suas identidades podem acessar após a autenticação, o Centro de Identidade do IAM correlaciona o conjunto de permissões a um perfil no IAM. Para obter informações sobre conjuntos de permissões, consulte Conjuntos de Permissões no Guia do Usuário do AWS IAM Identity Center .
- Permissões temporárias para usuários do IAM: um usuário ou um perfil do IAM pode presumir 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 conhecer a diferença entre perfis e políticas baseadas em recurso para acesso entre contas, consulte Acesso a recursos entre contas no IAM 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 aplicativos na Amazon EC2 ou armazene objetos no Amazon S3. Um serviço pode fazer isso usando as permissões da entidade principal da chamada, usando um perfil de serviço ou um perfil vinculado 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 Sessões de acesso direto.

- Perfil de serviço: um perfil de serviço é um perfil do IAM que um serviço assume para executar 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 presumir o perfil para 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 editar as permissões para perfis vinculados a serviço.

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

Gerenciar o 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 perfis não têm permissões. Para conceder permissão aos usuários para executar ações nos recursos que eles precisam, um administrador do IAM pode criar políticas do IAM. O administrador pode então adicionar as políticas do IAM aos 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, grupo de usuários ou perfil do IAM. Essas políticas controlam quais ações os usuários e perfis podem realizar, em quais recursos e em que condições. Para saber como criar uma política baseada em identidade, consulte Definir permissões personalizadas do IAM com as políticas gerenciadas pelo cliente no Guia do Usuário do IAM.

As políticas baseadas em identidade podem ser categorizadas 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 especificado 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 recursos são políticas em linha 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 diretores (membros da conta, usuários ou funções) têm permissões para acessar um recurso. ACLs são semelhantes às políticas baseadas em recursos, embora não usem o formato de documento de política JSON.

O Amazon S3 e o AWS WAF Amazon VPC são exemplos de serviços que oferecem suporte. ACLs Para saber mais ACLs, consulte a 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 (usuário ou perfil 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 com seus limites de permissões. As políticas baseadas em recurso 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 recursos em uma organização, poderá aplicar políticas de controle de serviço (SCPs) a qualquer uma ou a todas as suas 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 Organizations e SCPs, consulte Políticas de controle de serviços no Guia AWS Organizations do Usuário.
- Políticas de controle de recursos (RCPs) — RCPs são políticas JSON que você pode usar para definir o máximo de permissões disponíveis para recursos em suas contas sem atualizar as políticas do IAM anexadas a cada recurso que você possui. O RCP limita as permissões para recursos nas contas dos membros e pode afetar as permissões efetivas para identidades, incluindo a Usuário raiz da conta da AWS, independentemente de pertencerem à sua organização. Para obter mais informações sobre Organizations e RCPs, incluindo uma lista Serviços da

AWS desse suporte RCPs, consulte Políticas de controle de recursos (RCPs) no Guia AWS Organizations do usuário.

- 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 recursos. 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

Recurso do IAM	Compatibilidade com o Elastic Load Balancing
Políticas baseadas em identidade	Sim
Políticas baseadas em recurso	Não
Ações de políticas	Sim
Recursos de políticas	Sim
Chaves de condição de política (específicas do serviço)	Sim
ACLs	Não

Recurso do IAM	Compatibilidade com o Elastic Load Balancing
ABAC (tags 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

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 do IAM, grupo de usuários ou perfil. Essas políticas controlam quais ações os usuários e perfis podem realizar, em quais recursos e em que condições. Para saber como criar uma política baseada em identidade, consulte Definir permissões personalizadas do IAM com as políticas gerenciadas pelo cliente no Guia do Usuário do IAM.

Com as políticas baseadas em identidade do IAM, é possível especificar ações e recursos permitidos ou negados, assim como as condições sob as quais as ações são permitidas ou negadas. Você não pode especificar a entidade principal em uma política baseada em identidade porque ela se aplica ao usuário ou perfil ao qual ela está anexada. Para saber mais sobre todos os elementos que podem ser usados em uma política JSON, consulte Referência de elemento de política JSON do IAM no Guia do usuário do IAM.

Políticas baseadas em recursos no Elastic Load Balancing

Compatibilidade com 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 especificado 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 recursos. Adicionar uma entidade principal entre contas à política baseada em recurso 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 recurso conceder acesso a uma entidade principal na mesma conta, nenhuma política baseada em identidade adicional será necessária. Consulte mais informações em Acesso a recursos entre contas no IAM no Guia do usuário do IAM.

Ações de política para o Elastic Load Balancing

Compatível com 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 podem ser usadas 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. Algumas operações também 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 de ações do Elastic Load Balancing, consulte Ações definidas pelo Elastic Load Balancing V2 e Ações definidas pelo Elastic Load Balancing V1 na Referência de Autorização de 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

Recursos de política para o Elastic Load Balancing

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

Para ações que não oferecem compatibilidade com permissões em 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 ver uma lista dos tipos de recursos do Elastic Load Balancing e seus ARNs, consulte Recursos definidos pelo Elastic Load Balancing V2 e Recursos definidos pelo Elastic Load Balancing V1 na Referência de Autorização de Serviço. Para saber com quais ações você pode especificar o ARN de cada recurso, consulte Ações definidas pelo Elastic Load Balancing V2 e Ações definidas pelo Elastic Load Balancing V1.

Chaves de condição de política para o Elastic Load Balancing

Compatível com chaves de condição de política específicas de serviço: 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 `Condition` (ou bloco `Condition`) permite que você especifique condições nas quais uma instrução estiver em vigor. O elemento `Condition` é opcional. É possível criar expressões condicionais que usem 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 de `Condition` em uma declaração ou várias chaves em um único elemento de `Condition`, 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 antes que as permissões da instrução sejam concedidas.

Você também pode usar variáveis de espaço reservado ao especificar condições. Por exemplo, é possível conceder a um usuário do IAM permissão para acessar um recurso somente se ele estiver marcado com seu nome de usuário do IAM. Para obter mais informações, consulte Elementos da 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 das chaves de condição do Elastic Load Balancing, consulte Chaves de condição do Elastic Load Balancing V2 e Chaves de condição do Elastic Load Balancing V1 na Referência de Autorização de Serviço. Para saber com quais ações e recursos você pode usar uma chave de

condição, consulte Ações definidas pelo Elastic Load Balancing V2 e Ações definidas pelo Elastic Load Balancing V1.

Chaves de condição

- Chave da condição elasticloadbalancing:ListenerProtocol
- Chave da condição elasticloadbalancing:SecurityPolicy
- Chave da condição elasticloadbalancing:Scheme
- Chave da condição elasticloadbalancing:SecurityGroup
- Chave da condição elasticloadbalancing:Subnet
- Chave da condição elasticloadbalancing:ResourceTag

Chave da condição elasticloadbalancing:ListenerProtocol

A chave de condição `elasticloadbalancing:ListenerProtocol` é usada em condições que definem os tipos de receptores que podem ser criados e usados. A política está disponível para Application Load Balancers, Network Load Balancers e Classic Load Balancers. 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`

O exemplo de política a seguir exige que os usuários selecionem o protocolo HTTPS para os ouvintes para seus Application Load Balancers e o protocolo TLS para os ouvintes para seus Network Load Balancers.

```
{
    "Version": "2012-10-17",
    "Statement": {
        "Effect": "Allow",
```

```
        "Action": [
            "elasticloadbalancing:CreateListener",
            "elasticloadbalancing:ModifyListener"
        ],
        "Resource": "*",
        "Condition": {
            "StringEquals":{
                "elasticloadbalancing:ListenerProtocol": [
                    "HTTPS",
                    "TLS"
                ]
            },
        }
    }
}
```

Com um Classic Load Balancer, você pode especificar vários ouvintes em uma única chamada. Portanto, sua política deve usar uma chave de contexto de vários valores, conforme mostrado no exemplo a seguir.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Action": [
                "elasticloadbalancing:CreateLoadBalancer",
                "elasticloadbalancing:CreateLoadBalancerListeners"
            ],
            "Resource": "*",
            "Condition": {
                "ForAnyValue:StringEquals": {
                    "elasticloadbalancing:ListenerProtocol": [
                        "TCP",
                        "HTTP",
                        "HTTPS"
                    ]
                }
            }
        }
```

```
    ]
}
```

Chave da condição elasticloadbalancing:SecurityPolicy

A chave de condição `elasticloadbalancing:SecurityPolicy` é usada em condições que definem e impõem políticas de segurança específicas nos balanceadores de carga. A política está disponível para Application Load Balancers, Network Load Balancers e Classic Load Balancers. 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`

O exemplo de política a seguir exige que os usuários selecionem uma das políticas de segurança especificadas para seus Application Load Balancers e Network Load Balancers.

```
{
    "Version": "2012-10-17",
    "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 condição `elasticloadbalancing:Scheme` é usada em condições que definem qual esquema pode ser selecionado durante a criação do balanceador de carga. A política está disponível para Application Load Balancers, Network Load Balancers e Classic Load Balancers. 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`

O exemplo de política a seguir exige que os usuários selecionem o esquema especificado para seus balanceadores de carga.

```
{
    "Version": "2012-10-17",
    "Statement": {
        "Effect": "Allow",
        "Action":"elasticloadbalancing:CreateLoadBalancer",
        "Resource": "*",
        "Condition": {
            "StringEquals":{
                "elasticloadbalancing:Scheme": "internal"
            },
        }
    }
}
```

Chave da condição **elasticloadbalancing:SecurityGroup**

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

A chave de condição `elasticloadbalancing:SecurityGroup` é usada em condições que definem quais grupos de segurança podem ser aplicados aos balanceadores de carga. A política está disponível para Application Load Balancers, Network Load Balancers e Classic Load Balancers. 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`

O exemplo de política a seguir exige que os usuários selecionem um dos grupos de segurança especificados para seus balanceadores de carga.

```
"Version": "2012-10-17",
"Statement": {
    "Effect": "Allow",
    "Action": [
        "elasticloadbalancing:CreateLoadBalancer",
        "elasticloadbalancing:SetSecurityGroup"
    ],
    "Resource": "*",
    "Condition": {
        "ForAnyValue:StringEqualsIgnoreCase":{
            "elasticloadbalancing:SecurityGroup": [
```

```
                    "sg-51530134",
                    "sg-51530144",
                    "sg-51530139"
                ]
            },
        }
    }
}
```

Chave da condição elasticloadbalancing:Subnet

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

A chave de condição `elasticloadbalancing:Subnet` é usada em condições que definem quais sub-redes podem ser criadas e conectadas aos balanceadores de carga. A política está disponível para Application Load Balancers, Network Load Balancers, Gateway Load Balancers e Classic Load Balancers. 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`

O exemplo de política a seguir exige que os usuários selecionem uma das sub-redes especificadas para seus balanceadores de carga.

```
{
    "Version": "2012-10-17",
    "Statement": {
```

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

Chave da condição elasticloadbalancing:ResourceTag

A chave de *key* condição`elasticloadbalancing:ResourceTag`/é específica do Elastic Load Balancing. Todas as ações mutantes suportam essa chave de condição.

ACLs em Elastic Load Balancing

Suportes ACLs: Não

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

ABAC com o Elastic Load Balancing

Compatível com ABAC (tags em políticas): sim

O controle de acesso por atributo (ABAC) é uma estratégia de autorização que define as 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. Marcar 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 recurso que ela estiver 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 `aws:ResourceTag/key-name`, `aws:RequestTag/key-name` ou chaves de condição `aws:TagKeys`.

Se um serviço for compatível com as três chaves de condição para cada tipo de recurso, o valor será Sim para o serviço. Se um serviço for compatível com as três chaves de condição somente para alguns tipos de recursos, o valor será Parcial

Para obter mais informações sobre o ABAC, consulte Definir permissões com autorização do 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

Compatível com 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 funções, consulte Alternar para um perfil do IAM (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 obter mais informações, consulte Credenciais de segurança temporárias no IAM.

Permissões de entidade principal entre serviços para o Elastic Load Balancing

Compatibilidade com o recurso de 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 Sessões de acesso direto.

Perfis de serviço para o Elastic Load Balancing

Compatível com perfis de serviço: não

O perfil de serviço é um perfil do IAM que um serviço assume para executar 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

Compatibilidade com perfis vinculados a serviços: sim

Uma função vinculada ao serviço é um tipo de função de serviço vinculada a um. AWS service (Serviço da AWS) O serviço pode presumir o perfil para 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 editar as permissões para perfis vinculados a 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 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).

```
{
  "Version": "2012-10-17",
  "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).

```
{
  "Version": "2012-10-17",
  "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 Perfis vinculados ao 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 AWSServiceRoleForElasticLoadBalancing para chamar outros AWS serviços em seu nome.

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

A política de permissões de função éAWSElasticLoadBalancingServiceRolePolicy. Para visualizar as permissões para esta política, consulte AWSElasticLoadBalancingServiceRolePolicy na Referência de políticas gerenciadas pela AWS .

Criar a função vinculada ao serviço

Você não precisa criar manualmente a função AWSServiceRoleForElasticLoadBalancing. 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 Service-linked role permissions (Permissões de função vinculada a serviços) no Guia do usuário do IAM.

Editar a função vinculada ao serviço

Você pode editar a descrição do AWSServiceRoleForElasticLoadBalancing uso do IAM. Para obter mais informações, consulte Editar uma descrição de perfil vinculado ao serviço no Guia 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.

É possível 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 excluirAWSServiceRoleForElasticLoadBalancing, 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 as políticas gerenciadas pelo cliente que são 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 visualizar as permissões para esta política, consulte AWSElasticLoadBalancingClassicServiceRolePolicy na Referência de políticas gerenciadas pela AWS .

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 visualizar as permissões para esta política, consulte AWSElasticLoadBalancingServiceRolePolicy na Referência de políticas gerenciadas pela AWS .

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 visualizar as permissões para esta política, consulte ElasticLoadBalancingFullAccess na Referência de políticas gerenciadas pela AWS .

AWS política gerenciada: ElasticLoadBalancingReadOnly

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

Para visualizar as permissões para esta política, consulte ElasticLoadBalancingReadOnly na Referência de políticas gerenciadas pela AWS .

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
AWSElasticLoadBalancingServiceRolePolicy - Atualização em uma política existente	Foi adicionada a `ec2:AllocateIpamPoolCidr` ação para conceder permissões para alocar blocos CIDR de pools IPAM.	17 de fevereiro de 2025
ElasticLoadBalancingFullAccess - Atualização em uma política existente	Foram adicionadas as `arc-zonal-shift:*` ações para conceder as permissões necessárias para a mudança zonal.	28 de novembro de 2023
ElasticLoadBalancingReadOnly - Atualização em uma política existente	Foram adicionadas as seguintes ações para conceder as permissões necessárias para o deslocamento zonal: `arc-zonal-shift:GetManagedResource` `arc-zonal-shift:ListManagedResources` e. `arc-zonal-shift:ListZonalShifts`	28 de novembro de 2023
AWSElasticLoadBalancingServiceRolePolicy - Atualização em uma política existente	Foi adicionada a `ec2:DescribeVpcPeeringConnections` ação para conceder as permissões necessárias para conexões de emparelhamento.	11 de outubro de 2021
ElasticLoadBalancingFullAccess - Atualização em uma política existente	Foi adicionada a `ec2:DescribeVpcPeeringConnections` ação para conceder as permissões necessárias para conexões de emparelhamento.	11 de outubro de 2021

Alteração	Descrição	Data
ElasticLoadBalancingFullAccess: nova política	Fornece acesso total ao Elastic Load Balancing e aos serviços dependentes.	20 de setembro de 2018
ElasticLoadBalancingReadOnly: nova política	Fornece acesso somente leitura ao Elastic Load Balancing e a serviços dependentes.	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.	20 de setembro de 2018

Validação de conformidade para o Elastic Load Balancing

Para saber se um AWS service (Serviço da AWS) está dentro do escopo de programas de conformidade específicos, consulte Serviços da AWS Escopo por Programa de Conformidade Serviços da AWS e escolha o programa de conformidade em que você está interessado. Para obter informações gerais, consulte Programas de AWS conformidade Programas AWS de .

Você pode baixar relatórios de auditoria de terceiros usando AWS Artifact. Para obter mais informações, consulte Baixar relatórios em AWS Artifact .

Sua responsabilidade de conformidade ao usar Serviços da AWS é determinada pela confidencialidade de seus dados, pelos objetivos de conformidade de sua empresa e pelas leis e regulamentações aplicáveis. AWS fornece os seguintes recursos para ajudar na conformidade:

- Governança e conformidade de segurança: esses guias de implementação de solução abordam considerações sobre a arquitetura e fornecem etapas para implantar recursos de segurança e conformidade.
- Referência de serviços qualificados para HIPAA: lista os serviços qualificados para HIPAA. Nem todos Serviços da AWS são elegíveis para a HIPAA.
- AWS Recursos de https://aws.amazon.com/compliance/resources/ de conformidade — Essa coleção de pastas de trabalho e guias pode ser aplicada ao seu setor e local.
- AWS Guias de conformidade do cliente — Entenda o modelo de responsabilidade compartilhada sob a ótica da conformidade. Os guias resumem as melhores práticas de proteção 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)).

- Avaliação de recursos com regras no Guia do AWS Config desenvolvedor — O AWS Config serviço avalia o quão bem suas configurações de recursos estão em conformidade com as práticas internas, as diretrizes e os regulamentos do setor.
- AWS Security Hub— Isso AWS service (Serviço da AWS) fornece uma visão abrangente do seu estado de segurança interno AWS. O Security Hub usa controles de segurança para avaliar os recursos 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.
- Amazon GuardDuty — Isso AWS service (Serviço da AWS) detecta possíveis ameaças às suas cargas de trabalho Contas da AWS, contêineres e dados monitorando seu ambiente em busca de atividades suspeitas e maliciosas. GuardDuty pode ajudá-lo a atender a vários requisitos de conformidade, como o PCI DSS, atendendo aos requisitos de detecção de intrusões exigidos por determinadas estruturas de conformidade.
- AWS Audit Manager— Isso AWS service (Serviço da AWS) ajuda você a auditar continuamente seu AWS uso para simplificar a forma como você gerencia o risco e a conformidade com as regulamentações e os padrões do setor.

Resiliência no Elastic Load Balancing

A infraestrutura AWS global é construída em torno de AWS regiões e zonas de disponibilidade. 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 altamente disponíveis, tolerantes a falhas e escaláveis que uma ou várias infraestruturas de data center tradicionais.

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

Além da infraestrutura AWS global, o Elastic Load Balancing fornece os seguintes recursos para apoiar sua resiliência de dados:

- Distribui o tráfego de entrada entre várias instâncias em uma única zona de disponibilidade ou em várias zonas de disponibilidade.
- Você pode usar AWS Global Accelerator com seus Application Load Balancers para distribuir o tráfego de entrada entre vários balanceadores de carga em uma ou mais regiões. 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 Docker em um cluster de EC2 instâncias. 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 serviço Elastic Container da Amazon.

Segurança de infraestrutura no Elastic Load Balancing

Como um serviço gerenciado, o Elastic Load Balancing é protegido pela segurança de rede AWS global. Para obter informações sobre serviços AWS de segurança e como AWS proteger a infraestrutura, consulte AWS Cloud Security. Para projetar seu AWS ambiente usando as melhores práticas de segurança de infraestrutura, consulte Proteção de infraestrutura no Security Pillar AWS Well-Architected Framework.

Você usa chamadas de API AWS publicadas para acessar o Elastic Load Balancing pela rede. Os clientes devem oferecer compatibilidade com:

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

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

Isolamento de rede

Uma nuvem privada virtual (VPC) é uma rede virtual em sua própria área logicamente isolada 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

EC2 instâncias nas sub-redes da sua VPC e registrá-las no seu balanceador de carga. 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 obter 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 AWS Certificate Manager (ACM) ou AWS Identity and Access Management (IAM) para gerenciar os certificados do servidor instalados no seu 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 das suas EC2 instâncias da Amazon para aceitar 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 VPC de interface são alimentados por AWS PrivateLink um recurso que permite a comunicação entre seus aplicativos e o Serviços da AWS uso de endereços IP privados. Para obter mais informações, consulte AWS PrivateLink.

Limite

AWS PrivateLink não suporta balanceadores de carga de rede com mais de 50 ouvintes.

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 obter 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 Controlar o Acesso a Serviços Usando Políticas de Endpoint no Guia do AWS PrivateLink .

Limitação de solicitações para a API Elastic Load Balancing

O Elastic Load Balancing limita suas solicitações de API para cada AWS conta por região. Fazemos isso para ajudar no desempenho e na disponibilidade do serviço. A limitação garante que as solicitações para a API do Elastic Load Balancing não excedam os limites máximos permitidos de solicitações de API. As solicitações de API estão sujeitas aos limites de solicitação, quer você as chame ou elas sejam chamadas em seu nome (por exemplo, pelo aplicativo AWS Management Console ou por um aplicativo de terceiros).

Se você exceder o limite de limitação da API Elastic Load Balancing, receberá o código de erro e uma `ThrottlingException` mensagem de erro. `Rate exceeded`

Recomendamos que você se prepare para lidar com a limitação de velocidade sem problemas. Para obter mais informações, consulte Tempos limite, novas tentativas e recuo com variação de sinal. Se você tiver um alto nível de limitação, entre em contato AWS Support para ajudá-lo a avaliar o uso da API e as possíveis soluções. Cada caso é avaliado individualmente. Suporte pode aumentar seus limites dentro dos limites de segurança do sistema, para manter a alta disponibilidade e o desempenho previsível.

Como o controle de utilização é aplicado

O Elastic Load Balancing usa o algoritmo de token bucket para implementar a limitação de API. Com esse algoritmo, sua conta tem um bucket que contém um número específico de tokens. O número de tokens no bucket representa seu limite de limitação a qualquer segundo.

O Elastic Load Balancing fornece dois conjuntos de ações de API. A API ELB versão 2 oferece suporte aos seguintes tipos de balanceadores de carga: balanceadores de carga de aplicativos, balanceadores de carga de rede e balanceadores de carga de gateway. A versão 1 da API ELB é compatível com balanceadores de carga clássicos. Cada versão da API ELB tem seus próprios buckets e tokens.

Serviços que chamam a API do Elastic Load Balancing em seu nome, como Amazon, Amazon ECS, EC2 Amazon EC2 Auto Scaling, e AWS CloudFormation têm seus próprios buckets no nível da conta. Esses serviços não consomem tokens de seus buckets.

Limitação de intervalo de solicitações

Com a limitação da taxa de solicitação, o número de solicitações de API que você faz é limitado. Cada solicitação feita remove um token do bucket. Por exemplo, o tamanho do token bucket para ações de API sem mutação é de 40 tokens. Você pode fazer até 40 `Describe*` solicitações em um segundo. Se você exceder 40 `Describe*` solicitações em um segundo, você será limitado e as solicitações restantes nesse segundo falharão.

Os buckets são recarregados automaticamente a uma taxa definida. Se um bucket estiver abaixo de sua capacidade máxima, um determinado número de tokens será adicionado novamente a cada segundo até que o bucket atinja sua capacidade máxima. Se um balde estiver cheio quando os tokens de recarga chegarem, eles serão descartados. Um bucket não pode conter mais do que seu número máximo de tokens. Por exemplo, o tamanho do bucket para ações de API sem mutação é de 40 tokens e a taxa de recarga é de 10 tokens por segundo. Se você fizer 40 `DescribeLoadBalancers` solicitações em um segundo, o bucket será reduzido para zero (0) tokens. Adicionamos 10 fichas de recarga ao balde a cada segundo, até que ele atinja sua capacidade máxima de 40 fichas. Isso significa que são necessários 4 segundos para que um bucket vazio atinja sua capacidade máxima, se nenhuma solicitação for feita durante esse período.

Você não precisa esperar que um bucket esteja completamente cheio para poder fazer solicitações de API. Você pode usar tokens à medida que eles são adicionados a um bucket. Se você usar imediatamente os tokens de recarga, o bucket não atingirá sua capacidade máxima.

Há um limite de limitação no nível da conta que é compartilhado entre todas as ações da API Elastic Load Balancing. A capacidade do bucket no nível da conta é de 40 tokens e a taxa de recarga é de 10 tokens de solicitação por segundo.

Solicite tamanhos de baldes de tokens e taxas de recarga

Para fins de limitação da taxa de solicitação, as ações da API são agrupadas em categorias. Cada categoria tem seus próprios limites.

Categorias

- Ações de mutação — ações de API que criam, modificam ou excluem recursos. Essa categoria geralmente inclui todas as ações de API que não são categorizadas como ações sem mutação. Essas ações têm um limite de limitação menor do que as ações de API sem mutação.
- Ações não mutantes — ações de API que recuperam dados sobre recursos. Essas ações de API geralmente têm os maiores limites de limitação de API.

- Ações que consomem muitos recursos — ações de API que levam mais tempo e consomem mais recursos para serem concluídas. Essas ações têm um limite de limitação ainda menor do que as ações de mutação. Essas ações são limitadas separadamente de outras ações mutantes.
- Ações de registro — ações de API que registram ou cancelam o registro de alvos. Essas ações da API são limitadas separadamente de outras ações de mutação.
- Ações não categorizadas — Essas ações de API recebem seus próprios tamanhos de repositórios de tokens e taxas de recarga, embora se enquadrem em uma das outras categorias.

A tabela a seguir mostra a capacidade padrão e as taxas de recarga dos buckets de tokens de solicitação categorizados.

Categoria	ELBv2 ações	ELBv1 ações	Capacidade do balde	Taxa de recarga (por segundo)
Consome muitos recursos	`CreateLoadBalancer` , `SetSubnets`	`CreateLoadBalancer` , `AttachLoadBalancerToSubnets` , `DetachLoadBalancerFromSubnets` , `EnableAvailabilityZonesForLoadBalancer` , `DisableAvailabilityZonesForLoadBalancer`	10	0,2 †
Registro	`RegisterTargets` , `DeregisterTargets`	`RegisterInstancesWithLoadBalancer` , `DeregisterInstancesFromLoadBalancer`	20	4
Não mutante	`DescribeAccountLimits` , `DescribeListenerCertificate`	`Describe*`	40	10

Categoria	ELBv2 ações	ELBv1 ações	Capacidade do balde	Taxa de recarga (por segundo)
	`s` , `DescribeListeners` , `DescribeLoadBalancerAttributes` , `DescribeLoadBalancers` , `DescribeRules` , `DescribeSSLPolicies` , `DescribeTags` , `DescribeTargetGroupAttributes` , `DescribeTargetGroups` , `DescribeTargetHealth`			

Categoria	ELBv2 ações	ELBv1 ações	Capacidade do balde	Taxa de recarga (por segundo)
Mutando	`AddListenerCertificates`, `AddTags`, `CreateListener`, `CreateRule`, `CreateTargetGroup`, `DeleteListener`, `DeleteLoadBalancer`, `DeleteRule`, `DeleteTargetGroup`, `ModifyListener`, `ModifyLoadBalancerAttributes`, `ModifyRule`, `ModifyTargetGroup`, `ModifyTargetGroupAttributes`, `RemoveListenerCertificates`, `RemoveTags`, `SetIpAddressType`, `SetRulePriorities`, `SetSecurityGroups`	`AddTags`, `ApplySecurityGroupsToLoadBalancer`, `ConfigureHealthCheck`, `CreateAppCookieStickinessPolicy`, `CreateLbCookieStickinessPolicy`, `CreateLoadBalancerListener`, `CreateLoadBalancerPolicy`, `Delete*`, `ModifyLoadBalancerAttributes`, `RemoveTags`, `SetLoadBalancer*`	20	3

A tabela a seguir mostra a capacidade padrão e as taxas de recarga dos buckets de tokens de solicitação não categorizados para. ELBv2

ELBv2 ações	Capacidade do balde	Taxa de recarga (por segundo)
`CreateTrustStore`	10	0,2 †
`AddTrustStoreRevocations`, `DeleteSharedTrustStoreAssoc`	10	0,2 †

ELBv2 ações	Capacidade do balde	Taxa de recarga (por segundo)
`iation` , `DeleteTrustStore` , `ModifyTrustStore` , `RemoveTrustStoreRevocations`		
`GetTrustStoreCaCertificatesBundle` , `GetTrustStoreRevocationContent`	20	4
`DescribeTrustStoreAssociations` , `DescribeTrustStoreRevocations` , `DescribeTrustStores`	40	10

† As taxas de recarga fracionárias requerem vários segundos para gerar um token completo.

Monitoramento de solicitações de API

Você pode usar AWS CloudTrail para monitorar suas solicitações da API Elastic Load Balancing. Para obter mais informações, consulte Registre chamadas de API para o Elastic Load Balancing usando AWS CloudTrail.

Entenda os códigos do Elastic Load Balancing em relatórios de faturamento e uso

Quando você usa o Elastic Load Balancing, incluímos códigos relacionados em seus relatórios de AWS faturamento e uso. A revisão desses códigos ajuda você a entender os custos e os padrões de uso do balanceador de carga. Acompanhar e gerenciar suas despesas é essencial para otimizar seus custos.

Para obter mais informações, consulte Preço do Elastic Load Balancing.

As tabelas a seguir descrevem os códigos do Elastic Load Balancing que aparecem nos seus relatórios de faturamento e uso. As unidades são horas ou unidades de capacidade do balanceador de carga (LCU). Cada tipo de balanceador de carga tem uma definição específica de LCU. Para obter informações sobre cada tipo de balanceador de carga, consulte os preços do Elastic Load Balancing. LCUs Para obter uma lista dos códigos de região usados nos relatórios de faturamento e uso, consulte Códigos de cobrança AWS da região.

Application Load Balancers

Código	Descrição	Unidades
region-LoadBalancerUsage	O tempo de execução.	Horas
region-LCUUsage	O LCUs usado.	LCU
region-IdleProvisionedLBCapacity	O LCUs reservado, mas não usado.	LCU
region-TS-LoadBalancerUsage	A hora em que um armazenamento fiduciário é usado pela Mutual TLS.	Horas
region-Outposts-LoadBalancerUsage	O tempo de execução em Outposts.	Horas

Código	Descrição	Unidades
region-Outposts-LCUUsage	O LCUs usado em Outposts.	LCU
region-ReservedLCUUsage	O LCUs reservado.	LCU

Network Load Balancers

Código	Descrição	Unidades
region-LoadBalancerUsage	O tempo de execução.	Horas
region-LCUUsage	O LCUs usado.	LCU

Balanceadores de carga de gateway

Código	Descrição	Unidades
region-LoadBalancerUsage	O tempo de execução.	Horas
region-LCUUsage	O LCUs usado.	LCU

Classic Load Balancers

Código	Descrição	Unidades
region-LoadBalancerUsage	O tempo de execução.	Horas

Código	Descrição	Unidades
region-DataProcessing-Bytes	Os dados processados.	GB
region-IdleProvisionedLBCapacity	O LCUs reservado, mas não usado.	LCU

Registre chamadas de API para o Elastic Load Balancing usando AWS CloudTrail

O Elastic Load Balancing é integrado com AWS CloudTrail um serviço que fornece um registro das ações realizadas por um usuário, função ou serviço. AWS CloudTrail captura chamadas de API para o Elastic Load Balancing como eventos. As chamadas capturadas incluem chamadas de AWS Management Console e chamadas de código para as operações da API Elastic Load Balancing. Usando as informações coletadas por CloudTrail, você pode determinar a solicitação que foi feita ao Elastic Load Balancing, o endereço IP do qual a solicitação foi feita, quando foi feita e detalhes adicionais.

Cada entrada de log ou evento contém informações sobre quem gerou a solicitação. As informações de identidade ajudam a determinar o seguinte:

- Se a solicitação foi feita com credenciais de usuário raiz ou credenciais de usuário.
- Se a solicitação foi feita em nome de um usuário do Centro de Identidade do IAM.
- Se a solicitação foi feita com credenciais de segurança temporárias de um perfil ou de um usuário federado.
- Se a solicitação foi feita por outro AWS service (Serviço da AWS).

CloudTrail está ativo Conta da AWS quando você cria a conta e você tem acesso automático ao histórico de CloudTrail eventos. O histórico de CloudTrail eventos fornece um registro visível, pesquisável, baixável e imutável dos últimos 90 dias de eventos de gerenciamento registrados em um. Região da AWS Para obter mais informações, consulte Trabalhando com o histórico de CloudTrail eventos no Guia AWS CloudTrail do usuário. Não há CloudTrail cobrança pela visualização do histórico de eventos.

Para um registro contínuo dos eventos dos Conta da AWS últimos 90 dias, crie uma trilha ou um armazenamento de dados de eventos do CloudTrailLake.

CloudTrail trilhas

Uma trilha permite CloudTrail entregar arquivos de log para um bucket do Amazon S3. Todas as trilhas criadas usando o AWS Management Console são multirregionais. Só é possível criar uma trilha de região única ou de várias regiões usando a AWS CLI. É recomendável criar uma trilha multirregional porque você captura todas as atividades Regiões da AWS em sua conta. Ao criar

uma trilha de região única, é possível visualizar somente os eventos registrados na Região da AWS da trilha. Para obter mais informações sobre trilhas, consulte Criar uma trilha para a Conta da AWS e Criar uma trilha para uma organização no Guia do usuário do AWS CloudTrail .

Você pode entregar uma cópia dos seus eventos de gerenciamento em andamento para o bucket do Amazon S3 sem nenhum custo CloudTrail criando uma trilha. No entanto, existem taxas de armazenamento do Amazon S3. Para obter mais informações sobre CloudTrail preços, consulte AWS CloudTrail Preços. Para receber informações sobre a definição de preços do Amazon S3, consulte Definição de preços do Amazon S3.

CloudTrail Armazenamentos de dados de eventos em Lake

CloudTrail O Lake permite que você execute consultas baseadas em SQL em seus eventos. CloudTrail O Lake converte eventos existentes no formato JSON baseado em linhas para o formato Apache ORC. O ORC é um formato colunar de armazenamento otimizado para recuperação rápida de dados. Os eventos são agregados em armazenamentos de dados de eventos, que são coleções imutáveis de eventos baseados nos critérios selecionados com a aplicação de seletores de eventos avançados. Os seletores que aplicados a um armazenamento de dados de eventos controlam quais eventos persistem e estão disponíveis para consulta. Para obter mais informações sobre o CloudTrail Lake, consulte Trabalhando com o AWS CloudTrail Lake no Guia AWS CloudTrail do Usuário.

CloudTrail Os armazenamentos e consultas de dados de eventos em Lake incorrem em custos. Ao criar um armazenamento de dados de eventos, você escolhe a opção de preço que deseja usar para ele. A opção de preço determina o custo para a ingestão e para o armazenamento de eventos, e o período de retenção padrão e máximo para o armazenamento de dados de eventos. Para obter mais informações sobre CloudTrail preços, consulte AWS CloudTrail Preços.

Eventos de gerenciamento do Elastic Load Balancing em CloudTrail

Os eventos de gerenciamento fornecem informações sobre as operações de gerenciamento que são realizadas nos recursos do seu Conta da AWS. Também são conhecidas como operações de ambiente de gerenciamento. Por padrão, CloudTrail registra eventos de gerenciamento.

O Elastic Load Balancing registra em log as operações do ambiente de gerenciamento como eventos de gerenciamento. Para obter uma lista das operações do ambiente de gerenciamento, consulte os seguintes recursos:

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

Exemplos de eventos do Elastic Load Balancing

Um evento representa uma única solicitação de qualquer fonte e inclui informações sobre a operação de API solicitada, a data e a hora da operação, os parâmetros da solicitação e assim por diante. CloudTrail os arquivos de log não são um rastreamento de pilha ordenado das chamadas públicas de API, portanto, os eventos não aparecem em nenhuma ordem específica.

Os exemplos a seguir mostram CloudTrail eventos de um usuário que criou um balanceador de carga e o excluiu usando o. AWS CLI Você pode identificar a CLI usando os elementos `userAgent`. Você pode identificar as chamadas de APIs solicitadas usando os elementos `eventName`. Informações sobre o usuário (`Alice`) podem ser encontradas no elemento `userIdentity`.

Example Exemplo 1: CreateLoadBalancer da ELBv2 API


```
{
    "eventVersion": "1.03",
    "userIdentity": {
        "type": "IAMUser",
        "principalId": "123456789012",
        "arn": "arn:aws:iam::123456789012:user/Alice",
        "accountId": "123456789012",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
        "userName": "Alice"
    },
    "eventTime": "2016-04-01T15:31:48Z",
    "eventSource": "elasticloadbalancing.amazonaws.com",
    "eventName": "CreateLoadBalancer",
    "awsRegion": "us-west-2",
    "sourceIPAddress": "198.51.100.1",
    "userAgent": "aws-cli/1.10.10 Python/2.7.9 Windows/7 botocore/1.4.1",
    "requestParameters": {
        "subnets": ["subnet-8360a9e7","subnet-b7d581c0"],
        "securityGroups": ["sg-5943793c"],
        "name": "my-load-balancer",
        "scheme": "internet-facing"
```

```
    },
    "responseElements": {
        "loadBalancers":[{
            "type": "application",
            "loadBalancerName": "my-load-balancer",
            "vpcId": "vpc-3ac0fb5f",
            "securityGroups": ["sg-5943793c"],
            "state": {"code":"provisioning"},
            "availabilityZones": [
               {"subnetId":"subnet-8360a9e7","zoneName":"us-west-2a"},
               {"subnetId":"subnet-b7d581c0","zoneName":"us-west-2b"}
            ],
            "dNSName": "my-load-balancer-1836718677.us-west-2.elb.amazonaws.com",
            "canonicalHostedZoneId": "Z2P70J7HTTTPLU",
            "createdTime": "Apr 11, 2016 5:23:50 PM",
            "loadBalancerArn": "arn:aws:elasticloadbalancing:us-
west-2:123456789012:loadbalancer/app/my-load-balancer/ffcddace1759e1d0",
            "scheme": "internet-facing"
        }]
    },
    "requestID": "b9960276-b9b2-11e3-8a13-f1ef1EXAMPLE",
    "eventID": "6f4ab5bd-2daa-4d00-be14-d92efEXAMPLE",
    "eventType": "AwsApiCall",
    "apiVersion": "2015-12-01",
    "recipientAccountId": "123456789012"
}
```

Example Exemplo 2: DeleteLoadBalancer da ELBv2 API

```
{
    "eventVersion": "1.03",
    "userIdentity": {
        "type": "IAMUser",
        "principalId": "123456789012",
        "arn": "arn:aws:iam::123456789012:user/Alice",
        "accountId": "123456789012",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
        "userName": "Alice"
    },
    "eventTime": "2016-04-01T15:31:48Z",
    "eventSource": "elasticloadbalancing.amazonaws.com",
    "eventName": "DeleteLoadBalancer",
    "awsRegion": "us-west-2",
```

```
    "sourceIPAddress": "198.51.100.1",
    "userAgent": "aws-cli/1.10.10 Python/2.7.9 Windows/7 botocore/1.4.1",
    "requestParameters": {
        "loadBalancerArn": "arn:aws:elasticloadbalancing:us-
west-2:123456789012:loadbalancer/app/my-load-balancer/ffcddace1759e1d0"
    },
    "responseElements": null,
    "requestID": "349598b3-000e-11e6-a82b-298133eEXAMPLE",
    "eventID": "75e81c95-4012-421f-a0cf-babdaEXAMPLE",
    "eventType": "AwsApiCall",
    "apiVersion": "2015-12-01",
    "recipientAccountId": "123456789012"
}
```

Example Exemplo 3: CreateLoadBalancer da API ELB

```
{
    "eventVersion": "1.03",
    "userIdentity": {
        "type": "IAMUser",
        "principalId": "AIDAJDPLRKLG7UEXAMPLE",
        "arn": "arn:aws:iam::123456789012:user/Alice",
        "accountId": "123456789012",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
        "userName": "Alice"
    },
    "eventTime": "2016-04-01T15:31:48Z",
    "eventSource": "elasticloadbalancing.amazonaws.com",
    "eventName": "CreateLoadBalancer",
    "awsRegion": "us-west-2",
    "sourceIPAddress": "198.51.100.1",
    "userAgent": "aws-cli/1.10.10 Python/2.7.9 Windows/7 botocore/1.4.1",
    "requestParameters": {
        "subnets": ["subnet-12345678","subnet-76543210"],
        "loadBalancerName": "my-load-balancer",
        "listeners": [{
           "protocol": "HTTP",
           "loadBalancerPort": 80,
           "instanceProtocol": "HTTP",
           "instancePort": 80
        }]
    },
    "responseElements": {
```

```
        "dNSName": "my-loadbalancer-1234567890.elb.amazonaws.com"
    },
    "requestID": "b9960276-b9b2-11e3-8a13-f1ef1EXAMPLE",
    "eventID": "6f4ab5bd-2daa-4d00-be14-d92efEXAMPLE",
    "eventType": "AwsApiCall",
    "apiVersion": "2012-06-01",
    "recipientAccountId": "123456789012"
}
```

Example Exemplo 4: DeleteLoadBalancer da API ELB

```
{
    "eventVersion": "1.03",
    "userIdentity": {
        "type": "IAMUser",
        "principalId": "AIDAJDPLRKLG7UEXAMPLE",
        "arn": "arn:aws:iam::123456789012:user/Alice",
        "accountId": "123456789012",
        "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
        "userName": "Alice"
    },
    "eventTime": "2016-04-08T12:39:25Z",
    "eventSource": "elasticloadbalancing.amazonaws.com",
    "eventName": "DeleteLoadBalancer",
    "awsRegion": "us-west-2",
    "sourceIPAddress": "198.51.100.1",
    "userAgent": "aws-cli/1.10.10 Python/2.7.9 Windows/7 botocore/1.4.1",
    "requestParameters": {
        "loadBalancerName": "my-load-balancer"
    },
    "responseElements": null,
    "requestID": "f0f17bb6-b9ba-11e3-9b20-999fdEXAMPLE",
    "eventID": "4f99f0e8-5cf8-4c30-b6da-3b69fEXAMPLE"
    "eventType": "AwsApiCall",
    "apiVersion": "2012-06-01",
    "recipientAccountId": "123456789012"
}
```

Para obter informações sobre o conteúdo do CloudTrail registro, consulte o conteúdo do CloudTrail registro no Guia AWS CloudTrail do usuário.

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 Recursos do Elastic Load Balancing.

Também é possível 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

O uso de um Application Load Balancer em vez de um Classic Load Balancer oferece os seguintes benefícios:

Suporte para:

- Condições do caminho, condições do host e condições do cabeçalho HTTP.
- Redirecionamento de solicitações de um URL para outro e roteamento de solicitações para vários aplicativos em uma única instância. EC2
- Devolução de respostas HTTP personalizadas.
- Registro de destinos por endereço IP e registro de funções do Lambda como destinos. Inclusão de destinos fora da VPC para o balanceador de carga.
- Autenticação de usuários por meio de identidades corporativas ou sociais.
- Aplicações em contêineres do Amazon Elastic Container Service (Amazon ECS).
- Monitoramento independente da integridade de cada serviço.

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

Melhora no desempenho geral do balanceador de carga.

Network Load Balancer

O uso de um Network Load Balancer em vez de um Classic Load Balancer oferece os seguintes benefícios:

Suporte para:

- 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 EC2 instância.
- Aplicações 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.

Migrar 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 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 prefixo `aws:` em sua chave, essas tags não serão migradas.

Ao migrar para um Application Load Balancer

- Se o Classic Load Balancer tiver apenas uma sub-rede, você deve especificar uma segunda.
- Se o Classic Load Balancer tiver receptores HTTP/HTTPS que usam as verificações de integridade TCP, o protocolo de verificação de integridade será atualizado para HTTP e o caminho definido como "/".
- Se o Classic Load Balancer tiver receptores HTTPS usando uma política de segurança personalizada ou sem suporte, 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-alvo: C1,,, CC1, CC2, CG1, CG2, G1 CR1 CS1,,,, M1, M2 HI1 HS1, M3, T1
- Algumas configurações de verificação de integridade do Classic Load Balancer podem não ser transferíveis para o novo grupo de destino. Esses casos serão indicados como uma alteração na seção de resumo do assistente de migração.
- Se o Classic Load Balancer tiver receptores SSL, o assistente de migração cria um receptor TLS usando o certificado e a política de segurança do receptor SSL.

Processo do assistente de migração

Para migrar um Classic Load Balancer usando o assistente de migração

1. Abra o EC2 console da Amazon 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 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 Nomear novo balanceador de carga, em Nome do balanceador de carga, insira um nome para o novo balanceador de carga.
7. Em Nomear novo grupo de destino e revisar destinos, em Nome do grupo de destino, insira um nome para o novo grupo de destino.

8. (Opcional) Em Destinos, você pode revisar as instâncias de destino que serão registradas no novo grupo de destino.
9. (Opcional) Em Revisar tags, você pode revisar as tags que serão aplicadas ao novo balanceador de carga
10. Em Resumo do Application Load Balancer ou Resumo do Network Load Balancer, revise e verifique as opções de configuração atribuídas pelo assistente de migração.
11. Quando o resumo da configuração estiver do seu agrado, escolha Criar Application Load Balancer ou Criar Network Load Balancer para iniciar a migração.

Migrar 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

- Elastic Load Balancing Tools
- Classic Load Balancer to Application Load Balancer copy utility
- Classic Load Balancer to Network Load Balancer copy utility

Migrar o 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 obter 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 EC2 instâncias com seu grupo-alvo.
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 os 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.

Impedir que os usuários criem balanceadores de carga clássicos

Você pode criar uma política do IAM que impeça os usuários de criar balanceadores de carga clássicos na sua conta.

Tanto o Elastic Load Balancing V2 quanto o Elastic Load Balancing APIs V1 fornecem uma ação de API. `CreateLoadBalancer` Ao criar um Classic Load Balancer, você usa a ação da API V1, que cria o balanceador de carga e os ouvintes. Ao criar um Application Load Balancer, Network Load Balancer ou Gateway Load Balancer, você usa a ação da API V2, que cria somente o balanceador de carga. A API V2 fornece uma `CreateListener` ação, que você usa para criar ouvintes para um balanceador de carga depois de criá-lo.

A política a seguir nega aos usuários a permissão de criar um balanceador de carga se o protocolo do ouvinte for especificado. Como você deve configurar pelo menos um ouvinte ao criar um Classic Load Balancer, essa política impede que os usuários criem Classic Load Balancers. Isso não impede que os usuários criem outros tipos de balanceadores de carga, porque há ações de API separadas para criar esses balanceadores de carga e seus ouvintes.

```
{
    "Version": "2012-10-17",
    "Effect": "Deny",
    "Action": "elasticloadbalancing:CreateLoadBalancer",
    "Resource": [
        "arn:aws:elasticloadbalancing:*:*:loadbalancer/*"
    ],
    "Condition": {
        "Null": {
            "elasticloadbalancing:ListenerProtocol": false
```

```
            }
        }
}
```

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.