aws

Gateway Load Balancer

Elastic Load Balancing

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Elastic Load Balancing: Gateway Load Balancer

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon 的商标和商业外观不得用于任何非 Amazon 的商品或服务，也不得以任何可能引起客户混淆、贬低或诋毁 Amazon 的方式使用。所有非 Amazon 拥有的其他商标均为各自所有者的财产，这些所有者可能附属于 Amazon、与 Amazon 有关联或由 Amazon 赞助，也可能不是如此。

Table of Contents

什么是网关负载均衡器？ 1
网关负载均衡器概述 1
设备供应商 1
入门 2
定价 2
入门 3
概览 3
路由 5
先决条件 6
第 1 步：创建网关负载均衡器 6
第 2 步：创建网关负载均衡器端点服务 7
第 3 步：创建网关负载均衡器端点 8
第 4 步：配置路由 9
开始使用 CLI 11
概览 11
路由 5
先决条件 14
第 1 步：创建网关负载均衡器并注册目标 14
第 2 步：创建网关负载均衡器端点 15
步骤 3：配置路由 16
Gateway Load Balancer 18
负载均衡器状态 18
IP 地址类型 19
可用区 19
空闲超时 20
负载均衡器属性 20
网络 ACLs 21
非对称流量 21
网络最大传输单元（MTU） 21
创建负载均衡器 21
先决条件 22
创建负载均衡器 22
重要后续步骤 23
更新 IP 地址类型 23

编辑负载均衡器属性 23
删除保护 24
跨可用区负载均衡 24
为负载均衡器添加标签 25
删除负载均衡器 26
LCU 预订 26
申请预约 27
更新或终止预订 28
监控预约 29
侦听器 30
侦听器属性 30
更新侦听器目标组 30
更新空闲超时 31
目标组 32
路由配置 32
Target type 33
已注册目标 33
目标组属性 34
创建目标组 35
配置运行状况检查 36
运行状况检查设置 36
目标运行状况 37
运行状况检查原因代码 38
目标故障场景 39
检查目标的运行状况 40
修改运行状况检查设置 40
编辑目标组属性 41
目标失效转移 41
取消注册延迟 42
流量粘性 43
注册目标 44
注意事项 45
目标安全组 45
网络 ACLs 45
按实例 ID 注册目标 45
按 IP 地址注册目标 46

取消注册目标 46
为目标组添加标签 47
删除目标组 48
监控负载均衡器 49
CloudWatch 指标 49
网关负载均衡器指标 50
网关负载均衡器的指标维度 53
查看 Gateway Load Balancer 的 CloudWatch 指标 54
限额 56
负载均衡器 56
目标组 56
Load Balancer 容量单位 57
文档历史记录 58
..... lx

什么是网关负载均衡器？

弹性负载均衡会在一个或多个可用区中的多个目标之间自动分配传入的流量。它会监控已注册目标的运行状况，并仅将流量传输到运行状况良好的目标。弹性负载均衡 根据传入流量随时间的变化对负载均衡器进行扩展。它可以自动扩展来处理绝大部分工作负载。

弹性负载均衡 支持以下负载均衡器：Application Load Balancer、Network Load Balancer、Gateway Load Balancer 和经典负载均衡器。您可以选择最适合自己需求的负载均衡器类型。本指南讨论的是网关负载均衡器。有关其他负载均衡器的更多信息，请参阅 应用程序负载均衡器用户指南、网络负载均衡器用户指南 和 经典负载均衡器用户指南。

网关负载均衡器概述

Gateway Load Balancers 让您能够部署、扩展和管理虚拟设备，例如防火墙、入侵检测和防御系统以及深度数据包检测系统。它结合了一个透明的网络网关（即所有流量的单个入口和出口点），并分配流量，同时根据需求扩展虚拟设备。

Gateway Load Balancer 在 Open Systems Interconnection (OSI) 模型的第四层运行，网络层。它监听所有端口上的所有 IP 数据包，并将流量转发到监听程序规则中指定的目标组。使用 5 元组（默认）、3 元组或 2 元组来保持流向特定目标设备的流量粘性。网关负载均衡器及其注册的虚拟设备实例使用端口 6081 上的 GENEVE 协议交换应用程序流量。

Gateway Load Balancers 使用 Gateway Load Balancer 端点来安全地跨 VPC 边界交换流量。Gateway Load Balancer 端点是在服务提供商 VPC 中的虚拟设备与服务使用者 VPC 中的应用程序服务器之间提供私有连接的 VPC 端点。您将 Gateway Load Balancer 部署在与虚拟设备相同的 VPC 中。向 Gateway Load Balancer 的目标组注册虚拟设备。

进出网关负载均衡器端点的流量是使用路由表配置的。流量从服务使用者 VPC 通过网关负载均衡器端点流向服务提供者 VPC 中的网关负载均衡器，然后返回到服务使用者 VPC。您必须在不同的子网中创建网关负载均衡器端点和应用程序服务器。这将使您能够将网关负载均衡器端点配置为应用程序子网的路由表中的下一跃点。

有关更多信息，请参阅《AWS PrivateLink 指南》中的 通过 AWS PrivateLink访问虚拟设备。

设备供应商

您负责选择设备供应商的软件并进行资格鉴定。您必须信任设备软件才能让其检查或修改来自负载均衡器的流量。被列为 Elastic Load Balancing Par tners 的设备供应商已将其设备软件集成并通过认证

AWS。您可以对该列表中的供应商提供的设备软件给予更高的信任度。但是，AWS 不能保证这些供应商提供的软件的安全性或可靠性。

入门

要使用创建 Gateway Load Balancer AWS Management Console，请参阅入门。要使用创建 Gateway Load Balancer AWS Command Line Interface，请参阅开始使用 CLI。

定价

利用负载均衡器，您可以按实际用量付费。有关更多信息，请参阅 弹性负载均衡 定价。

网关负载均衡器入门

借助网关负载均衡器可以轻松部署、扩展和管理第三方虚拟设备，例如安全设备。

在本教程中，我们将使用一个网关负载均衡器和一个网关负载均衡器端点，实现一个检验系统。

内容

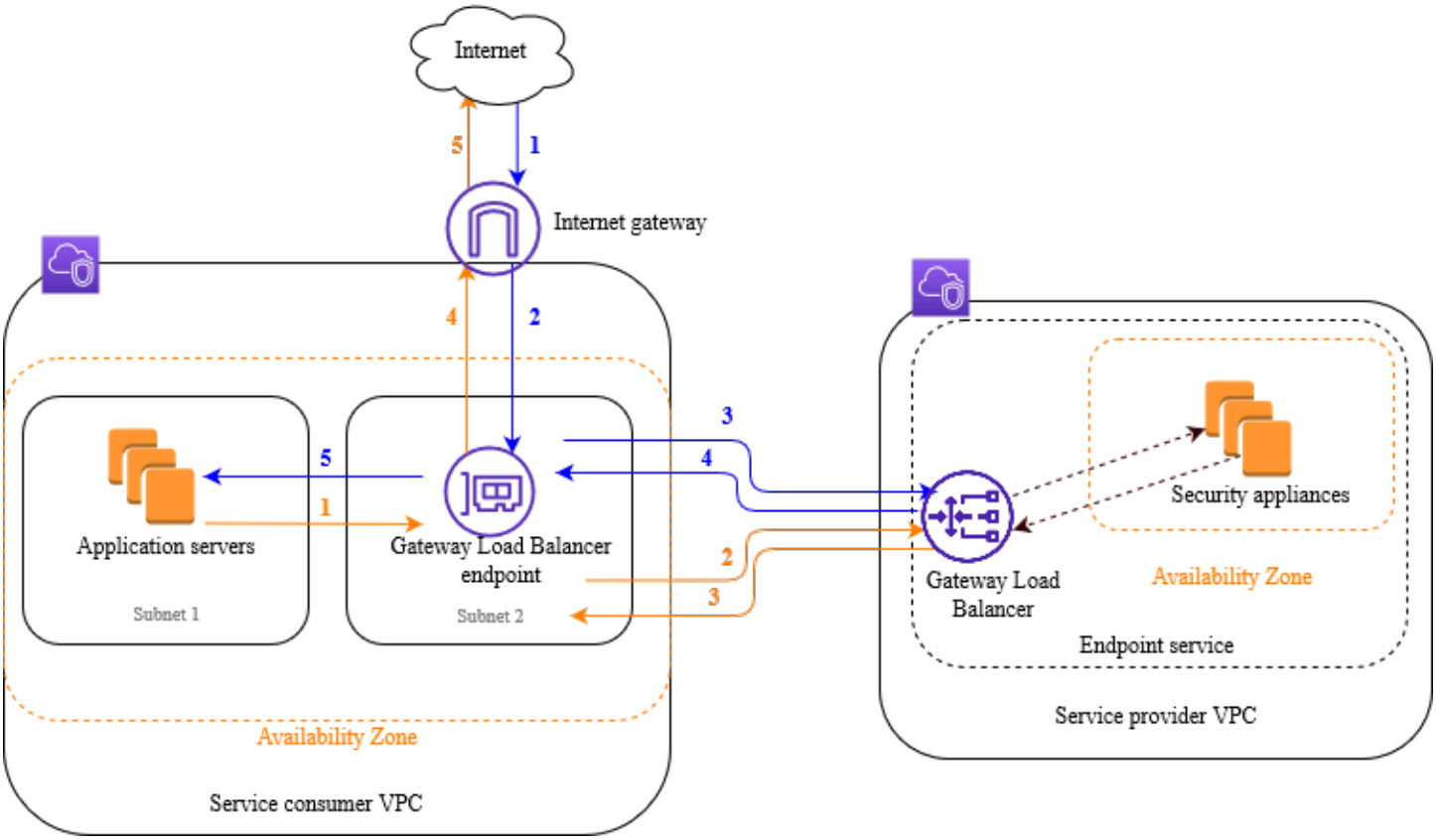
- 概览
- 先决条件
- 第 1 步：创建网关负载均衡器
- 第 2 步：创建网关负载均衡器端点服务
- 第 3 步：创建网关负载均衡器端点
- 第 4 步：配置路由

概览

网关负载均衡器端点是在服务提供者 VPC 中的虚拟设备与服务使用者 VPC 中的应用程序服务器之间提供私有连接的 VPC 端点。网关负载均衡器将部署在与虚拟设备相同的 VPC 中。这些设备将会注册到网关负载均衡器的目标组。

应用程序服务器与服务使用者 VPC 在同一个子网（目的地子网）中运行，而网关负载均衡器端点位于另一个子网中。通过互联网网关进入服务使用者 VPC 的所有流量首先会路由到网关负载均衡器端点，然后再路由到目标子网。

同样，离开应用程序服务器（目的地子网）的所有流量会首先路由到网关负载均衡器端点，然后再路由回互联网。以下网络图形象地演示了如何使用网关负载均衡器访问端点服务。



随后的编号项突出显示并解释了上面网络图中显示的元素。

从互联网到应用程序的流量（蓝色箭头）：

1. 流量通过互联网网关进入服务使用者 VPC。
2. 根据入口路由将流量发送到网关负载均衡器端点。
3. 将流量发送到网关负载均衡器，然后由后者将流量分配到其中的一个安全设备。
4. 安全设备完成检查后，将流量发送回网关负载均衡器端点。
5. 将流量发送到应用程序服务器（目的地子网）。

从应用程序到互联网的流量（橙色箭头）：

1. 根据应用程序服务器子网上配置的默认路由表，将流量发送到网关负载均衡器端点。
2. 将流量发送到网关负载均衡器，然后由后者将流量分配到其中的一个安全设备。
3. 安全设备完成检查后，将流量发送回网关负载均衡器端点。
4. 根据路由表配置，将流量发送到互联网网关。
5. 流量被路由回互联网。

路由

互联网网关的路由表必须具有将发往应用程序服务器的流量路由到网关负载均衡器端点的条目。要指定网关负载均衡器端点，请使用 VPC 端点的 ID。以下示例显示了 dualstack 配置的路由。

目标位置	目标
VPC IPv4 CIDR	本地
VPC IPv6 CIDR	本地
Subnet 1 IPv4 CIDR	*vpc-endpoint-id*
Subnet 1 IPv6 CIDR	*vpc-endpoint-id*

应用程序服务器所在子网的路由表必须具有将来自应用程序服务器的所有流量路由到网关负载均衡器端点的条目。

目标位置	目标
VPC IPv4 CIDR	本地
VPC IPv6 CIDR	本地
0.0.0.0/0	*vpc-endpoint-id*
::/0	*vpc-endpoint-id*

网关负载均衡器端点所在子网的路由表必须将从检查返回的流量路由到最终目的地。对于来自互联网的流量，本地路由将确保其会到达应用程序服务器。对于来自应用程序服务器的流量，则添加会将所有流量路由到互联网网关的条目。

目标位置	目标
VPC IPv4 CIDR	本地
VPC IPv6 CIDR	本地

目标位置	目标
0.0.0.0/0	*internet-gateway-id*
::/0	*internet-gateway-id*

先决条件

- 确保服务使用者 VPC 在包含应用程序服务器的每个可用区中至少有两个子网。其中一个子网用于网关负载均衡器端点，另一个用于应用程序服务器。
- 网关负载均衡器和目标可以位于同一子网中。
- 您不能使用其他账户共享的子网来部署网关负载均衡器。
- 在服务提供者 VPC 的每个安全设备子网中启动至少一个安全设备实例。这些实例的安全组必须允许端口 6081 上的 UDP 流量。

第 1 步：创建网关负载均衡器

按照以下过程创建负载均衡器、侦听器和目标组。

使用控制台创建负载均衡器、侦听器和目标组

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中的负载平衡下，选择负载均衡器。
3. 选择创建负载均衡器。
4. 在网关负载均衡器下，选择创建。
5. 基本配置
   a. 对于Load balancer name（负载均衡器名称），输入负载均衡器的名称。
   b. 对于 IP 地址类型，请选择IPv4仅支持 IPv4 地址，或者选择 Dualstack 以同时支持 IPv4 和 IPv6地址。
6. 网络映射
   a. 对于 VPC，请选择服务提供者 VPC。
   b. 对于映射，请选择您启动了安全设备实例的所有可用区，然后为每个可用区选择一个子网。
7. IP 侦听器路由

a. 对于默认操作，选择一个现有的目标组以用来接收流量。该目标组必须使用 GENEVE 协议。

   如果您没有目标组，请选择创建目标组，这时将在浏览器中打开一个新选项卡。选择一个目标组，输入目标组的名称，并保持选择 GENEVE 协议。选择您的安全设备实例所在的 VPC。根据需要修改运行状况检查设置，并添加需要的任何标签。选择下一步。您可以立即向目标组注册安全设备实例，也可以在完成此过程之后再注册。选择创建目标组，然后返回上一个浏览器选项卡。

b. （可选）展开侦听器标签并添加所需的标签。

8. （可选）展开负载均衡器标签并添加所需的标签。
9. 选择创建负载均衡器。

第 2 步：创建网关负载均衡器端点服务

按照以下过程创建将使用您的网关负载均衡器的端点服务。

创建网关负载均衡器端点服务

1. 打开位于 https://console.aws.amazon.com/vpc/ 的 Amazon VPC 控制台。
2. 在导航窗格中，选择 Endpoint services（端点服务）。
3. 选择创建端点服务，然后执行以下操作：

   a. 在 Load balancer type（负载均衡器类型）选项中选择 Gateway（网关）。

   b. 对于 Available load balancers（可用负载均衡器），选择您的网关负载均衡器。

   c. 对于需要接受以使用端点，选择需要接受以手动接受对端点服务的连接请求。否则这些请求将被自动接受。

   d. 对于 Supported IP address types（支持的 IP 地址类型），执行以下任一操作：

      - 选择 IPv4-启用终端节点服务以接受 IPv4 请求。
      - 选择 IPv6-启用终端节点服务以接受 IPv6 请求。
      - 选择IPv4和 IPv6-使终端节点服务能够同时接受 IPv4 和 IPv6 请求。

   e. （可选）若要添加标签，请选择 Add new tag（添加新标签），然后输入标签键和标签值。

   f. 选择创建。记下服务名称；您在创建端点时将需要此名称。

4. 选择新的端点服务，然后选择操作、允许主体。输入允许 ARNs 为服务创建终端节点的服务使用者。服务使用者可以是用户、IAM 角色或 AWS 账户。选择 Allow principals（允许委托人）。

第 3 步：创建网关负载均衡器端点

按照以下过程，创建可连接到您的网关负载均衡器端点服务的网关负载均衡器端点。网关负载均衡器端点是基于可用区的。我们建议您在每个可用区中创建一个网关负载均衡器端点。有关更多信息，请参阅《AWS PrivateLink 指南》中的 通过 AWS PrivateLink访问虚拟设备。

要创建网关负载均衡器终端节点

1. 打开位于 https://console.aws.amazon.com/vpc/ 的 Amazon VPC 控制台。
2. 在导航窗格中，选择端点。
3. 选择创建端点，然后执行以下操作：
   a. 在 Service category（服务类别）选项中，选择 Other endpoint services（其他端点服务）。
   b. 对于服务名称，输入您之前记下的服务名称，然后选择验证服务。
   c. 对于 VPC，请选择服务使用者 VPC。
   d. 对于子网，选择网关负载均衡器端点的子网。

      注意：创建 Gateway Load Balancer 终端节点时，您只能在每个可用区域内选择一个子网。
   e. 对于 IP address type（IP 地址类型），可从以下选项中进行选择：
      - IPv4— 为您的端点网络接口分配 IPv4 地址。仅当所有选定的子网都有 IPv4 地址范围时，才支持此选项。
      - IPv6— 为您的端点网络接口分配 IPv6 地址。仅当所有选定的子网仅为子网时， IPv6才支持此选项。
      - Dualstack — 将 IPv4 和 IPv6 地址分配给您的端点网络接口。仅当所有选定的子网同时具有 IPv4 和 IPv6 地址范围时，才支持此选项。
   f. （可选）若要添加标签，请选择 Add new tag（添加新标签），然后输入标签键和标签值。
   g. 选择创建端点。初始状态为 `pending acceptance`。

要接受端点连接请求，请按以下过程操作。

1. 在导航窗格中，选择 Endpoint services（端点服务）。
2. 选择端点服务。
3. 从 Endpoint connections（端点连接）选项卡中，选择端点连接。

4. 要接受连接请求，依次选择 Actions（操作）、Accept endpoint connection request（接受端点连接请求）。提示进行确认时，输入 **accept**，然后选择 Accept（接受）。

第 4 步：配置路由

按如下说明为服务使用者 VPC 配置路由表。这将使安全设备能够对发往应用程序服务器的入站流量执行安全检查。

配置路由

1. 打开位于 https://console.aws.amazon.com/vpc/ 的 Amazon VPC 控制台。
2. 在导航窗格中，选择 Route tables（路由表）。
3. 为互联网网关选择路由表，并执行以下操作：
   a. 依次选择 Actions（操作）、Edit routes（编辑路由）。
   b. 选择 Add route（添加路由）。在目标中，输入应用程序服务器的子网的 IPv4 CIDR 块。在 Target（目标）选项中，选择 VPC 端点。
   c. 如果您支持 IPv6，请选择添加路由。在目标中，输入应用程序服务器的子网的 IPv6 CIDR 块。在 Target（目标）选项中，选择 VPC 端点。
   d. 选择保存更改。
4. 为包含应用程序服务器的子网选择路由表，并执行以下操作：
   a. 依次选择 Actions（操作）、Edit routes（编辑路由）。
   b. 选择 Add route（添加路由）。在 Destination（目标位置）字段，输入 **0.0.0.0/0**。在 Target（目标）选项中，选择 VPC 端点。
   c. 如果您支持 IPv6，请选择添加路由。在 Destination（目标位置）字段，输入 **::/0**。在 Target（目标）选项中，选择 VPC 端点。
   d. 选择保存更改。
5. 为包含网关负载均衡器端点的子网选择路由表，并执行以下操作：
   a. 依次选择 Actions（操作）、Edit routes（编辑路由）。
   b. 选择 Add route（添加路由）。在 Destination（目标位置）字段，输入 **0.0.0.0/0**。在 Target（目标）选项中，选择互联网网关。
   c. 如果您支持 IPv6，请选择添加路由。在目标位置字段，输入 **::/0**。在 Target（目标）选项中，选择互联网网关。

d. 选择保存更改。

开始使用 Gateway 负载均衡器 AWS CLI

借助网关负载均衡器可以轻松部署、扩展和管理第三方虚拟设备，例如安全设备。

在本教程中，我们将使用一个网关负载均衡器和一个网关负载均衡器端点，实现一个检验系统。

内容

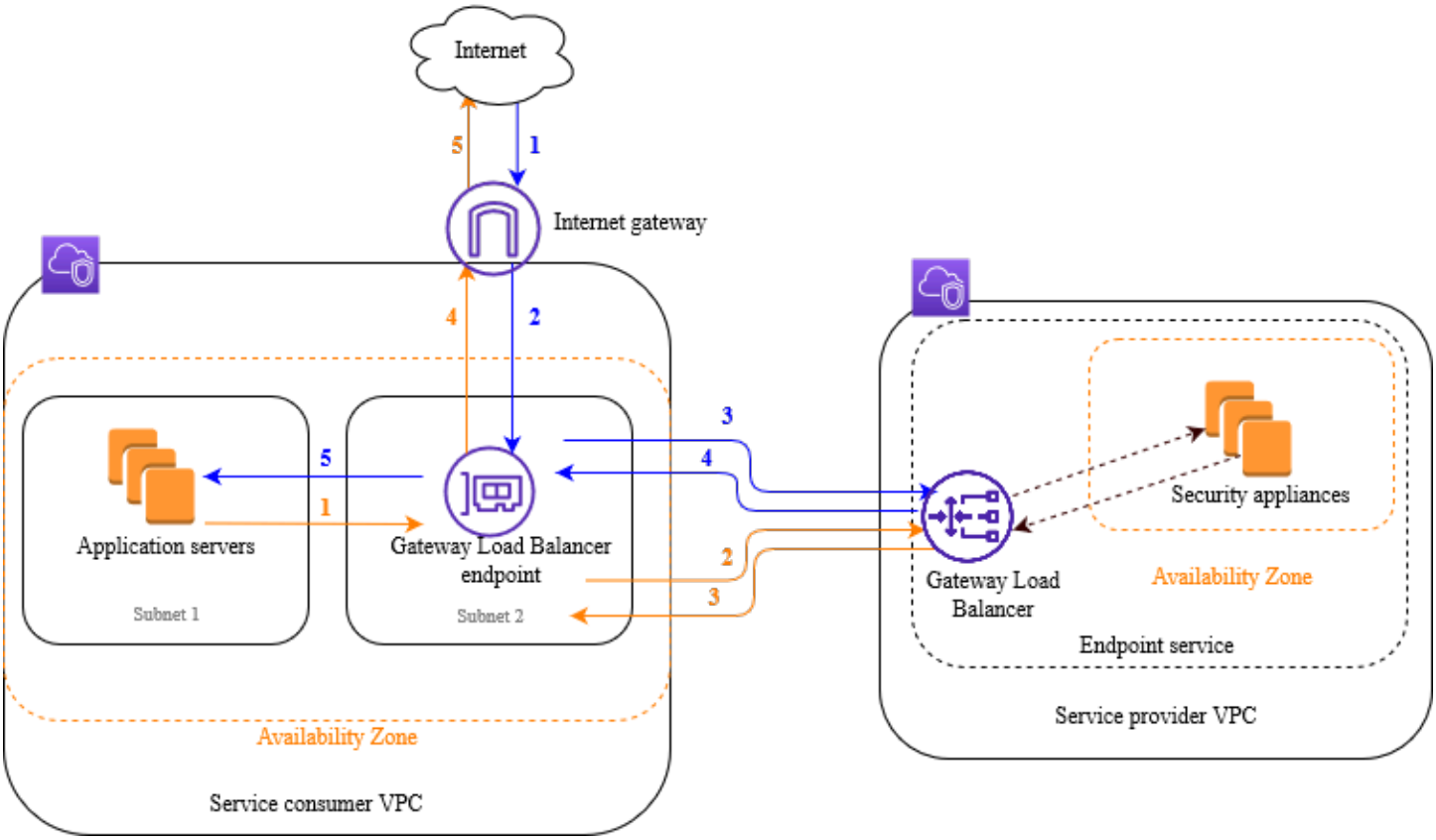
- 概览
- 先决条件
- 第 1 步：创建网关负载均衡器并注册目标
- 第 2 步：创建网关负载均衡器端点
- 步骤 3：配置路由

概览

网关负载均衡器端点是在服务提供者 VPC 中的虚拟设备与服务使用者 VPC 中的应用程序服务器之间提供私有连接的 VPC 端点。网关负载均衡器将部署在与虚拟设备相同的 VPC 中。这些设备将会注册到网关负载均衡器的目标组。

应用程序服务器与服务使用者 VPC 在同一个子网（目的地子网）中运行，而网关负载均衡器端点位于另一个子网中。通过互联网网关进入服务使用者 VPC 的所有流量首先会路由到网关负载均衡器端点，然后再路由到目标子网。

同样，离开应用程序服务器（目的地子网）的所有流量会首先路由到网关负载均衡器端点，然后再路由回互联网。以下网络图形象地演示了如何使用网关负载均衡器访问端点服务。



随后的编号项突出显示并解释了上面网络图中显示的元素。

从互联网到应用程序的流量（蓝色箭头）：

1. 流量通过互联网网关进入服务使用者 VPC。
2. 根据入口路由将流量发送到网关负载均衡器端点。
3. 将流量发送到网关负载均衡器，然后由后者将流量分配到其中的一个安全设备。
4. 安全设备完成检查后，将流量发送回网关负载均衡器端点。
5. 将流量发送到应用程序服务器（目的地子网）。

从应用程序到互联网的流量（橙色箭头）：

1. 根据应用程序服务器子网上配置的默认路由表，将流量发送到网关负载均衡器端点。
2. 将流量发送到网关负载均衡器，然后由后者将流量分配到其中的一个安全设备。
3. 安全设备完成检查后，将流量发送回网关负载均衡器端点。
4. 根据路由表配置，将流量发送到互联网网关。
5. 流量被路由回互联网。

路由

互联网网关的路由表必须具有将发往应用程序服务器的流量路由到网关负载均衡器端点的条目。要指定网关负载均衡器端点，请使用 VPC 端点的 ID。以下示例显示了 dualstack 配置的路由。

目标位置	目标
VPC IPv4 CIDR	本地
VPC IPv6 CIDR	本地
Subnet 1 IPv4 CIDR	*vpc-endpoint-id*
Subnet 1 IPv6 CIDR	*vpc-endpoint-id*

应用程序服务器所在子网的路由表必须具有将来自应用程序服务器的所有流量路由到网关负载均衡器端点的条目。

目标位置	目标
VPC IPv4 CIDR	本地
VPC IPv6 CIDR	本地
0.0.0.0/0	*vpc-endpoint-id*
::/0	*vpc-endpoint-id*

网关负载均衡器端点所在子网的路由表必须将从检查返回的流量路由到最终目的地。对于来自互联网的流量，本地路由将确保其会到达应用程序服务器。对于来自应用程序服务器的流量，则添加会将所有流量路由到互联网网关的条目。

目标位置	目标
VPC IPv4 CIDR	本地
VPC IPv6 CIDR	本地

目标位置	目标
0.0.0.0/0	*internet-gateway-id*
::/0	*internet-gateway-id*

先决条件

- AWS CLI 如果您使用的版本不支持 Gateway Load Balancers，请安装 AWS CLI 或更新到当前版本的。有关更多信息，请参阅AWS Command Line Interface 《用户指南》中的安装 AWS CLI。
- 确保服务使用者 VPC 在包含应用程序服务器的每个可用区中至少有两个子网。其中一个子网用于网关负载均衡器端点，另一个用于应用程序服务器。
- 确保服务提供者 VPC 在包含安全设备实例的每个可用区中至少有两个子网。一个子网用于网关负载均衡器端点，另一个用于实例。
- 在服务提供者 VPC 的每个安全设备子网中启动至少一个安全设备实例。这些实例的安全组必须允许端口 6081 上的 UDP 流量。

第 1 步：创建网关负载均衡器并注册目标

安装以下过程创建负载均衡器、侦听器和目标组，并将安全设备实例注册为目标。

创建网关负载均衡器并注册目标

1. 使用create-load-balancer命令创建类型的负载均衡器gateway。您可以为您启动了安全设备实例的每个可用区指定一个子网。

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

   默认设置为仅支持 IPv4 地址。要同时支持 IPv4 和 IPv6 地址，请添加--ip-address-type dualstack选项。

   输出包含负载均衡器的 Amazon 资源名称（ARN），格式如下例所示：

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

2. 使用create-target-group命令创建目标组，指定您在其中启动实例的服务提供商 VPC。

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

输出包含目标组的 ARN，格式如下。

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

3. 使用 register-targets 命令将您的实例注册到目标组。

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

4. 使用 create-listener 命令为您的负载均衡器创建一个侦听器，该侦听器带有将请求转发到目标组的默认规则。

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

输出包含侦听器的 ARN，格式如下。

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

5. （可选）您可以使用以下describe-target-health命令验证目标组的注册目标的运行状况。

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

第 2 步：创建网关负载均衡器端点

按照以下过程创建网关负载均衡器端点。网关负载均衡器端点是基于可用区的。我们建议您在每个可用区中创建一个网关负载均衡器端点。有关更多信息，请参阅 通过 AWS PrivateLink访问虚拟设备。

要创建网关负载均衡器终端节点

1. 使用 create-vpc-endpoint-service-configuration 命令使用您的 Gateway Load Balancer 创建终端节点服务配置。

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

要同时支持 IPv4 和 IPv6 地址，请添加`--supported-ip-address-types ipv4 ipv6`选项。

输出包含服务 ID（例如 vpce-svc-12345678901234567）和服务名称（例如，com.amazonaws.vpce.us-east-2.vpce-svc-12345678901234567）。

2. 使用 modify-vpc-endpoint-service-permis sions 命令允许服务使用者为您的服务创建终端节点。服务使用者可以是用户、IAM 角色或 AWS 账户。以下示例添加了对指定项的权限 AWS 账户。

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

3. 使用create-vpc-endpoint命令为您的服务创建 Gateway Load Balancer 终端节点。

```
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
```

要同时支持 IPv4 和 IPv6 地址，请添加`--ip-address-type dualstack`选项。

输出包含 网关负载均衡器端点的 ID（例如，vpce-01234567890abcdef）。

步骤 3：配置路由

按如下说明为服务使用者 VPC 配置路由表。这将使安全设备能够对发往应用程序服务器的入站流量执行安全检查。

配置路由

1. 使用 create-route 命令将向互联网网关的路由表添加会将发往应用程序服务器的流量路由到网关负载均衡器端点的条目。

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

如果您支持 IPv6，请添加以下路由。

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

2. 使用 create-route 命令将向应用程序服务器所在子网的路由表添加一条会将来自应用程序服务器的所有流量路由到网关负载均衡器端点的条目。

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

如果您支持 IPv6，请添加以下路由。

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

3. 使用 create-route 命令将向网关负载均衡器端点所在子网的路由表添加一条会将来自应用程序服务器的所有流量路由到互联网网关的条目。

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

如果您支持 IPv6，请添加以下路由。

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

4. 对每个可用区中的每个应用程序子网路由表重复此操作。

Gateway Load Balancer

使用网关负载均衡器部署和管理支持 GENEVE 协议的虚拟设备实例集。

网关负载均衡器在开放系统互联（OSI）模型的第三层运行。它会侦听所有端口上的所有 IP 数据包，并使用 GENEVE 协议通过端口 6081 将流量转发到侦听器规则中指定的目标组。

您可以根据需求变化在负载均衡器中添加或移除目标，而不会中断整体请求流。弹性负载均衡 根据传输到应用程序的流量随时间的变化对负载均衡器进行扩展。弹性负载均衡 能够自动扩展来处理绝大部分工作负载。

内容

- 负载均衡器状态
- IP 地址类型
- 可用区
- 空闲超时
- 负载均衡器属性
- 网络 ACLs
- 非对称流量
- 网络最大传输单元（MTU）
- 创建网关负载均衡器
- 更新网关负载均衡器的 IP 地址类型
- 编辑网关负载均衡器的属性
- 为网关负载均衡器添加标签
- 删除网关负载均衡器
- 为您的网关 Load Balancer 预留容量

负载均衡器状态

网关负载均衡器的可能状态如下：

`provisioning`

正在设置网关负载均衡器。

`active`

已完全设置网关负载均衡器并可立即路由流量。

`failed`

无法设置网关负载均衡器。

IP 地址类型

您可以设置应用程序服务器可用于访问网关负载均衡器的 IP 地址类型。

网关负载均衡器支持以下 IP 地址类型：

`ipv4`

IPv4 仅支持。

`dualstack`

两者 IPv6 都支持 IPv4 和。

注意事项

- 您为负载均衡器指定的虚拟私有云 (VPC) 和子网必须具有关联的 IPv6 CIDR 块。
- 服务使用者 VPC 中子网的路由表必须路由 IPv6 流量，并且这些子网 ACLs 的网络必须允许 IPv6 流量。
- Gateway Load Balanc IPv4 er 使用 IPv4 GENEVE 标头封装 IPv6 客户端流量并将其发送到设备。设备使用 IPv4 GENEVE 标头封装 IPv4 IPv6 客户端流量并将其发送回 Gateway Load Balancer。

有关 IP 地址类型的更多信息，请参阅更新网关负载均衡器的 IP 地址类型。

可用区

创建网关负载均衡器时，您可以启用一个或多个可用区，并指定与每个可用区对应的子网。如果启用多个可用区，将可以确保即使某个可用区不可用，负载均衡器也可以继续路由流量。您指定的每个子网必须具有至少 8 个可用 IP 地址。创建负载均衡器之后，将无法移除子网。要移除子网，必须创建新的负载均衡器。

空闲超时

对于通过网关负载均衡器发出的每个 TCP 请求，都将跟踪该连接的状态。如果客户端或目标通过连接发送数据的间隔超过空闲超时期限，则连接将关闭。空闲超时期限过后，负载均衡器会考虑将下一个 TCP SYN 作为新的流量，并将其路由到新的目标。但是，在空闲超时期限过后发送的数据包会被丢弃。

TCP 流量的默认空闲超时值为 350 秒，但可以更新为 60-6000 秒之间的任何值。客户端或目标可以使用 TCP keepalive 数据包重置空闲超时值。

> **粘性限制**
>
> 只有在使用 5 元组粘性时，才能更新您的网关负载均衡器空闲超时。使用 3 元组或 2 元组粘性时，将使用默认的空闲超时值。有关更多信息，请参阅 流量粘性。

虽然 UDP 无连接，但是负载均衡器将根据源和目标 IP 地址和端口保持 UDP 流状态。这可确保属于同一个流中的数据包始终发送到相同的目标。空闲超时期限后，负载均衡器会考虑将传入的 UDP 数据包作为新流，并路由到新的目标。Elastic Load Balancing 将 UDP 流的空闲超时值设置为 120 秒。无法对其进行更改。

EC2 实例必须在 30 秒内响应新请求才能建立返回路径。

有关更多信息，请参阅 更新空闲超时。

负载均衡器属性

网关负载均衡器的负载均衡器属性如下：

`deletion_protection.enabled`

指示是否启用删除保护。默认为 `false`。

`load_balancing.cross_zone.enabled`

指示是否启用了跨区域负载均衡。默认值为 `false`。

有关更多信息，请参阅 编辑负载均衡器属性。

网络 ACLs

如果应用程序服务器和网关负载均衡器端点在同一个子网中，则会对从应用程序服务器到网关负载均衡器端点的流量评估 NACL 规则。

非对称流量

当负载均衡器处理初始流量数据包并且响应流数据包未通过负载均衡器路由时，网关负载均衡器支持非对称流量。不建议使用非对称路由，因为这会导致网络性能降低。当负载均衡器不处理初始流量数据包但响应流数据包通过负载均衡器路由时，网关负载均衡器不支持非对称流量。

网络最大传输单元（MTU）

最大传输单元（MTU）是能够通过网络传输的最大数据包大小。网关负载均衡器接口 MTU 支持最大 8500 字节的数据包。到达网关负载均衡器的数据包如果超过 8500 字节，则将被丢弃。

网关负载均衡器使用 GENEVE 标头封装 IP 流量并将其转发到设备。GENEVE 封装过程在原始数据包的基础上增加 68 个字节。因此，要支持最大 8,500 字节的数据包，请确保设备的 MTU 设置支持至少 8,568 字节的数据包。

网关负载均衡器不支持 IP 分段。此外，网关负载均衡器不会生成 ICMP 消息“无法到达目的地：需要分段和设置 DF”。由于这一原因，不支持路径 MTU 发现（PMTUD）。

创建网关负载均衡器

Gateway Load Balancer 接受来自客户端的请求，并将其分配到目标组中的目标（例如 EC2 实例）。

要使用创建 Gateway Load Balancer AWS Management Console，请完成以下任务。或者，要使用创建 Gateway Load Balancer AWS CLI，请参阅开始使用 CLI。

任务

- 先决条件
- 创建负载均衡器
- 重要后续步骤

先决条件

在开始之前，请确保网关负载均衡器的虚拟私有云（VPC）在目标所在的每个可用区中至少有一个子网。

创建负载均衡器

按照以下过程创建网关负载均衡器。提供负载均衡器的基本配置信息，例如名称和 IP 地址类型。然后提供有关网络以及要将流量路由到目标组的侦听器的信息。网关负载均衡器要求目标组使用 GENEVE 协议。

使用控制台创建负载均衡器和侦听器

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中的负载平衡下，选择负载均衡器。
3. 选择创建负载均衡器。
4. 在网关负载均衡器下，选择创建。
5. 基本配置
   a. 对于Load balancer name（负载均衡器名称），输入负载均衡器的名称。例如：**my-glb**。网关负载均衡器的名称在该区域的所有负载均衡器中必须唯一。名称最多可包含 32 个字符，只能包含字母数字字符和连字符，并且不能以连字符开头或结尾。
   b. 对于 IP 地址类型，请选择IPv4仅支持 IPv4 地址，或者选择 Dualstack 以同时支持 IPv4 和 IPv6地址。
6. 网络映射
   a. 对于 VPC，请选择服务提供者 VPC。
   b. 对于映射，请选择您启动了安全设备实例的所有可用区，以及对应的公有子网。
7. IP 侦听器路由
   a. 对于默认操作，选择用来接收流量的目标组。如果您没有目标组，请选择创建目标组。有关更多信息，请参阅 创建目标组。
   b. （可选）展开侦听器标签并添加所需的标签。
8. （可选）展开负载均衡器标签并添加所需的标签。
9. 检查配置，然后选择创建负载均衡器。

重要后续步骤

创建负载均衡器后，请验证您的 EC2 实例是否已通过初始运行状况检查。要测试负载均衡器，您必须创建一个网关负载均衡器端点，并更新路由表以将该网关负载均衡器端点设置为下一个跃点。这些配置可在 Amazon VPC 控制台中设置。有关更多信息，请参阅入门教程。

更新网关负载均衡器的 IP 地址类型

您可以配置 Gateway Load Balancer，以便应用程序服务器只能使用地址访问您的负载均衡器，也可以同时使用 IPv4 IPv6 地址 IPv4 和地址（双堆栈）访问您的负载均衡器。负载均衡器根据目标组的 IP 地址类型与目标进行通信。有关更多信息，请参阅 IP 地址类型。

使用控制台更新 IP 地址类型

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的 Load Balancing（负载均衡）下，选择 Load Balancers（负载均衡器）。
3. 选择负载均衡器。
4. 选择 Actions (操作) 和 Edit IP address type (编辑 IP 地址类型)。
5. 对于 IP 地址类型，选择 ip v4 以仅支持 IPv4 地址，或选择 dualstack 以同时支持和 IPv4 地址。IPv6
6. 选择保存。

要更新 IP 地址，请使用 AWS CLI

使用 set-ip-address-type 命令。

编辑网关负载均衡器的属性

创建网关负载均衡器后，可以编辑其负载均衡器属性。

负载均衡器属性

- 删除保护
- 跨可用区负载均衡

删除保护

为防止您的网关负载均衡器被意外删除，您可以启用删除保护。默认情况下，将禁用删除保护。

如果您为网关负载均衡器启用了删除保护，则必须先禁用删除保护，然后才能删除网关负载均衡器。

使用控制台启用删除保护

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中的 Load Balancing (负载平衡) 下，选择 Load Balancers (负载均衡器)。
3. 选择网关负载均衡器。
4. 选择操作、编辑属性。
5. 在编辑负载均衡器属性页面上，为删除保护选择启用，然后选择保存。

使用控制台禁用删除保护

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中的 Load Balancing (负载平衡) 下，选择 Load Balancers (负载均衡器)。
3. 选择网关负载均衡器。
4. 选择操作、编辑属性。
5. 在编辑负载均衡器属性页面上，为删除保护清除启用，然后选择保存。

要启用或禁用删除保护，请使用 AWS CLI

使用带 `deletion_protection.enabled` 属性的 modify-load-balancer-attributes 命令。

跨可用区负载均衡

默认情况下，每个负载均衡器节点仅在其可用区中的已注册目标之间分配流量。如果您启用了跨可用区负载均衡，则每个网关负载均衡器节点都会在所有启用的可用区中的已注册目标之间分配流量。有关更多信息，请参阅 Elastic Load Balancing 用户指南中的跨可用区负载均衡。

使用控制台启用跨可用区负载均衡

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中的 Load Balancing (负载平衡) 下，选择 Load Balancers (负载均衡器)。
3. 选择网关负载均衡器。

4. 选择操作、编辑属性。
5. 在编辑负载均衡器属性页面上，对于跨可用区负载均衡请选择启用，然后选择保存。

要启用跨区域负载均衡，请使用 AWS CLI

使用带 `load_balancing.cross_zone.enabled` 属性的 modify-load-balancer-attributes 命令。

为网关负载均衡器添加标签

使用标签可帮助您按各种标准对负载均衡器进行分类，例如按用途、所有者或环境。

您最多可以为每个负载均衡器添加多个标签。每个网关负载均衡器的标签键必须唯一。如果您添加的标签中的键已经与负载均衡器关联，它将更新该标签的值。

使用完标签后，您可以将其从网关负载均衡器中移除。

限制

- 每个资源的标签数上限 – 50
- 最大密钥长度 - 127 个 Unicode 字符
- 最大值长度 - 255 个 Unicode 字符
- 标签键和值区分大小写。允许使用的字符包括可用 UTF-8 格式表示的字母、空格和数字，以及以下特殊字符： + - = 。 _ : / @。请不要使用前导空格或尾随空格。
- 请勿在标签名称或值中使用`aws:`前缀，因为它已保留供 AWS 使用。您无法编辑或删除带此前缀的标签名称或值。具有此前缀的标签不计入每个资源的标签数限制。

使用控制台更新网关负载均衡器的标签

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中的 Load Balancing (负载平衡) 下，选择 Load Balancers (负载均衡器)。
3. 选择网关负载均衡器。
4. 选择 Tags、Add/Edit Tags，然后执行下列一个或多个操作：
   a. 要更新标签，请编辑 Key 和 Value 的值。
   b. 要添加新标签，请选择 Create Tag。对于 Key（键）和 Value（值），输入值。
   c. 要删除标签，请选择标签旁边的删除图标 (X)。

5. 完成更新标签后，选择保存。

要更新 Gateway Load Balancer 的标签，请使用 AWS CLI

使用 add-tags 和 remove-tags 命令。

删除网关负载均衡器

网关负载均衡器变为可用之后，您需要按其保持运行的小时数或部分小时数付费。当您不再需要该网关负载均衡器时，可将其删除。当网关负载均衡器被删除之后，您便不再需要为其付费。

如果其他服务正在使用网关负载均衡器，则无法将其删除。例如，假设该网关负载均衡器与某个 VPC 端点服务关联，则必须先删除端点服务配置，然后才能删除关联的网关负载均衡器。

删除网关负载均衡器也将删除其侦听器。删除网关负载均衡器不会影响其已注册的目标。例如，您的 EC2 实例会继续运行，并且仍会注册到其目标组。要删除目标组，请参阅删除网关负载均衡器的目标组。

使用控制台删除网关负载均衡器

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中的 Load Balancing (负载平衡) 下，选择 Load Balancers (负载均衡器)。
3. 选择网关负载均衡器。
4. 依次选择 Actions（操作）和 Delete（删除）。
5. 当系统提示进行确认时，选择 Yes, Delete（是，删除）。

要使用 Gateway Load Balancer 删除 AWS CLI

使用 delete-load-balancer 命令。

为您的网关 Load Balancer 预留容量

负载均衡器容量单位 (LCU) 预留允许您为负载均衡器预留静态的最小容量。网关负载均衡器会自动扩展以支持检测到的工作负载并满足容量需求。配置最低容量后，您的负载均衡器会根据收到的流量继续向上或向下扩展，但也可以防止容量低于配置的最小容量。

在以下情况下，可以考虑使用 LCU 预留：

- 您即将举办一个活动，该活动会突然出现异常的高流量，并希望确保您的负载均衡器能够支持活动期间的突然流量激增。
- 由于工作量的性质，您在短时间内会遇到不可预测的尖峰流量。
- 您正在将负载均衡器设置为在特定的开始时间加入或迁移服务，并且需要从高容量开始，而不是等待 auto-scaling 生效。
- 您需要保持最低容量以满足服务级别协议或合规性要求。
- 您正在负载均衡器之间迁移工作负载，并希望配置目标以匹配源的规模。

估算您需要的容量

在确定应为负载均衡器预留的容量时，我们建议您执行负载测试或查看代表预期即将到来的流量的历史工作负载数据。使用 Elastic Load Balancing 控制台，您可以根据审核的流量估算需要预留多少容量。

或者，您可以参考 CloudWatch 指标ProcessedBytes来确定正确的容量级别。您的负载均衡器的容量已预留在中 LCUs，每个 LCU 等于 2.2Mbps。您可以使用该PeakBytesPerSecond指标来查看负载均衡器上每分钟的最大吞吐量流量，然后将该吞吐量转换为 LCUs 使用 2.2Mbps 等于 1 LCU 的转换率。

如果您没有历史工作负载数据可供参考且无法执行负载测试，则可以使用 LCU 预留计算器估算所需的容量。LCU 预留计算器使用基于 AWS 观察到的历史工作负载的数据，可能无法代表您的特定工作负载。有关更多信息，请参阅 Load Balancer 容量单位预留计算器。

LCU 预留配额

您的账户有与之相关的配额 LCUs。有关更多信息，请参阅 the section called “Load Balancer 容量单位”。

为您的网关 Load Balancer 申请负载均衡器容量单位预留

在使用 LCU 预留之前，请查看以下内容：

- LCU 预留仅支持为网关负载均衡器预留吞吐容量。申请 LCU 预留时，请将您的容量需求从 Mbps 转换为 LCUs 使用 1 LCU 到 2.2 Mbps 的转换速率。
- 容量在区域层面预留，在可用区之间均匀分配。在启用 LCU 预留之前，请确认每个可用区中都有足够的均匀分布目标。
- LCU 预订请求按先到先得的原则满足，具体取决于当时某个区域的可用容量。大多数请求通常在一小时内完成，但最多可能需要几个小时。
- 要更新现有预留，必须配置之前的请求或失败。您可以根据需要多次增加预留容量，但是每天只能减少两次预留容量。

申请 LCU 预约

此过程中的步骤说明了如何在负载均衡器上申请 LCU 预留。

使用控制台申请 LCU 预留

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中，选择负载均衡器。
3. 选择负载均衡器名称。
4. 在容量选项卡上，选择编辑 LCU 预留。
5. 选择 “基于历史参考的估计”，然后从下拉列表中选择负载均衡器。
6. 选择参考时段以查看推荐的预留 LCU 水平。
7. 如果您没有历史参考工作量，则可以选择手动估算并输入 LCUs 要保留的数量。
8. 选择保存。

要申请 LCU 预订，请使用 AWS CLI

使用 modify-capacity-reservation 命令。

更新或终止 Gateway Load Balancer 的负载均衡器容量单位预留

更新或终止 LCU 预留

此过程中的步骤说明了如何更新或终止负载均衡器上的 LCU 预留。

使用控制台更新或终止 LCU 预留

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中，选择负载均衡器。
3. 选择负载均衡器名称。
4. 在 “容量” 选项卡上，确认预留状态为 “已配置”。
    a. 要更新 LCU 预留，请选择编辑 LCU 预留。
    b. 要终止 LCU 预留，请选择 “取消容量”。

要更新或终止 LCU 预留，请使用 AWS CLI

使用 modify-capacity-reservation 命令。

监控您的 Gateway Load Balancer 的负载均衡器容量单位

预订状态

LCU 预留有四种可用状态：

- pending-表示它正在置备的预留。
- 已配置-表示预留容量已准备就绪，可供使用。
- failed-表示当时无法完成请求。
- rebalancing-表示已添加可用区，且负载均衡器正在重新平衡容量。

预留 LCU

要确定预留 LCU 利用率，您可以将每分钟PeakBytesPerSecond指标与每小时总和（预留）进行比较。LCUs要将每分钟字节数转换为每小时 LCU，请使用（每分钟字节数）*8/60/ (10^6) /2.2。

监控预留容量

此过程中的步骤说明了如何检查负载均衡器上的 LCU 预留状态。

使用控制台查看 LCU 预留的状态

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中，选择负载均衡器。
3. 选择负载均衡器名称。
4. 在容量选项卡上，您可以查看预留状态和预留 LCU 值。

使用监控 LCU 预留的状态 AWS CLI

使用 describe-capacity-reservation 命令。

网关负载均衡器的侦听器

在创建网关负载均衡器时，您需要添加一个侦听器。侦听器是用于检查连接请求的进程。

网关负载均衡器的侦听器用于侦听所有端口上的所有 IP 数据包。不能在为网关负载均衡器创建侦听器时指定协议或端口。

在创建侦听器时，将会指定用于路由请求的规则。该规则将请求转发到指定的目标组。您可以更新侦听器规则，以将请求转发到其他目标组。

侦听器属性

网关负载均衡器的侦听器属性如下：

`tcp.idle_timeout.seconds`

tcp 空闲超时值（以秒为单位）。有效范围为 60-6000 秒。默认值为 350 秒。

有关更多信息，请参阅 更新空闲超时。

更新网关负载均衡器侦听器的目标组

在创建侦听器时，将会指定用于路由请求的规则。该规则将请求转发到指定的目标组。您可以更新侦听器规则，以将请求转发到其他目标组。

使用控制台更新侦听器

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中的 Load Balancing (负载平衡) 下，选择 Load Balancers (负载均衡器)。
3. 选择负载均衡器，然后选择 Listeners。
4. 选择编辑侦听器。
5. 对于转发到目标组，选择一个目标组。
6. 选择保存。

要更新您的听众，请使用 AWS CLI

使用 modify-listener 命令。

更新网关负载均衡器侦听器的 TCP 空闲超时

对于通过网关负载均衡器发出的每个 TCP 请求，都将跟踪该连接的状态。如果客户端或目标通过连接发送数据的间隔超过空闲超时期限，则连接将关闭。TCP 流量的默认空闲超时值为 350 秒，但可以更新为 60-6000 秒之间的任何值。

使用控制台更新 TCP 空闲超时

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格中的 Load Balancing (负载平衡) 下，选择 Load Balancers (负载均衡器)。
3. 选择网关负载均衡器。
4. 在侦听器选项卡上，依次选择操作和查看侦听器详细信息。
5. 在侦听器详细信息页面的属性选项卡中，选择编辑。
6. 在编辑侦听器属性页的侦听器属性部分中，输入 TCP 空闲超时的值。
7. 选择 Save changes（保存更改）

要更新 TCP 空闲超时，请使用 AWS CLI

使用带有`tcp.idle_timeout.seconds`属性的modify-listener-attributes命令。

网关负载均衡器的目标组

每个目标组均用于将请求路由到一个或多个已注册的目标。创建侦听器时，您为其默认操作指定目标组。流量将转发到在侦听器规则中指定的目标组。您可以为不同类型的请求创建不同的目标组。

您按目标组定义的网关负载均衡器运行状况检查设置。每个目标组均使用默认运行状况检查设置，除非您在创建目标组时将其覆盖或稍后对其进行修改。在侦听器规则中指定一个目标组后，对于已为网关负载均衡器启用的可用区中的目标组，网关负载均衡器将持续监控已注册到该目标组的所有目标的运行状况。网关负载均衡器将请求路由到运行正常的已注册目标。有关更多信息，请参阅 网关负载均衡器目标组的运行状况检查。

目录

- 路由配置
- Target type
- 已注册目标
- 目标组属性
- 为网关负载均衡器创建目标组
- 网关负载均衡器目标组的运行状况检查
- 编辑网关负载均衡器的目标组属性
- 为网关负载均衡器注册目标
- 为网关负载均衡器的目标组添加标签
- 删除网关负载均衡器的目标组

路由配置

网关负载均衡器的目标组支持以下协议和端口：

- 协议：GENEVE
- 端口：6081

Gateway Load Balancer 使用 GENEVE 封装原始数据包。GENEVE 标头使用 Type-Length-Value (TLV) 格式存储信息，使用选项类 0x0108。设备必须解封 TLV 对才能处理原始数据包。有关更多信息，请参阅以下博客文章：将您的设备与网关负载均衡器集成。

Target type

在创建目标组时，应指定其目标类型，这决定您如何指定其目标。创建目标组后，将无法更改其目标类型。

以下是可能的目标类型：

`instance`

这些目标通过实例 ID 指定。

`ip`

这些目标通过 IP 地址指定。

当目标类型为 `ip` 时，您可以指定来自以下 CIDR 块之一的 IP 地址：

- 目标组的 VPC 的子网
- 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)

> **Important**
>
> 不能指定可公开路由的 IP 地址。

已注册目标

您的网关负载均衡器充当客户端的单一接触点，并跨其运行正常的已注册目标分配传入流量。每个目标组在为网关负载均衡器启用的每个可用区中必须至少有一个已注册目标。您可以将每个目标注册到一个或多个目标组中。

如果需求增加，您可以向一个或多个目标组注册其他目标以便满足该需求。注册过程完成后，网关负载均衡器会立即开始将流量路由到新注册的目标。

如果需求减少或者您需要为目标提供服务，您可以从目标组中注销目标。取消注册目标将从目标组中删除目标，但不会影响目标。注销目标后，网关负载均衡器会立即停止将流量路由到目标。目标将进入 draining 状态，直至进行中请求完成。当您准备好恢复接收流量时，可以再次向目标组注册目标。

目标组属性

您可以对目标组使用以下属性：

`deregistration_delay.timeout_seconds`

Elastic Load Balancing 在将取消注册目标的状态从 `draining` 更改为 unused 之前需等待的时间。范围为 0-3600 秒。默认值为 300 秒。

`stickiness.enabled`

指示是否为目标组起用了可配置的流量粘性。可能的值为 `true` 或 `false`。默认值为 false。当该属性设置为 `false` 时，将使用 5_tuple。

`stickiness.type`

指示流量粘性的类型。对于与网关负载均衡器关联的目标组，可能的值为：

- `source_ip_dest_ip`
- `source_ip_dest_ip_proto`

`target_failover.on_deregistration`

指示当注销某个目标时，网关负载均衡器将如何处理现有的流量。可能的值为 `rebalance` 和 `no_rebalance`。默认为 `no_rebalance`。这两项属性（`target_failover.on_deregistration` 和 `target_failover.on_unhealthy`）不可分别设置。您必须为这两项属性设置相同的值。

`target_failover.on_unhealthy`

指示当某个目标运行不正常时，网关负载均衡器将如何处理现有的流量。可能的值为 `rebalance` 和 `no_rebalance`。默认为 `no_rebalance`。这两项属性（`target_failover.on_deregistration` 和 `target_failover.on_unhealthy`）不可分别设置。您必须为这两项属性设置相同的值。

有关更多信息，请参阅 编辑目标组属性。

为网关负载均衡器创建目标组

使用目标组为网关负载均衡器注册目标。

要将流量路由到目标组中的目标，请创建侦听器，并在侦听器的默认操作中指定目标组。有关更多信息，请参阅 侦听器。

您可以随时在目标组中添加或删除目标。有关更多信息，请参阅 注册目标。您也可以修改目标组的运行状况检查设置。有关更多信息，请参阅 修改运行状况检查设置。

使用控制台创建目标组

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的负载均衡下，选择目标组。
3. 选择 Create target group (创建目标组)。
4. 基本配置
   a. 对于选择目标类型，选择实例以按实例 ID 指定目标，或选择 IP 地址以按 IP 地址指定目标。
   b. 对于 Target group name (目标组名称)，输入目标组的名称。此名称在每个区域的每个账户中必须唯一，最多可以有 32 个字符，只能包含字母数字字符或连字符，不得以连字符开头或结尾。
   c. 验证协议是否为 GENEVE，端口是否为 6081。不支持任何其他协议或端口。
   d. 对于 VPC，选择具有要包含在目标组中的安全设备实例的虚拟私有云（VPC）。
5. （可选）对于运行状况检查，请根据需要修改设置和高级设置。如果运行状况检查连续超过不正常运行阈值计数，负载均衡器将使目标停止服务。如果运行状况检查连续超过运行状况正常阈值计数，负载均衡器将使目标恢复使用。有关更多信息，请参阅 网关负载均衡器目标组的运行状况检查。
6. （可选）展开标签并添加您需要的标签。
7. 选择下一步。
8. 对于注册目标，按如下方式添加一个或多个目标：
   - 如果目标类型为实例，请选择一个或多个实例，输入一个或多个端口，然后选择在下面以待注册的形式添加。
   - 如果目标类型为 IP addresses (IP 地址)，请选择网络，输入 IP 地址和端口，然后选择 Include as pending below (在下面以待注册的形式添加)。
9. 选择 Create target group。

要创建目标群组，请使用 AWS CLI

使用create-target-group命令创建目标组，使用 add -t ags 命令标记目标组，使用 register-targets 命令添加目标。

网关负载均衡器目标组的运行状况检查

您可以将目标注册到一个或多个目标组中。注册过程完成后，网关负载均衡器会立即开始将请求路由到新注册的目标。完成注册过程和开始运行状况检查可能需要几分钟时间。

网关负载均衡器会定期向每个已注册的目标发送请求以检查其状态。在完成每次运行状况检查后，网关负载均衡器将关闭为运行状况检查而建立的连接。

运行状况检查设置

您可以使用以下设置为目标组中的目标配置主动运行状况检查。如果运行状况检查超过指定的UnhealthyThresholdCount连续失败次数，则 Gateway Load Balancer 会使目标停止服务。当运行状况检查超过指定的HealthyThresholdCount连续成功次数时，Gateway Load Balancer 会将目标重新投入使用。

设置	描述
HealthCheckProtocol	对目标执行运行状况检查时负载均衡器使用的协议。可能的协议有 HTTP、HTTPS 和 TCP。默认值为 TCP。
HealthCheckPort	对目标执行运行状况检查时网关负载均衡器使用的端口。范围为 1 至 65535。默认值为 80。
HealthCheckPath	[HTTP/HTTPS 运行状况检查] 进行运行状况检查的目标上的目的地的运行状况检查路径。默认值为 /。
HealthCheckTimeoutSeconds	以秒为单位的时间长度，在此期间内，没有来自目标的响应意味着无法通过运行状况检查。范围为 2 至 120。默认值为 5。
HealthCheckIntervalSeconds	各个目标的运行状况检查之间的大约时间量(以秒为单位)。范围为 5 至 300。默认值为 10

设置	描述
	秒。此值必须大于或等于HealthCheckTimeout Seconds。 **Important** 网关负载均衡器的运行状况检查是分布式的，使用共识机制来确定目标运行状况。因此，预计目标设备将在配置的时间间隔内收到几次运行状况检查。
HealthyThresholdCount	将不正常目标视为正常运行之前所需的连续运行状况检查成功次数。范围为 2 至 10。默认值为 5。
UnhealthyThresholdCount	将目标视为不正常之前所需的连续运行状况检查失败次数。范围为 2 至 10。默认值为 2。
Matcher	[HTTP/HTTPS 运行状况检查] 检查来自目标的成功响应时使用的 HTTP 代码。该值必须为 200-399。

目标运行状况

您必须首先将目标注册到目标组，在侦听器规则中指定其目标组，并确保已为网关负载均衡器启用目标所在的可用区，然后网关负载均衡器才会向目标发送运行状况检查请求。

下表描述已注册目标的正常状态的可能值。

值	描述
`initial`	网关负载均衡器正在注册目标或正在对目标执行初始运行状况检查。 相关原因代码：`Elb.RegistrationInProgress` | `Elb.InitialHealthChecking`

值	描述
healthy	目标正常。 相关原因代码：无
unhealthy	目标未响应运行状况检查或未通过运行状况检查。 相关原因代码：Target.FailedHealthChecks
unused	目标未注册到目标组，侦听器规则中未使用目标组，或者目标在没有启用的可用区中，或者目标处于停止或终止状态。 相关原因代码：Target.NotRegistered |Target.NotInUse |Target.InvalidState |Target.IpUnusable
draining	目标正在取消注册，连接即将耗尽。 相关原因代码：Target.DeregistrationInProgress
unavailable	目标运行状况不可用。 相关原因代码：Elb.InternalError

运行状况检查原因代码

如果目标的状态是 Healthy 以外的任何值，则 API 将返回问题的原因代码和描述，并且控制台将显示相同的描述。以 Elb 开头的原因代码源自网关负载均衡器端，以 Target 开头的原因代码源自目标端。

原因代码	说明
Elb.InitialHealthChecking	正在进行初始运行状况检查
Elb.InternalError	由于内部错误，运行状况检查失败

原因代码	说明
Elb.RegistrationInProgress	目标注册正在进行中
Target.DeregistrationInProgress	目标取消注册正在进行中
Target.FailedHealthChecks	运行状况检查失败
Target.InvalidState	目标处于停止状态 目标处于终止状态 目标处于终止或停止状态 目标处于无效状态
Target.IpUnusable	该 IP 地址正被负载均衡器使用，因此无法用作目标
Target.NotInUse	没有将目标组配置为接收来自网关负载均衡器的流量 目标位于尚未为网关负载均衡器启用的可用区
Target.NotRegistered	目标未注册到目标组

网关负载均衡器目标故障场景

现有流量：默认情况下，无论目标的运行状况和注册状态如何，除非流量超时或被重置，否则现有流量会流向同一个目标。这种方法有助于连接耗尽，并且可以容纳有时由于 CPU 使用率过高而无法响应运行状况检查的第三方防火墙。有关更多信息，请参阅 Target failover。

新流量：新流量将发送到运行正常的目标。在对流量做出负载均衡决策后，即使该目标运行不正常或其他目标变为运行正常，网关负载均衡器也会将流量发送到同一个目标。

当所有目标都运行不正常时，网关负载均衡器会随机选择一个目标，并在流量生命周期内将流量转发给该目标，直到该目标被重置或超时为止。由于流量被转发到运行不正常的目标，因此流量会被丢弃，直到该目标恢复正常为止。

TLS 1.3：如果目标组配置了 HTTPS 运行状况检查，则如果其注册目标仅支持 TLS 1.3，则无法通过运行状况检查。这些目标必须支持 TLS 的早期版本，例如 TLS 1.2。

跨可用区负载均衡：默认情况下，跨可用区负载均衡处于禁用状态。如果启用跨可用区负载均衡，则每个网关负载均衡器都能看到所有可用区中的所有目标，并且无论位于哪个可用区，这些目标都将受到同等对待。

可用区之间的负载均衡和运行状况检查决策始终是独立的。即使启用了跨可用区负载均衡，现有流量和新流量的行为也与上述相同。有关更多信息，请参阅 Elastic Load Balancing 用户指南中的跨可用区负载均衡。

检查目标的运行状况

您可以检查已注册到目标组的目标的运行状况。

使用控制台检查目标的运行状况

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的负载均衡下，选择目标组。
3. 选择目标组的名称以打开其详细信息页面。
4. 在 Targets (目标) 选项卡上，Status (状态) 列指示每个目标的状态。
5. 如果目标状态是 `Healthy` 以外的任何值，则 Status details (状态详细信息) 列将包含更多信息。

要检查目标的生命值，请使用 AWS CLI

使用 describe-target-health 命令。此命令的输出包含目标运行状况。如果状态是 `Healthy` 以外的任何值，则它包括原因代码。

接收有关运行状况不佳的目标的电子邮件通知

使用 CloudWatch 警报触发 Lambda 函数以发送有关不健康目标的详细信息。有关 step-by-step说明，请参阅以下博客文章：识别负载均衡器的运行状况不佳的目标。

修改运行状况检查设置

您可以修改目标组的部分运行状况检查设置。

使用控制台修改目标组的运行状况检查设置

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。

2. 在导航窗格上的负载均衡下，选择目标组。
3. 选择目标组的名称以打开其详细信息页面。
4. 在组详细信息选项卡的运行状况检查设置部分中，选择编辑。
5. 在 Edit health check settings (编辑运行状况检查设置) 页面上，根据需要修改设置，然后选择 Save changes (保存更改)。

要修改目标群体的健康检查设置，请使用 AWS CLI

使用 modify-target-group 命令。

编辑网关负载均衡器的目标组属性

为网关负载均衡器创建目标组后，可以编辑其目标组属性。

目标组属性

- 目标失效转移
- 取消注册延迟
- 流量粘性

目标失效转移

借助目标失效转移功能，您可以指定当目标运行不正常或注销时，网关负载均衡器将如何处理现有的流量。默认情况下，即使目标未通过运行状况检查或已注销，网关负载均衡器仍会继续将现有流量发送到同一个目标。您可以通过重新哈希处理这些流量（`rebalance`）或将其保留为默认状态（`no_rebalance`）来管理这些流量。

无再平衡：

网关负载均衡器继续将现有流量发送到未通过运行状况检查或耗尽的目标。如果网关负载均衡器无法到达目标，则流量将被丢弃。

但新流量会发送到运行正常的目标。这是默认行为。

再平衡：

网关负载均衡器会重新哈希现有流量，并在注销延迟超时到期后将其发送到运行正常的目标。

对于已注销的目标，失效转移的最短时间将取决于注销延迟。在注销延迟到期之前，目标不会被标记为已注销。

对于运行不正常的目标，失效转移的最短时间将取决于目标组的运行状况检查配置（间隔时间阈值）。这是目标在被标记为运行不正常前将经过的最短时间。超过此时间后，由于需要额外的传播时间和 TCP 重传回退，网关负载均衡器可能需要在几分钟后才能将新流量重新路由到运行正常的目标。

使用控制台更新目标失效转移属性

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的负载均衡下，选择目标组。
3. 选择目标组的名称以打开其详细信息页面。
4. 在组详细信息页面的属性部分中，选择编辑。
5. 在编辑属性页面上，根据需要更改失效转移的值。
6. 选择保存更改。

要更新目标故障转移属性，请使用 AWS CLI

使用带有以下键值对的modify-target-group-attributes命令：

- 键 = `target_failover.on_deregistration`，值 = `no_rebalance`（默认）或 `rebalance`
- 键 = `target_failover.on_unhealthy`，值 = `no_rebalance`（默认）或 `rebalance`

> **Note**
>
> 这两个属性（`target_failover.on_deregistration` 和 `target_failover.on_unhealthy`）的值必须具有相同。

取消注册延迟

当注销某个目标时，网关负载均衡器会按如下方式管理流向该目标的流量：

新流量

网关负载均衡器将停止发送新流量。

现有流量

网关负载均衡器按照协议来处理现有流量：

- TCP：如果现有流量的空闲时间超过 350 秒，则会将其关闭。
- 其他协议：如果现有流量的空闲时间超过 120 秒，则会将其关闭。

为帮助耗尽现有流量，您可以为目标组启用流量再平衡。有关更多信息，请参阅 the section called “目标失效转移”。

在超时到期之前，已注销的目标将显示处于 `draining` 状态。注销延迟超时到期后，目标的状态将变为 `unused`。

使用控制台更新取消注册延迟属性

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的负载均衡下，选择目标组。
3. 选择目标组的名称以打开其详细信息页面。
4. 在组详细信息页面的属性部分中，选择编辑。
5. 在编辑属性页面上，根据需要更改注销延迟的值。
6. 选择保存更改。

要更新取消注册延迟属性，请使用 AWS CLI

使用 modify-target-group-attributes 命令。

流量粘性

默认情况下，网关负载均衡器使用 5 元组（对于 TCP/UDP 流量）来保持流向特定目标设备的流量粘性。5 元组包括源 IP、源端口、目标 IP、目标端口和传输协议。您可以使用粘性类型属性来修改默认值（5 元组），然后选择 3 元组（源 IP、目标 IP 和传输协议）或 2 元组（源 IP 和目标 IP）。

流量粘性注意事项

- 流量粘性是在目标组级别配置和应用的，并且适用于所有流向目标组的流量。
- AWS Transit Gateway 设备模式开启时，不支持 2 元组和 3 元组流量粘性。要在您的设备上使用设备模式 AWS Transit Gateway，请在 Gateway Load Balancer 上使用 5 元组流量粘性

- 流量粘性可能会导致连接和流量分布不均，并相应影响目标的可用性。建议您在修改目标组的流量粘性类型之前，先终止或耗尽所有现有的流量。

使用控制台更新流量粘性属性

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的负载均衡下，选择目标组。
3. 选择目标组的名称以打开其详细信息页面。
4. 在组详细信息页面的属性部分中，选择编辑。
5. 在编辑属性页面上，根据需要更改流量粘性的值。
6. 选择保存更改。

要更新 flow 粘性属性，请使用 AWS CLI

使用带有`stickiness.enabled`和`stickiness.type`目标组属性的modify-target-group-attributes命令。

为网关负载均衡器注册目标

当您的目标准备好处理请求时，您将其注册到一个或多个目标组。您可以通过实例 ID 或 IP 地址注册目标。注册过程完成并且目标通过初始运行状况检查后，网关负载均衡器会立即开始将请求路由至目标。完成注册过程和开始运行状况检查可能需要几分钟时间。有关更多信息，请参阅 网关负载均衡器目标组的运行状况检查。

如果当前已注册目标的需求增加，您可以注册其他目标以便满足该需求。如果对已注册目标的需求减少，您可以从目标组中取消注册目标。完成注销过程并让网关负载均衡器停止将请求路由到目标可能需要几分钟时间。如果需求随后增加，您可以再次向目标组注册已取消注册的目标。如果您需要为目标提供服务，您可以取消注册，然后在服务完成后重新注册。

内容

- 注意事项
- 目标安全组
- 网络 ACLs
- 按实例 ID 注册目标
- 按 IP 地址注册目标

- 取消注册目标

注意事项

- 每个目标组在为网关负载均衡器启用的每个可用区中必须至少有一个已注册目标。
- 您的目标组的目标类型将确定如何向该目标组注册目标。有关更多信息，请参阅 Target type。
- 无法通过区域间 VPC 对等注册目标。
- 无法通过区域内 VPC 对等按实例 ID 注册实例，但可以通过 IP 地址注册这些实例。

目标安全组

将 EC2 实例注册为目标时，必须确保这些实例的安全组允许端口 6081 上的入站和出站流量。

网关负载均衡器没有关联任何安全组。因此，您的目标的安全组必须使用 IP 地址以允许来自负载均衡器的流量。

网络 ACLs

将 EC2 实例注册为目标时，必须确保实例子网的网络访问控制列表 (ACL) 允许端口 6081 上的流量。VPC 的默认网络 ACL 会允许所有入站和出站流量。如果您创建自定义网络 ACLs，请确认它们允许适当的流量。

按实例 ID 注册目标

当您注册实例时，实例必须处于 `running` 状态。

使用控制台按实例 ID 注册目标

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的负载均衡下，选择目标组。
3. 选择目标组的名称以打开其详细信息页面。
4. 在目标选项卡上，选择注册目标。
5. 选择实例，然后选择包含如下待处理事项。
6. 添加完实例后，选择注册待注册目标。

要按实例 ID 注册目标，请使用 AWS CLI

对实例使用 register- target IDs s 命令。

按 IP 地址注册目标

您注册的 IP 地址必须来自下列 CIDR 块之一：

- 目标组的 VPC 的子网
- 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)

使用控制台按 IP 地址注册目标

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的 Load Balancing（负载均衡）下，选择 Target Groups（目标组）。
3. 选择目标组的名称以打开其详细信息页面。
4. 在目标选项卡上，选择注册目标。
5. 选择网络、IP 地址和端口，然后选择在包含如下待处理事项。
6. 指定完地址后，选择注册待注册目标。

要使用 IP 地址注册目标 AWS CLI

使用带有目标的 IP 地址的 register-targets 命令。

取消注册目标

在取消注册目标时，Elastic Load Balancing 会一直等待，直到进行中的请求完成。这称作连接耗尽。在连接耗尽期间，目标的状态为 `draining`。在取消注册完成后，目标的状态将更改为 unused。有关更多信息，请参阅 取消注册延迟。

使用控制台取消注册目标

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的负载均衡下，选择目标组。
3. 选择目标组的名称以打开其详细信息页面。

4. 选择目标选项卡。
5. 选择 IP 地址，然后选择取消注册。

要取消注册目标，请使用 AWS CLI

使用 deregister-targets 命令以移除目标。

为网关负载均衡器的目标组添加标签

标签有助于按各种标准 (例如用途、所有者或环境) 对目标组进行分类。

您可以为每个目标组添加多个标签。每个目标组的标签键必须是唯一的。如果您添加的标签中的键已经与目标组关联，它将更新该标签的值。

用完标签后可以将其删除。

限制

- 每个资源的标签数上限 – 50
- 最大密钥长度 - 127 个 Unicode 字符
- 最大值长度 - 255 个 Unicode 字符
- 标签键和值区分大小写。允许使用的字符包括可用 UTF-8 格式表示的字母、空格和数字，以及以下特殊字符： + - = 。_ : / @。请不要使用前导空格或尾随空格。
- 请勿在标签名称或值中使用aws:前缀，因为它已保留供 AWS 使用。您无法编辑或删除带此前缀的标签名称或值。具有此前缀的标签不计入每个资源的标签数限制。

使用控制台更新目标组的标签

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的负载均衡下，选择目标组。
3. 选择目标组的名称以打开其详细信息页面。
4. 在标签选项卡上，选择管理标签，然后执行以下一项或多项操作：
    a. 要更新标签，请为键和值输入新值。
    b. 要添加标签，请选择添加标签，然后为键和值输入值。
    c. 要删除标签，请选择标签旁边的删除。

5. 更新完标签后，选择保存更改。

要更新目标群组的标签，请使用 AWS CLI

使用 add-tags 和 remove-tags 命令。

删除网关负载均衡器的目标组

如果目标组未由任何侦听器规则的转发操作引用，则可以删除该目标组。删除目标组不会影响已注册到目标组的目标。如果您不再需要已注册的 EC2 实例，则可以停止或终止该实例。

使用控制台删除目标组

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的负载均衡下，选择目标组。
3. 选择目标组，然后依次选择操作、删除。
4. 当系统提示进行确认时，选择是，删除。

要删除目标组，请使用 AWS CLI

使用 delete-target-group 命令。

监控网关负载均衡器

您可以使用以下功能来监控网关负载均衡器，以分析流量模式，排查相关问题。但网关负载均衡器属于透明的第 3 层负载均衡器，不会终止流量，因此不会生成访问日志。要接收访问日志，您必须在网关负载均衡器目标设备（例如防火墙、IDS/IPS 和安全设备）上启用访问日志记录。此外，您还可以选择在网关负载均衡器上启用 VPC 流日志。

CloudWatch 指标

您可以使用 Amazon CloudWatch 以一组有序的时间序列数据（称为指标）的形式检索有关网关负载均衡器和目标的数据点的统计数据。您可使用这些指标来验证您的系统是否按预期运行。有关更多信息，请参阅 CloudWatch 网关 Load Balancer 的指标。

Amazon VPC 流日志

您可以使用 VPC 流日志来捕获有关进出网关负载均衡器的流量的详细信息。有关更多信息，请参阅 Amazon VPC 用户指南中的 VPC 流日志。

为网关负载均衡器的每个网络接口创建流日志。每个子网都有一个网络接口。要确定某个网关负载均衡器的网络接口，请在网络接口的描述字段中查找该网关负载均衡器的名称。

通过网关负载均衡器的每个连接都有两个条目，一个用于客户端与网关负载均衡器之间的前端连接，另一个用于网关负载均衡器和目标之间的后端连接。如果目标由实例 ID 注册，连接将作为来自客户端的实例向实例显示。如果实例的安全组不允许来自客户端的连接，但子网的网络 ACLs 允许连接，则 Gateway Load Balancer 的网络接口日志显示前端和后端连接的 “ACCEPT OK”，而实例网络接口的日志显示该连接的 “拒绝确定”。

CloudTrail 日志

您可以使用 AWS CloudTrail 捕获有关对 Elastic Load Balancing API 的调用的详细信息，并将其作为日志文件存储在 Amazon S3 中。您可以使用这些 CloudTrail 日志来确定拨打了哪些呼叫、呼叫来自哪个源 IP 地址、谁拨打了电话、何时拨打了呼叫等。有关更多信息，请参阅使用记录 Elastic Load Balancing 的 API 调用 CloudTrail。

CloudWatch 网关 Load Balancer 的指标

Elastic Load Balancing 将您的网关负载均衡器和目标的数据点发布到亚马逊 CloudWatch 。CloudWatch 允许您以一组有序的时间序列数据（称为指标）的形式检索有关这些数据点的统计信息。可将指标视为要监控的变量，而将数据点视为该变量随时间变化的值。例如，您可以在指定时间段内监控网关负载均衡器的运行正常的目标总数。每个数据点都有相关联的时间戳和可选测量单位。

您可使用指标来验证系统是否正常运行。例如，您可以创建 CloudWatch 警报来监控指定的指标，并在该指标超出您认为可接受的范围时启动操作（例如向电子邮件地址发送通知）。

CloudWatch 只有当请求流经网关负载均衡器时，Elastic Load Balancing 才会向其报告指标。如果有请求流经负载均衡器，则弹性负载均衡会进行测量并以 60 秒的间隔发送指标。如果没有请求流经负载均衡器或指标无数据，则不报告指标。

有关更多信息，请参阅 Amazon CloudWatch 用户指南。

内容

- 网关负载均衡器指标
- 网关负载均衡器的指标维度
- 查看 Gateway Load Balancer 的 CloudWatch 指标

网关负载均衡器指标

AWS/GatewayELB 命名空间包括以下指标。

指标	描述
ActiveFlowCount	客户端至目标的并发流（或连接）的总数。 报告标准：有非零值 统计数据：最有用的统计工具是 Average、Maximum 和 Minimum。 Dimensions • LoadBalancer • AvailabilityZone ，LoadBalancer
ConsumedLCUs	负载均衡器使用的负载均衡器容量单位 (LCU) 数量。您需要为每小时 LCUs 的使用量付费。有关更多信息，请参阅 Elastic Load Balancing 定价。 报告标准：始终报告 统计数据：全部

指标	描述
	Dimensions • LoadBalancer
HealthyHostCount	被视为正常运行的目标数量。 报告标准：在启用了运行状况检查时报告 统计数据：最有用的统计工具为 Maximum 和 Minimum。 Dimensions • LoadBalancer ，TargetGroup • AvailabilityZone ，LoadBalancer ，TargetGroup
NewFlowCount	时段内建立的客户端至目标的新流（或连接）的总数。 报告标准：有非零值 Statistics：最有用的统计工具是 Sum。 Dimensions • LoadBalancer • AvailabilityZone ，LoadBalancer
PeakBytesPerSecond	每秒处理的最高平均字节数，在采样窗口期间每 10 秒计算一次。此指标不包括运行状况检查流量。 报告标准：始终报告 Statistics：最有用的统计工具是 Maximum。 Dimensions • LoadBalancer • AvailabilityZone ，LoadBalancer

指标	描述
PeakPacketsPerSecond	最高平均数据包速率（每秒处理的数据包数），在采样窗口期间每 10 秒计算一次。此指标包含运行状况检查流量。 报告标准：始终报告 Statistics：最有用的统计工具是 Maximum。 Dimensions • LoadBalancer • AvailabilityZone ，LoadBalancer
ProcessedBytes	负载均衡器处理的总字节数。此计数包括进出目标的流量，但不包含运行状况检查流量。 报告标准：有非零值 Statistics：最有用的统计工具是 Sum。 Dimensions • LoadBalancer • AvailabilityZone ，LoadBalancer
RejectedFlowCount	遭负载均衡器拒绝的流量（或连接）的总数。 报告标准：始终报告。 统计数据：最有用的统计工具是 Average、Maximum 和 Minimum。 Dimensions • LoadBalancer • AvailabilityZone ，LoadBalancer

指标	描述
RejectedFlowCount_TCP	遭负载均衡器拒绝的 TCP 流量（或连接）的数量。 报告标准：有非零值。 Statistics：最有用的统计工具是 Sum。 Dimensions • LoadBalancer • AvailabilityZone ，LoadBalancer
UnHealthyHostCount	被视为未正常运行的目标数量。 报告标准：在启用了运行状况检查时报告 统计数据：最有用的统计工具为 Maximum 和 Minimum。 Dimensions • LoadBalancer ，TargetGroup • AvailabilityZone ，LoadBalancer ，TargetGroup

网关负载均衡器的指标维度

要筛选网关负载均衡器的指标，请使用以下维度。

维度	描述
AvailabilityZone	按可用区筛选指标数据。
LoadBalancer	按网关负载均衡器筛选指标数据。按如下方式指定 Gateway Load Balancer：gateway load-balancer-name/1234567890123456（ARN 的最后一部分）。
TargetGroup	按目标组筛选指标数据。按如下方式指定目标组：targetgroup target-group-name/1234567890123456（目标组 ARN 的最后一部分）。

查看 Gateway Load Balancer 的 CloudWatch 指标

您可以使用 Amazon EC2 控制台查看网关负载均衡器的 CloudWatch 指标。这些指标显示为监控图表。如果网关负载均衡器处于活动状态并且正在接收请求，则监控图表会显示数据点。

或者，您可以使用 CloudWatch 控制台查看 Gateway Load Balancer 的指标。

使用控制台查看指标

1. 打开 Amazon EC2 控制台，网址为https://console.aws.amazon.com/ec2/。
2. 要查看按目标组筛选的指标，请执行以下操作：
   a. 在导航窗格中，选择 Target Groups。
   b. 选择目标组并选择 Monitoring。
   c. (可选) 要按时间筛选结果，请从 Showing data for 中选择时间范围。
   d. 要获得单个指标的一个较大视图，请选择其图形。
3. 要查看按网关负载均衡器筛选的指标，请执行以下操作：
   a. 在导航窗格中，选择负载均衡器。
   b. 选择您的网关负载均衡器，然后选择监控。
   c. (可选) 要按时间筛选结果，请从 Showing data for 中选择时间范围。
   d. 要获得单个指标的一个较大视图，请选择其图形。

使用 CloudWatch 控制台查看指标

1. 打开 CloudWatch 控制台，网址为https://console.aws.amazon.com/cloudwatch/。
2. 在导航窗格中，选择指标。
3. 选择 GatewayELB 命名空间。
4. （可选）要跨所有维度查看某个指标，请在搜索字段中输入其名称。

要查看指标，请使用 AWS CLI

使用以下 list-metrics 命令列出可用指标：

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

要获取指标的统计数据，请使用 AWS CLI

使用以下get-metric-statistics命令获取指定指标和维度的统计信息。请注意，CloudWatch 将每个唯一的维度组合视为一个单独的指标。您无法使用未专门发布的维度组合检索统计数据。您必须指定创建指标时使用的同一维度。

```
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
```

下面是示例输出。

```
{
    "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"
}
```

网关负载均衡器的配额

您的 AWS 账户对每项 AWS 服务都有默认配额（以前称为限制）。除非另有说明，否则，每个限额是区域特定的。您可以请求增加某些配额，但其他一些配额无法增加。

要请求提高配额，请参阅《服务配额用户指南》中的请求提高配额。如果 Service Quotas 中尚无配额，请提交增加服务配额的请求。

限额

- 负载均衡器
- 目标组
- Load Balancer 容量单位

负载均衡器

您的 AWS 账户具有以下与网关负载均衡器相关的配额。

名称	默认值	可调整
每区域的网关负载均衡器数	100	是
每 VPC 的网关负载均衡器数	100	是
ENIs 每个 VPC 的网关 Load Balancer	300 *	是
每网关负载均衡器的侦听器数	1	否

* 每个网关负载均衡器在每个可用区使用一个网络接口。

目标组

以下配额适用于目标组。

名称	默认值	可调整
每区域的 GENEVE 目标组数	100	是

名称	默认值	可调整
每个目标组的目标	1000	是
每可用区每 GENEVE 目标组的目标数	300	否
每可用区每网关负载均衡器的目标数	300	否
每网关负载均衡器的目标数	300	否

Load Balancer 容量单位

以下配额适用于 Load Balancer 容量单位 (LCUs)。

名称	默认值	可调整
每个区域的预留网关 Load Balancer 容量单位 (LCU)	0	是

网关负载均衡器的文档历史记录

下表介绍了网关负载均衡器的版本。

变更	说明	日期
容量单位预留	此版本增加了对为负载均衡器设置最低容量的支持。	2025 年 4 月 10 日
IPv6 支持	您可以将 Gateway Load Balancer 配置为同时支持 IPv4 和 IPv6 地址。	2022 年 12 月 12 日
流量再平衡	此版本添加了对在目标失败或取消注册时定义网关负载均衡器的流量处理行为的支持。	2022 年 10 月 13 日
可配置的流量粘性	您可以配置哈希算法，以保持指向特定目标设备的流量粘性。	2022 年 8 月 25 日
在新区域中可用	此版本增加了对 AWS GovCloud (US) 各区域网关负载均衡器的支持。	2021 年 6 月 17 日
在新区域中可用	此版本在加拿大（中部）、亚太地区（首尔）和亚太地区（大阪）区域中添加了对网关负载均衡器的支持。	2021 年 3 月 31 日
在新区域中可用	此版本在美国西部（北加利福尼亚）、欧洲地区（伦敦）、欧洲地区（巴黎）、欧洲地区（米兰）、非洲（开普敦）、中东（巴林）、亚太地区（香港）、亚太地区（新加坡）和亚太地区（孟买）区域中添加了对网关负载均衡器的支持。	2021 年 3 月 19 日

初始版本	此版本的弹性负载均衡器引入了网关负载均衡器。	2020 年 11 月 10 日

本文属于机器翻译版本。若本译文内容与英语原文存在差异，则一律以英文原文为准。