aws

Gateway Load Balancers

# Elastic Load Balancing

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

# Elastic Load Balancing: Gateway Load Balancers

Copyright © 2026 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

I marchi e l’immagine commerciale di Amazon non possono essere utilizzati in relazione a prodotti o servizi che non siano di Amazon, in una qualsiasi modalità che possa causare confusione tra i clienti o in una qualsiasi modalità che denigri o discrediti Amazon. Tutti gli altri marchi non di proprietà di Amazon sono di proprietà dei rispettivi proprietari, che possono o meno essere affiliati, collegati o sponsorizzati da Amazon.

# Table of Contents

Cos'è un Gateway Load Balancer? ... 1
Panoramica di Gateway Load Balancer ... 1
Fornitori di appliance ... 2
Nozioni di base ... 2
Prezzi ... 2
Nozioni di base ... 3
Panoramica di ... 3
Routing ... 5
Prerequisiti ... 6
Fase 1: creare un Gateway Load Balancer ... 6
Fase 2: creazione di un servizio endpoint del Gateway Load Balancer ... 8
Fase 3: creare un endpoint del Gateway Load Balancer ... 9
Fase 4: configurazione del routing ... 10
Nozioni di base per l'utilizzo della CLI ... 12
Panoramica di ... 12
Routing ... 5
Prerequisiti ... 15
Fase 1: creare un Gateway Load Balancer e registrare le destinazioni ... 15
Fase 2: creare un endpoint del Gateway Load Balancer ... 17
Fase 3: configurazione dell'instradamento ... 18
Gateway Load Balancers ... 20
Stato del sistema di bilanciamento del carico ... 20
Tipo di indirizzo IP ... 21
Zone di disponibilità ... 22
Tempo di inattività ... 22
Attributi del sistema di bilanciamento del carico ... 23
Liste di controllo accessi (ACL) di rete ... 23
Flussi asimmetrici ... 24
Unità di trasmissione massima (MTU) ... 24
Creazione di un sistema di bilanciamento del carico ... 24
Prerequisiti ... 25
Creazione del sistema di bilanciamento del carico ... 25
Fasi successive importanti ... 26
Aggiornare il tipo di indirizzo IP ... 26

Modifica gli attributi del load balancer ... 27
Protezione da eliminazione ... 27
Cross-zone bilanciamento del carico ... 28
Assegna un tag a un load balancer ... 28
Eliminazione di un sistema di bilanciamento del carico ... 30
Prenotazioni LCU ... 30
Richiedere una prenotazione ... 32
Aggiornare o terminare la prenotazione ... 33
Monitora la prenotazione ... 34
Listener ... 35
attributi del listener ... 35
Aggiorna il gruppo target dei listener ... 35
Aggiorna il timeout di inattività ... 36
Invia il ripristino TCP in caso di timeout di inattività ... 37
Gruppi di destinazione ... 39
Configurazione dell'instradamento ... 39
Target type (Tipo di destinazione) ... 40
Destinazioni registrate ... 41
Attributi dei gruppi di destinazione ... 41
Creazione di un gruppo target ... 42
Configurazione dei controlli dello stato ... 44
Impostazioni del controllo dello stato ... 44
Stato di integrità della destinazione ... 46
Codici di motivo di controllo dello stato ... 47
Scenari di errore della destinazione ... 48
Controllo dello stato delle destinazioni ... 49
Modifica le impostazioni del controllo dell'integrità ... 50
Modifica gli attributi del gruppo target ... 50
Failover della destinazione ... 51
Reimpostazione del protocollo TCP in caso di errore e annullamento della registrazione del target ... 52
Ritardo di annullamento della registrazione ... 54
Persistenza del flusso ... 54
Registrazione di destinazioni ... 55
Considerazioni ... 56
Gruppi di sicurezza target ... 56

Liste di controllo accessi (ACL) di rete ..... 57
Registra le destinazioni in base all'ID dell'istanza ..... 57
Registra le destinazioni in base all'indirizzo IP ..... 57
Annulla la registrazione degli obiettivi ..... 58
Contrassegna un gruppo target ..... 59
Eliminazione di un gruppo target ..... 60
Monitoraggio dei sistemi di bilanciamento del carico ..... 61
CloudWatch metriche ..... 62
Parametri di Gateway Load Balancer ..... 62
Dimensioni di parametro per Gateway Load Balancer ..... 69
Visualizza le metriche per CloudWatch il tuo Gateway Load Balancer ..... 69
Quote ..... 72
Sistemi di load balancer ..... 72
Gruppi di destinazione ..... 72
Unità di capacità Load Balancer ..... 73
Cronologia dei documenti ..... 74
..... lxxvi

# Cos'è un Gateway Load Balancer?

Il servizio di bilanciamento del carico elastico distribuisce automaticamente il traffico in ingresso su più destinazioni, in una o più zone di disponibilità. Monitora lo stato di integrità delle destinazioni registrate e instrada il traffico solo verso le destinazioni integre. Elastic Load Balancing ridimensiona il load balancer di volta in volta, in quanto il traffico in ingresso varia nel corso del tempo. Può ridimensionare le risorse per la maggior parte dei carichi di lavoro automaticamente.

Elastic Load Balancing supporta i seguenti bilanciatori del carico: Application Load Balancer, Network Load Balancer, Gateway Load Balancer e Classic Load Balancer. È possibile selezionare il tipo di load balancer più adatto alle proprie esigenze. In questa guida vengono illustrati i Gateway Load Balancer. Per ulteriori informazioni sugli altri sistemi di bilanciamento del carico, consulta la Guida per l'utente sugli Application Load Balancer, la Guida per l'utente sui Network Load Balancer, e la Guida per l'utente sui Classic Load Balancer.

## Panoramica di Gateway Load Balancer

I Gateway Load Balancer consentono di implementare, dimensionare e gestire appliance virtuali, come firewall, sistemi di prevenzione e rilevamento delle intrusioni e sistemi di ispezione approfondita dei pacchetti. Combina un gateway di rete trasparente (ovvero un unico punto di entrata e uscita per tutto il traffico) e distribuisce il traffico dimensionando le appliance virtuali in base alla domanda.

Un Gateway Load Balancer funziona al terzo livello del modello Open Systems Interconnection (OSI), il livello di rete. È in ascolto per tutti i pacchetti IP attraverso tutte le porte e inoltra il traffico al gruppo di destinazione specificato nella regola del listener. Mantiene l'aderenza del flusso a uno specifico dispositivo di destinazione utilizzando 5 tuple (impostazione predefinita), 3 tuple o 2 tuple. Il Gateway Load Balancer e le istanze dell'appliance virtuale registrate scambiano il traffico delle applicazioni utilizzando il protocollo GENEVE sulla porta 6081.

I Gateway Load Balancer utilizzano gli endpoint Gateway Load Balancer per scambiare in modo sicuro il traffico attraverso i confini VPC. Un endpoint Gateway Load Balancer è un endpoint VPC che fornisce connettività privata tra appliance virtuali nel VPC del provider di servizi e server delle applicazioni nel VPC consumer del servizio. Distribuisci il Gateway Load Balancer nello stesso VPC delle appliance virtuali. Registra le appliance virtuali con un gruppo di destinazione per il Gateway Load Balancer.

Il traffico da e verso un endpoint del Gateway Load Balancer viene configurato tramite le tabelle di routing. Il traffico scorre dal VPC consumatore del servizio sull'endpoint del Gateway Load Balancer

al Gateway Load Balancer nel VPC fornitore del servizio, per poi tornare al VPC consumatore. È necessario creare l'endpoint del Gateway Load Balancer e i server delle applicazioni in sottoreti diverse. In questo modo è possibile configurare l'endpoint del Gateway Load Balancer come hop successivo nella tabella di routing per la sottorete dell'applicazione.

Per ulteriori informazioni, consulta Accesso alle appliance virtuali tramite AWS PrivateLink nella Guida di AWS PrivateLink .

## Fornitori di appliance

L'utente è responsabile della scelta e della qualificazione del software proveniente dai fornitori di appliance. È necessario affidarsi al software dell'appliance per ispezionare o modificare il traffico dal sistema di bilanciamneto del carico. I fornitori di appliance elencati come Elastic Load Balancing Partner hanno integrato e qualificato il loro software di appliance con. AWSÈ possibile attribuire un grado di affidabilità maggiore ai software di appliance offerti dai fornitori presenti in questo elenco. Tuttavia, AWS non garantisce la sicurezza o l'affidabilità dei software di questi fornitori.

## Nozioni di base

Per creare un Gateway Load Balancer utilizzando il Console di gestione AWS, vedere. Nozioni di base Per creare un Gateway Load Balancer utilizzando il AWS Command Line Interface, vedere. Nozioni di base per l'utilizzo della CLI

## Prezzi

Con il load balancer paghi solo in base all'uso effettivo. Per ulteriori informazioni, consultare Prezzi di Elastic Load Balancing.

# Nozioni di base sui Gateway Load Balancer.

I Gateway Load Balancer semplificano distribuzione, dimensionamento e gestione delle appliance virtuali di terze parti, ad esempio quelle di sicurezza.

In questo tutorial implementeremo un sistema di ispezione utilizzando un Gateway Load Balancer e un endpoint del Gateway Load Balancer.

Indice

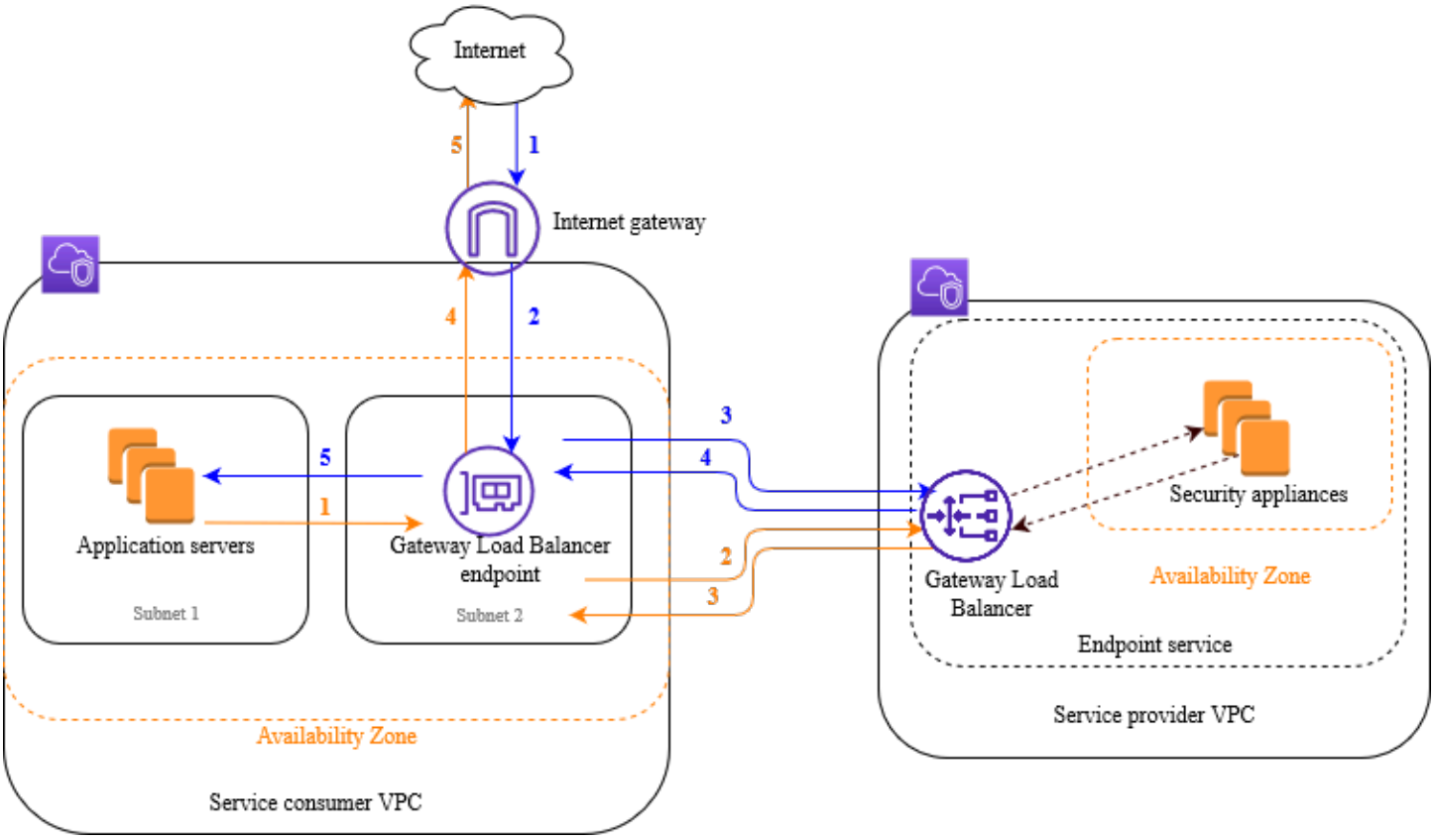

- Panoramica di
- Prerequisiti
- Fase 1: creare un Gateway Load Balancer
- Fase 2: creazione di un servizio endpoint del Gateway Load Balancer
- Fase 3: creare un endpoint del Gateway Load Balancer
- Fase 4: configurazione del routing


## Panoramica di

Un endpoint Gateway Load Balancer è un endpoint VPC che fornisce connettività privata tra appliance virtuali nel VPC del provider di servizi e server delle applicazioni nel VPC consumer del servizio. Gateway Load Balancer viene distribuito nello stesso VPC delle appliance virtuali. Queste appliance vengono registrate come gruppo di destinazione del Gateway Load Balancer.

I server delle applicazioni vengono eseguiti in una sottorete (di destinazione) nel VPC del consumatore di servizi, mentre l'endpoint Gateway Load Balancer si trova in un'altra sottorete dello stesso VPC. Tutto il traffico che entra nel VPC dell'utente del servizio attraverso il gateway Internet viene innanzitutto instradato all'endpoint Gateway Load Balancer e poi instradato alla sottorete di destinazione.

Analogamente, tutto il traffico che esce dai server dell'applicazione (sottorete di destinazione) viene instradato sull'endpoint Gateway Load Balancer prima di essere instradato nuovamente attraverso Internet. Il seguente diagramma di rete è una rappresentazione visiva di come un endpoint del Gateway Load Balancer viene utilizzato per accedere a un servizio endpoint.




Gli articoli numerati di seguito evidenziano e spiegano gli elementi mostrati nell'immagine precedente.

Traffico in transito da Internet ai server dell'applicazione (frecce blu):

1. Il traffico entra nel VPC dell'utente del servizio attraverso il gateway Internet.
2. Il traffico viene inviato all'endpoint Gateway Load Balancer come risultato del routing in ingresso.
3. Il traffico viene inviato al Gateway Load Balancer che lo distribuisce a una delle appliance di sicurezza.
4. Il traffico viene inviato nuovamente all'endpoint Gateway Load Balancer dopo l'ispezione da parte dell'appliance di sicurezza.
5. Il traffico viene inviato ai server dell'applicazione (sottorete di destinazione).

Traffico in transito dall'applicazione a Internet (frecce arancioni):

1. Il traffico viene inviato all'endpoint Gateway Load Balancer come risultati dell'instradamento predefinito configurato nella sottorete del server dell'applicazione.

2. Il traffico viene inviato al Gateway Load Balancer che lo distribuisce a una delle appliance di sicurezza.
3. Il traffico viene inviato nuovamente all'endpoint Gateway Load Balancer dopo l'ispezione da parte dell'appliance di sicurezza.
4. Il traffico viene inviato al gateway Internet in base alla configurazione della tabella di instradamento.
5. Il traffico viene reindirizzato a Internet.

## Routing

Questa tabella di routing per il gateway Internet deve disporre di una voce che invia il traffico destinato ai server dell'applicazione all'endpoint Gateway Load Balancer. Per specificare l'endpoint del Gateway Load Balancer utilizza l'ID dell'endpoint VPC. L'esempio seguente mostra le route per una configurazione dualstack.

| Destinazione | Target |
| --- | --- |
| *VPC IPv4 CIDR* | Locale |
| *VPC IPv6 CIDR* | Locale |
| *Subnet 1 IPv4 CIDR* | *vpc-endpoint-id* |
| *Subnet 1 IPv6 CIDR* | *vpc-endpoint-id* |

Questa tabella di routing per la sottorete con i server dell'applicazione deve disporre di voci che indirizzino tutto il traffico dai server dell'applicazione all'endpoint del Gateway Load Balancer.

| Destinazione | Target |
| --- | --- |
| *VPC IPv4 CIDR* | Locale |
| *VPC IPv6 CIDR* | Locale |
| 0.0.0. 0/0 | *vpc-endpoint-id* |
| ::/0 | *vpc-endpoint-id* |

La tabella di routing per la sottorete con l'endpoint del Gateway Load Balancer deve instradare il traffico dall'ispezione alla destinazione finale. Per il traffico proveniente da Internet, la route locale garantisce che raggiunga i server dell'applicazione. Per il traffico proveniente dai server dell'applicazione, aggiungi voci che indirizzino tutto il traffico al gateway Internet.

| Destinazione | Target |
| --- | --- |
| *VPC IPv4 CIDR* | Locale |
| *VPC IPv6 CIDR* | Locale |
| 0,0,0. 0/0 | *internet-gateway-id* |
| ::/0 | *internet-gateway-id* |

## Prerequisiti

- Assicurati che il VPC consumatore del servizio disponga di almeno due sottoreti per ogni zona di disponibilità che contiene i server dell'applicazione. Una sottorete è destinata all'endpoint del Gateway Load Balancer e l'altra ai server dell'applicazione.
- Il Gateway Load Balancer e le destinazioni possono trovarsi nella stessa sottorete.
- Non è possibile utilizzare una sottorete condivisa da un altro account per distribuire il Gateway Load Balancer.
- Avvia almeno un'istanza dell'appliance di sicurezza in ogni sottorete nel VPC del provider di servizi. I gruppi di sicurezza per queste istanze devono permettere il traffico UDP sulla porta 6081.

## Fase 1: creare un Gateway Load Balancer

Utilizza la procedura seguente per creare sistema di bilanciamento del carico, ascoltatore e gruppo di destinazione.

Per creare il sistema di bilanciamento del carico, l'ascoltatore e il gruppo target utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, in Bilanciamento del carico, scegli Sistemi di bilanciamento del carico.

3. Selezionare Create Load Balancer (Crea sistema di bilanciamento del carico).
4. Sotto a Gateway Load Balancer scegli Crea.
5. Configurazione di base
   a. In Nome del sistema di bilanciamento del carico immetti un nome univoco per il sistema di bilanciamento del carico.
   b. In IP address type (Tipo di indirizzo IP), seleziona ipv4 per supportare i soli indirizzi IPv4 oppure dualstack per supportare sia gli indirizzi IPv4 che IPv6.
6. Mappatura della rete
   a. Per VPC, seleziona il VPC del fornitore del servizio.
   b. Per le Mappature, seleziona tutte le zone di disponibilità in cui sono state avviate le istanze delle appliance di sicurezza e una sottorete per zona di disponibilità.
7. Routing dell'ascoltatore IP
   a. Per Azione predefinita, seleziona un gruppo di destinazione esistente per ricevere il traffico. Questo gruppo di destinazione deve utilizzare il protocollo GENEVE.

      Se non hai un gruppo di destinazione, seleziona Crea gruppo di destinazione, si aprirà una nuova scheda nel browser. Selezione un tipo di destinazione, inserisci un nome per il gruppo di destinazione e mantieni il protocollo GENEVE. Seleziona il VPC con le istanze dell'appliance di sicurezza. Modifica le impostazioni del controllo dell'integrità secondo necessità e aggiungi i tag di cui hai bisogno. Scegli Next (Successivo). È possibile registrare le istanze dell'appliance di sicurezza con il gruppo di destinazione ora o dopo aver completato questa procedura. Seleziona Crea gruppo di destinazione, quindi torna alla scheda precedente del browser.
   b. (Facoltativo) Espandi Tag dell'ascoltatore e aggiungi i tag di cui hai bisogno.
8. (Facoltativo) Espandi Tag del sistema di bilanciamento del carico e aggiungi i tag di cui hai bisogno.
9. Selezionare Create Load Balancer (Crea sistema di bilanciamento del carico).

# Fase 2: creazione di un servizio endpoint del Gateway Load Balancer

Utilizza la procedura seguente per creare un servizio endpoint utilizzando un Gateway Load Balancer.

Creare di un servizio endpoint Gateway Load Balancer

1. Apri la console Amazon VPC all'indirizzo https://console.aws.amazon.com/vpc/.
2. Nel pannello di navigazione scegli Endpoint Services (Servizi endpoint).
3. Seleziona Crea servizio endpoint ed effettua le seguenti operazioni:
   a. Per Load balancer type (Tipo di load balancer), scegli Gateway.
   b. In Available load balancers (Load balancer disponibili), seleziona il Gateway Load Balancer.
   c. In Richiedi accettazione per l'endpoint, seleziona Accettazione richiesta per accettare manualmente le richieste di connessione al servizio endpoint. In caso contrario, queste vengono accettate automaticamente.
   d. Per Supported IP address types (Tipi di indirizzo IP supportati), esegui una delle operazioni seguenti:
      - Seleziona IPv4 per consentire al servizio endpoint di accettare richieste IPv4.
      - Seleziona IPv6 per consentire al servizio endpoint di accettare richieste IPv6.
      - Seleziona IPv4 e IPv6 per consentire al servizio endpoint di accettare richieste IPv4 e IPv6.
   e. (Facoltativo) Per aggiungere un tag, scegli Add new tag (Aggiungi nuovo tag) e inserisci la chiave e il valore del tag.
   f. Scegli Create (Crea). Il traffico viene inviato all'endpoint del Gateway Load Balancer.
4. Seleziona il propnuovo rio servizio endpoint e scegli Operazioni, Autorizza principali. Inserisci gli ARN dei consumatori del servizio che sono autorizzati a creare un endpoint per il servizio. Un consumatore del servizio può essere un utente, un ruolo IAM o Account AWS. Seleziona Allow principals (Consenti entità principali).

# Fase 3: creare un endpoint del Gateway Load Balancer

Utilizza la procedura seguente per creare un endpoint Gateway Load Balancer che si connette al servizio endpoint del Gateway Load Balancer. Gli endpoint del Gateway Load Balancer sono zonali. È consigliabile creare un endpoint del Gateway Load Balancer per zona. Per ulteriori informazioni, consulta Accesso alle appliance virtuali tramite AWS PrivateLink nella Guida di AWS PrivateLink.

Per creare un endpoint Gateway Load Balancer

1. Apri la console Amazon VPC all'indirizzo https://console.aws.amazon.com/vpc/.
2. Nel pannello di navigazione, seleziona Endpoint.
3. Seleziona Crea endpoint ed effettua le seguenti operazioni:
   a. In Service category (Categoria del servizio), scegli Other endpoint services (Altri servizi endpoint).
   b. In Nome servizio, specifica il nome del servizio annotato precedentemente, quindi seleziona Verifica servizio.
   c. Per VPC, seleziona il VPC del consumatore del servizio.
   d. Per Sottoreti, seleziona una sottorete per l'endpoint del Gateway Load Balancer.

      Nota: è possibile selezionare solo una sottorete all'interno di ciascuna zona di disponibilità durante la creazione di un endpoint Gateway Load Balancer.
   e. Per IP address type (Tipo di indirizzo IP), seleziona una delle opzioni seguenti:
      - IPv4: consente di assegnare indirizzi IPv4 alle interfacce di rete dell'endpoint. Questa opzione è supportata solo se tutte le sottoreti selezionate dispongono di intervalli di indirizzi IPv4.
      - IPv6: consente di assegnare indirizzi IPv6 alle interfacce di rete dell'endpoint. Questa opzione è supportata solo se tutte le sottoreti selezionate sono sottoreti solo IPv6.
      - Dualstack: consente di assegnare sia indirizzi IPv4 che IPv6 alle interfacce di rete dell'endpoint. Questa opzione è supportata solo se tutte le sottoreti selezionate dispongono di intervalli di indirizzi IPv4 e IPv6.
   f. (Facoltativo) Per aggiungere un tag, scegli Add new tag (Aggiungi nuovo tag) e inserisci la chiave e il valore del tag.
   g. Seleziona Crea endpoint. Lo stato iniziale è `pending acceptance`.

Per accettare la richiesta di connessione all'endpoint, utilizza la seguente procedura.

1. Nel pannello di navigazione scegli Endpoint Services (Servizi endpoint).
2. Selezionare il servizio endpoint.
3. Dalla scheda Endpoint connections (Connessioni endpoint), seleziona la connessione endpoint.
4. Per accettare la richiesta di connessione, scegli Actions (Operazioni), Accept endpoint connection request (Accetta richiesta di connessione endpoint). Quando viene richiesta la conferma, immetti **`accept`** e seleziona Accept (Accetta).

# Fase 4: configurazione del routing

Configura le tabelle di routing per il VPC consumatore del servizio come segue. Ciò consente alle appliance di sicurezza di eseguire ispezioni sul traffico in entrata destinato ai server dell'applicazione.

Configurazione del routing

1. Apri la console Amazon VPC all'indirizzo https://console.aws.amazon.com/vpc/.
2. Nel riquadro di navigazione, seleziona Tabelle di routing.
3. Seleziona la tabella di instradamento per il gateway Internet ed esegui le operazioni seguenti:
   a. Seleziona Operazioni, Modifica route.
   b. Seleziona Aggiungi route. Per Destination (Destinazione), immetti il blocco CIDR IPv4 della sottorete per i server dell'applicazione. Per Target, seleziona l'endpoint VPC.
   c. Se supporti IPv6, scegli Add route (Aggiungi route). Per Destination (Destinazione), immetti il blocco CIDR IPv6 della sottorete per i server dell'applicazione. Per Target, seleziona l'endpoint VPC.
   d. Scegli Save changes (Salva modifiche).
4. Seleziona la tabella di instradamento per la sottorete con i server dell'applicazione ed esegui le operazioni seguenti:
   a. Seleziona Operazioni, Modifica route.
   b. Seleziona Aggiungi route. In Destination (Destinazione), immettere **`0.0.0.0/0`**. Per Target, seleziona l'endpoint VPC.
   c. Se supporti IPv6, scegli Add route (Aggiungi route). In Destination (Destinazione), immettere **`::/0`**. Per Target, seleziona l'endpoint VPC.

d. Scegli Save changes (Salva modifiche).

5. Seleziona la tabella di instradamento per la sottorete con l'endpoint Gateway Load Balancer ed esegui le operazioni seguenti:

   a. Seleziona Operazioni, Modifica route.

   b. Seleziona Aggiungi route. In Destination (Destinazione), immettere **`0.0.0.0/0`**. Per Target, seleziona il gateway Internet.

   c. Se supporti IPv6, scegli Add route (Aggiungi route). In Destination (Destinazione), immettere **`::/0`**. Per Target, seleziona il gateway Internet.

   d. Scegli Save changes (Salva modifiche).

# Guida introduttiva ai sistemi Gateway Load Balancer utilizzando AWS CLI

I Gateway Load Balancer semplificano distribuzione, dimensionamento e gestione delle appliance virtuali di terze parti, ad esempio quelle di sicurezza.

In questo tutorial implementeremo un sistema di ispezione utilizzando un Gateway Load Balancer e un endpoint del Gateway Load Balancer.

Indice

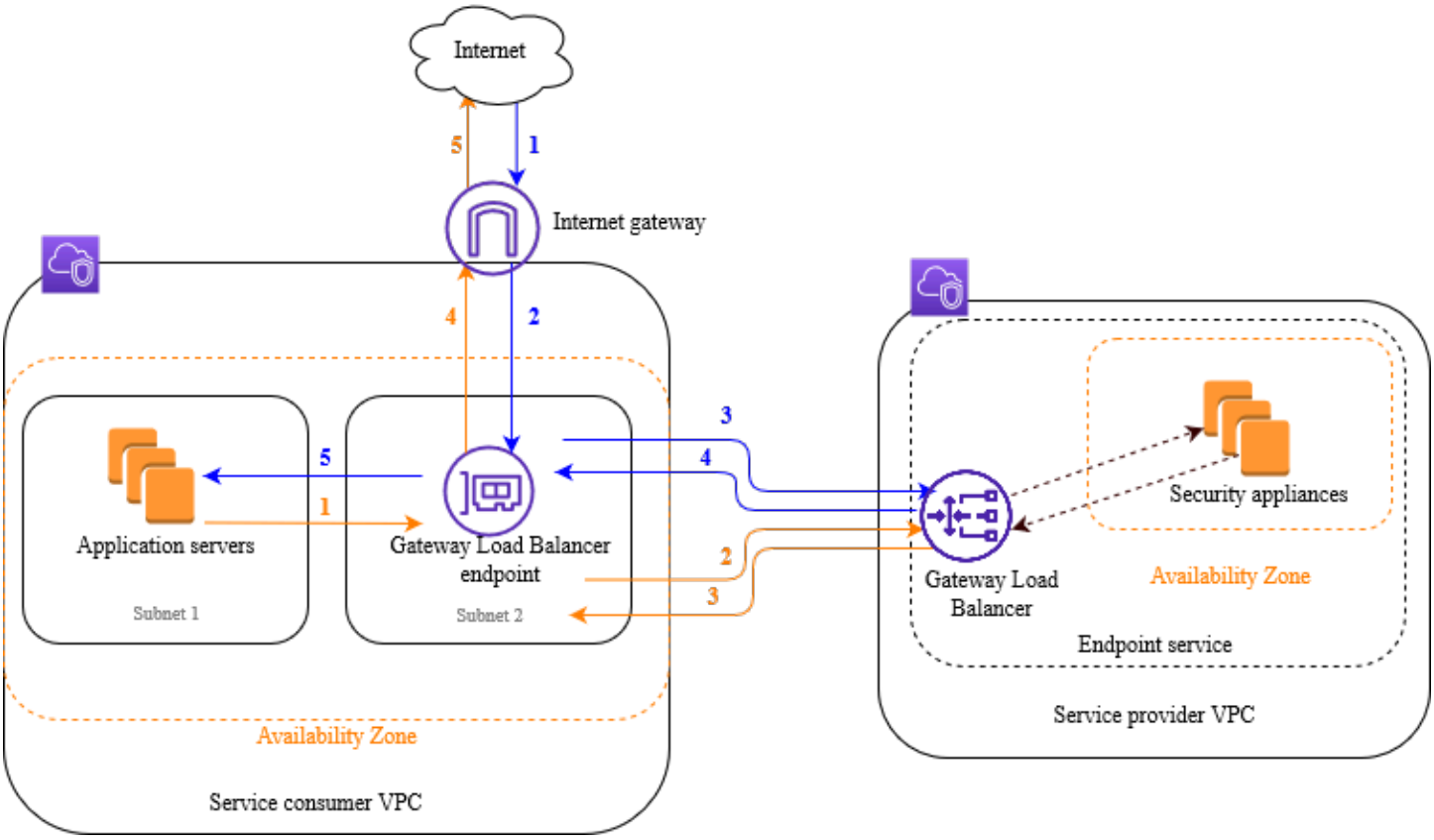

- Panoramica di
- Prerequisiti
- Fase 1: creare un Gateway Load Balancer e registrare le destinazioni
- Fase 2: creare un endpoint del Gateway Load Balancer
- Fase 3: configurazione dell'instradamento


## Panoramica di

Un endpoint Gateway Load Balancer è un endpoint VPC che fornisce connettività privata tra appliance virtuali nel VPC del provider di servizi e server delle applicazioni nel VPC consumer del servizio. Gateway Load Balancer viene distribuito nello stesso VPC delle appliance virtuali. Queste appliance vengono registrate come gruppo di destinazione del Gateway Load Balancer.

I server delle applicazioni vengono eseguiti in una sottorete (di destinazione) nel VPC del consumatore di servizi, mentre l'endpoint Gateway Load Balancer si trova in un'altra sottorete dello stesso VPC. Tutto il traffico che entra nel VPC dell'utente del servizio attraverso il gateway Internet viene innanzitutto instradato all'endpoint Gateway Load Balancer e poi instradato alla sottorete di destinazione.

Analogamente, tutto il traffico che esce dai server dell'applicazione (sottorete di destinazione) viene instradato sull'endpoint Gateway Load Balancer prima di essere instradato nuovamente attraverso Internet. Il seguente diagramma di rete è una rappresentazione visiva di come un endpoint del Gateway Load Balancer viene utilizzato per accedere a un servizio endpoint.




Gli articoli numerati di seguito evidenziano e spiegano gli elementi mostrati nell'immagine precedente.

Traffico in transito da Internet ai server dell'applicazione (frecce blu):

1. Il traffico entra nel VPC dell'utente del servizio attraverso il gateway Internet.
2. Il traffico viene inviato all'endpoint Gateway Load Balancer come risultato del routing in ingresso.
3. Il traffico viene inviato al Gateway Load Balancer che lo distribuisce a una delle appliance di sicurezza.
4. Il traffico viene inviato nuovamente all'endpoint Gateway Load Balancer dopo l'ispezione da parte dell'appliance di sicurezza.
5. Il traffico viene inviato ai server dell'applicazione (sottorete di destinazione).

Traffico in transito dall'applicazione a Internet (frecce arancioni):

1. Il traffico viene inviato all'endpoint Gateway Load Balancer come risultati dell'instradamento predefinito configurato nella sottorete del server dell'applicazione.

2. Il traffico viene inviato al Gateway Load Balancer che lo distribuisce a una delle appliance di sicurezza.
3. Il traffico viene inviato nuovamente all'endpoint Gateway Load Balancer dopo l'ispezione da parte dell'appliance di sicurezza.
4. Il traffico viene inviato al gateway Internet in base alla configurazione della tabella di instradamento.
5. Il traffico viene reindirizzato a Internet.

## Routing

Questa tabella di routing per il gateway Internet deve disporre di una voce che invia il traffico destinato ai server dell'applicazione all'endpoint Gateway Load Balancer. Per specificare l'endpoint del Gateway Load Balancer utilizza l'ID dell'endpoint VPC. L'esempio seguente mostra le route per una configurazione dualstack.

| Destinazione | Target |
| --- | --- |
| *VPC IPv4 CIDR* | Locale |
| *VPC IPv6 CIDR* | Locale |
| *Subnet 1 IPv4 CIDR* | *vpc-endpoint-id* |
| *Subnet 1 IPv6 CIDR* | *vpc-endpoint-id* |

Questa tabella di routing per la sottorete con i server dell'applicazione deve disporre di voci che indirizzino tutto il traffico dai server dell'applicazione all'endpoint del Gateway Load Balancer.

| Destinazione | Target |
| --- | --- |
| *VPC IPv4 CIDR* | Locale |
| *VPC IPv6 CIDR* | Locale |
| 0.0.0. 0/0 | *vpc-endpoint-id* |
| ::/0 | *vpc-endpoint-id* |

La tabella di routing per la sottorete con l'endpoint del Gateway Load Balancer deve instradare il traffico dall'ispezione alla destinazione finale. Per il traffico proveniente da Internet, la route locale garantisce che raggiunga i server dell'applicazione. Per il traffico proveniente dai server dell'applicazione, aggiungi voci che indirizzino tutto il traffico al gateway Internet.

| Destinazione | Target |
| --- | --- |
| *VPC IPv4 CIDR* | Locale |
| *VPC IPv6 CIDR* | Locale |
| 0,0,0. 0/0 | *internet-gateway-id* |
| ::/0 | *internet-gateway-id* |

## Prerequisiti

- Installa AWS CLI o esegui l'aggiornamento alla versione corrente di AWS CLI se utilizzi una versione che non supporta Gateway Load Balancer. Per ulteriori informazioni, consulta Installazione dell'AWS CLI nella Guida per l'utente dell'AWS Command Line Interface.
- Assicurati che il VPC consumatore del servizio disponga di almeno due sottoreti per ogni zona di disponibilità che contiene i server dell'applicazione. Una sottorete è destinata all'endpoint del Gateway Load Balancer e l'altra ai server dell'applicazione.
- Assicurati che il VPC del fornitore di servizi disponga di almeno due sottoreti per ogni zona di disponibilità che contiene i server delle appliance. Una sottorete è destinata al Gateway Load Balancer e l'altra alle istanze.
- Avvia almeno un'istanza dell'appliance di sicurezza in ogni sottorete nel VPC del provider di servizi. I gruppi di sicurezza per queste istanze devono permettere il traffico UDP sulla porta 6081.

## Fase 1: creare un Gateway Load Balancer e registrare le destinazioni

Utilizza la procedura seguente per creare il sistema di bilanciamento del carico, l'ascoltatore, gruppi di destinazione e per registrare le istanze dell'appliance di sicurezza come destinazioni.

Creare un Gateway Load Balancer e registrare le destinazioni

1. Utilizzare il comando create-load-balancer per creare un sistema di bilanciamento del carico del tipo `gateway`. È possibile specificare una sottorete per ogni zona di disponibilità nella quale sono state lanciate le istanze dell'appliance di sicurezza.

   ```
   aws elbv2 create-load-balancer --name my-load-balancer --type gateway --
   subnets provider-subnet-id
   ```

   La modalità predefinita supporta solamente gli indirizzi IPv4. Per il supporto di entrambi gli indirizzi IPv4 e IPv6, aggiungi l'opzione `--ip-address-type dualstack`.

   L'output include il nome della risorsa Amazon (ARN) del sistema di bilanciamento del carico, con il formato illustrato nel seguente esempio.

   ```
   arn:aws:elasticloadbalancing:us-east-2:123456789012:loadbalancer/gwy/my-load-
   balancer/1234567890123456
   ```

2. Utilizza il comando create-target-group per creare un gruppo di destinazione, specificando il VPC del provider di servizi in cui hai avviato le tue istanze.

   ```
   aws elbv2 create-target-group --name my-targets --protocol GENEVE --port 6081 --
   vpc-id provider-vpc-id
   ```

   L'output include l'ARN del gruppo di destinazione, con il seguente formato.

   ```
   arn:aws:elasticloadbalancing:us-east-2:123456789012:targetgroup/my-
   targets/0123456789012345
   ```

3. Utilizzare il comando register-target (registrare target) per registrare le istanze con il gruppo target.

   ```
   aws elbv2 register-targets --target-group-arn targetgroup-arn --targets
    Id=i-1234567890abcdef0 Id=i-0abcdef1234567890
   ```

4. Utilizza il comando create-listener per creare un ascoltatore per il sistema di bilanciamento del carico con una regola predefinita che inoltra le richieste verso il gruppo target:

```
aws elbv2 create-listener --load-balancer-arn loadbalancer-arn --default-actions
 Type=forward,TargetGroupArn=targetgroup-arn
```

L'output contiene l'ARN dell'ascoltatore, con il formato seguente.

```
arn:aws:elasticloadbalancing:us-east-2:123456789012:listener/gwy/my-load-
balancer/1234567890123456/abc1234567890123
```

5. (Facoltativo) È possibile verificare l'integrità delle destinazioni registrate per il gruppo target utilizzando il seguente comando describe-target-health.

```
aws elbv2 describe-target-health --target-group-arn targetgroup-arn
```

## Fase 2: creare un endpoint del Gateway Load Balancer

Utilizza la procedura seguente per creare un endpoint del Gateway Load Balancer. Gli endpoint del Gateway Load Balancer sono zonali. È consigliabile creare un endpoint del Gateway Load Balancer per zona. Per ulteriori informazioni, consulta Accesso alle appliance virtuali tramite AWS PrivateLink nella Guida di .

Per creare un endpoint Gateway Load Balancer

1. Usa il comando create-vpc-endpoint-service-configuration per creare una configurazione del servizio di endpoint utilizzando il Gateway Load Balancer.

```
aws ec2 create-vpc-endpoint-service-configuration --gateway-load-balancer-
arns loadbalancer-arn --no-acceptance-required
```

Per il supporto di entrambi gli indirizzi IPv4 e IPv6, aggiungi l'opzione `--supported-ip-address-types ipv4 ipv6`.

L'output contiene l'ID del servizio (ad esempio, vpce-svc-12345678901234567) e il nome del servizio (ad esempio, com.amazonaws.vpce.us-east-2.vpce-svc-12345678901234567).

2. Usa il comando modify-vpc-endpoint-service-permissions per consentire ai consumatori del servizio di creare un endpoint sul tuo servizio. Un consumatore del servizio può essere un utente, un ruolo IAM o Account AWS. L'esempio seguente aggiunge l'autorizzazione per il file specificato.Account AWS

```
aws ec2 modify-vpc-endpoint-service-permissions --service-id vpce-
svc-12345678901234567 --add-allowed-principals arn:aws:iam::123456789012:root
```

3. Usa il comando create-vpc-endpoint per creare l'endpoint del Gateway Load Balancer per il tuo servizio.

```
aws ec2 create-vpc-endpoint --vpc-endpoint-type GatewayLoadBalancer --service-
name com.amazonaws.vpce.us-east-2.vpce-svc-12345678901234567 --vpc-id consumer-vpc-
id --subnet-ids consumer-subnet-id
```

   Per il supporto di entrambi gli indirizzi IPv4 e IPv6, aggiungi l'opzione `--ip-address-type dualstack`.

   L'output contiene l'ID dell'endpoint del Gateway Load Balancer (ad esempio, vpce-01234567890abcdef).

## Fase 3: configurazione dell'instradamento

Configura le tabelle di routing per il VPC consumatore del servizio come segue. Ciò consente alle appliance di sicurezza di eseguire ispezioni sul traffico in entrata destinato ai server dell'applicazione.

Configurazione del routing

1. Usa il comando create-route per aggiungere voci alla tabella di routing per il gateway Internet che instrada il traffico destinato ai server dell'applicazione all'endpoint del Gateway Load Balancer.

```
aws ec2 create-route --route-table-id gateway-rtb --destination-cidr-block Subnet 1
 IPv4 CIDR --vpc-endpoint-id vpce-01234567890abcdef
```

   Se supporti IPv6, aggiungi la route seguente.

```
aws ec2 create-route --route-table-id gateway-rtb --destination-cidr-block Subnet 1
 IPv6 CIDR --vpc-endpoint-id vpce-01234567890abcdef
```

2. Usa il comando create-route per aggiungere una voce alla tabella di routing per la sottorete con i server dell'applicazione che instradano tutto il traffico dai server dell'applicazione all'endpoint del Gateway Load Balancer.

```
aws ec2 create-route --route-table-id application-rtb --destination-cidr-block
 0.0.0.0/0 --vpc-endpoint-id vpce-01234567890abcdef
```

Se supporti IPv6, aggiungi la route seguente.

```
aws ec2 create-route --route-table-id application-rtb --destination-cidr-block ::/0
 --vpc-endpoint-id vpce-01234567890abcdef
```

3. Usa il comando create-route per aggiungere una voce alla tabella di routing per la sottorete con l'endpoint del Gateway Load Balancer che instrada tutto il traffico proveniente dai server dell'applicazione al gateway Internet.

```
aws ec2 create-route --route-table-id endpoint-rtb --destination-cidr-block
 0.0.0.0/0 --gateway-id igw-01234567890abcdef
```

Se supporti IPv6, aggiungi la route seguente.

```
aws ec2 create-route --route-table-id endpoint-rtb --destination-cidr-block ::/0 --
gateway-id igw-01234567890abcdef
```

4. Ripeti l'operazione per ogni tabella di routing della sottorete dell'applicazione in ogni zona.

# Gateway Load Balancers

Utilizza un Gateway Load Balancer per implementare e gestire una parco istanze di appliance virtuali che supportano il protocollo GENEVE.

Un Gateway Load Balancer funziona al terzo livello del modello Open Systems Interconnection (OSI). È in ascolto per tutti i pacchetti IP attraverso tutte le porte e inoltra il traffico al gruppo di destinazione specificato nella regola dell'ascoltatore, usando il protocollo GENEVE sulla porta 6081.

È possibile aggiungere o rimuovere le destinazioni dal sistema di bilanciamento del carico in base alle proprie esigenze, senza interrompere il flusso di richieste. Elastic Load Balancing ridimensiona il load balancer di volta in volta, in quanto il traffico verso l'applicazione varia nel corso del tempo. Elastic Load Balancing è in grado di ridimensionare automaticamente le risorse per la maggior parte dei carichi di lavoro.

Indice


- Stato del sistema di bilanciamento del carico
- Tipo di indirizzo IP
- Zone di disponibilità
- Tempo di inattività
- Attributi del sistema di bilanciamento del carico
- Liste di controllo accessi (ACL) di rete
- Flussi asimmetrici
- Unità di trasmissione massima (MTU)
- Crea un Gateway Load Balancer
- Aggiorna i tipi di indirizzi IP per il tuo Gateway Load Balancer
- Modifica gli attributi per il tuo Gateway Load Balancer
- Etichetta un Gateway Load Balancer
- Elimina un Gateway Load Balancer
- Prenotazioni di capacità per il tuo Gateway Load Balancer


## Stato del sistema di bilanciamento del carico

Un Gateway Load Balancer può avere uno dei seguenti stati:

`provisioning`

Il Gateway Load Balancer è in fase di configurazione.

`active`

Il Gateway Load Balancer è completamente configurato e pronto a instradare il traffico.

`failed`

Il Gateway Load Balancer non può essere configurato.

## Tipo di indirizzo IP

È possibile impostare i tipi di indirizzi IP che i server dell'applicazione possono utilizzare per accedere ai Gateway Load Balancer.

I Gateway Load Balancer supportano i seguenti tipi di indirizzi IP:

**`ipv4`**

È supportato solo IPv4.

**`dualstack`**

Sono supportati sia IPv4 che IPv6.

Considerazioni

- Il cloud privato virtuale (VPC, Virtual Private Cloud) e le sottoreti specificate per il sistema di bilanciamento del carico devono avere blocchi CIDR IPv6 associati.
- Le tabelle di routing per le sottoreti nel VPC del consumatore del servizio devono instradare il traffico IPv6 e le ACL di rete per queste sottoreti devono consentire il traffico IPv6.
- Un Gateway Load Balancer incapsula il traffico client IPv4 e IPv6 con un'intestazione IPv4 GENEVE e lo invia all'appliance. L'appliance incapsula il traffico client IPv4 e IPv6 con un'intestazione IPv4 GENEVE e lo invia al Gateway Load Balancer.

Per ulteriori informazioni sui tipi di indirizzi IP, vedere. Aggiorna i tipi di indirizzi IP per il tuo Gateway Load Balancer

# Zone di disponibilità

Quando si crea un Gateway Load Balancer, si abilitano una o più zone di disponibilità e si specifica la sottorete che corrisponde a ciascuna zona. Quando si abilitano più zone di disponibilità, si assicura che il sistema di bilanciamento del carico possa continuare a instradare il traffico anche se una zona di disponibilità non è più disponibile. Le sottoreti specificate devono avere a disposizione almeno 8 indirizzi IP ciascuna. Le sottoreti non possono essere rimosse dopo la creazione del load balancer. Per rimuovere una sottorete, è necessario creare un nuovo sistema di bilanciamento del carico.

# Tempo di inattività

Per ogni richiesta TCP effettuata tramite un Gateway Load Balancer, viene tracciato lo stato di quella connessione. Se non vengono inviati dati tramite la connessione dal client o dal target per un periodo superiore al tempo di inattività, la connessione viene chiusa. Trascorso il periodo di timeout di inattività, il load balancer considera il successivo TCP SYN come un nuovo flusso e lo indirizza verso una nuova destinazione. Tuttavia, i pacchetti di dati inviati dopo la scadenza del periodo di timeout di inattività vengono eliminati.

Il valore di timeout di inattività predefinito per i flussi TCP è 350 secondi, ma può essere aggiornato a qualsiasi valore compreso tra 60-6000 secondi. I client o i target possono utilizzare i pacchetti keepalive TCP per ripristinare il tempo di inattività.

⚠ Mancata corrispondenza del timeout di tracciamento della connessione ENI

Se configuri un valore di timeout di inattività TCP superiore a 350 secondi, assicurati che il timeout di inattività per il tracciamento della connessione ENI delle istanze di destinazione (`TcpEstablishedTimeout`) sia impostato su un valore uguale o superiore al timeout di inattività del Gateway Load Balancer. Nelle istanze Nitro di sesta generazione (Nitro V6+), il timeout di tracciamento della connessione ENI predefinito è di 350 secondi. Se il timeout di inattività del Gateway Load Balancer supera il timeout di tracciamento della connessione del target, l'interfaccia di rete del target potrebbe eliminare la voce di tracciamento della connessione mentre il load balancer considera ancora la connessione attiva, con conseguente perdita di pacchetti sulla connessione. Puoi configurare il timeout di tracciamento della connessione della destinazione utilizzando. `ModifyNetworkInterfaceAttribute` Per ulteriori informazioni, consulta Best practice per la gestione delle connessioni TCP su EC2.

⚠ Limitazione della viscosità

Il timeout di inattività del Gateway Load Balancer può essere aggiornato solo quando si utilizza la viscosità a 5 tuple. Quando si utilizza lo stickness a 3 o 2 tuple, viene utilizzato il valore di timeout di inattività predefinito. Per ulteriori informazioni, consulta Persistenza del flusso

Quando UDP è senza connessione, il sistema di bilanciamento del carico mantiene lo stato del flusso UDP basandosi sugli indirizzi IP di origine e di destinazione e sulle porte. Ciò garantisce che i pacchetti appartenenti allo stesso flusso siano regolarmente inviati alla stessa destinazione. Una volta trascorso il periodo di timeout di inattività, il sistema di bilanciamento del carico considera il pacchetto UDP in entrata come un nuovo flusso e lo instrada a una nuova destinazione. Elastic Load Balancing imposta il valore di timeout di inattività per i flussi UDP su 120 secondi. Non possono essere modificate.

Le istanze EC2 devono rispondere a una nuova richiesta entro 30 secondi per stabilire un percorso di ritorno.

Per ulteriori informazioni, consulta Aggiorna il timeout di inattività.

## Attributi del sistema di bilanciamento del carico

I seguenti sono gli attributi del sistema di bilanciamento del carico per il Gateway Load Balancer

`deletion_protection.enabled`

Indica se è abilitata la protezione da eliminazione. Il valore predefinito è `false`.

`load_balancing.cross_zone.enabled`

Indica se è abilitato il bilanciamento del carico tra zone. Il valore predefinito è `false`.

Per ulteriori informazioni, consulta Modifica gli attributi del load balancer.

## Liste di controllo accessi (ACL) di rete

Se i server delle applicazioni e l'endpoint Gateway Load Balancer si trovano nella stessa sottorete, le regole NACL vengono valutate per il traffico dai server delle applicazioni all'endpoint Gateway Load Balancer.

# Flussi asimmetrici

I Gateway Load Balancer supportano flussi asimmetrici quando il sistema di bilanciamento del carico elabora il pacchetto di flusso iniziale e il pacchetto di flusso di risposta non viene instradato attraverso il sistema di bilanciamento del carico. Il routing asimmetrico non è consigliato, in quanto può comportare una riduzione delle prestazioni di rete. I Gateway Load Balancer non supportano flussi asimmetrici quando il sistema di bilanciamento del carico non elabora il pacchetto di flusso iniziale e il pacchetto di flusso di risposta viene instradato attraverso il sistema di bilanciamento del carico.

# Unità di trasmissione massima (MTU)

L'unità massima di trasmissione (MTU) è la dimensione del pacchetto di dati più grande che può essere trasmesso attraverso la rete. La MTU dell'interfaccia del Gateway Load Balancer supporta pacchetti fino a 8.500 byte. I pacchetti con una dimensione maggiore di 8.500 byte che arrivano al Gateway Load Balancer vengono eliminati.

Un Gateway Load Balancer incapsula il traffico client IP con un'intestazione GENEVE e lo invia all'appliance. Il processo di incapsulamento GENEVE aggiunge 68 byte al pacchetto originale. Pertanto, per supportare pacchetti fino a 8.500 byte, assicuratevi che l'impostazione MTU del dispositivo supporti pacchetti di almeno 8.568 byte.

I Gateway Load Balancer non supportano la frammentazione IP. Inoltre, i Gateway Load Balancer non generano il messaggio ICMP "Destinazione irraggiungibile: frammentazione necessaria e set DF". A causa di ciò, il percorso MTU Discovery (PMTUD) non è supportato.

# Crea un Gateway Load Balancer

Un Gateway Load Balancer accetta richieste dai client e le distribuisce nei target di un gruppo target, ad esempio nelle istanze EC2.

Per creare un Gateway Load Balancer utilizzando Console di gestione AWS, completare le seguenti attività. In alternativa, per creare un Gateway Load Balancer utilizzando il AWS CLI, vedere. Nozioni di base per l'utilizzo della CLI

Processi

- Prerequisiti

- Creazione del sistema di bilanciamento del carico
- Fasi successive importanti


## Prerequisiti

Prima di iniziare, accertati che il cloud privato virtuale (VPC) per il tuo Gateway Load Balancer disponga di almeno una sottorete in ciascuna zona di disponibilità in cui si dispone di destinazioni.

## Creazione del sistema di bilanciamento del carico

Utilizza la procedura seguente per creare il Gateway Load Balancer. Fornisci alcune informazioni di configurazione di base per il sistema di bilanciamento del carico, ad esempio un nome e un tipo di indirizzo IP. Successivamente, fornisci alcune informazioni relative alla rete e all'ascoltatore che indirizza il traffico verso le istanze. I Gateway Load Balancer richiedono gruppi di destinazione che utilizzano il protocollo GENEVE.

Per creare il load balancer e il listener utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, in Bilanciamento del carico, scegli Sistemi di bilanciamento del carico.
3. Selezionare Create Load Balancer (Crea sistema di bilanciamento del carico).
4. Sotto a Gateway Load Balancer scegli Crea.
5. Configurazione di base
    a. In Nome del sistema di bilanciamento del carico immetti un nome univoco per il sistema di bilanciamento del carico. Ad esempio, **`my-glb`**. Il nome del Gateway Load Balancer deve essere univoco all'interno del set di sistemi di bilanciamento del carico per la regione. I nomi possono avere un massimo di 32 caratteri e contenere solo caratteri alfanumerici e trattini. Inoltre non devono iniziare o finire con un carattere di sottolineatura.
    b. Per tipo di indirizzo IP, seleziona IPv4 per supportare solo gli indirizzi IPv4, oppure Dualstack per supportare entrambi gli indirizzi IPv4 e IPv6.
6. Mappatura della rete
    a. Per VPC, seleziona il VPC del fornitore del servizio.
    b. Per le Mappature, seleziona tutte le zone di disponibilità in cui sono state avviate le istanze delle appliance di sicurezza e le sottoreti pubbliche corrispondenti.

7. Routing dell'ascoltatore IP

    a. Per Azione predefinita, seleziona il gruppo di destinazione per ricevere il traffico. Se non hai un gruppo di destinazione, scegli Crea un gruppo di destinazione. Per ulteriori informazioni, consulta Creazione di un gruppo target.

    b. (Facoltativo) Espandi Tag dell'ascoltatore e aggiungi i tag di cui hai bisogno.

8. (Facoltativo) Espandi Tag del sistema di bilanciamento del carico e aggiungi i tag di cui hai bisogno.

9. Controlla la configurazione e scegli Crea sistema di bilanciamento del carico.

## Fasi successive importanti

Dopo aver creato il sistema di bilanciamento del carico, verifica che le istanze EC2 abbiano superato il controllo dell'integrità iniziale. Per testare il sistema di bilanciamento del carico, devi creare un endpoint del Gateway Load Balancer e aggiornare la tabella di routing per fare in modo che l'endpoint del Gateway Load Balancer sia la fase successiva. Queste configurazioni sono impostate all'interno della console Amazon VPC. Per maggiori informazioni, vedi il tutorial Nozioni di base.

# Aggiorna i tipi di indirizzi IP per il tuo Gateway Load Balancer

È possibile configurare il Gateway Load Balancer in modo che i server delle applicazioni possano accedere al sistema di bilanciamento del carico utilizzando solo indirizzi IPv4 o utilizzando entrambi gli indirizzi IPv4 e IPv6 (dualstack). Il sistema di bilanciamento del carico comunica con le destinazioni in base al tipo di indirizzo IP del gruppo di destinazione. Per ulteriori informazioni, consulta Tipo di indirizzo IP.

Per aggiornare il tipo di indirizzo IP utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel pannello di navigazione, sotto Load Balancing (Bilanciamento del carico), scegli Load Balancers (Load balancer).
3. Selezionare il load balancer.
4. Scegli Actions (Azioni), Edit IP address type (Modifica tipo di indirizzo IP).
5. In IP address type (Tipo di indirizzo IP), seleziona ipv4 per supportare i soli indirizzi IPv4 oppure dualstack per supportare sia gli indirizzi IPv4 che IPv6.
6. Scegli Save (Salva).

Per aggiornare il tipo di indirizzo IP utilizzando AWS CLI

Utilizza il comando set-ip-address-type.

# Modifica gli attributi per il tuo Gateway Load Balancer

Dopo aver creato un Gateway Load Balancer, puoi modificarne gli attributi del load balancer.

Attributi del sistema di bilanciamento del carico

- Protezione da eliminazione
- Cross-zone bilanciamento del carico

## Protezione da eliminazione

Per evitare che il Gateway Load Balancer venga eliminato accidentalmente, è possibile abilitare la protezione da eliminazione. Per impostazione predefinita, la protezione da eliminazione è disabilitata.

Se abiliti la protezione da eliminazione per il Gateway Load Balancer, devi disabilitarla prima di poterlo eliminare.

Per abilitare la protezione da eliminazione tramite la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, in Bilanciamento del carico, scegli Sistemi di bilanciamento del carico.
3. Seleziona il Gateway Load Balancer.
4. Seleziona Operazioni, Modifica attributi.
5. Nella pagina Edit load balancer attributes (Modifica attributi del sistema di bilanciamento del carico) seleziona Enable (Abilita) in Delete Protection (Elimina protezione) e poi scegli Save (Salva).

Per disabilitare la protezione da eliminazione tramite la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, in Bilanciamento del carico, scegli Sistemi di bilanciamento del carico.
3. Seleziona il Gateway Load Balancer.

4. Seleziona Operazioni, Modifica attributi.
5. Nella pagina Edit load balancer attributes (Modifica attributi del sistema di bilanciamento del carico) deseleziona Enable (Abilita) in Delete Protection (Elimina protezione) e poi scegli Save (Salva).

Per abilitare o disabilitare la protezione da eliminazione utilizzando il AWS CLI

Utilizza il comando modify-load-balancer-attributes con l'attributo `deletion_protection.enabled`.

## Cross-zone bilanciamento del carico

Per impostazione predefinita, ogni nodo del sistema di bilanciamento del carico distribuisce il traffico solo tra i target registrati nella zona di disponibilità del sistema. Se attivi il bilanciamento del carico su più zone, ogni Gateway Load Balancer distribuisce le richieste nelle destinazioni registrate in tutte le zone di disponibilità attivate. Per ulteriori informazioni, consulta il bilanciamento Cross-zone del carico nella Elastic Load Balancing User Guide.

Per abilitare il bilanciamento del carico su più zone utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, in Bilanciamento del carico, scegli Sistemi di bilanciamento del carico.
3. Seleziona il Gateway Load Balancer.
4. Seleziona Operazioni, Modifica attributi.
5. Nella pagina Modifica gli attributi del load balancer, seleziona Abilita per il bilanciamento Cross-Zone del carico, quindi scegli Salva.

Per abilitare il bilanciamento del carico tra zone utilizzando AWS CLI

Utilizza il comando modify-load-balancer-attributes con l'attributo `load_balancing.cross_zone.enabled`.

# Etichetta un Gateway Load Balancer

I tag ti aiutano a classificare i bilanciatori del carico in modi diversi, ad esempio in base a scopo, proprietario o ambiente.

È possibile aggiungere più tag a ciascun sistema di bilanciamento del carico. Le chiavi dei tag devono essere univoche per ogni Gateway Load Balancer. Se aggiungi un tag con una chiave già associata al load balancer, il valore del tag viene aggiornato.

Quando il tag non è più necessario, è possibile eliminarlo dal Gateway Load Balancer.

Restrizioni

- Numero massimo di tag per risorsa: 50
- Lunghezza massima della chiave: 127 caratteri Unicode
- Lunghezza massima del valore: 255 caratteri Unicode
- Per le chiavi e i valori dei tag viene fatta la distinzione tra maiuscole e minuscole. I caratteri consentiti sono lettere, spazi e numeri rappresentabili in UTF-8, più i seguenti caratteri speciali: + - =. _ : / @. Non utilizzare spazi iniziali o finali.
- Non utilizzate il `aws:` prefisso nei nomi o nei valori dei tag perché èAWS riservato all'uso. Non è possibile modificare né eliminare i nomi o i valori di tag con tale prefisso. I tag con questo prefisso non vengono conteggiati per il limite del numero di tag per risorsa.

Per aggiornare i tag di un Gateway Load Balancer tramite la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, in Bilanciamento del carico, scegli Sistemi di bilanciamento del carico.
3. Seleziona il Gateway Load Balancer.
4. Scegliete Add/Edit Tag, Tag, quindi effettuate una o più delle seguenti operazioni:
   a. Per aggiornare un tag, modificare i valori di Chiave e Valore.
   b. Per aggiungere un nuovo tag, scegliere Create Tag (Crea tag). Digita un valore per Chiave e Valore.
   c. Per eliminare un tag, scegliere l'icona di eliminazione (X) accanto al tag.
5. Una volta completato l'aggiornamento dei tag, scegli Salva.

Per aggiornare i tag per un Gateway Load Balancer utilizzando il AWS CLI

Utilizza i comandi add-tags e remove-tags.

# Elimina un Gateway Load Balancer

Non appena il Gateway Load Balancer diventa disponibile, ti verrà addebitata ogni ora o frazione di ora in cui lo mantieni in esecuzione. Se il Gateway Load Balancer non ti è più utile, puoi eliminarlo. Non appena il Gateway Load Balancer viene eliminato, i relativi addebiti vengono bloccati.

Non è possibile eliminare un Gateway Load Balancer, se è in uso da un altro servizio. Ad esempio, se il Gateway Load Balancer è associato a un servizio endpoint VPC, è necessario eliminare la configurazione del servizio endpoint prima di poter eliminare il Gateway Load Balancer associato.

L'eliminazione di un Gateway Load Balancer elimina anche i relativi ascoltatori. L'eliminazione di un Gateway Load Balancer non influisce sulle relative destinazioni registrate. Ad esempio, le istanze EC2 proseguono l'esecuzione e sono comunque registrate nei loro gruppi target. Per eliminare i gruppi target, consulta Elimina un gruppo target per il tuo Gateway Load Balancer.

Per eliminare un Gateway Load Balancer utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, in Bilanciamento del carico, scegli Sistemi di bilanciamento del carico.
3. Seleziona il Gateway Load Balancer.
4. Scegli Operazioni > Elimina.
5. Quando viene richiesta la conferma, seleziona Sì, elimina.

Per eliminare un Gateway Load Balancer utilizzando il AWS CLI

Utilizza il comando delete-load-balancer.

# Prenotazioni di capacità per il tuo Gateway Load Balancer

Le prenotazioni di Load Balancer Capacity Unit (LCU) consentono di riservare una capacità minima statica per il sistema di bilanciamento del carico. I Gateway Load Balancer si ridimensionano automaticamente per supportare i carichi di lavoro rilevati e soddisfare le esigenze di capacità. Quando viene configurata la capacità minima, il sistema di bilanciamento del carico continua a scalare verso l'alto o verso il basso in base al traffico ricevuto, ma impedisce anche che la capacità scenda al di sotto della capacità minima configurata.

Prendi in considerazione l'utilizzo della prenotazione LCU nelle seguenti situazioni:

- Hai un evento imminente che avrà un traffico improvviso e insolito e vuoi assicurarti che il sistema di bilanciamento del carico sia in grado di supportare l'improvviso picco di traffico durante l'evento.
- La natura del carico di lavoro comporta picchi di traffico imprevedibili per un breve periodo.
- Stai configurando il tuo sistema di bilanciamento del carico per integrare o migrare i tuoi servizi a un orario di avvio specifico e devi iniziare con una capacità elevata invece di aspettare che l'auto-scaling abbia effetto.
- Stai migrando i carichi di lavoro tra i sistemi di bilanciamento del carico e desideri configurare la destinazione in modo che corrisponda alla scala dell'origine.

### Stima la capacità di cui hai bisogno

Per determinare la quantità di capacità da riservare al sistema di bilanciamento del carico, consigliamo di eseguire test di carico o di esaminare i dati storici sul carico di lavoro che rappresentano il traffico imminente previsto. Utilizzando la console Elastic Load Balancing, puoi stimare la capacità da riservare in base al traffico esaminato.

In alternativa, puoi fare riferimento alla CloudWatch metrica ProcessedBytesper determinare il giusto livello di capacità. La capacità del sistema di bilanciamento del carico è riservata alle unità LCU, con ciascuna LCU pari a 2,2 Mbps. È possibile utilizzare la PeakBytesPerSecondmetrica per visualizzare il traffico di throughput massimo al minuto sul sistema di bilanciamento del carico, quindi convertire tale throughput in unità LCU utilizzando un tasso di conversione di 2,2 Mbps pari a 1 LCU.

Se non disponi di dati storici sul carico di lavoro a cui fare riferimento e non puoi eseguire test di carico, puoi stimare la capacità necessaria utilizzando il calcolatore di prenotazione LCU. Il calcolatore delle prenotazioni LCU utilizza dati basati sui carichi di lavoro storici AWS osservati e potrebbe non rappresentare il carico di lavoro specifico dell'utente. Per ulteriori informazioni, consulta Load Balancer Capacity Unit Reservation Calculator.

### Regioni supportate

Questa funzionalità è disponibile solo nelle seguenti regioni:

- Stati Uniti orientali (Virginia settentrionale)
- Stati Uniti orientali (Ohio)
- Stati Uniti occidentali (Oregon)
- Asia Pacifico (Hong Kong)

- Asia Pacifico (Singapore)
- Asia Pacifico (Sydney)
- Asia Pacifico (Tokyo)
- Europa (Francoforte)
- Europa (Irlanda)
- Europa (Stoccolma)

Valori minimi e massimi per una prenotazione LCU

La richiesta di prenotazione totale deve essere di almeno 2.750 LCU per zona di disponibilità. Il valore massimo è determinato dalle quote del tuo account. Per ulteriori informazioni, consulta the section called “Unità di capacità Load Balancer”.

## Richiedi la prenotazione della Load Balancer Capacity Unit per il tuo Gateway Load Balancer

Prima di utilizzare la prenotazione LCU, verifica quanto segue:

- La prenotazione LCU supporta solo la prenotazione della capacità di throughput per i Gateway Load Balancer. Quando richiedi una prenotazione LCU, converti le tue esigenze di capacità da Mbps a LCU utilizzando il tasso di conversione di 1 LCU a 2,2 Mbps.
- La capacità è riservata a livello regionale ed è distribuita uniformemente tra le zone di disponibilità. Verifica di avere un numero sufficiente di obiettivi distribuiti in modo uniforme in ciascuna zona di disponibilità prima di attivare la prenotazione LCU.
- Le richieste di prenotazione LCU vengono soddisfatte in base al principio «primo arrivato, primo servito» e dipendono dalla capacità disponibile per una zona in quel momento. La maggior parte delle richieste viene in genere soddisfatta entro un'ora, ma può richiedere fino a qualche ora.
- Per aggiornare una prenotazione esistente, è necessario che la richiesta precedente sia stata effettuata o non sia andata a buon fine. Puoi aumentare la capacità riservata tutte le volte che vuoi, tuttavia puoi diminuirla solo due volte al giorno.

Richiedi una prenotazione LCU

I passaggi di questa procedura spiegano come richiedere una prenotazione LCU sul sistema di bilanciamento del carico.

Per richiedere una prenotazione LCU utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Seleziona Sistemi di bilanciamento del carico nel riquadro di navigazione.
3. Selezionare il nome del sistema di bilanciamento del carico.
4. Nella scheda Capacità, scegli Modifica prenotazione LCU.
5. Seleziona Stima storica basata su riferimenti, quindi seleziona il sistema di bilanciamento del carico dall'elenco a discesa.
6. Seleziona il periodo di riferimento per visualizzare il livello LCU riservato consigliato.
7. Se non disponi di un carico di lavoro di riferimento storico, puoi scegliere Stima manuale e inserire il numero di LCU da prenotare.
8. Scegli Save (Salva).

Per richiedere una prenotazione LCU utilizzando AWS CLI

Utilizza il comando modify-capacity-reservation.

## Aggiorna o termina le prenotazioni Load Balancer Capacity Unit per il tuo Gateway Load Balancer

Aggiornare o terminare una prenotazione LCU

I passaggi di questa procedura spiegano come aggiornare o terminare una prenotazione LCU sul sistema di bilanciamento del carico.

Per aggiornare o terminare una prenotazione LCU utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Seleziona Sistemi di bilanciamento del carico nel riquadro di navigazione.
3. Selezionare il nome del sistema di bilanciamento del carico.
4. Nella scheda Capacità, conferma che lo stato della prenotazione è Fornito.
    a. Per aggiornare la prenotazione LCU, scegli Modifica prenotazione LCU.
    b. Per terminare la prenotazione LCU, scegli Annulla capacità.

Per aggiornare o terminare una prenotazione LCU utilizzando il AWS CLI

Utilizza il comando modify-capacity-reservation.

## Monitora la prenotazione della Load Balancer Capacity Unit per il tuo Gateway Load Balancer

Stato della prenotazione

La prenotazione LCU ha quattro stati disponibili:

- in sospeso - Indica che la prenotazione è in fase di approvvigionamento.
- fornito - Indica che la capacità riservata è pronta e disponibile per l'uso.
- fallito - Indica che la richiesta non può essere completata in quel momento.
- ribilanciamento - Indica che è stata aggiunta una zona di disponibilità e il bilanciamento del carico sta riequilibrando la capacità.

LCU riservata

Per determinare l'utilizzo della LCU riservata, è possibile confrontare la PeakBytesPerSecondmetrica al minuto con la somma oraria (ReservedLCUS). Per convertire i byte al minuto in LCU all'ora, usa (byte per min) */(10^6) /2.2. 8/60

Monitora la capacità riservata

I passaggi di questo processo spiegano come verificare lo stato di una prenotazione LCU sul sistema di bilanciamento del carico.

Per visualizzare lo stato di una prenotazione LCU utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Seleziona Sistemi di bilanciamento del carico nel riquadro di navigazione.
3. Selezionare il nome del sistema di bilanciamento del carico.
4. Nella scheda Capacità, è possibile visualizzare lo stato della prenotazione e il valore della LCU riservata.

Per monitorare lo stato della prenotazione LCU utilizzando AWS CLI

Utilizzate il comando describe-capacity-reservation.

# Ascoltatori dei Gateway Load Balancer

Quando crei il tuo Gateway Load Balancer, aggiungi un ascoltatore. Si definisce listener il processo che verifica la presenza di richieste di connessione.

Gli ascoltatori per Gateway Load Balancer ascoltano tutti i pacchetti IP su tutte le porte. Non è possibile specificare un protocollo o una porta quando si crea un ascoltatore per un Gateway Load Balancer.

Quando si crea un listener, è necessario specificare una regola per instradare le richieste. Questa regola inoltra le richieste verso il gruppo target indicato. È possibile aggiornare la regola dell'ascoltatore per inoltrare le richieste a un gruppo target diverso.

## attributi del listener

Di seguito sono riportati gli attributi del listener per Gateway Load Balancer:

`tcp.idle_timeout.seconds`

Il valore di timeout di inattività tcp, in secondi. L'intervallo valido è 60-6000 secondi. L'impostazione predefinita è 350 secondi.

`send_tcp_reset.on_idle_timeout.enabled`

Indica se il Gateway Load Balancer invia un ripristino TCP (RST) al mittente del traffico. Se abilitato, Gateway Load Balancer invia l'RST quando scade il timeout di inattività di un flusso TCP o quando riceve un pacchetto TCP non SYN per un flusso che non è nella tabella di flusso. I valori possibili sono `true` e `false`. Il valore di default è `false`. Dopo aver inviato l'RST, il Gateway Load Balancer rimuove la voce di flusso dalla sua tabella di flusso.

Per ulteriori informazioni, consulta Aggiorna il timeout di inattività.

## Aggiorna il gruppo target per il tuo listener Gateway Load Balancer

Quando si crea un listener, è necessario specificare una regola per instradare le richieste. Questa regola inoltra le richieste verso il gruppo target indicato. È possibile aggiornare la regola dell'ascoltatore per inoltrare le richieste a un gruppo target diverso.

Per aggiornare il listener utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, in Bilanciamento del carico, scegli Sistemi di bilanciamento del carico.
3. Selezionare il sistema del bilanciamento del carico e scegliere Listeners (Listener).
4. Scegli Modifica ascoltatore.
5. Per l'Inoltro al gruppo di destinazione, scegli un gruppo di destinazione.
6. Selezionare Salva.

Per aggiornare il tuo ascoltatore usando il AWS CLI

Utilizza il comando modify-listener.

# Aggiorna il timeout di inattività del TCP per il listener Gateway Load Balancer

Per ogni richiesta TCP effettuata tramite un Gateway Load Balancer, viene monitorato lo stato di tale connessione. Se non vengono inviati dati tramite la connessione dal client o dal target per un periodo superiore al tempo di inattività, la connessione viene chiusa. Il valore di timeout di inattività predefinito per i flussi TCP è 350 secondi, ma può essere aggiornato a qualsiasi valore compreso tra 60 e 6000 secondi.

- Se imposti un timeout di inattività superiore a 350 secondi, verifica che il timeout di inattività per il tracciamento della connessione ENI (`TcpEstablishedTimeout`) dei tuoi target sia uguale o superiore al valore di timeout di inattività del Gateway Load Balancer. Sulle istanze Nitro V6+, il timeout predefinito per il tracciamento della connessione è di 350 secondi. Una mancata corrispondenza fa sì che l'interfaccia di rete del bersaglio interrompa silenziosamente lo stato della connessione prima che il load balancer chiuda la connessione. Per ulteriori informazioni, consulta Best Practices for TCP Connection Management su EC2.

Per aggiornare il timeout di inattività del TCP utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, in Bilanciamento del carico, scegli Sistemi di bilanciamento del carico.

3. Seleziona il Gateway Load Balancer.
4. Nella scheda listener scegli Azioni, Visualizza i dettagli dei listener.
5. Nella pagina dei dettagli del listener, nella scheda Attributi, seleziona Modifica.
6. Nella pagina Modifica gli attributi del listener, nella sezione Attributi del listener, inserisci un valore per il timeout di inattività TCP.
7. Scegliere Salva modifiche.

Per aggiornare il timeout di inattività del TCP utilizzando AWS CLI

Utilizzate il comando https://docs.aws.amazon.com/cli/latest/reference/elbv2/modify-listener-attributes.html modify-listener-attributes con l'attributo. `tcp.idle_timeout.seconds`

## Invia il ripristino TCP in caso di timeout di inattività

Se abilitato, il Gateway Load Balancer invia un ripristino del TCP (RST) al mittente del traffico allo scadere del timeout di inattività. Ciò consente ai mittenti di segnalare il traffico di interrompere immediatamente le connessioni inattive. Per controllare questo comportamento, usa l'attributo. `send_tcp_reset.on_idle_timeout.enabled`

Per abilitare Send TCP reset in caso di timeout di inattività utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, in Bilanciamento del carico, scegli Sistemi di bilanciamento del carico.
3. Seleziona il tuo Gateway Load Balancer.
4. Nella scheda Listener, scegli Azioni, Visualizza i dettagli dei listener.
5. Nella pagina dei dettagli del listener, nella scheda Attributi, scegli Modifica.
6. Scegli Invia ripristino TCP in caso di timeout di inattività.
7. Scegli Save changes (Salva modifiche).

Per abilitare Invia ripristino TCP in caso di timeout di inattività utilizzando AWS CLI

Utilizzate il comando https://docs.aws.amazon.com/cli/latest/reference/elbv2/modify-listener-attributes.html modify-listener-attributes con l'attributo impostato su. `send_tcp_reset.on_idle_timeout.enabled` `true`

```
aws elbv2 modify-listener-attributes --listener-arn listener-arn --attributes
 Key=send_tcp_reset.on_idle_timeout.enabled,Value=true
```

# Quota per i Gateway Load Balancer

Ogni gruppo target viene utilizzato per instradare le richieste a uno o più target registrati. Quando si crea un listener, si specifica un gruppo di destinazione per l'operazione predefinita. Il traffico viene inoltrato al gruppo di destinazione specificato nella regola dell'ascoltatore. È possibile creare diversi gruppi target per diversi tipi di richieste.

È possibile definire le impostazioni di controllo dell'integrità per il Gateway Load Balancer per ciascun gruppo target. Ogni gruppo target utilizza le impostazioni di controllo dello stato predefinite, a meno che non vengano sostituite al momento della creazione del gruppo target o modificate in un secondo momento. Dopo aver specificato un gruppo target in una regola per un ascoltatore, il Gateway Load Balancer monitora continuamente lo stato di tutti i target registrati con il gruppo di destinazione che si trovano in una zona di disponibilità abilitata per il Gateway Load Balancer. Il Gateway Load Balancer instrada le richieste alle destinazioni registrate con stato integro. Per ulteriori informazioni, consulta Controlli sanitari per i gruppi target di Gateway Load Balancer.

Indice


- Configurazione dell'instradamento
- Target type (Tipo di destinazione)
- Destinazioni registrate
- Attributi dei gruppi di destinazione
- Creare un gruppo di destinazione per il Gateway Load Balancer
- Controlli sanitari per i gruppi target di Gateway Load Balancer
- Modifica gli attributi del gruppo target per il tuo Gateway Load Balancer
- Registra gli obiettivi per il tuo Gateway Load Balancer
- Contrassegna un gruppo target per il tuo Gateway Load Balancer
- Elimina un gruppo target per il tuo Gateway Load Balancer


## Configurazione dell'instradamento

I gruppi di destinazione per Gateway Load Balancer supportano i seguenti protocolli e porte:

- Protocollo: GENEVE
- Porta: 6081

Il Gateway Load Balancer incapsula i pacchetti originali utilizzando GENEVE. L'intestazione GENEVE utilizza un formato Type-Length-Value (TLV) per memorizzare le informazioni, utilizzando la classe di opzione 0x0108. Le apparecchiature devono decapsulare le coppie TLV per elaborare i pacchetti originali. Per ulteriori informazioni, consulta il seguente post del blog: Integra il tuo dispositivo con un sistema di bilanciamento del carico Gateway.

## Target type (Tipo di destinazione)

Quando crei un gruppo target, devi specificare il tipo di target, che determina come vengono specificati i relativi oggetti target. Dopo aver creato un gruppo target, non è possibile modificarne il tipo di target.

I tipi di target possibili sono i seguenti:

`instance`

I target vengono specificati in base all'ID istanza.

`ip`

I target vengono specificati in base all'indirizzo IP.

Quando il tipo di target è `ip`, è possibile specificare gli indirizzi IP da uno dei blocchi CIDR seguenti:

- Sottoreti del VPC per il gruppo target
- 10.0.0. 0/8 (RFC 1918)
- 100.64.0. 0/10 (RFC 6598) https://tools.ietf.org/html/rfc6598
- 172.16.0. 0/12 (RFC 1918)
- 192.168.0. 0/16 (RFC 1918)

> **Important**
>
> Non è possibile specificare indirizzi IP instradabili pubblicamente.

# Destinazioni registrate

Il Gateway Load Balancer funge da singolo punto di contatto per i client e distribuisce il traffico in entrata tra le destinazioni registrati con stato integro. Ogni gruppo target deve avere almeno un target registrato in ciascuna zona di disponibilità abilitata per il Gateway Load Balancer. È possibile registrare ogni target con uno o più gruppi target.

Se il carico di richieste per l'applicazione aumenta, puoi registrare destinazioni aggiuntive con uno o più gruppi di destinazione al fine di gestire le richieste. Il Gateway Load Balancer inizia a instradare il traffico a un nuovo target registrato non appena viene completato il processo di registrazione.

Se il carico di richieste diminuisce o devi eseguire la manutenzione delle destinazioni, puoi annullare la loro registrazione dai gruppi target. L'annullamento della registrazione di un target rimuove il target dal gruppo target, ma non influisce in altro modo sul target stesso. Il Gateway Load Balancer arresta l'instradamento del traffico a una destinazione non appena la sua registrazione viene annullata. Il target passa allo stato `draining` fino a quando non vengono completate le richieste in transito. Puoi registrare di nuovo il target con il gruppo target quando è possibile riprendere la ricezione del traffico.

# Attributi dei gruppi di destinazione

Puoi utilizzare i seguenti attributi con i gruppi di destinazione:

`deregistration_delay.timeout_seconds`

La quantità di tempo che Elastic Load Balancing attende prima di modificare lo stato di un target di cui viene annullata la registrazione da `draining` a `unused`. L'intervallo è compreso tra 0 e 3600 secondi. Il valore predefinito è 300 secondi.

`stickiness.enabled`

Indica se la persistenza configurabile del flusso è abilitata per il gruppo di destinazione. I valori possibili sono `true` o `false`. Il valore predefinito è false. Quando l'attributo è impostato su `false`, viene utilizzato 5_tuple.

`stickiness.type`

Indica il tipo di persistenza del flusso. I valori possibili per i gruppi di destinazione associati a Gateway Load Balancers sono:

- `source_ip_dest_ip`

- `source_ip_dest_ip_proto`

`target_failover.on_deregistration`

Indica il modo in cui il Gateway Load Balancer gestisce i flussi esistenti quando viene annullata la registrazione di una destinazione. I valori possibili sono `rebalance` e `no_rebalance`. Il valore di default è `no_rebalance`. I due attributi (`target_failover.on_deregistration` e `target_failover.on_unhealthy`) non possono essere impostati separatamente. Il valore impostato per entrambi gli attributi deve essere lo stesso.

`target_failover.on_unhealthy`

Indica il modo in cui il Gateway Load Balancer gestisce i flussi esistenti quando una destinazione non è integra. I valori possibili sono `rebalance` e `no_rebalance`. Il valore di default è `no_rebalance`. I due attributi (`target_failover.on_deregistration` e `target_failover.on_unhealthy`) non possono essere impostati separatamente. Il valore impostato per entrambi gli attributi deve essere lo stesso.

Per ulteriori informazioni, consulta Modifica gli attributi del gruppo target.

## Creare un gruppo di destinazione per il Gateway Load Balancer

Registra le destinazioni per il vostro Gateway Load Balancer utilizzando un gruppo di destinazione.

Per instradare il traffico verso i target in un gruppo target, crea un listener e specifica il gruppo target nell'operazione predefinita per il listener. Per ulteriori informazioni, consulta Listener.

È possibile aggiungere o rimuovere target dal gruppo target in qualsiasi momento. Per ulteriori informazioni, consulta Registrazione di destinazioni. È anche possibile modificare le impostazioni di controllo dello stato per il gruppo target. Per ulteriori informazioni, consulta Modifica le impostazioni del controllo dell'integrità.

Per creare un gruppo target tramite la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, sotto Bilanciamento del carico, scegli Gruppi di destinazioni.
3. Scegliere Crea gruppo target.

4. Configurazione di base

   a. Per Scegli un tipo di destinazione, seleziona Istanze per specificare le destinazioni per ID dell'istanza, oppure seleziona Indirizzi IP per specificare le destinazioni per indirizzo IP.

   b. Per Nome del gruppo di destinazione, inserisci un nome per il gruppo di destinazione. Questo nome deve essere unico per regione per ogni account, può avere un massimo di 32 caratteri, deve contenere solo caratteri alfanumerici o trattini e non deve iniziare o terminare con un trattino.

   c. Verifica che il protocollo sia `GENEVE` e che la porta sia `6081`. Non sono supportati altri protocolli o porte.

   d. Per VPC, seleziona il cloud privato virtuale (VPC) con le istanze dell'appliance di sicurezza da includere nel gruppo di destinazione.

5. (Facoltativo) Per Controlli dell'integrità, modifica le impostazioni e le impostazioni avanzate secondo necessità. Se i controlli dell'integrità eccedono in modo consecutivo la Soglia di non integrità, il sistema di bilanciamento del carico considererà la destinazione fuori servizio. Se durante i controlli dell'integrità il numero di successi consecutivi supera la Soglia di integrità, il sistema di bilanciamento del carico considererà la destinazione nuovamente in servizio. Per ulteriori informazioni, consulta Controlli sanitari per i gruppi target di Gateway Load Balancer.

6. (Facoltativo) Espandi Tag e aggiungi i tag di cui hai bisogno.

7. Scegli Avanti.

8. Per Registra destinazioni aggiungi una o più destinazioni come segue:

   - Se il tipo di destinazione è Istanze, seleziona una o più istanze, inserisci una o più porte e in seguito scegli Includi come in sospeso di seguito.
   - Se il tipo di destinazione è Indirizzi IP, seleziona la rete, inserisci indirizzo IP e porte, quindi seleziona Includi come in sospeso di seguito.

9. Scegliere Crea gruppo target.

Per creare un gruppo target utilizzando il AWS CLI

Utilizza il comando create-target-group per creare il gruppo target, il comando add-tags per aggiungere un tag al gruppo target e il comando register-targets per aggiungere target.

# Controlli sanitari per i gruppi target di Gateway Load Balancer

È possibile registrare i target con uno o più gruppi target. Il Gateway Load Balancer inizia a instradare le richieste a un nuovo target registrato non appena viene completato il processo di registrazione. Il completamento del processo di registrazione e l'avvio dei controlli dell'integrità possono richiedere alcuni minuti.

Il Gateway Load Balancer invia periodicamente una richiesta ad ogni destinatario registrato per controllare lo stato. Dopo il completamento di ciascun controllo dell'integrità, il Gateway Load Balancer chiude la connessione definita per il controllo dello stato.

## Impostazioni del controllo dello stato

È possibile configurare controlli dell'integrità attivi per i target in un gruppo target utilizzando le seguenti impostazioni. Se i controlli di integrità superano il numero specificato di errori UnhealthyThresholdCountconsecutivi, il Gateway Load Balancer mette fuori servizio il target. Quando i controlli di integrità superano il numero specificato di successi HealthyThresholdCountconsecutivi, il Gateway Load Balancer riattiva l'obiettivo.

| Impostazione | Descrizione |
| --- | --- |
| HealthCheckProtocol | Il protocollo utilizzato dal sistema di bilanciam ento del carico durante l'esecuzione dei controlli dell'integrità sulle destinazioni. I protocolli possibili sono HTTP, HTTPS e TCP. Il valore predefinito è TCP. |
| HealthCheckPort | La porta utilizzata dal Gateway Load Balancer durante l'esecuzione dei controlli dell'integrità sulle destinazioni. L'intervallo è compreso tra 1 e 65535. Il predefinito è 80. |
| HealthCheckPath | [Controlli di integrità HTTP/HTTPS] Il percorso dei controlli di integrità che è la destinazione degli obiettivi per i controlli sanitari. Il valore di default è /. |
| HealthCheckTimeoutSeconds | Il periodo di tempo, in secondi, durante il quale l'assenza di risposta da un target indica che un |

| Impostazione | Descrizione |
| --- | --- |
| | controllo dello stato non è riuscito. L'intervallo è compreso tra 2 e 120. Il predefinito è 5. |
| HealthCheckIntervalSeconds | Il periodo di tempo approssimativo, in secondi, tra i controlli dell'integrità di una singola destinazione. L'intervallo è compreso tra 5 e 300. Il valore predefinito è 10 secondi. Questo valore deve essere maggiore o uguale a. HealthCheckTimeoutSeconds<br><br>**Important**<br>I controlli dell'integrità per i Gateway Load Balancer vengono distribuiti e utilizzano un meccanismo di consenso per determinare l'integrità della destinazione. Pertanto, è necessario aspettarsi che i dispositivi di destinazi one ricevano diversi controlli dell'inte grità entro l'intervallo di tempo configura to. |
| HealthyThresholdCount | Il numero di controlli dello stato andati a buon fine consecutivi necessari prima di considera re integro un target non integro. L'intervallo è compreso tra 2 e 10. Il predefinito è 5. |
| UnhealthyThresholdCount | Numero di controlli dello stato consecutivi non andati a buon fine necessari prima di considera re un target non integro. L'intervallo è compreso tra 2 e 10. Il valore predefinito è 2. |

| Impostazione | Descrizione |
| --- | --- |
| Matcher | [Controlli dello stato HTTP/HTTPS] I codici HTTP da utilizzare durante la verifica di una risposta con esito positivo ricevuta da un target. Questo valore deve essere 200-399. |

## Stato di integrità della destinazione

Prima che il Gateway Load Balancer invii una richiesta di controllo dell'integrità a una destinazione, è necessario registrarlo con un gruppo di destinazione, specificare il gruppo di destinazione in una regola dell'ascoltatore e assicurarsi che la zona di disponibilità del target sia abilitata per il Gateway Load Balancer.

La tabella seguente descrive i valori possibili per lo stato di un target registrato.

| Valore | Descrizione |
| --- | --- |
| `initial` | È in corso il processo di registrazione della destinazione o di esecuzione dei controlli dello stato iniziali della destinazione da parte del Gateway Load Balancer.<br><br>Codici di motivo correlati: `Elb.RegistrationInProgress` \| `Elb.InitialHealthChecking` |
| `healthy` | Il target è integro.<br><br>Codici di motivo correlati: Nessuno |
| `unhealthy` | Il target non ha risposto a un controllo di stato o il controllo dello stato non è andato a buon fine.<br><br>Codice di motivo correlato: `Target.FailedHealthChecks` |
| `unused` | La destinazione non è registrata con un gruppo di destinazione, il gruppo di destinazione non è utilizzato in una regola del listener, la destinazione è in una zona di |

| Valore | Descrizione |
| --- | --- |
| | disponibilità non abilitata oppure è nello stato arrestato o terminato.<br><br>Codici di motivo correlati: `Target.NotRegistered` \| `Target.NotInUse` \| `Target.InvalidState` \| `Target.IpUnusable` |
| `draining` | Il target viene revocato e la connection draining è in corso.<br><br>Codice di motivo correlato: `Target.DeregistrationInProgress` |
| `unavailable` | Lo stato della destinazione non è disponibile.<br><br>Codice di motivo correlato: `Elb.InternalError` |

## Codici di motivo di controllo dello stato

Se lo stato di una destinazione è un valore diverso da `Healthy`, l'API restituisce un codice di motivo e una descrizione del problema e la console visualizza la stessa descrizione. I codici di motivo che iniziano con `Elb` vengono creati nella parte relativa al Gateway Load Balancer e i codici di motivo che iniziano con `Target` vengono creati nella parte relativa ai target.

| Codice di motivo | Descrizione |
| --- | --- |
| `Elb.InitialHealthChecking` | Controlli dello stato iniziali in corso |
| `Elb.InternalError` | I controlli dello stato non andati a buon fine a causa di un errore interno |
| `Elb.RegistrationInProgress` | La registrazione del target è in corso |
| `Target.DeregistrationInProgress` | La revoca del target è in corso |

| Codice di motivo | Descrizione |
| --- | --- |
| `Target.FailedHealthChecks` | Controlli dello stato non andati a buon fine |
| `Target.InvalidState` | La destinazione è in stato di arresto<br><br>La destinazione è in stato terminato<br><br>I target sono in stato di arresto o terminato<br><br>Il target è in uno stato non valido |
| `Target.IpUnusable` | L'indirizzo IP non può essere utilizzato come destinazione, poiché è in uso in un sistema di bilanciamento del carico. |
| `Target.NotInUse` | Il gruppo di destinazione non è configurato per la ricezione del traffico dal Gateway Load Balancer.<br><br>Il target si trova in una zona di disponibilità che non è abilitata per il Gateway Load Balancer |
| `Target.NotRegistered` | Il target non è registrato nel gruppo target |

## Scenari di errore della destinazione del Gateway Load Balancer

Flussi esistenti: per impostazione predefinita, i flussi esistenti vanno alla stessa destinazione a meno che il flusso non scada o venga ripristinato, indipendentemente dallo stato di integrità e di registrazione del target. Questo approccio facilita lo svuotamento della connessione e supporta firewall di terze parti che a volte non sono in grado di rispondere ai controlli dell'integrità a causa dell'elevato utilizzo della CPU. Per ulteriori informazioni, consulta the section called “Failover della destinazione”.

Nuovi flussi: i nuovi flussi vengono inviati a una destinazione integra. Una volta presa una decisione di bilanciamento del carico per un flusso, il Gateway Load Balancer invierà il flusso alla stessa destinazione anche se non integra o in presenza di altre destinazioni integre.

Quando tutti le destinazioni non sono integre, il Gateway Load Balancer sceglie una destinazione casuale e inoltra il traffico per l'intera durata del flusso, fino a quando non viene ripristinato o non

scade il timeout. Poiché il traffico viene inoltrato a una destinazione non integra, il traffico viene interrotto finché tale destinazione non ricomincia a funzionare.

TLS 1.3: se un gruppo di destinazione è configurato con i controlli di integrità HTTPS, le destinazioni registrate non superano i controlli di integrità se supportano solo TLS 1.3. Queste destinazioni devono supportare una versione precedente di TLS, come TLS 1.2.

Bilanciamento del carico del carico tra zone: per impostazione predefinita, il bilanciamento del carico tra le zone di disponibilità è disabilitato. Se il bilanciamento del carico tra le zone è abilitato, ogni Gateway Load Balancer è in grado di vedere tutte le destinazioni in tutte le zone di disponibilità e vengono tutte trattate allo stesso modo, indipendentemente dalla zona.

Le decisioni relative al bilanciamento del carico e al controllo dell'integrità sono sempre indipendenti tra le zone. Anche quando il bilanciamento del carico tra le zone è abilitato, il comportamento dei flussi esistenti e dei nuovi flussi è lo stesso descritto sopra. Per ulteriori informazioni, consulta Bilanciamento del carico su più zone nella Guida per l'utente di Elastic Load Balancing.

## Controllo dello stato delle destinazioni

È possibile controllare lo stato dei target registrato con i gruppi target.

Per controllare lo stato dei target utilizzando la console

1. Apri la EC2 console Amazon all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, sotto Bilanciamento del carico, scegli Gruppi di destinazioni.
3. Scegli il nome del gruppo di destinazione per visualizzarne i dettagli.
4. Nella scheda Target, la colonna Stato indica lo stato di ogni destinazione.
5. Se lo stato di destinazione è un valore diverso da `Healthy`, la colonna Dettagli dello stato contiene ulteriori informazioni.

Per verificare lo stato di salute dei tuoi bersagli, utilizza il AWS CLI

Utilizza il comando describe-target-health. L'output di questo comando contiene lo stato del target. Include un codice di motivo, se lo stato è un valore diverso da `Healthy`.

Per ricevere notifiche via e-mail su destinazioni non integre

Usa gli CloudWatch allarmi per attivare una funzione Lambda per inviare dettagli su obiettivi non sani. Per step-by-step istruzioni, consulta il seguente post sul blog: Identificazione degli obiettivi non integri del sistema di bilanciamento del carico.

## Modifica le impostazioni del controllo dell'integrità

È anche possibile modificare alcune delle impostazioni di controllo dello stato per il gruppo target.

Per modificare le impostazioni di controllo dello stato per un gruppo target utilizzando la console

1. Apri la EC2 console Amazon all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, sotto Bilanciamento del carico, scegli Gruppi di destinazioni.
3. Scegli il nome del gruppo di destinazione per visualizzarne i dettagli.
4. Nella scheda Dettagli del gruppo, nella sezione Impostazioni del controllo dell'integrità, scegli Modifica.
5. Nella pagina Modifica le impostazioni del controllo dell'integrità, modificare le impostazioni secondo necessità, quindi scegliere Salva modifiche.

Per modificare le impostazioni del controllo dello stato di salute per un gruppo target utilizzando il AWS CLI

Utilizza il comando modify-target-group.

# Modifica gli attributi del gruppo target per il tuo Gateway Load Balancer

Dopo aver creato un gruppo target per il tuo Gateway Load Balancer, puoi modificarne gli attributi.

Attributi dei gruppi di destinazione

- Failover della destinazione
- Reimpostazione del protocollo TCP in caso di errore e annullamento della registrazione del target
- Ritardo di annullamento della registrazione
- Persistenza del flusso

# Failover della destinazione

Con il failover della destinazione, si specifica il modo in cui il Gateway Load Balancer gestisce i flussi di traffico esistenti dopo che una destinazione diventa non integra o quando viene annullata la registrazione della destinazione. Per impostazione predefinita, il Gateway Load Balancer continua a inviare flussi esistenti alla stessa destinazione, anche se la destinazione ha avuto esito negativo o è stata annullata la registrazione. È possibile gestire questi flussi eseguendone un nuovo hashing (`rebalance`) o lasciandoli allo stato predefinito (`no_rebalance`).

Nessun ribilanciamento:

Il Gateway Load Balancer continua a inviare i flussi esistenti a destinazioni non integre o esaurite. Se il Gateway Load Balancer non riesce a raggiungere l'obiettivo, il traffico viene interrotto.

Tuttavia, i nuovi flussi vengono inviati a destinazioni integre. Questo è il comportamento che segue di default.

Ribilanciamento:

Il Gateway Load Balancer rielabora i flussi esistenti e li invia a destinazioni integre dopo il timeout del ritardo di annullamento della registrazione.

Per le destinazioni annullate, il tempo minimo di failover dipenderà dal ritardo di annullamento della registrazione. La destinazione non viene contrassegnata come annullata fino al completamento del ritardo di annullamento della registrazione.

Per le destinazioni non integre, il tempo minimo di failover dipenderà dalla configurazione del controllo dello stato del gruppo di destinazione (soglia dei tempi di intervallo). Si tratta del periodo minimo prima del quale una destinazione viene contrassegnata come non integro. Trascorso questo periodo, il Gateway Load Balancer può impiegare diversi minuti a causa del tempo di propagazione aggiuntivo e del backoff della ritrasmissione TCP prima di reindirizzare i nuovi flussi verso destinazioni integre.

Per aggiornare l'attributo di failover di destinazione utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel pannello di navigazione, in Bilanciamento del carico scegli Gruppi di destinazione.
3. Scegli il nome del gruppo di destinazione per visualizzarne i dettagli.
4. Nella pagina Dettagli del gruppo, all'interno della Attributi, scegli Modifica.

5. Nella pagina Modifica attributi, modifica il valore di Failover della destinazione secondo necessità.
6. Scegli Save changes (Salva modifiche).

Per aggiornare l'attributo di failover di destinazione utilizzando AWS CLI

Utilizza il comando modify-target-group-attributes, con le seguenti coppie chiave-valore:

- Chiave=`target_failover.on_deregistration` e Valore= `no_rebalance` (predefinito) o `rebalance`
- Chiave=`target_failover.on_unhealthy` e Valore= `no_rebalance` (predefinito) o `rebalance`

> **Note**
>
> Entrambi gli attributi (`target_failover.on_deregistration` e `target_failover.on_unhealthy`) devono avere lo stesso valore.

## Reimpostazione del protocollo TCP in caso di errore e annullamento della registrazione del target

Se abilitato, il Gateway Load Balancer invia un ripristino TCP (RST) al mittente del traffico quando una destinazione non è integra o viene annullata la registrazione, al termine del tempo di esaurimento della connessione. Grazie a questa funzionalità, i mittenti possono recuperare rapidamente il traffico stabilendo nuove connessioni verso destinazioni integre.

`send_tcp_reset.on_unhealthy.enabled`

Indica se il Gateway Load Balancer invia un ripristino del TCP quando un target non è integro. I valori possibili sono `true` e `false`. Il valore di default è `false`.

`send_tcp_reset.on_deregistration.enabled`

Indica se il Gateway Load Balancer invia un ripristino del TCP quando una destinazione viene annullata la registrazione e il tempo di esaurimento della connessione è scaduto. I valori possibili sono `true` e `false`. Il valore di default è `false`.

Per abilitare la reimpostazione del protocollo TCP in caso di errore e annullamento della registrazione della destinazione utilizzando AWS CLI

Usa il comando https://docs.aws.amazon.com/cli/latest/reference/elbv2/modify-target-group-attributes.html modify-target-group-attributes, con entrambi gli attributi impostati su. `true`

```
aws elbv2 modify-target-group-attributes --target-group-arn target-group-
arn --attributes Key=send_tcp_reset.on_unhealthy.enabled,Value=true
 Key=send_tcp_reset.on_deregistration.enabled,Value=true
```

Per abilitare il ripristino del protocollo TCP in caso di errore del target e annullamento della registrazione utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, sotto Bilanciamento del carico, scegli Gruppi di destinazioni.
3. Seleziona il tuo gruppo target, scegli la  scheda Attributi, quindi scegli Modifica.
4. Scegli il riquadro  Nessun ribilanciamento e invia il ripristino del protocollo TCP (consigliato).
5. Scegli  Invia ripristino TCP in caso di errore e Invia ripristino TCP in caso di  annullamento della registrazione.
6. Scegli Save changes (Salva modifiche).

Questa funzionalità ha i seguenti requisiti:

- Questa funzionalità richiede una resistenza del flusso a 5 tuple, che il gruppo target utilizza per impostazione predefinita quando `stickiness.enabled` è impostata su. `false`
- Non puoi abilitare `send_tcp_reset.on_unhealthy.enabled` o `send_tcp_reset.on_deregistration.enabled` quando `stickiness.enabled` è impostato su. `true`
- Non è possibile abilitare questa funzionalità quando gli attributi di failover di destinazione (`target_failover.on_unhealthyotarget_failover.on_deregistration`) sono impostati su. `rebalance`
- Questa funzionalità si applica solo al traffico TCP. UDP e altri protocolli non sono interessati.
- Questa funzionalità è compatibile con i Gateway Load Balancer nuovi ed esistenti.

## Ritardo di annullamento della registrazione

Quando si annulla la registrazione di una destinazione, il Gateway Load Balancer gestisce i flussi verso tale destinazione nel modo seguente:

Nuovi flussi

Il Gateway Load Balancer interrompe l'invio di nuovi flussi.

Flussi esistenti

Il Gateway Load Balancer gestisce i flussi esistenti in base al protocollo:

- TCP: i flussi esistenti vengono chiusi se rimangono inattivi per più di 350 secondi.
- Altri protocolli: i flussi esistenti vengono chiusi se rimangono inattivi per più di 120 secondi.

Per favorire il drenaggio dei flussi esistenti, puoi abilitare il riequilibrio dei flussi per il tuo gruppo di destinazione. Per ulteriori informazioni, consulta the section called “Failover della destinazione”.

Un obiettivo annullato indica che è `draining` fino alla scadenza del timeout. Dopo la scadenza del timeout, la destinazione passa allo stato `unused`.

Per aggiornare l'attributo del ritardo di annullamento della registrazione utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel pannello di navigazione, in Bilanciamento del carico scegli Gruppi di destinazione.
3. Scegli il nome del gruppo di destinazione per visualizzarne i dettagli.
4. Nella pagina Dettagli del gruppo, all'interno della Attributi, scegliere Modifica.
5. Nella pagina Modifica attributi, modificare il valore di Intervallo annullamento registrazione secondo necessità.
6. Scegli Save changes (Salva modifiche).

Per aggiornare l'attributo del ritardo di annullamento della registrazione utilizzando AWS CLI

Utilizza il comando modify-target-group-attributes.

## Persistenza del flusso

Per impostazione predefinita, il Gateway Load Balancer mantiene l'aderenza dei flussi a uno specifico dispositivo di destinazione utilizzando 5 tuple (per i TCP/UDP flussi). La 5-tuple include IP di origine,

porta di origine, IP di destinazione, porta di destinazione e protocollo di trasporto. È possibile utilizzare l'attributo di tipo di persistenza e per modificare l'impostazione predefinita (5-tuple) e scegliere 3-tuple (IP di origine, IP di destinazione e protocollo di trasporto) o 2-tuple (IP di origine e IP di destinazione).

Considerazioni sulla persistenza del flusso

- La persistenza del flusso viene configurata e applicata a livello di gruppo di destinazione e si applica a tutto il traffico diretto al gruppo di destinazione.
- La persistenza del flusso a 2-tuple e 3-tuple non è supportata quando la modalità appliance AWS Transit Gateway è attivata. Per utilizzare la modalità appliance sul tuo Gateway Load Balancer AWS Transit Gateway, utilizza 5 tuple Flow Stickiness
- La persistenza del flusso può portare a una distribuzione non uniforme di connessioni e flussi, che potrebbe influire sulla disponibilità delle destinazioni. Si consiglia di terminare o svuotare tutti i flussi esistenti prima di modificare il tipo di persistenza del gruppo di destinazione.

Per aggiornare l'attributo flow stickiness utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel pannello di navigazione, in Bilanciamento del carico scegli Gruppi di destinazione.
3. Scegli il nome del gruppo di destinazione per visualizzarne i dettagli.
4. Nella pagina Dettagli del gruppo, all'interno della Attributi, scegli Modifica.
5. Nella pagina Modifica attributi, modifica il valore di Persistenza del flusso secondo necessità.
6. Scegli Save changes (Salva modifiche).

Per aggiornare l'attributo di aderenza del flusso utilizzando AWS CLI

Utilizza il comando modify-target-group-attributes con gli attributi del gruppo di destinazione `stickiness.enabled` e `stickiness.type`.

# Registra gli obiettivi per il tuo Gateway Load Balancer

Quando la destinazione è pronta per gestire le richieste, è possibile registrarla con uno o più gruppi di destinazione. È possibile registrare i target in base a ID istanza o indirizzo IP. Il Gateway Load balancer inizia a instradare le richieste alla destinazione non appena viene completato il processo di registrazione e la destinazione supera i controlli dell'integrità iniziali. Il completamento del processo di

registrazione e l'avvio dei controlli dello stato può richiedere alcuni minuti. Per ulteriori informazioni, consulta Controlli sanitari per i gruppi target di Gateway Load Balancer.

Se il carico di richieste per i target attualmente registrati aumenta, puoi registrare target aggiuntivi al fine di gestire le richieste. Se la richiesta sulle destinazioni registrate diminuisce, è possibile annullare la registrazione delle destinazioni dal gruppo di destinazione. Il completamento del processo di annullamento della registrazione e l'interruzione delle richieste di instradamento alla destinazione da parte del Gateway Load balancer può richiedere alcuni minuti. Se successivamente la domanda aumenta, è possibile registrare nuovamente le destinazioni di cui si era annullata la registrazione con il gruppo di destinazione. Se è necessario eseguire la manutenzione di una destinazione, è possibile annullarne la registrazione e registrarla nuovamente al termine della manutenzione.

Indice


- Considerazioni
- Gruppi di sicurezza target
- Liste di controllo accessi (ACL) di rete
- Registra le destinazioni in base all'ID dell'istanza
- Registra le destinazioni in base all'indirizzo IP
- Annulla la registrazione degli obiettivi


## Considerazioni

- Ogni di destinazione deve avere almeno una destinazione registrata in ciascuna zona di disponibilità abilitata per il Gateway Load Balancer.
- Il tipo di destinazione del gruppo di destinazioni determina il modo in cui si registrano le destinazioni con quel gruppo di destinazioni. Per ulteriori informazioni, consulta Target type (Tipo di destinazione).
- Non è possibile registrare obiettivi su un peering VPC interregionale.
- Non puoi registrare le istanze in base all'ID dell'istanza su un peering VPC interregionale, ma puoi registrarle tramite indirizzo IP.

## Gruppi di sicurezza target

Quando esegui la registrazione delle istanze EC2 come destinazione, devi assicurarti che i gruppi di sicurezza per queste istanze permettano il traffico in ingresso e in uscita sulla porta 6081.

I Gateway Load Balancer non hanno gruppi di sicurezza associati. Pertanto, i gruppi di sicurezza per i target devono utilizzare gli indirizzi IP per permettere il traffico dal sistema di bilanciamento del carico.

## Liste di controllo accessi (ACL) di rete

Quando esegui la registrazione delle istanze EC2 come destinazione, devi assicurarti che le liste di controllo degli accessi (ACL) di rete per le sottoreti delle istanze permettano il traffico sulla porta 6081. L'ACL di rete predefinita per un VPC consente tutto il traffico in entrata e in uscita. Se crei liste di controllo accessi di rete personalizzate, verifica che consentano il traffico appropriato.

## Registra le destinazioni in base all'ID dell'istanza

Quando viene registrata, un'istanza deve essere nello stato `running`.

Per registrare le destinazioni in base all'ID dell'istanza utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel pannello di navigazione, in Bilanciamento del carico scegli Gruppi di destinazione.
3. Scegli il nome del gruppo di destinazione per visualizzarne i dettagli.
4. Nella scheda  Obiettivi, scegli  Registra obiettivi.
5. Seleziona le istanze, quindi scegli  Includi come in sospeso di seguito.
6. Dopo aver finito di aggiungere le istanze, scegli Registra destinazioni in sospeso.

Per registrare le destinazioni in base all'ID dell'istanza utilizzando AWS CLI

Utilizzate il  comando  register-targets con gli ID delle istanze.

## Registra le destinazioni in base all'indirizzo IP

Un indirizzo IP registrato deve provenire da uno dei seguenti blocchi CIDR:

- Sottoreti del VPC per il gruppo target
- 10.0.0. 0/8 (RFC 1918)
- 100.64.0. 0/10 (RFC 6598)
- 172.16.0. 0/12 (RFC 1918)
- 192.168.0. 0/16 (RFC 1918)

Per registrare le destinazioni in base all'indirizzo IP utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel pannello di navigazione, in Bilanciamento del carico scegli Gruppi di destinazione.
3. Scegli il nome del gruppo di destinazione per visualizzarne i dettagli.
4. Nella scheda  Obiettivi, scegli  Registra obiettivi.
5. Scegli la rete, gli indirizzi IP e le porte, quindi scegli  Includi come in sospeso di seguito.
6. Dopo aver finito di specificare gli indirizzi, scegli Registra destinazioni in sospeso.

Per registrare le destinazioni in base all'indirizzo IP utilizzando AWS CLI

Utilizzate il  comando  register-targets con gli indirizzi IP delle destinazioni.

## Annulla la registrazione degli obiettivi

Quando annulli la registrazione di una destinazione, Elastic Load Balancing attende il completamento delle richieste in transito. Questo comportamento è noto come Connection Draining. Lo stato di un target è `draining` durante la fase di Connection Draining. Una volta completata l'annullamento della registrazione, lo stato del target diventa `unused`. Per ulteriori informazioni, consulta Ritardo di annullamento della registrazione.

Per annullare la registrazione degli obiettivi utilizzando la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel pannello di navigazione, in Bilanciamento del carico scegli Gruppi di destinazione.
3. Scegli il nome del gruppo di destinazione per visualizzarne i dettagli.
4. Scegliere la scheda Destinazioni.
5. Seleziona le destinazioni, quindi scegli  Annulla registrazione.

Per annullare la registrazione degli obiettivi utilizzando il AWS CLI

Utilizzate il comando https://docs.aws.amazon.com/cli/latest/reference/elbv2/deregister-targets.html deregister-targets per rimuovere le destinazioni.

# Contrassegna un gruppo target per il tuo Gateway Load Balancer

I tag ti aiutano a classificare i gruppi target in modi diversi, ad esempio in base a scopo, proprietario o ambiente.

È possibile aggiungere più tag a ciascun gruppo target. Le chiavi dei tag devono essere univoche per ogni gruppo target. Se aggiungi un tag con una chiave già associata al gruppo target, il valore del tag viene aggiornato.

Quando un tag non serve più, è possibile rimuoverlo.

Restrizioni

- Numero massimo di tag per risorsa: 50
- Lunghezza massima della chiave: 127 caratteri Unicode
- Lunghezza massima del valore: 255 caratteri Unicode
- I valori e le chiavi dei tag rispettano la distinzione tra maiuscole e minuscole. I caratteri consentiti sono lettere, spazi e numeri rappresentabili in UTF-8, più i seguenti caratteri speciali: + - =. _ : / @. Non utilizzare spazi iniziali o finali.
- Non utilizzate il `aws:` prefisso nei nomi o nei valori dei tag perché è AWS riservato all'uso. Non è possibile modificare né eliminare i nomi o i valori di tag con tale prefisso. I tag con questo prefisso non vengono conteggiati per il limite del numero di tag per risorsa.

Per aggiornare i tag per un gruppo target tramite la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel pannello di navigazione, in Bilanciamento del carico scegli Gruppi di destinazione.
3. Scegli il nome del gruppo di destinazione per visualizzarne i dettagli.
4. Nella scheda Tag, scegli Aggiungi/modifica tag ed eseguire una o più delle operazioni seguenti:
   a. Per aggiornare un tag, inserisci nuovi valori per Chiave e Valore.
   b. Per aggiungere un tag, scegli Aggiungi tag e inserire valori per Chiave e Valore.
   c. Per eliminare un tag, scegli Rimuovi accanto al tag.
5. Una volta completato l'aggiornamento dei tag, scegli Salva.

Per aggiornare i tag per un gruppo target utilizzando AWS CLI

Utilizza i comandi add-tags e remove-tags.

# Elimina un gruppo target per il tuo Gateway Load Balancer

È possibile eliminare un gruppo di destinazioni se non ci sono operazioni di inoltro di alcuna regola dell'ascoltatore che vi fanno riferimento. L'eliminazione di un gruppo target non influisce sui target registrati con il gruppo target. Se non hai più bisogno di un'istanza EC2 registrata, puoi arrestarla o terminarla.

Per eliminare un gruppo target tramite la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Nel riquadro di navigazione, sotto Bilanciamento del carico, scegli Gruppi di destinazioni.
3. Selezionare il gruppo target e scegliere Operazioni, Elimina.
4. Quando viene richiesta la conferma, seleziona Sì, elimina.

Per eliminare un gruppo target utilizzando il AWS CLI

Utilizza il comando delete-target-group.

# Monitora i Gateway Load Balancer

È possibile utilizzare le seguenti funzionalità per monitorare i Gateway Load Balancer per analizzare i modellid i traffico e risolvere i problemi. Tuttavia, il Gateway Load Balancer non genera log di accesso poiché è un sistema di bilanciamento del carico di livello 3 trasparente che non interrompe i flussi. Per ricevere i log degli accessi, è necessario abilitare la registrazione degli accessi sui dispositivi di destinazione Gateway Load Balancer, IDS/IPS come firewall e dispositivi di sicurezza. Inoltre, puoi anche scegliere di abilitare i log di flusso VPC sui Gateway Load Balancer.

CloudWatch metriche

Puoi usare Amazon CloudWatch per recuperare statistiche sui punti dati per i tuoi Gateway Load Balancer e target sotto forma di un set ordinato di dati di serie temporali, noti come metriche. È possibile utilizzare questi parametri per verificare che le prestazioni del sistema siano quelle previste. Per ulteriori informazioni, consulta CloudWatch metriche per il tuo Gateway Load Balancer.

Log di flusso VPC

Puoi utilizzare i log di flusso VPC per acquisire informazioni dettagliate sul traffico in entrata e in uscita dal Gateway Load Balancer. Per ulteriori informazioni, consulta Log di flusso VPC nella Guida per l'utente di Amazon VPC.

Crea un log di flusso per ciascuna interfaccia di rete del Gateway Load Balancer. Esiste una sola interfaccia di rete per ogni sottorete. Per identificare le interfacce di rete di un Gateway Load Balancer, cerca il nome del Gateway Load Balancer nel campo descrizione dell'interfaccia di rete.

Per ogni connessione tramite Gateway Load Balancer sono disponibili due voci: una per la connessione front-end tra il client e il Gateway Load Balancer e l'altra per la connessione back-end tra il Gateway Load Balancer e la destinazione. Se il target è registrato dall'ID dell'istanza, la connessione appare sull'istanza come connessione dal client. Se il gruppo di sicurezza dell'istanza non consente le connessioni dal client ma queste sono permesse dalle liste di controllo degli accessi di rete della sottorete, i log dell'interfaccia di rete del Gateway Load Balancer mostrano "ACCEPT OK" per le connessioni front-end e back-end, mentre i log dell'interfaccia di rete dell'istanza mostrano "REJECT OK" per la connessione.

CloudTrail log

Puoi utilizzarli AWS CloudTrail per acquisire informazioni dettagliate sulle chiamate effettuate all'API Elastic Load Balancing e archiviarle come file di registro in Amazon S3. Puoi utilizzare

questi CloudTrail registri per determinare quali chiamate sono state effettuate, l'indirizzo IP di origine da cui proviene la chiamata, chi ha effettuato la chiamata, quando è stata effettuata la chiamata e così via. Per ulteriori informazioni, consulta Log API calls for Elastic Load Balancing using CloudTrail.

# CloudWatch metriche per il tuo Gateway Load Balancer

Elastic Load Balancing pubblica i punti dati su Amazon CloudWatch per i tuoi Gateway Load Balancer e i tuoi obiettivi. CloudWatch ti consente di recuperare le statistiche su tali punti dati come un insieme ordinato di dati di serie temporali, noti come metriche. Pensa a un parametro come a una variabile da monitorare e ai dati di utilizzo come ai valori di questa variabile nel tempo. Ad esempio, puoi monitorare il numero totale di target integri per un Gateway Load Balancer in un periodo di tempo specifico. A ogni punto di dati sono associati un timestamp e un'unità di misura facoltativa.

Puoi utilizzare le metriche per verificare che le prestazioni del sistema siano quelle previste. Ad esempio, puoi creare un CloudWatch allarme per monitorare una metrica specificata e avviare un'azione (come l'invio di una notifica a un indirizzo email) se la metrica non rientra nell'intervallo che consideri accettabile.

Elastic Load Balancing riporta le metriche CloudWatch solo quando le richieste fluiscono attraverso il Gateway Load Balancer. Se ci sono delle richieste che passano, Elastic Load Balancing ne misura e invia i parametri a intervalli di 60 secondi. Se non vi sono richieste o in assenza di dati per un parametro, questa non viene segnalata.

Per ulteriori informazioni, consulta la Amazon User Guide. CloudWatch

Indice


- Parametri di Gateway Load Balancer
- Dimensioni di parametro per Gateway Load Balancer
- Visualizza le metriche per CloudWatch il tuo Gateway Load Balancer


## Parametri di Gateway Load Balancer


- Flussi
- Produttività
- Connessioni e ripristini

- I pacchetti cadono
- CLU
- Integrità del gruppo di destinazioni

Il `AWS/GatewayELB` namespace include le seguenti metriche per i flussi.

| Metrica | Description |
| --- | --- |
| `ActiveFlowCount` | Il numero totale di flussi simultanei (o connessioni) da client a target.<br><br>Criteri di segnalazione: è presente un valore diverso da zero<br><br>Statistiche: le statistiche più utili sono `Average`, `Maximum` e `Minimum`.<br><br>Dimensioni<br><br>• `LoadBalancer`<br>• `AvailabilityZone` , `LoadBalancer` |
| `NewFlowCount` | Il numero totale di nuovi flussi (o connessioni) stabiliti da client a target nel periodo di tempo.<br><br>Criteri di segnalazione: è presente un valore diverso da zero<br><br>Statistiche: la statistica più utile è `Sum`.<br><br>Dimensioni<br><br>• `LoadBalancer`<br>• `AvailabilityZone` , `LoadBalancer` |
| `RejectedFlowCount` | Il numero totale di flussi (o connessioni) rifiutati dal load balancer.<br><br>Criteri di segnalazione: sempre segnalati.<br><br>Statistiche: le statistiche più utili sono `Average`, `Maximum` e `Minimum`. |

| Metrica | Description |
| --- | --- |
| | Dimensioni<br>• LoadBalancer<br>• AvailabilityZone , LoadBalancer |
| RejectedFlowCount_TCP | Il numero di flussi (o connessioni) TCP rifiutati dal load balancer.<br>Criteri di segnalazione: vi è un valore diverso da zero.<br>Statistiche: la statistica più utile è Sum.<br>Dimensioni<br>• LoadBalancer<br>• AvailabilityZone , LoadBalancer |

Il AWS/GatewayELB namespace include le seguenti metriche per la velocità effettiva.

| Metrica | Description |
| --- | --- |
| ProcessedBytes | Il numero totale di byte elaborati dal sistema di bilanciamento del carico. Questo conteggio include il traffico da e verso le destinazioni, ma non il traffico di controllo dell'integrità.<br>Criteri di segnalazione: è presente un valore diverso da zero<br>Statistiche: la statistica più utile è Sum.<br>Dimensioni<br>• LoadBalancer<br>• AvailabilityZone , LoadBalancer |
| ProcessedPackets | Il numero totale di pacchetti elaborati dal sistema di bilanciamento del carico. Questo conteggio include il traffico da e verso le destinazioni, incluso il traffico del controllo dell'integrità. |

| Metrica | Description |
| --- | --- |
| | Criteri di segnalazione: sempre segnalati.<br><br>Statistiche: la statistica più utile è `Sum`.<br><br>Dimensioni<br><br>• `LoadBalancer`<br>• `AvailabilityZone` , `LoadBalancer` |
| `PeakBytesPerSecond` | I byte medi più alti elaborati al secondo, calcolati ogni 10 secondi durante la finestra di campionamento. Questa metrica non include il traffico relativo ai controlli sullo stato di salute.<br><br>Criteri di segnalazione: sempre segnalati<br><br>Statistiche: la statistica più utile è `Maximum`.<br><br>Dimensioni<br><br>• `LoadBalancer`<br>• `AvailabilityZone` , `LoadBalancer` |
| `PeakPacketsPerSecond` | La più alta frequenza media di pacchetti (pacchetti elaborati al secondo), calcolata ogni 10 secondi durante la finestra di campionamento. Questo parametro include il traffico relativo ai controlli dell'integrità.<br><br>Criteri di segnalazione: sempre segnalati<br><br>Statistiche: la statistica più utile è `Maximum`.<br><br>Dimensioni<br><br>• `LoadBalancer`<br>• `AvailabilityZone` , `LoadBalancer` |

Il `AWS/GatewayELB` namespace include le seguenti metriche per le connessioni e le reimpostazioni.

| Metrica | Description |
| --- | --- |
| TCP_GWLB_Reset_Count | Il numero totale di pacchetti TCP reset (RST) generati dal Gateway Load Balancer quando la funzione di ripristino TCP è attiva.<br><br>Criteri di segnalazione: sempre segnalati<br><br>Statistiche: la statistica più utile è Sum.<br><br>Dimensioni<br><br>• LoadBalancer<br>• AvailabilityZone , LoadBalancer |

Il AWS/GatewayELB namespace include le seguenti metriche per le perdite di pacchetti.

| Metrica | Description |
| --- | --- |
| PacketsDroppedCount_Invalid GeneveTunnel | Quando si restituisce il pacchetto a GWLB, l'appliance di destinazione deve scambiare gli indirizzi IP di origine e destinazione del tunnel GENEVE e utilizzare la porta di destinazione GENEVE corretta (6081). Se il pacchetto non è conforme alle linee guida di cui sopra, GWLB eliminerà il pacchetto e aumenterà questa metrica.<br><br>Criteri di segnalazione: sempre segnalati<br><br>Statistiche: la statistica più utile è Sum.<br><br>Dimensioni<br><br>• LoadBalancer<br>• AvailabilityZone , LoadBalancer |
| PacketsDroppedCount_Invalid GwlbEndpointId | L'appliance deve restituire l'in nel TLV quando risponde a GWLB GwlbeEniId . Se manca questo TLV, GWLB eliminerà il pacchetto e incrementerà questa metrica.<br><br>Criteri di segnalazione: sempre segnalati |

| Metrica | Description |
| --- | --- |
| | Statistiche: la statistica più utile è `Sum`.<br><br>Dimensioni<br><br>• `LoadBalancer`<br>• `AvailabilityZone` , `LoadBalancer` |
| `PacketsDroppedCount_InvalidGwlbFlowCookie` | L'appliance deve restituire il FlowCookie TLV così com'è quando risponde a GWLB. Questa metrica viene incrementata se il cookie di flusso per un determinato flusso non corrisponde.<br><br>Criteri di segnalazione: sempre segnalati<br><br>Statistiche: la statistica più utile è `Sum`.<br><br>Dimensioni<br><br>• `LoadBalancer`<br>• `AvailabilityZone` , `LoadBalancer` |
| `PacketsDroppedCount_MissingFlowEntry` | GWLB tiene traccia di una voce di flusso per ogni flusso attivo che elabora. Se manca la voce di flusso per un pacchetto del percorso di ritorno, GWLB eliminerà il pacchetto e incrementerà questa metrica. Una voce di flusso può mancare se il flusso è scaduto, il pacchetto non appartiene a un flusso noto o il flusso è stato eliminato a causa del carico del sistema.<br><br>Criteri di segnalazione: sempre segnalati<br><br>Statistiche: la statistica più utile è `Sum`.<br><br>Dimensioni<br><br>• `LoadBalancer`<br>• `AvailabilityZone` , `LoadBalancer` |

Il AWS/GatewayELB namespace include le seguenti metriche per le unità di capacità di bilanciamento del carico (LCU).

| Metrica | Description |
| --- | --- |
| ConsumedLCUs | Il numero di unità di capacità del sistema di bilanciamento del carico (LCU) utilizzate dal tuo sistema di bilanciamento del carico. Paghi per il numero di LCU che usi all'ora. Per ulteriori informazioni, consulta Prezzi di Elastic Load Balancing.<br><br>Criteri di segnalazione: sempre segnalati<br><br>Statistiche: tutte<br><br>Dimensioni<br><br>• LoadBalancer |

Lo spazio dei nomi AWS/GatewayELB include i seguenti parametri per i l'integrità del gruppo di destinazioni.

| Metrica | Description |
| --- | --- |
| HealthyHostCount | Il numero di target considerati integri.<br><br>Criteri di segnalazione: segnalati se sono abilitati i controlli dello stato<br><br>Statistiche: le statistiche più utili sono Maximum e Minimum.<br><br>Dimensioni<br><br>• LoadBalancer , TargetGroup<br>• AvailabilityZone , LoadBalancer , TargetGroup |
| UnHealthyHostCount | Il numero di target considerati non integri.<br><br>Criteri di segnalazione: segnalati se sono abilitati i controlli dello stato<br><br>Statistiche: le statistiche più utili sono Maximum e Minimum. |

| Metrica | Description |
| --- | --- |
| | Dimensioni<br>• `LoadBalancer` , `TargetGroup`<br>• `AvailabilityZone` , `LoadBalancer` , `TargetGroup` |

## Dimensioni di parametro per Gateway Load Balancer

Per filtrare i parametri relativi al tuo Gateway Load Balancer, usa le seguenti dimensioni.

| Dimensione | Description |
| --- | --- |
| `AvailabilityZone` | Consente di filtrare i dati del parametro per zona di disponibilità. |
| `LoadBalancer` | Consente di filtrare i dati del parametro per Gateway Load Balancer. Specifica il Gateway Load Balancer nel modo seguente: load-balancer-name/1234567890123456 (la parte finale dell'ARN). |
| `TargetGroup` | Consente di filtrare i dati del parametro per gruppo target. Specifica il gruppo target nel modo seguente: targetgroup/target-group-name/1234567890123456 (la parte finale dell'ARN del gruppo target). |

## Visualizza le metriche per CloudWatch il tuo Gateway Load Balancer

Puoi visualizzare le CloudWatch metriche per i tuoi Gateway Load Balancer utilizzando la console Amazon EC2. Tali parametri vengono visualizzati come grafici di monitoraggio. I grafici di monitoraggio mostrano punti di dati se il Gateway Load Balancer è attivo e riceve richieste.

In alternativa, puoi visualizzare le metriche per il tuo Gateway Load Balancer utilizzando la console. CloudWatch

Per visualizzare i parametri tramite la console

1. Apri la console Amazon EC2 all'indirizzo https://console.aws.amazon.com/ec2/.
2. Per visualizzare i parametri filtrati per gruppo target, procedi nel seguente modo:

a. Seleziona Gruppi di destinazioni nel riquadro di navigazione.
b. Scegliere il gruppo target e selezionare Monitoring (Monitoraggio).
c. (Opzionale) Per filtrare i risultati in base al tempo, seleziona un intervallo di tempo in Visualizzazione dati per.
d. Per ingrandire la visualizzazione di un singolo parametro, selezionarne il grafico.

3. Per visualizzare i parametri filtrati in base al Gateway Load Balancer, procedi nel seguente modo:

a. Selezionare Sistemi di bilanciamento del carico nel riquadro di navigazione.
b. Scegli il Gateway Load Balancer e seleziona Monitoraggio.
c. (Opzionale) Per filtrare i risultati in base al tempo, seleziona un intervallo di tempo in Visualizzazione dati per.
d. Per ingrandire la visualizzazione di un singolo parametro, selezionarne il grafico.

Per visualizzare le metriche utilizzando la console CloudWatch

1. Apri la CloudWatch console all'indirizzo. https://console.aws.amazon.com/cloudwatch/
2. Nel riquadro di navigazione, seleziona Parametri.
3. Seleziona lo spazio dei nomi GatewayELB.
4. (Facoltativo) Per visualizzare tutte le dimensioni di un parametro, inserirne il nome nel campo di ricerca.

Per visualizzare le metriche utilizzando AWS CLI

Utilizza il seguente comando list-metrics per elencare i parametri disponibili:

```
aws cloudwatch list-metrics --namespace AWS/GatewayELB
```

Per ottenere le statistiche per una metrica utilizzando AWS CLI

Utilizza il seguente comando get-metric-statistics per ottenere statistiche su un parametro e una dimensione specifici. Nota che CloudWatch considera ogni combinazione univoca di dimensioni come una metrica separata. Non si possono recuperare le statistiche utilizzando combinazioni di dimensioni che non siano state specificamente pubblicate. Occorre specificare le stesse dimensioni utilizzate al momento della creazione dei parametri.

```
aws cloudwatch get-metric-statistics --namespace AWS/GatewayELB \
--metric-name UnHealthyHostCount --statistics Average  --period 3600 \
--dimensions Name=LoadBalancer,Value=net/my-load-balancer/50dc6c495c0c9188 \
Name=TargetGroup,Value=targetgroup/my-targets/73e2d6bc24d8a067 \
--start-time 2017-04-18T00:00:00Z --end-time 2017-04-21T00:00:00Z
```

Di seguito è riportato un output di esempio.

```
{
    "Datapoints": [
        {
            "Timestamp": "2020-12-18T22:00:00Z",
            "Average": 0.0,
            "Unit": "Count"
        },
        {
            "Timestamp": "2020-12-18T04:00:00Z",
            "Average": 0.0,
            "Unit": "Count"
        },
        ...
    ],
    "Label": "UnHealthyHostCount"
}
```

# Quota per i load balancer del gateway

L' AWS account dispone di quote predefinite, in precedenza denominate limiti, per ogni servizio. AWS Salvo diversa indicazione, ogni quota si applica a una Regione specifica. Se per alcune quote è possibile richiedere aumenti, altre quote non possono essere modificate.

Per richiedere un aumento delle quote, consultare Richiesta di aumento delle quote nella Guida dell'utente di Service Quotas. Se la quota non è ancora disponibile in Service Quotas, invia una richiesta di aumento della quota di servizio.

Quote

- Sistemi di load balancer
- Gruppi di destinazione
- Unità di capacità Load Balancer

## Sistemi di load balancer

Il tuo AWS account ha le seguenti quote relative ai Gateway Load Balancers.

| Name | Predefinita | Adattabile |
|---|---|---|
| Gateway Load Balancer per regione | 100 | Sì |
| Gateway Load Balancer per VPC | 100 | Sì |
| Load Balancer ENIs del gateway per VPC | 300 | Sì |
| Ascoltatori per Gateway Load Balancer | 1 | No |

* Ogni Gateway Load Balancer utilizza un'interfaccia di rete per zona.

## Gruppi di destinazione

Le quote elencate di seguito sono per i gruppi di destinazione.

| Name | Predefinita | Adattabile |
| --- | --- | --- |
| Gruppi di destinazione GENEVE per regione | 100 | Sì |
| Destinazioni per zona di disponibilità per gruppo di destinazione GENEVE | 300 | No |
| Destinazioni per Gateway Load Balancer e per zona di disponibilità | 300 | No |
| Destinazioni per Gateway Load Balancer | 300 | No |

## Unità di capacità Load Balancer

Le seguenti quote si riferiscono alle Load Balancer Capacity Units LCUs ().

| Name | Predefinita | Adattabile |
| --- | --- | --- |
| Unità di capacità (LCU) Gateway Load Balancer Riservate per regione | 0 | Sì |

# Cronologia dei documenti per i Gateway Load Balancer

La tabella seguente descrive i rilasci dei Gateway Load Balancer.

| Modifica | Descrizione | Data |
| --- | --- | --- |
| Prenotazione dell'unità di capacità | Questa versione aggiunge il supporto per impostare una capacità minima per il sistema di bilanciamento del carico. | 10 aprile 2025 |
| IPv6 supporto | Puoi configurare il tuo Gateway Load Balancer per supportare entrambi IPv4 gli IPv6 indirizzi. | 12 dicembre 2022 |
| Ribilanciamento del flusso | Questa versione aggiunge il supporto per definire il comportamento di gestione del flusso per Gateway Load Balancer in caso di errore o annullamento della registraz ione degli obiettivi. | 13 ottobre 2022 |
| Persistenza del flusso configurabile | È possibile configurare l'hashing che mantiene la persistenza dei flussi verso una specifica appliance di destinazione. | 25 agosto 2022 |
| Disponibile in nuove regioni | Questa versione aggiunge il supporto per i Gateway Load Balancer nelle AWS GovCloud (US) regioni. | 17 giugno 2021 |
| Disponibile in nuove regioni | Questa versione aggiunge il supporto per Gateway Load Balancer nelle regioni Canada | 31 marzo 2021 |

| | | |
|---|---|---|
| | (Centrale), Asia Pacifico (Seoul) e Asia Pacifico (Osaka). | |
| Disponibile in nuove regioni | Questa versione aggiunge il supporto per Gateway Load Balancer negli Stati Uniti occidentali (Californ ia settentrionale), Europa (Londra), Europa (Parigi), Europa (Milano), Africa (Città del Capo), Medio Oriente (Bahrein), Asia Pacifico (Hong Kong), Asia Pacifico (Singapor e) e Asia Pacifico (Mumbai). | 19 marzo 2021 |
| Versione iniziale | Questo rilascio di sistema di bilanciamento del carico elastico introduce i Gateway Load Balancer. | 10 novembre 2020 |

Le traduzioni sono generate tramite traduzione automatica. In caso di conflitto tra il contenuto di una traduzione e la versione originale in Inglese, quest'ultima prevarrà.