Elastic Load Balancing

Application Load Balancer

Elastic Load Balancing: Application Load Balancer

Table of Contents

什么是应用程序负载均衡器？

Elastic Load Balancing 支持三种类型的负载均衡器：Application Load Balancer、Network Load Balancer 和 Classic Load Balancer。本指南讨论 Application Load Balancer。有关 Network Load Balancer 的更多信息，请参阅 Network Load Balancer 用户指南。有关 Classic Load Balancer 的更多信息，请参阅 Classic Load Balancer 用户指南。

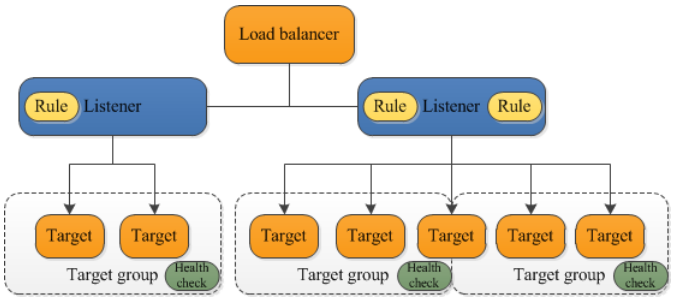
应用程序负载均衡器 组件

负载均衡器 充当客户端的单一接触点。负载均衡器在多个可用区中的多个目标 (例如 EC2 实例) 间分配应用程序的传入流量。这将提高应用程序的可用性。可以向您的负载均衡器添加一个或多个侦听器。

侦听器 使用您配置的协议和端口检查来自客户端的连接请求，并根据您定义的规则将请求转发到一个或多个目标组。每个规则指定一个目标组、条件和优先级。满足条件时，流量会转发到目标组。您必须为每个侦听器定义一个默认规则，然后，您可以添加规则来根据请求内容指定不同目标组 (也称为基于内容的路由)。

每个目标组 使用您指定的协议和端口号将请求路由到一个或多个注册目标，例如 EC2 实例。您可以向多个目标组注册一个目标。您可以对每个目标组配置运行状况检查。在注册到目标组 (它是使用负载均衡器的侦听器规则指定的) 的所有目标上，执行运行状况检查。

下图介绍基本组成部分。请注意，每个侦听器包含一个默认规则，并且一个侦听器包含将请求路由到不同目标组的另一条规则。向两个目标组注册一个目标。



有关更多信息，请参阅以下文档：

- 个负载均衡器 (p. 13)
- 侦听器 (p. 22)
- 目标组 (p. 39)

应用程序负载均衡器概述

应用程序负载均衡器在应用程序层正常工作，该层是开放系统互连 (OSI) 模型的第 7 层。负载均衡器收到请求后，将按照优先级顺序评估侦听器规则以确定应用哪个规则，然后从目标组中选择规则操作目标。可以配置侦听器规则，以根据应用程序流量的内容，将请求路由至不同的目标组。每个目标组的路由都是单独进行的，即使某个目标已在多个目标组中注册。

可以根据需求变化在负载均衡器中添加和删除目标，而不会中断应用程序的整体请求流。Elastic Load Balancing 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。Elastic Load Balancing 能够自动扩展来处理绝大部分工作负载。

您可以配置运行状况检查，这些检查可用来监控注册目标的运行状况，以便负载均衡器只能将请求发送到正常运行的目标。

有关更多信息，请参阅 Elastic Load Balancing 用户指南 中的 Elastic Load Balancing 工作原理。

从 传统负载均衡器迁移的好处

使用应用程序负载均衡器而不是传统负载均衡器具有以下优势：

- 支持基于路径的路由。对于根据请求中的 URL 转发请求的侦听器，您可以为它配置规则。这让您可以将应用程序构造为较小的服务，并根据 URL 内容将请求路由到正确的服务。
- 支持基于主机的路由。对于基于 HTTP 标头中主机字段转发请求的侦听器，您可以为它配置规则。这使您能够使用单个负载均衡器将请求路由到多个域。
- 支持将请求路由到单个 EC2 实例上的多个应用程序。可以使用多个端口向同一个目标组注册每个实例或 IP 地址。
- 支持通过 IP 地址注册目标，包括位于负载均衡器的 VPC 之外的目标。
- 支持容器化的应用程序。计划任务时，Amazon Elastic Container Service (Amazon ECS) 可以选择一个未使用的端口，并可以使用此端口向目标组注册该任务。这样可以高效地使用您的群集。
- 支持单独监控每个服务的运行状况，因为运行状况检查是在目标组级别定义的，并且许多 CloudWatch 指标是在目标组级别报告的。将目标组挂载到 Auto Scaling 组的功能使您能够根据需求动态扩展每个服务。
- 访问日志包含附加信息，并以压缩格式存储。
- 已改进负载均衡器性能。

有关每个负载均衡器类型支持的功能的更多信息，请参阅 Elastic Load Balancing 产品比较。

如何开始

要创建应用程序负载均衡器，请尝试以下教程之一：

- Elastic Load Balancing 用户指南中的 Elastic Load Balancing 入门。
- 教程：对您的应用程序负载均衡器使用基于路径的路由 (p. 6)
- 教程：使用微服务作为您的应用程序负载均衡器的目标 (p. 8)

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 Elastic Load Balancing 定价。

Application Load Balancer 入门

本教程介绍通过 AWS 管理控制台（基于 Web 的界面）创建 Application Load Balancer 的实际操作。要创建第一个应用程序负载均衡器，请完成以下步骤。

任务

此外，要创建 网络负载均衡器，请参阅 Network Load Balancer 用户指南 中的 Network Load Balancer 入门。若要创建 传统负载均衡器，请参阅 Classic Load Balancer 用户指南 中的创建 传统负载均衡器。

开始前的准备工作

- 决定您将用于 EC2 实例的两个可用区。在每个这些可用区中配置至少带有一个公有子网的 Virtual Private Cloud (VPC)。这些公有子网用于配置负载均衡器。您可以改为在这些可用区的其他子网中启动您的 EC2 实例。
- 在每个可用区中至少启动一个 EC2 实例。请确保在每个 EC2 实例上安装 Web 服务器，例如 Apache 或 Internet Information Services (IIS)。确保这些实例的安全组允许端口 80 上的 HTTP 访问。

步骤 1：选择负载均衡器类型

Elastic Load Balancing 支持三种负载均衡器。在本教程中，您将创建一个 应用程序负载均衡器。

创建应用程序负载均衡器

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航栏上，选择您的负载均衡器的区域。请确保选择用于 EC2 实例的同一个区域。
3. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
4. 选择 Create Load Balancer。
5. 对于 Application Load Balancer，选择 Create。

步骤 2：配置负载均衡器和侦听器

在 Configure Load Balancer 页面上，完成以下过程。

配置负载均衡器和侦听器

1. 对于 Name，键入负载均衡器的名称。

在区域的 Application Load Balancer 和 Network Load Balancer 集内，应用程序负载均衡器 的名称必须唯一，最多可以有 32 个字符，只能包含字母数字字符和连字符，不能以连字符开头或结尾，并且不能以“internal-”开头。

2. 对于 Scheme 和 IP address type，请保留默认值。
3. 对于 Listeners，保留默认值，默认侦听器负责接收端口 80 上的 HTTP 流量。
4. 对于 Availability Zones，选择用于 EC2 实例的 VPC。对于用于启动 EC2 实例的每个可用区，选择一个可用区，然后为该可用区选择公有子网。
5. 选择 Next: Configure Security Settings。
6. 在本教程中，将不创建 HTTPS 侦听器。选择 Next: Configure Security Groups。

步骤 3：为负载均衡器配置安全组

您负载均衡器的安全组必须允许其通过侦听器端口和运行状况检查端口与已注册目标进行通信。控制台可以代表您创建负载均衡器的安全组，其中包括指定正确协议和端口的规则。如果您愿意，也可以自行创建和选择安全组。有关更多信息，请参阅 推荐的规则 (p. 19)。

在 Configure Security Groups 页面，完成以下步骤即可令 Elastic Load Balancing 代表您为负载均衡器创建安全组。

为负载均衡器配置安全组

1. 选择 Create a new security group。
2. 为安全组键入名称和描述，或者保留默认名称和描述。此新安全组包含一条规则，该规则允许将流量传送到在 Configure Load Balancer 页面上选择的负载均衡器侦听器端口。
3. 选择 Next: Configure Routing。

步骤 4：配置目标组

创建一个要在请求路由中使用的目标组。您侦听器的默认规则将请求路由到此目标组中的已注册目标。负载均衡器使用为目标组定义的运行状况检查设置来检查此目标组中目标的运行状况。在 Configure Routing 页面上，完成以下过程。

配置目标组

1. 对于 Target group，保留默认值 New target group。
2. 对于 Name，键入新目标组的名称。
3. 将 Protocol 保留为“HTTP”，Port 为“80”，Target type 为“instance”。
4. 对于 Health checks，保留默认协议和 ping 路径。
5. 选择 Next: Register Targets。

步骤 5：向您的目标组注册目标

在 Register Targets 页面上，完成以下过程。

向目标组注册目标

1. 对于 Instances，选择一个或多个实例。
2. 保留默认端口 80，并选择 Add to registered。

3. 当您完成选择实例后，选择 Next: Review。

步骤 6：创建并测试您的负载均衡器

在创建负载均衡器之前，请检查所选的设置。在创建负载均衡器之后，可以验证其是否将流量发送到您的 EC2 实例。

创建并测试您的负载均衡器

1. 在 Review 页面上，选择 Create 。
2. 在您收到已成功创建负载均衡器的通知后，选择 Close。
3. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。
4. 选择新创建的目标组。
5. 在 Targets 选项卡中，验证您的实例是否已准备就绪。如果实例状态是 `initial`，很可能是因为，实例仍在注册过程中，或者未通过视为正常运行所需的运行状况检查最小数量。在您的至少一个实例的状态为 `healthy` 后，便可测试负载均衡器。
6. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
7. 选择新创建的负载均衡器。
8. 在 Description 选项卡中，复制负载均衡器 (例如，my-load-balancer-1234567890.us-west-2.elb.amazonaws.com) 的 DNS 名称。将 DNS 名称粘贴到已连接 Internet 的 Web 浏览器的地址栏中。如果一切正常，浏览器会显示您服务器的默认页面。
9. （可选）要定义其他侦听器规则，请参阅添加规则 (p. 28)。

步骤 7：删除您的负载均衡器 (可选)

在您的负载均衡器可用之后，您需要为保持其运行的每小时或部分小时支付费用。当您不再需要负载均衡器时，可将其删除。当负载均衡器被删除之后，您便不再需要支付负载均衡器费用。请注意，删除负载均衡器不会影响在负载均衡器中注册的目标。例如，您的 EC2 实例会继续运行。

删除您的负载均衡器

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选中负载均衡器的复选框，然后依次选择 Actions、Delete 。
4. 当系统提示进行确认时，选择 Yes, Delete。

应用程序负载均衡器教程

以下 Elastic Load Balancing 教程介绍了如何使用应用程序负载均衡器执行常见任务。

- Elastic Load Balancing 入门 (Elastic Load Balancing 用户指南)
- 教程：对您的应用程序负载均衡器使用基于路径的路由 (p. 6)
- 教程：使用微服务作为您的应用程序负载均衡器的目标 (p. 8)
- 教程：使用 AWS CLI 创建应用程序负载均衡器 (p. 9)

教程：对您的应用程序负载均衡器使用基于路径的路由

您可以创建一个侦听器，其中包含根据 URL 路径转发请求的规则。这称为基于路径的路由。如果您运行的是微服务，可以使用基于路径的路由将流量路由给多个后端服务。例如，您可以将一般请求路由到一个目标组，并将图像呈现请求路由到另一个目标组。

在您开始之前

- 启动 Virtual Private Cloud (VPC) 中您的 EC2 实例。确保这些实例的安全组允许访问侦听器端口和运行状况检查端口。有关更多信息，请参阅 目标安全组 (p. 47)。
- 验证您的微服务是否部署在您计划注册的 EC2 实例。

创建负载均衡器

创建使用基于路径的路由的负载均衡器

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航栏上，选择您为 EC2 实例选择的同一个区域。
3. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。
4. 为第一组目标创建目标组，如下所示：
   a. 选择 Create target group。
   b. 为目标组指定名称、协议、端口和 VPC，然后选择 Create。
   c. 选择新目标组。
   d. 在 Targets 选项卡上，选择 Edit。
   e. 对于 Instances，选择一个或多个实例。指定实例的端口，选择 Add to registered，然后选择 Save。

      请注意，实例的状态为 `initial`，直至实例注册并通过运行状况检查，然后，实例的状态将成为 `unused`，直至您配置目标组从负载均衡器接收流量。
5. 为第二组目标创建目标组，如下所示：
   a. 选择 Create target group。
   b. 为目标组指定名称、协议、端口和 VPC，然后选择 Create。
   c. 在 Targets 选项卡上，选择 Edit。

d. 对于 Instances，选择一个或多个实例。指定实例的端口，选择 Add to registered，然后选择 Save。

   请注意，实例的状态为 `initial`，直至实例注册并通过运行状况检查，然后，实例的状态将成为 `unused`，直至您配置目标组从负载均衡器接收流量。

6. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
7. 选择 Create Load Balancer。
8. 对于 Select load balancer type，选择 Application Load Balancer。
9. 选择 Continue。
10. 完成 Configure Load Balancer 页面，如下所示：

    a. 对于 Name，键入负载均衡器的名称。

       在当前区域的 Application Load Balancer 和 Network Load Balancer 集内，应用程序负载均衡器 的名称必须唯一，最多可以有 32 个字符，只能包含字母数字字符和连字符，不能以连字符开头或结尾。

    b. 对于 Scheme，面向 Internet 的负载均衡器将来自客户端的请求通过 Internet 路由到目标。内部负载均衡器使用私有 IP 地址将请求路由到目标。

    c. 对于 Listeners，默认值是负责接收端口 80 上的 HTTP 流量的侦听器。您可以保留默认侦听器设置，修改侦听器的协议或端口，或者选择 Add 以添加另一个侦听器。

    d. 对于 Availability Zones，选择用于 EC2 实例的 VPC。选择至少两个可用区。如果可用区有一个子网，则将选择此子网。如果可用区有多个子网，请选择子网之一。请注意，您只能为每个可用区选择一个子网。

    e. 选择 Next: Configure Security Settings。

11. (可选) 如果在上一步中创建了一个安全侦听器，请完成 Configure Security Settings 页面，如下所示：

    a. 如果使用 AWS Certificate Manager 创建或导入了证书，请选择 Choose an existing certificate from AWS Certificate Manager (ACM)，然后从 Certificate name 中选择证书。

    b. 如果使用 IAM 上传了证书，请选择 Choose an existing certificate from AWS Identity and Access Management (IAM)，然后从 Certificate name 中选择证书。

    c. 如果要上传证书，但您所在区域不支持 ACM，请选择 Upload a new SSL Certificate to AWS Identity and Access Management (IAM)。对于 Certificate name，键入证书的名称。对于 Private Key，复制并粘贴私钥文件 (PEM 编码的文件) 的内容。在 Public Key Certificate 中，复制并粘贴公钥证书文件 (PEM 编码的文件) 的内容。在 Certificate Chain 中，复制并粘贴证书链文件 (PEM 编码的文件) 的内容，除非您使用的是自签名证书并且浏览器是否隐式接受证书并不重要。

    d. 对于 Select policy，请保持默认的安全策略。

12. 选择 Next: Configure Security Groups。
13. 完成 Configure Security Groups 页面，如下所示：

    a. 选择 Create a new security group。

    b. 为安全组键入名称和描述，或者保留默认名称和描述。此新安全组包含一条规则，该规则允许将流量传送到在 Configure Load Balancer 页面上为负载均衡器选择的端口。

    c. 选择 Next: Configure Routing。

14. 完成 Configure Routing 页面，如下所示：

    a. 对于 Target group，选择 `Existing target group`。

    b. 对于 Name，选择您创建的第一个目标组。

    c. 选择 Next: Register Targets。

15. 在 Register Targets 页面上，在 Registered instances 下会显示向目标组注册的实例。在完成向导之前，您无法修改向目标组注册的目标。选择 Next: Review。
16. 在 Review 页面上，选择 Create 。
17. 在您收到已成功创建负载均衡器的通知后，选择 Close。

18. 选择新创建的负载均衡器。
19. 在 Listeners 选项卡中，使用箭头查看侦听器的规则，然后选择 Add rule。按如下所示指定规则：

    a. 对于 Target group name，选择您创建的第二个目标组。
    b. 对于 Path pattern，指定针对基于路径的路由使用的准确模式 (例如，`/img/*`)。有关更多信息，请参阅 侦听器规则 (p. 22)。
    c. 选择 Save。

教程：使用微服务作为您的应用程序负载均衡器的目标

您可以使用微服务架构来构造应用程序，作为您可以独立开发和部署的服务。您可以在各 EC2 实例上安装一个或多个这样的服务，每个服务在不同端口上接受连接。您可以使用单个应用程序负载均衡器将请求路由到应用程序的所有服务。在您将 EC2 实例注册到目标组时，可以多次注册；对于每个服务，使用该服务的端口注册实例。

> Important
>
> 使用 Amazon Elastic Container Service (Amazon ECS) 部署服务时，您可以使用动态端口映射来支持同一个容器实例上单个服务的多个任务。Amazon ECS 使用各个容器的实例 ID 和端口，自动对您的目标组注册和注销容器，从而管理对您服务的更新。有关更多信息，请参阅 Amazon Elastic Container Service Developer Guide 中的服务负载均衡。

在您开始之前

- 启动 EC2 实例。确保实例的安全组允许负载均衡器安全组在侦听器端口和运行状况检查端口上进行访问。有关更多信息，请参阅 目标安全组 (p. 47)。
- 将服务部署到您的 EC2 实例 (例如，使用容器)。

创建负载均衡器

创建使用多个服务作为目标的负载均衡器

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航栏上，选择您为 EC2 实例选择的同一个区域。
3. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
4. 选择 Create Load Balancer。
5. 对于 Select load balancer type，选择 Application Load Balancer。
6. 选择 Continue。
7. 完成 Configure Load Balancer 页面，如下所示：

    a. 对于 Name，键入负载均衡器的名称。

    在当前区域的 Application Load Balancer 和 Network Load Balancer 集内，应用程序负载均衡器 的名称必须唯一，最多可以有 32 个字符，只能包含字母数字字符和连字符，不能以连字符开头或结尾。
    b. 对于 Scheme，面向 Internet 的负载均衡器将来自客户端的请求通过 Internet 路由到目标。内部负载均衡器使用私有 IP 地址将请求路由到目标。
    c. 对于 Listeners，默认值是负责接收端口 80 上的 HTTP 流量的侦听器。您可以保留默认侦听器设置，修改侦听器的协议或端口，或者选择 Add 以添加另一个侦听器。

d. 对于 Availability Zones，选择用于 EC2 实例的 VPC。选择至少两个可用区。如果可用区有一个子网，则将选择此子网。如果可用区有多个子网，请选择子网之一。请注意，您只能为每个可用区选择一个子网。
e. 选择 Next: Configure Security Settings。

8. (可选) 如果在上一步中创建了一个安全侦听器，请完成 Configure Security Settings 页面，如下所示：
   a. 如果使用 AWS Certificate Manager 创建或导入了证书，请选择 Choose an existing certificate from AWS Certificate Manager (ACM)，然后从 Certificate name 中选择证书。
   b. 如果使用 IAM 上传了证书，请选择 Choose an existing certificate from AWS Identity and Access Management (IAM)，然后从 Certificate name 中选择证书。
   c. 如果要上传证书，但您所在区域不支持 ACM，请选择 Upload a new SSL Certificate to AWS Identity and Access Management (IAM)。对于 Certificate name，键入证书的名称。对于 Private Key，复制并粘贴私钥文件 (PEM 编码的文件) 的内容。在 Public Key Certificate 中，复制并粘贴公钥证书文件 (PEM 编码的文件) 的内容。在 Certificate Chain 中，复制并粘贴证书链文件 (PEM 编码的文件) 的内容，除非您使用的是自签名证书并且浏览器是否隐式接受证书并不重要。
   d. 对于 Select policy，请保持默认的安全策略。
9. 选择 Next: Configure Security Groups。
10. 完成 Configure Security Groups 页面，如下所示：
    a. 选择 Create a new security group。
    b. 为安全组键入名称和描述，或者保留默认名称和描述。此新安全组包含一条规则，该规则允许将流量传送到在 Configure Load Balancer 页面上为负载均衡器选择的端口。
    c. 选择 Next: Configure Routing。
11. 完成 Configure Routing 页面，如下所示：
    a. 对于 Target group，保留默认值 `New target group`。
    b. 对于 Name，键入新目标组的名称。
    c. 根据需要设置 Protocol 和 Port。
    d. 对于 Health checks，保留默认运行状况检查设置。
    e. 选择 Next: Register Targets。
12. 对于 Register Targets，请执行以下操作：
    a. 对于 Instances，选择 EC2 实例。
    b. 键入服务使用的端口，然后选择 Add to registered。
    c. 针对每个服务重复上述操作进行注册。完成后，选择 Next: Review。
13. 在 Review 页面上，选择 Create 。
14. 在您收到已成功创建负载均衡器的通知后，选择 Close。

教程：使用 AWS CLI 创建应用程序负载均衡器

本教程介绍通过 AWS CLI 创建 Application Load Balancer 的实际操作。

开始前的准备工作

- 使用以下命令可验证您运行的是否是支持 Application Load Balancer 的 AWS CLI 版本。

```
aws elbv2 help
```

如果您获取的错误消息指示 elbv2 不是有效选择，请更新您的 AWS CLI。有关更多信息，请参阅 AWS Command Line Interface 用户指南 中的安装 AWS 命令行界面。

- 启动 Virtual Private Cloud (VPC) 中您的 EC2 实例。确保这些实例的安全组允许访问侦听器端口和运行状况检查端口。有关更多信息，请参阅 目标安全组 (p. 47)。

创建负载均衡器

要创建第一个负载均衡器，请完成以下步骤。

创建负载均衡器

1. 使用 create-load-balancer 命令创建负载均衡器。您必须指定来自不同可用区的两个子网。

```
aws elbv2 create-load-balancer --name my-load-balancer  \
--subnets subnet-12345678 subnet-23456789 --security-groups sg-12345678
```

输出包含负载均衡器的 Amazon 资源名称 (ARN)，格式如下：

```
arn:aws-cn:elasticloadbalancing:us-west-2:123456789012:loadbalancer/app/my-load-
balancer/1234567890123456
```

2. 使用 create-target-group 命令创建目标组，并指定用于 EC2 实例的相同 VPC：

```
aws elbv2 create-target-group --name my-targets --protocol HTTP --port 80 \
--vpc-id vpc-12345678
```

输出包含目标组的 ARN，格式如下：

```
arn:aws-cn:elasticloadbalancing:us-west-2:123456789012:targetgroup/my-
targets/1234567890123456
```

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-12345678 Id=i-23456789
```

4. 使用 create-listener 命令为您的负载均衡器创建侦听器，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 \
--protocol HTTP --port 80  \
--default-actions Type=forward,TargetGroupArn=targetgroup-arn
```

输出包含侦听器的 ARN，格式如下：

```
arn:aws-cn:elasticloadbalancing:us-west-2:123456789012:listener/app/my-load-
balancer/1234567890123456/1234567890123456
```

5. (可选) 您可以使用此 describe-target-health 命令验证目标组的已注册目标的运行状况：

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

添加 HTTPS 侦听器

如果您拥有带 HTTP 侦听器的负载均衡器，则可按如下方式添加 HTTPS 侦听器。

向您的负载均衡器添加 HTTPS 侦听器

1. 使用下列方法之一创建要用于负载均衡器的 SSL 证书：
   - 使用 AWS Certificate Manager (ACM) 创建或导入证书。有关更多信息，请参阅 AWS Certificate Manager 用户指南中的请求证书或导入证书。
   - 使用 AWS Identity and Access Management (IAM) 上传证书。有关更多信息，请参阅 IAM 用户指南中的使用服务器证书。
2. 使用 create-listener 命令创建侦听器，该侦听器带有将请求转发到目标组的默认规则。在创建 HTTPS 侦听器时，您必须指定 SSL 证书。请注意，您可以使用 `--ssl-policy` 选项指定默认值之外的 SSL 策略。

```
aws elbv2 create-listener --load-balancer-arn loadbalancer-arn \
--protocol HTTPS --port 443  \
--certificates CertificateArn=certificate-arn \
--default-actions Type=forward,TargetGroupArn=targetgroup-arn
```

使用端口覆盖添加目标

如果您在一个实例中有多个 ECS 容器，则每个容器均通过不同的端口接受连接。您可以将实例注册到目标组多次，每次使用不同的端口进行注册。

使用端口覆盖添加目标

1. 使用 create-target-group 命令创建目标组：

```
aws elbv2 create-target-group --name my-targets --protocol HTTP --port 80 \
--vpc-id vpc-12345678
```

2. 使用 register-targets 命令将您的实例注册到目标组。请注意，每个容器的实例 ID 相同，但端口不同。

```
aws elbv2 register-targets --target-group-arn targetgroup-arn  \
--targets Id=i-12345678,Port=80 Id=i-12345678,Port=766
```

3. 使用 create-rule 命令向侦听器添加可将请求转发到目标组的规则：

```
aws elbv2 create-rule --listener-arn listener-arn --priority 10 \
--actions Type=forward,TargetGroupArn=targetgroup-arn
```

添加基于路径的路由

如果您的侦听器具有可将请求转发到一个目标组的默认规则，则可添加一个将请求转发到另一个基于 URL 的目标组的规则。例如，您可以将一般请求路由到一个目标组，并将图像显示请求路由到另一个目标组。

将规则添加到带路径模式的侦听器

1. 使用 create-target-group 命令创建目标组：

```
aws elbv2 create-target-group --name my-targets --protocol HTTP --port 80 \
--vpc-id vpc-12345678
```

2. 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-12345678 Id=i-23456789
```

3. 使用 create-rule 命令向侦听器添加一个可在 URL 包含指定模式时将请求转发到目标组的规则：

```
aws elbv2 create-rule --listener-arn listener-arn --priority 10 \
--conditions Field=path-pattern,Values='/img/*' \
--actions Type=forward,TargetGroupArn=targetgroup-arn
```

删除负载均衡器

当您不再需要负载均衡器和目标组时，可以将其删除，如下所示：

```
aws elbv2 delete-load-balancer --load-balancer-arn loadbalancer-arn
aws elbv2 delete-target-group --target-group-arn targetgroup-arn
```

应用程序负载均衡器

负载均衡器 充当客户端的单一接触点。客户端将请求发送到负载均衡器，然后负载均衡器将请求发送到两个或更多可用区中的目标 (例如 EC2 实例)。要配置您的负载均衡器，可以创建目标组 (p. 39)，然后将目标注册到目标组。您还可以创建侦听器 (p. 22)来检查来自客户端的连接请求，并创建侦听器规则以将来自客户端的请求路由到一个或多个目标组中的目标。

内容

负载均衡器安全组

安全组 起到防火墙的作用，可控制允许往返于负载均衡器的流量。您可以选择端口和协议以允许入站和出站流量。

与负载均衡器安全组关联的安全组的规则必须允许侦听器和运行状况检查端口上的双向流量。当您将侦听器添加到负载均衡器或更新目标组的运行状况检查端口时，您必须检查您的安全组规则，确保它们允许新端口上的双向流量。有关更多信息，请参阅 推荐的规则 (p. 19)。

负载均衡器状态

负载均衡器可能处于下列状态之一：

`provisioning`

正在设置负载均衡器。

`active`

负载均衡器已完全设置并准备好路由流量。

`failed`

负载均衡器无法设置。

负载均衡器属性

以下是负载均衡器属性：

`access_logs.s3.enabled`

指示是否启用存储在 Amazon S3 中的访问日志。默认为 `false`。

`access_logs.s3.bucket`

访问日志所用的 S3 存储桶的名称。如果启用访问日志，则此属性是必需的。有关更多信息，请参阅 存储桶权限 (p. 63)。

`access_logs.s3.prefix`

S3 存储桶中位置的前缀。

`deletion_protection.enabled`

指示是否启用删除保护。默认为 `false`。

`idle_timeout.timeout_seconds`

空闲超时值 (以秒为单位)。默认值为 60 秒。

`routing.http2.enabled`

指示是否启用了 HTTP/2。默认为 `true`。

IP 地址类型

在创建面向 Internet 的负载均衡器时或在该负载均衡器处于活动状态后，可以设置其 IP 地址类型。请注意，内部负载均衡器必须使用 IPv4 地址。

以下是负载均衡器 IP 地址类型：

`ipv4`

负载均衡器仅支持 IPv4 地址 (例如，192.0.2.1)

`dualstack`

负载均衡器支持 IPv4 和 IPv6 地址 (例如，2001:0db8:85a3:0:0:8a2e:0370:7334)。

使用 IPv4 地址与负载均衡器通信的客户端将解析 A 记录，使用 IPv6 地址与负载均衡器通信的客户端将解析 AAAA 记录。但是，无论客户端以何种方式与负载均衡器通信，负载均衡器均使用 IPv4 地址与其目标通信。

有关更多信息，请参阅 Application Load Balancer 的 IP 地址类型 (p. 20)。

删除保护

为了防止您的负载均衡器被意外删除，您可以启用删除保护。默认情况下，已为负载均衡器禁用删除保护。

如果您为负载均衡器启用删除保护，则必须先禁用删除保护，然后才能删除负载均衡器。

使用控制台启用删除保护

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器。

4. 在 Description 选项卡上，选择 Edit attributes。
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。
6. 选择 Save。

使用控制台禁用删除保护

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器。
4. 在 Description 选项卡上，选择 Edit attributes。
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。
6. 选择 Save。

使用 AWS CLI 启用或禁用删除保护

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

连接空闲超时

对于客户端通过负载均衡器发出的每个请求，负载均衡器将维护两个连接。前端连接是客户端和负载均衡器之间的连接，后端连接是负载均衡器和目标之间的连接。负载均衡器管理空闲超时，当在指定时间段内没有通过连接发送任何数据时，将触发空闲超时。超过空闲超时期限后，如果没有发送或接收任何数据，负载均衡器将关闭连接。

默认情况下，Elastic Load Balancing 将空闲超时值设为 60 秒。因此，如果在请求过程中，目标未至少每 60 秒发送一些数据，负载均衡器可以关闭前端连接。为确保长时间运行的操作 (例如文件上传) 有足够时间来完成，请在到达每个空闲超时期限前发送至少 1 个字节的数据，并根据需要增大空闲超时期限的长度。

对于后端连接，我们建议您对 EC2 实例启用 HTTP 保持连接选项。您可以在 EC2 实例的 Web 服务器设置中启用 HTTP 保持活动选项。如果您启用 HTTP 保持活动选项，负载均衡器即可重复使用后端连接，直到保持活动超时过期。此外，我们建议您将应用程序的空闲超时配置为大于负载均衡器的空闲超时的值。

使用控制台更新空闲超时值

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器。
4. 在 Description 选项卡上，选择 Edit attributes。
5. 在 Edit load balancer attributes 页面上，键入 Idle timeout 的值 (以秒为单位)。有效范围为 1-4000。默认值为 60 秒。
6. 选择 Save。

使用 AWS CLI 更新空闲超时值

使用带 `idle_timeout.timeout_seconds` 属性的 modify-load-balancer-attributes 命令。

Application Load Balancer 和 AWS WAF

您可以使用 AWS WAF 和 应用程序负载均衡器 以根据 Web 访问控制列表 (web ACL) 中的规则允许或阻止请求。有关更多信息，请参阅 AWS WAF 开发人员指南 中的使用 Web ACL。

创建 Application Load Balancer

负载均衡器接收来自客户端的请求，并将请求分发给目标组中的目标。

开始前，请确保您有一个满足以下要求的 Virtual Private Cloud (VPC)：在目标使用的每个可用区中至少有一个公有子网。

要使用 AWS CLI 创建负载均衡器，请参阅教程：使用 AWS CLI 创建应用程序负载均衡器 (p. 9)。

要使用 AWS 管理控制台 创建负载均衡器，请完成以下任务。

任务

步骤 1：配置负载均衡器和侦听器

首先，为负载均衡器提供一些基本配置信息，如名称、网络及一个或多个侦听器。侦听器是用于检查连接请求的进程。它配置了用于从客户端连接到负载均衡器的协议和端口。有关受支持的协议和端口的更多信息，请参阅侦听器配置 (p. 22)。

配置负载均衡器和侦听器

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择 Create Load Balancer。
4. 对于 Application Load Balancer，选择 Create。
5. 对于 Name，键入负载均衡器的名称。例如，**`my-alb`**。
6. 对于 Scheme，面向 Internet 的负载均衡器将来自客户端的请求通过 Internet 路由到目标。内部负载均衡器使用私有 IP 地址将请求路由到目标。
7. 对于 Listeners，默认值是负责接收端口 80 上的 HTTP 流量的侦听器。您可保留默认侦听器设置，修改协议或修改端口。选择 Add 可添加其他侦听器 (例如，HTTPS 侦听器)。
8. 对于可用区，请从您的 VPC 中选择至少两个可用区。如果可用区有一个子网，则将选择此子网。如果可用区有多个子网，请选择子网之一。请注意，您只能为每个可用区选择一个子网。
9. 选择 Next: Configure Security Settings。

步骤 2：为 HTTPS 侦听器配置安全设置

如果在上一步中创建了 HTTPS 侦听器，则配置必需的安全设置。否则，请转至向导中的下一页。

当您对负载均衡器侦听器使用 HTTPS 时，必须在负载均衡器上部署 SSL 证书。负载均衡器先使用此证书终止连接，然后解密来自客户端的请求，最后再将请求发送到目标。有关更多信息，请参阅 SSL 证书 (p. 26)。您还必须指定负载均衡器用于协商与客户端的 SSL 连接的安全策略。有关更多信息，请参阅安全策略 (p. 26)。

配置证书和安全策略

1. 对于 Select default certificate，执行下列操作之一：

- 如果使用 AWS Certificate Manager 创建或导入了证书，请选择 Choose a certificate from ACM，然后从 Certificate name 中选择证书。
- 如果使用 IAM 上传了证书，请选择 Choose a certificate from IAM，然后从 Certificate name 中选择证书。

2. 对于 Security policy，我们建议保留默认安全策略。
3. 选择 Next: Configure Security Groups。

步骤 3：配置安全组

您负载均衡器的安全组必须允许其通过侦听器端口和运行状况检查端口与已注册目标进行通信。控制台可代表您为负载均衡器创建一个具有允许此通信的规则的安全组。如果您愿意，也可创建一个安全组然后选择它。有关更多信息，请参阅 推荐的规则 (p. 19)。

为负载均衡器配置安全组

1. 选择 Create a new security group。
2. 为安全组键入名称和描述，或者保留默认名称和描述。此新安全组包含一条规则，该规则允许将流量传送到在 Configure Load Balancer 页面上为负载均衡器选择的端口。
3. 选择 Next: Configure Routing。

步骤 4：配置目标组

将目标注册到目标组。您在此步骤中配置的目标组将用作将请求转发到目标组的默认侦听器规则中的目标组。有关更多信息，请参阅 应用程序负载均衡器的目标组 (p. 39)。

配置目标组

1. 对于 Target group，保留默认值 New target group。
2. 对于 Name，键入目标组的名称。
3. 根据需要设置 Protocol 和 Port。
4. 对于目标类型，请选择 `instance` 通过实例 ID 指定目标或选择 `ip` 通过 IP 地址指定目标。
5. 对于 Health checks，保留默认运行状况检查设置。
6. 选择 Next: Register Targets。

步骤 5：配置目标组的目标

使用 应用程序负载均衡器，您可通过实例 ID 或 IP 地址注册目标，具体取决于您为目标组选择的目标类型。

通过实例 ID 注册目标

1. 对于 Instances，选择一个或多个实例。
2. 键入实例侦听器端口，然后选择 Add to registered。
3. 当您注册完实例后，选择 Next: Review。

通过 IP 地址注册目标

1. 对于每个要注册的 IP 地址，请执行以下操作：

a. 对于 Network，如果 IP 地址来自目标组 VPC 的子网，则选择该 VPC。否则，请选择 Other private IP address。
b. 对于 IP，键入 IP 地址。
c. 对于 Port，键入端口。
d. 选择 Add to list。

2. 在将 IP 地址添加到列表中后，选择 Next: Review。

步骤 6：创建负载均衡器

在创建负载均衡器之后，您可验证您的目标是否通过了初始运行状况检查，然后测试负载均衡器是否会将流量发送至您的目标。使用完负载均衡器之后，您可将其删除。有关更多信息，请参阅 删除负载均衡器 (p. 21)。

创建负载均衡器

1. 在 Review 页面上，选择 Create 。
2. 创建负载均衡器之后，选择 Close。
3. (可选) 要定义基于路径模式或主机名转发请求的其他侦听器规则，请参阅添加规则 (p. 28)。

应用程序负载均衡器的可用区

您可随时启用或禁用负载均衡器的可用区。在启用一个可用区后，负载均衡器会开始将请求路由到该可用区中的已注册目标。如果您确保每个启用的可用区均具有至少一个注册目标，则负载均衡器将具有最高效率。

在禁用一个可用区后，该可用区中的目标仍将注册到负载均衡器，但负载均衡器不会向这些目标路由请求。

使用控制台更新可用区

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器。
4. 在 Description 选项卡上的 Basic Configuration 下，选择 Edit Availability Zones。
5. 要启用一个可用区，请选中该可用区的复选框。如果该可用区有一个子网，则将选择此子网。如果该可用区有多个子网，请选择其中一个子网。请注意，您只能为每个可用区选择一个子网。
6. 要更改已启用的可用区的子网，请选中 Change subnet，然后选择其他子网之一。
7. 要删除可用区，请清除该可用区的复选框。
8. 选择 Save。

使用 AWS CLI 更新可用区

使用 set-subnets 命令。

应用程序负载均衡器的安全组

您必须确保负载均衡器可同时在侦听器端口和运行状况检查端口上与已注册目标进行通信。当您将侦听器添加到负载均衡器或更新负载均衡器所使用的目标组的运行状况检查端口来路由请求时，您必须验证与负载均衡器关联的安全组是否允许新端口上的双向流量。如果它们不允许，您可以编辑当前关联的安全组的规则或将其他安全组与负载均衡器关联。

推荐的规则

推荐规则取决于负载均衡器的类型 (面向 Internet 或内部)。

面向 Internet 的负载均衡器

入站		
源	端口范围	评论
0.0.0.0/0	*###*	在负载均衡器侦听器端口上允许所有入站流量
出站		
目的地	端口范围	评论
#####	*#####*	在实例侦听器端口上允许流向实例的出站流量
#####	*######*	在运行状况检查端口上允许流向实例的出站流量

内部负载均衡器

入站		
源	端口范围	评论
VPC CIDR	*###*	在负载均衡器侦听器端口上允许来自 VPC CIDR 的入站流量
出站		
目的地	端口范围	评论
#####	*#####*	在实例侦听器端口上允许流向实例的出站流量
#####	*######*	在运行状况检查端口上允许流向实例的出站流量

我们还建议您允许入站 ICMP 流量以支持路径 MTU 发现。有关更多信息，请参阅 Amazon EC2 用户指南（适用于 Linux 实例）中的路径 MTU 发现。

更新关联的安全组

您可以随时更新与负载均衡器关联的安全组。

使用控制台更新安全组

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器。
4. 在 Description 选项卡上的 Security 下，选择 Edit security groups。
5. 要将一个安全组与负载均衡器关联，请选择此安全组。要从负载均衡器中删除一个安全组，请清除该安全组。
6. 选择 Save。

使用 AWS CLI 更新安全组

使用 set-security-groups 命令。

Application Load Balancer 的 IP 地址类型

可以将 应用程序负载均衡器 配置为仅路由 IPv4 流量或同时路由 IPv4 和 IPv6 流量。有关更多信息，请参阅 IP 地址类型 (p. 14)。

IPv6 要求

- 一个面向 Internet 的负载均衡器。
- 您的 Virtual Private Cloud (VPC) 具有带关联的 IPv6 CIDR 块的子网。有关更多信息，请参阅 Amazon EC2 用户指南 中的 IPv6 地址。

使用控制台更新 IP 地址类型

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器。
4. 选择 Actions 和 Edit IP address type。
5. 对于 IP address type，选择 ipv4 可仅支持 IPv4 地址，选择 dualstack 可同时支持 IPv4 和 IPv6 地址。
6. 选择 Save。

使用 AWS CLI 更新 IP 地址类型

使用 set-ip-address-type 命令。

应用程序负载均衡器的标签

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。

当您用完标签时，可以从负载均衡器中将其删除。

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
- 标签密钥和值要区分大小写。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，因为它专为 AWS 使用预留。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。

使用控制台更新负载均衡器的标签

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。

3. 选择负载均衡器。
4. 在 Tags 选项卡上，选择 Add/Edit Tags，然后执行以下一项或多项操作：
   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，然后为 Key 和 Value 键入值。
   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 Save。

使用 AWS CLI 更新负载均衡器的标签

使用 add-tags 和 remove-tags 命令。

删除负载均衡器

在您的负载均衡器可用之后，您需要为保持其运行的每小时或部分小时支付费用。当您不再需要该负载均衡器时，可将其删除。当负载均衡器被删除之后，您便不再需要支付负载均衡器费用。

如果已启用删除保护，则无法删除负载均衡器。有关更多信息，请参阅 删除保护 (p. 14)。

请注意，删除负载均衡器不会影响其注册目标。例如，您的 EC2 实例将继续运行并仍注册到其目标组。要删除目标组，请参阅删除目标组 (p. 50)。

使用控制台删除负载均衡器

1. 如果您有一个指向负载均衡器的域的一个别名记录，请将它指向新的位置并等待 DNS 更改生效，然后再删除您的负载均衡器。
2. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
3. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
4. 选择负载均衡器，然后选择 Actions 和 Delete。
5. 当系统提示进行确认时，选择 Yes, Delete。

使用 AWS CLI 删除负载均衡器

使用 delete-load-balancer 命令。

应用程序负载均衡器的侦听器

在开始使用应用程序负载均衡器之前，您必须添加一个或多个侦听器。侦听器是一个使用您配置的协议和端口检查连接请求的进程。为侦听器定义的规则可以确定负载均衡器将请求路由到一个或多个目标组中的目标的方式。

内容

侦听器配置

侦听器支持以下协议和端口：

- 协议：HTTP、HTTPS
- 端口：1-65535

可以使用 HTTPS 侦听器将加密和解密的工作交给负载均衡器完成，以便应用程序可以专注于其业务逻辑。如果侦听器协议为 HTTPS，您必须在侦听器上确切地部署一个 SSL 服务器证书。有关更多信息，请参阅 Application Load Balancer 的 HTTPS 侦听器 (p. 25)。

Application Load Balancer 为 WebSockets 提供本机支持。您可以对 HTTP 和 HTTPS 侦听器同时使用 WebSocket。

Application Load Balancer 为将 HTTP/2 与 HTTPS 侦听器一起使用提供本机支持。使用一个 HTTP/2 连接最多可以并行发送 128 个请求。负载均衡器将这些请求转换为单独的 HTTP/1.1 请求，并将它们分配给目标组中的正常目标。由于 HTTP/2 可以更高效地使用前端连接，您可能注意到客户端与负载均衡器之间的连接较少。无法使用 HTTP/2 的服务器推送功能。

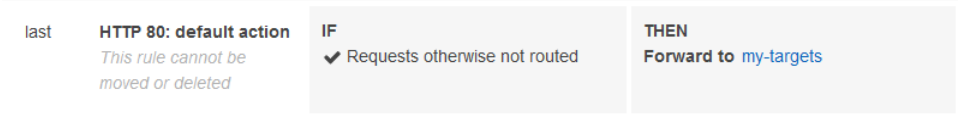
侦听器规则

每个侦听器都具有默认规则，您也可以选择定义其他规则。每个规则都包含优先级、转发操作、可选身份验证操作（针对 HTTPS 侦听器）、可选主机条件和可选路径条件。

默认规则

创建侦听器时，请为默认规则定义操作。默认规则不能有条件。如果未满足侦听器的任一规则条件，则将执行默认规则的操作。

下面是控制台中所示的默认规则的示例：



规则优先级

每个规则都有一个优先级。规则是按优先级顺序 (从最低值到最高值) 计算的。最后评估默认规则。您可以随时更改非默认规则的优先级。您不能更改默认规则的优先级。有关更多信息，请参阅 重新排序规则 (p. 30)。

规则操作

每个规则操作都具有执行操作所需的类型、顺序和信息。以下是支持的操作类型：

`authenticate-cognito`

[HTTPS 侦听器] 使用 Amazon Cognito 验证用户身份。

`authenticate-oidc`

[HTTPS 侦听器] 使用符合 OpenID Connect (OIDC) 条件的身份提供商验证用户身份。

`forward`

将请求转发到指定目标组。

先执行顺序值最小的操作。每个规则必须包含一个 `forward` 操作。`forward` 操作必须放在最后执行。可以随时编辑规则。有关更多信息，请参阅 编辑规则 (p. 29)。

规则条件

规则条件有两种类型：主机和路径。每个规则最多可以具有一个主机条件和一个路径条件。当规则的条件满足时，将执行其操作。

主机条件

您可使用主机条件定义一些规则，用于根据主机标头中的主机名将请求转发到不同的目标组 (也称为基于主机的路由)。这使您能够使用单个负载均衡器支持多个域。

每个主机条件有一个主机名。如果主机标头中的主机名与侦听器规则中的主机名完全匹配，则将使用该规则来路由请求。

主机名不区分大小写，长度上限为 128 个字符，并且可包含以下任何字符。可包含多达三个通配符。

- A-Z、a-z、0-9
- - .
- * (匹配 0 个或多个字符)
- ?(完全匹配 1 个字符)

主机名示例

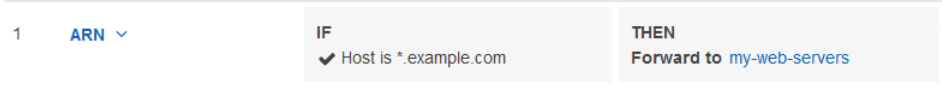
- **`example.com`**

- **test.example.com**
- ***.example.com**

规则 ***.example.com** 与 **test.example.com** 匹配，但与 **example.com** 不匹配。

控制台示例

下面是具有控制台中所示的主机条件的规则的示例。如果主机标头中的主机名与 ***.example.com** 匹配，则请求将转发到名为 **my-web-servers** 的目标组。有关更多信息，请参阅 添加规则 (p. 28)。



路径条件

您可使用路径条件定义根据请求中的 URL 将请求转发到不同的目标组 (也称为基于路径的路由) 的规则。

每个路径条件都有一个路径模式。如果请求中的 URL 与侦听器规则中的路径模式完全匹配，则将使用该规则来路由请求。

路径模式区分大小写，长度最多为 128 个字符，并且可包含以下任何字符。可包含多达三个通配符。

- A-Z、a-z、0-9
- _ - . $ / ~ " ' @ : +
- & (使用 &)
- * (匹配 0 个或多个字符)
- ?(完全匹配 1 个字符)

示例路径模式

- `/img/*`
- `/js/*`

路径模式用于路由请求，而不是更改请求。例如，如果一个规则具有路径模式 `/img/*`，此规则会将 `/img/picture.jpg` 的请求作为 `/img/picture.jpg` 的请求转发给指定目标组。

控制台示例

下面是具有控制台中所示的路径条件的规则的示例。如果请求中的 URL 与 `/img/*` 匹配，则会将请求转发到名为 **my-targets** 的目标组。有关更多信息，请参阅 添加规则 (p. 28)。

为 Application Load Balancer 创建侦听器

侦听器是用于检查连接请求的进程。您可在创建负载均衡器时定义侦听器，并可随时向负载均衡器添加侦听器。

先决条件

- 您必须为默认侦听器规则指定目标组。有关更多信息，请参阅 创建目标组 (p. 43)。
- 如果您创建了 HTTPS 侦听器，则必须指定证书和安全策略。负载均衡器先使用证书终止连接，然后解密来自客户端的请求，最后再将请求路由到目标。有关更多信息，请参阅 SSL 证书 (p. 26)。负载均衡器在协商与客户端的 SSL 连接时会使用安全策略。有关更多信息，请参阅 安全策略 (p. 26)。

添加侦听器

您为侦听器配置用于从客户端连接到负载均衡器的协议和端口，并为默认侦听器规则配置目标组。有关更多信息，请参阅 侦听器配置 (p. 22)。

使用控制台添加侦听器

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器，然后依次选择 Listeners 和 Add listener。
4. 对于 Protocol : port (协议: 端口)，选择 HTTP 或 HTTPS。保留默认端口或键入其他端口。
5. （可选，HTTPS 侦听器）要验证用户的身份，对于 Default actions (默认操作)，选择 Add action (添加操作)、Authenticate (身份验证) 并提供请求的信息。要保存操作，请选择选中标记图标。有关更多信息，请参阅 使用 应用程序负载均衡器 验证用户身份 (p. 33)。
6. 对于 Default actions (默认操作)，选择 Add action (添加操作)、Forward to (转发至) 并选择目标组。要保存操作，请选择选中标记图标。
7. [HTTPS 侦听器] 对于 Security policy (安全策略)，建议保留默认安全策略。
8. [HTTPS 侦听器] 对于 Default SSL certificate (默认 SSL 证书)，执行下列操作之一：
   - 如果使用 AWS Certificate Manager 创建或导入了证书，请选择 From ACM (从 ACM) 并选择证书。
   - 如果使用 IAM 上传了证书，则选择 From IAM (从 IAM) 并选择证书。
9. 选择 Save (保存)。
10. (可选) 要定义基于路径模式或主机名转发请求的其他侦听器规则，请参阅添加规则 (p. 28)。

使用 AWS CLI 添加侦听器

使用 create-listener 命令可创建侦听器和默认规则，使用 create-rule 命令可定义更多侦听器规则。

Application Load Balancer 的 HTTPS 侦听器

您可以创建使用加密连接的侦听器 (也称为 SSL 卸载)。此功能支持在您的负载均衡器与启动 SSL 或 TLS 会话的客户端之间进行流量加密。

要使用 HTTPS 侦听器，您必须在负载均衡器上部署 SSL/TLS 服务器证书。负载均衡器先使用此证书终止连接，然后解密来自客户端的请求并将请求发送到目标。

Elastic Load Balancing 使用安全套接字层 (SSL) 协商配置 (称为安全策略) 在客户端与负载均衡器之间协商 SSL 连接。安全策略是协议和密码的组合。 协议在客户端与服务器之间建立安全连接，确保在客户端与负载均衡器之间传递的所有数据都是私密数据。密码是使用加密密钥创建编码消息的加密算法。协议使用多种密码对 Internet 上的数据进行加密。在 连接协商过程中，客户端和负载均衡器会按首选项顺序提供各自支持的密码和协议的列表。默认情况下，会为安全连接选择服务器列表中与任何一个客户端的密码匹配的第一个密码。

对于客户端连接或目标连接，Application Load Balancer 不支持 SSL 重新协商。

SSL 证书

负载均衡器使用 X.509 证书 (SSL/TLS 服务器证书)。证书是由证书颁发机构 (CA) 颁发的数字化身份。证书包含标识信息、有效期限、公有密钥、序列号以及发布者的数字签名。

在创建用于负载均衡器的证书时，您必须指定域名。

我们建议您使用 AWS Certificate Manager (ACM) 为您的负载均衡器创建证书。ACM 与 Elastic Load Balancing 集成，以便您可以在负载均衡器上部署证书。有关更多信息，请参阅 AWS Certificate Manager 用户指南。

此外，还可以使用 SSL/TLS 工具创建证书签名请求 (CSR)，然后由 CA 签署此 CSR 以生成证书，并将证书导入 ACM，或将证书上传至 AWS Identity and Access Management (IAM)。有关将证书导入 ACM 中的更多信息，请参阅 AWS Certificate Manager 用户指南 中的导入证书。有关将证书上传到 IAM 的更多信息，请参阅IAM 用户指南中的使用服务器证书。

> **Important**
>
> 您无法通过与 ACM 集成在负载均衡器上安装带有 4096 位 RSA 密钥或 EC 密钥的证书。您必须将带有 4096 位 RSA 密钥或 EC 密钥的证书上传到 IAM，以便将它们与负载均衡器结合使用。

创建 HTTPS 侦听器时，请指定默认证书。您可以通过添加更多证书来为侦听器创建可选证书列表。这使负载均衡器能够在同一端口上支持多个域并为每个域提供一个不同的证书。默认情况下，侦听器的默认证书不会添加到证书列表中。有关更多信息，请参阅 更新服务器证书 (p. 31)。

客户端可以使用服务器名称标识 (SNI) 协议扩展来指定其尝试连接的主机名。如果此主机名与证书列表中的证书不匹配，则负载均衡器将选择默认证书。如果此主机名与证书列表中的一个证书匹配，则负载均衡器将选择此证书。如果客户端提供的主机名与证书列表中的多个证书匹配，则负载均衡器将选择客户端可支持的最佳证书。根据以下标准，按下面的顺序选择证书：

- 公有密钥算法 (ECDSA 优先于 RSA)
- 哈希算法 (SHA 优先于 MD5)
- 密钥长度 (首选最大值)
- 有效期

负载均衡器访问日志条目指示客户端指定的主机名和向客户端提供的证书。有关更多信息，请参阅 访问日志条目 (p. 60)。

安全策略

您可以选择用于前端连接的安全策略。`ELBSecurityPolicy-2016-08` 安全策略始终用于后端连接。Application Load Balancer 不支持自定义安全策略。

Elastic Load Balancing 为 Application Load Balancer 提供以下安全策略：

- `ELBSecurityPolicy-2016-08`
- `ELBSecurityPolicy-FS-2018-06`
- `ELBSecurityPolicy-TLS-1-2-2017-01`
- `ELBSecurityPolicy-TLS-1-2-Ext-2018-06`
- `ELBSecurityPolicy-TLS-1-1-2017-01`
- `ELBSecurityPolicy-2015-05`
- `ELBSecurityPolicy-TLS-1-0-2015-04`

建议将 `ELBSecurityPolicy-2016-08` 策略用于一般用途。如果需要向前保密 (FS)，可使用 `ELBSecurityPolicy-FS-2018-06` 策略。您可以使用 `ELBSecurityPolicy-TLS` 策略之一，以满足需要禁用特定 TLS 协议版本的合规性和安全标准，或者支持需要已弃用密码的旧客户端。只有一小部分 Internet 客户端需要 TLS 版本 1.0。要查看针对负载均衡器的请求的 TLS 协议版本，请为负载均衡器启用访问日志记录并查看访问日志。有关更多信息，请参阅访问日志 (p. 59)。

下表描述了为 Application Load Balancer 定义的安全策略。

安全策略	2016-08 *	FS-2018-0	TLS-1-2	TLS-1-2-Ext	TLS-1-1	TLS-1-0 †
TLS 协议						
Protocol-TLSv1	♦	♦				♦
Protocol-TLSv1.1	♦	♦			♦	♦
Protocol-TLSv1.2	♦	♦	♦	♦	♦	♦
TLS 密码						
ECDHE-ECDSA-AES128- GCM-SHA256	♦	♦	♦	♦	♦	♦
ECDHE-RSA-AES128- GCM-SHA256	♦	♦	♦	♦	♦	♦
ECDHE-ECDSA-AES128-SHA256	♦	♦	♦	♦	♦	♦
ECDHE-RSA-AES128-SHA256	♦	♦	♦	♦	♦	♦
ECDHE-ECDSA-AES128-SHA	♦	♦		♦	♦	♦
ECDHE-RSA-AES128-SHA	♦	♦		♦	♦	♦
ECDHE-ECDSA-AES256- GCM-SHA384	♦	♦	♦	♦	♦	♦
ECDHE-RSA-AES256- GCM-SHA384	♦	♦	♦	♦	♦	♦
ECDHE-ECDSA-AES256-SHA384	♦	♦	♦	♦	♦	♦
ECDHE-RSA-AES256-SHA384	♦	♦	♦	♦	♦	♦
ECDHE-RSA-AES256-SHA	♦	♦		♦	♦	♦
ECDHE-ECDSA-AES256-SHA	♦	♦		♦	♦	♦
AES128-GCM-SHA256	♦		♦	♦	♦	♦
AES128-SHA256	♦		♦	♦	♦	♦
AES128-SHA	♦			♦	♦	♦
AES256-GCM-SHA384	♦		♦	♦	♦	♦
AES256-SHA256	♦		♦	♦	♦	♦
AES256-SHA	♦			♦	♦	♦
DES-CBC3-SHA						♦

* Application Load Balancer 的 `ELBSecurityPolicy-2016-08` 和 `ELBSecurityPolicy-2015-05` 安全策略相同。

† 除非您必须支持需要 DES-CBC3-SHA 密码 (这是一种弱密码) 的旧客户端，否则不要使用此安全策略。

要使用 AWS CLI 查看 Application Load Balancer 的安全策略的配置，请使用 describe-ssl-policies 命令。

更新安全策略

在创建 HTTPS 侦听器时，您可以选择满足您的需求的安全策略。添加新的安全策略后，您可以将 HTTPS 侦听器更新为使用此新安全策略。Application Load Balancer 不支持自定义安全策略。

使用控制台更新安全策略

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器，然后选择 Listeners。
4. 选中 HTTPS 侦听器对应的复选框，然后选择 Edit (编辑)。
5. 对于 Security policy (安全策略)，选择安全策略。
6. 选择 Update。

使用 AWS CLI 更新安全策略

使用 modify-listener 命令。

Application Load Balancer 的侦听器规则

为侦听器定义的规则可确定负载均衡器如何将请求路由到一个或多个目标组中的目标。

每个规则都包含优先级、一个或多个操作、可选主机条件和可选路径条件。有关更多信息，请参阅 侦听器规则 (p. 22)。

> Note
>
> 控制台会显示每个规则的相对序列号，而不是规则优先级。您可以使用 AWS CLI 或 Elastic Load Balancing API 对规则进行描述，以获得规则优先级。

先决条件

规则会将请求路由至其目标组。在创建规则或更新规则的目标组之前，请创建目标组并为其添加目标。有关更多信息，请参阅 创建目标组 (p. 43)。

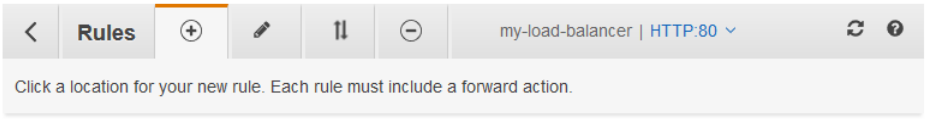
添加规则

您可在创建侦听器时定义默认规则，并可随时定义其他非默认规则。

使用控制台添加规则

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器，然后选择 Listeners。
4. 对于要更新的侦听器，选择 View/edit rules。

5. 选择菜单栏中的 Add rules 图标 (加号) 以在您可按优先级顺序插入规则的位置添加 Insert Rule 图标。



6. 按下面所示定义规则：

   a. 选择 Insert Rule。

   b. （可选）要配置基于主机的路由，请选择 Add condition (添加条件)、Host is (主机为) 并键入主机名（例如，`*.example.com`）。要保存此条件，请选择选中标记图标。

   c. （可选）要配置基于路径的路由，请选择 Add condition (添加条件)、Path is (路径为) 并键入路径模式（例如，`/img/*`）。要保存此条件，请选择选中标记图标。

   d. （可选，HTTPS 侦听器）要验证用户的身份，请选择 Add action (添加操作)、Authenticate (身份验证) 并提供请求的信息。有关更多信息，请参阅 使用 应用程序负载均衡器 验证用户身份 (p. 33)。

   e. 要添加转发操作，请选择 Add action (添加操作)、Forward to (转发至) 并选择目标组。每个规则必须刚好具有一个转发操作。

   f. (可选) 要更改规则的顺序，请使用箭头。默认规则始终具有最低优先级。

   g. 选择 Save。

7. 要离开此屏幕，请选择菜单栏中的 Back to the load balancer 图标 (后退按钮)。

使用 AWS CLI 添加规则

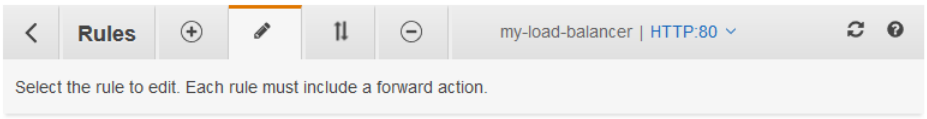
使用 create-rule 命令创建规则。使用 describe-rules 命令查看规则的相关信息。

编辑规则

您可随时编辑规则的操作和条件。

使用控制台编辑规则

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器，然后选择 Listeners。
4. 对于要更新的侦听器，选择 View/edit rules。
5. 选择菜单栏中的 Edit rules 图标 (铅笔)。



6. 对于要编辑的规则，选择 Edit rules 图标 (铅笔)。
7. （可选）按需修改条件和操作。例如，可以编辑条件或操作（铅笔图标）、添加路径条件（如果还没有）、添加主机条件（如果还没有）、为 HTTPS 侦听器添加规则的身份验证操作，或删除条件或操作（垃圾桶图标）。无法向默认规则添加条件。每个规则必须刚好具有一个转发操作。

8. 选择 Update。
9. 要离开此屏幕，请选择菜单栏中的 Back to the load balancer 图标 (后退按钮)。

使用 AWS CLI 编辑规则

使用 modify-rule 命令。

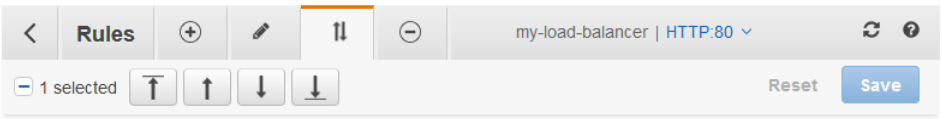
重新排序规则

规则是按优先级顺序 (从最低值到最高值) 计算的。最后评估默认规则。您可以随时更改非默认规则的优先级。您不能更改默认规则的优先级。

> **Note**
>
> 控制台会显示每个规则的相对序列号，而不是规则优先级。使用控制台以重新排序规则时，规则将根据现有规则优先级，获得新规则优先级。要将规则的优先级设为指定值，请使用 AWS CLI 或 Elastic Load Balancing API。

使用控制台为规则重新排序

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器，然后选择 Listeners。
4. 对于要更新的侦听器，选择 View/edit rules。
5. 选择菜单栏中的 Reorder rules图标 (箭头)。



6. 选中规则旁的复选框，然后使用箭头为规则指定新的优先级。默认规则始终具有最低优先级。
7. 为规则重新排序之后，选择 Save。
8. 要离开此屏幕，请选择菜单栏中的 Back to the load balancer 图标 (后退按钮)。

使用 AWS CLI 更新规则优先级

使用 set-rule-priorities 命令。

删除规则

您可以随时删除侦听器的非默认规则。您不能删除侦听器的默认规则。当您删除侦听器时，也会删除它的所有规则。

使用控制台删除规则

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器，然后选择 Listeners。
4. 对于要更新的侦听器，选择 View/edit rules。
5. 在菜单栏中选择 Delete rules (删除规则) 图标（减号）。
6. 选中规则对应的复选框，然后选择 Delete (删除)。无法删除侦听器的默认规则。
7. 要离开此屏幕，请选择菜单栏中的 Back to the load balancer 图标 (后退按钮)。

使用 AWS CLI 删除规则

使用 delete-rule 命令。

更新服务器证书

创建 HTTPS 侦听器时，请指定默认证书。您也可以通过添加其他证书来为侦听器创建证书列表。

每个证书都有有效期限。您必须确保在其有效期限结束前续订或替换证书。续订或替换证书不影响负载均衡器节点已收到的进行中的请求，并暂停指向正常运行的目标的路由。续订证书之后，新的请求将使用续订后的证书。更换证书之后，新的请求将使用新证书。

您可以按如下方式管理证书续订和替换：

- 由 AWS Certificate Manager 提供、部署在负载均衡器上的证书可以自动续订。ACM 会尝试在到期之前续订证书。有关更多信息，请参阅 AWS Certificate Manager 用户指南 中的托管续订。
- 如果您将证书导入 ACM，则必须监视证书的到期日期并在到期前续订。有关更多信息，请参阅 AWS Certificate Manager 用户指南 中的导入证书。
- 如果您已将证书导入 IAM 中，则必须创建一个新证书，将该新证书导入 ACM 或 IAM 中，将该新证书添加到负载均衡器，并从负载均衡器删除过期的证书。

限制

ACM 支持具有 4096 密钥长度的 RSA 证书和 EC 证书。但是，您无法通过与 ACM 集成在负载均衡器上安装这些证书。您必须将这些证书上传到 IAM，以便将它们与负载均衡器结合使用。

添加证书

您可使用以下过程将证书添加到侦听器的证书列表。虽然侦听器的默认证书在默认情况下不会添加到证书列表中，但您可以将默认证书添加到该证书列表中。

使用控制台添加证书

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。

2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器，然后选择 Listeners。
4. 对于要更新的 HTTPS 侦听器，请选择 View/edit certificates (查看/编辑证书)，这将显示默认证书，后跟已添加到侦听器的任何其他证书。
5. 选择菜单栏中的 Add certificates 图标 (加号)，这将显示默认证书，后跟由 ACM 和 IAM 管理的任何其他证书。如果已将证书添加到侦听器，则其复选框处于选中和禁用状态。
6. 要添加已由 ACM 或 IAM 管理的证书，请选中证书对应的复选框并选择 Add (添加)。
7. 如果您有一个未由 ACM 或 IAM 管理的证书，则按如下方式将其导入 ACM 中，并将其添加到侦听器：

   a. 选择 Import certificate。
   b. 对于 Certificate private key，粘贴证书的 PEM 编码的未加密私有密钥。
   c. 对于 Certificate body，粘贴 PEM 编码的证书。
   d. (可选) 对于 Certificate chain，粘贴 PEM 编码的证书链。
   e. 选择 Import。新导入的证书将显示在可用证书列表中并处于选中状态。
   f. 选择 Add。
8. 要离开此屏幕，请选择菜单栏中的 Back to the load balancer 图标 (后退按钮)。

使用 AWS CLI 添加证书

使用 add-listener-certificates 命令。

替换默认证书

您可以使用以下过程替换侦听器的默认证书。

使用控制台更改默认证书

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器，然后选择 Listeners。
4. 选中侦听器对应的复选框，然后选择 Edit (编辑)。
5. 对于 Default SSL certificate (默认 SSL 证书)，请执行下列操作之一：

   - 如果使用 AWS Certificate Manager 创建或导入了证书，请选择 From ACM (从 ACM) 并选择证书。
   - 如果使用 IAM 上传了证书，则选择 From IAM (从 IAM) 并选择证书。
6. 选择 Save。

使用 AWS CLI 更改默认证书

使用 modify-listener 命令。

删除证书

可以随时删除 HTTPS 侦听器的非默认证书。不能使用此过程删除 HTTPS 侦听器的默认证书。

使用控制台删除证书

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器，然后选择 Listeners。

4. 对于要更新的侦听器，请选择 View/edit certificates (查看/编辑证书)，这将显示默认证书，后跟已添加到侦听器的任何其他证书。
5. 在菜单栏中选择 Remove certificates 图标 (减号)。
6. 选中证书对应的复选框，然后选择 Remove (删除)。
7. 要离开此屏幕，请选择菜单栏中的 Back to the load balancer 图标 (后退按钮)。

使用 AWS CLI 删除证书

使用 remove-listener-certificates 命令。

使用 应用程序负载均衡器 验证用户身份

可以将 应用程序负载均衡器 配置为在用户访问应用程序时安全验证用户的身份。这使您可以将验证用户身份的工作交给负载均衡器完成，以便应用程序可以专注于其业务逻辑。

支持以下使用案例：

- 通过符合 OpenID Connect (OIDC) 条件的身份提供商 (IdP) 验证用户身份。
- 通过知名社交 IdP（如 Amazon、Facebook 或 Google）、通过 Amazon Cognito 支持的用户池验证用户的身份。
- 通过企业身份、使用 SAML、LDAP 或 Microsoft AD、通过 Amazon Cognito 支持的用户池验证用户身份。

准备使用符合 OIDC 条件的 IdP

如果要将符合 OIDC 条件的 IdP 与 应用程序负载均衡器 一起使用，请执行以下操作：

- 使用 IdP 创建新的 OIDC 应用程序。必须配置客户端 ID 和客户端密钥。
- 获取 IdP 发布的以下终端节点：授权终端节点、令牌终端节点和用户信息终端节点。可以在清晰的配置中找到此信息。
- 在 IdP 中将以下重定向 URL 之一列入白名单（无论您的用户将使用哪种 IdP 应用程序），其中 DNS 是负载均衡器的域名，CNAME 是应用程序的 DNS 别名：
  - https://*DNS*/oauth2/idpresponse
  - https://*CNAME*/oauth2/idpresponse

准备使用 Amazon Cognito

如果要将 Amazon Cognito 用户池 应用程序负载均衡器 一起使用，请执行以下操作：

- 创建用户池。有关更多信息，请参阅 Amazon Cognito 开发者指南 中的 Amazon Cognito 用户池。
- 创建用户池客户端。必须将客户端配置为生成客户端密钥，使用代码授予流程并支持与负载均衡器所用相同的 OAuth 范围。有关更多信息，请参阅 Amazon Cognito 开发者指南 中的配置用户池应用程序客户端。
- 创建用户池域。有关更多信息，请参阅 Amazon Cognito 开发者指南 中的为用户池添加域名。
- 要与社交或企业 IdP 联合，请在联合身份验证部分中启用 IdP。有关更多信息，请参阅 Amazon Cognito 开发者指南 中的将社交登录添加到用户池或将“使用 SAML IdP 登录”添加到用户池。
- 在 Amazon Cognito 的回调 URL 字段中将以下重定向 URL 列入白名单，其中 DNS 是负载均衡器的域名，CNAME 是应用程序（如果正在使用）的 DNS 别名：
  - https://*DNS*/oauth2/idpresponse
  - https://*CNAME*/oauth2/idpresponse

- 在 IdP 应用程序的回调 URL 中将您的用户池域列入白名单。使用 IdP 的格式。例如：
  - https://*domain-prefix*.auth.*region*.amazoncognito.com/saml2/idpresponse
  - https://*user-pool-domain*/oauth2/idpresponse

要启用 IAM 用户以将负载均衡器配置为使用 Amazon Cognito 验证用户身份，必须授予调用 `cognito-idp:DescribeUserPoolClient` 操作的用户权限。

准备使用 Amazon CloudFront

如果您在应用程序负载均衡器前使用 CloudFront 分配，请启用以下设置：

- 查询字符串转发和缓存（全部）
- Cookie 转发（全部）
- 基于缓存的请求标头（全部）

配置用户身份验证

通过为一个或多个侦听器规则创建身份验证操作来配置用户身份验证。HTTPS 侦听器仅支持 `authenticate-cognito` 和 `authenticate-oidc` 操作类型。有关对应字段的描述，请参阅 Elastic Load Balancing API 参考第 2015-12-01 版 中的 AuthenticateCognitoActionConfig 和 AuthenticateOidcActionConfig。

默认情况下，`SessionTimeout` 字段设置为 7 日。如果需要更短的会话，可将会话超时配置为短至 1 秒。有关更多信息，请参阅 身份验证注销和会话超时 (p. 37)。

视应用程序的情况设置 `OnUnauthenticatedRequest` 字段。例如：

- 需要用户使用社交或企业身份登录的应用程序 - 这由默认选项 `authenticate` 支持。如果用户未登录，则负载均衡器会将请求重定向到 IdP 授权终端节点并且 IdP 将提示用户使用其用户界面登录。
- 为已登录用户提供个性化视图或为未登录用户提供常规视图的应用程序 - 要支持此类型的应用程序，请使用 `allow` 选项。如果用户已登录，则负载均衡器将提供用户索赔并且应用程序可以提供个性化视图。如果用户未登录，则负载均衡器将转发请求而不提供用户索赔并且应用程序可以提供常规视图。
- 具有每隔几秒就加载的 JavaScript 的单页面应用程序 - 默认情况下，在身份验证会话 Cookie 到期后，AJAX 调用将重定向至 IdP 并受阻。如果使用 `deny` 选项，则负载均衡器将针对这些 AJAX 调用返回“HTTP 401 未授权”错误。

负载均衡器必须能够与 IdP 令牌终端节点 (`TokenEndpoint`) 和 IdP 用户信息终端节点 (`UserInfoEndpoint`) 通信。验证负载均衡器的安全组和 VPC 的网络 ACL 是否允许至这些终端节点的出站访问。

使用以下 create-rule 命令配置用户身份验证。

```
aws elbv2 create-rule --listener-arn listener-arn --priority 10 \
--conditions Field=path-pattern,Values="/login" --actions file://actions.json
```

下面是指定 `authenticate-oidc` 操作和 `forward` 操作的 `actions.json` 文件的示例。

```
[{
    "Type": "authenticate-oidc",
    "AuthenticateOidcConfig": {
        "Issuer": "https://idp-issuer.com",
        "AuthorizationEndpoint": "https://authorization-endpoint.com",
        "TokenEndpoint": "https://token-endpoint.com",
```

```
        "UserInfoEndpoint": "https://user-info-endpoint.com",
        "ClientId": "abcdefghijklmnopqrstuvwxyz123456789",
        "ClientSecret": "12345678901234567890123456789O",
        "SessionCookieName": "my-cookie",
        "SessionTimeout": 3600,
        "Scope": "email",
        "AuthenticationRequestExtraParams": {
            "display": "page",
            "prompt": "login"
        },
        "OnUnauthenticatedRequest": "deny"
    },
    "Order": 1
},
{
    "Type": "forward",
    "TargetGroupArn": "arn:aws-cn:elasticloadbalancing:region-code:account-
id:targetgroup/target-group-name/target-group-id",
    "Order": 2
}]
```

下面是指定 authenticate-cognito 操作和 forward 操作的 actions.json 文件的示例。

```
[{
    "Type": "authenticate-cognito",
    "AuthenticateCognitoConfig": {
        "UserPoolArn": "arn:aws-cn:cognito-idp:region-code:account-id:userpool/user-pool-
id",
        "UserPoolClientId": "abcdefghijklmnopqrstuvwxyz123456789",
        "UserPoolDomain": "userPoolDomain1",
        "SessionCookieName": "my-cookie",
        "SessionTimeout": 3600,
        "Scope": "email",
        "AuthenticationRequestExtraParams": {
            "display": "page",
            "prompt": "login"
        },
        "OnUnauthenticatedRequest": "deny"
    },
    "Order": 1
},
{
    "Type": "forward",
    "TargetGroupArn": "arn:aws-cn:elasticloadbalancing:region-code:account-
id:targetgroup/target-group-name/target-group-id",
    "Order": 2
}]
```

有关更多信息，请参阅 侦听器规则 (p. 22)。

身份验证流程

Elastic Load Balancing 使用 OIDC 授权代码流程，其中包括以下步骤。

1. 当满足具有身份验证操作的规则的条件时，负载均衡器将检查请求标头中的身份验证会话 Cookie。如果 Cookie 不存在，则负载均衡器会将用户重定向到 IdP 授权终端节点，以便 IdP 可对用户进行身份验证。
2. 验证用户身份之后，IdP 会使用授权授予代码将用户重定向回负载均衡器。负载均衡器会将此代码发送给 IdP 令牌终端节点，以获取 ID 令牌和访问令牌。
3. 负载均衡器验证 ID 令牌之后，它将访问令牌与 IdP 用户信息终端节点交换以获取用户索赔。
4. 负载均衡器将创建身份验证会话 Cookie 并将其发送到客户端，以便客户端的用户代理可在发出请求时将 Cookie 发送到负载均衡器。由于大多数浏览器将 Cookie 限制为 4K 大小，因此负载均衡器会将超

出 4K 大小的 Cookie 分片为多个 Cookie。如果从 IdP 收到的用户索赔和访问令牌的总大小超出 11K 大小，则负载均衡器将返回错误。

5. 负载均衡器将用户索赔发送到 HTTP 标头中的目标。有关更多信息，请参阅 用户索赔编码和签名验证 (p. 36)。
6. 如果 IdP 在 ID 令牌中提供了有效的刷新令牌，则负载均衡器将保存刷新令牌并在访问令牌过期时使用刷新令牌刷新用户索赔，直至会话超时或 IdP 刷新失败。如果用户注销，刷新将失败并且负载均衡器会将用户重定向到 IdP 授权终端节点。这使负载均衡器能够在用户注销后删除会话。有关更多信息，请参阅 身份验证注销和会话超时 (p. 37)。

用户索赔编码和签名验证

在负载均衡器成功验证用户身份之后，它会将从 IdP 收到的用户索赔发送给目标。负载均衡器先为用户索赔签名，以便应用程序可以验证该签名并验证索赔是负载均衡器发送的。

负载均衡器添加以下 HTTP 标头：

`x-amzn-oidc-accesstoken`

令牌终端节点中的访问令牌（明文格式）。

`x-amzn-oidc-identity`

用户信息终端节点中的主题字段 (sub)（明文格式）。

`x-amzn-oidc-data`

用户索赔（JSON Web 令牌 (JWT) 格式）。

需要完整用户索赔的应用程序可使用任何标准 JWT 库。JWT 格式包括 base64 URL 编码的标头、负载和签名。JWT 签名为 ECDSA + P-256 + SHA256。

JWT 标头为具有以下字段的 JSON 对象：

```
{
   "alg": "algorithm",
   "kid": "12345678-1234-1234-1234-123456789012",
   "signer": "arn:aws-cn:elasticloadbalancing:region-code:account-id:loadbalancer/app/load-balancer-name/load-balancer-id",
   "iss": "url",
   "client": "client-id",
   "exp": "expiration"
}
```

JWT 负载是一个 JSON 对象，该对象包含从 IdP 用户信息终端节点接收的用户索赔。

```
{
   "sub": "1234567890",
   "name": "name",
   "email": "alias@example.com",
   ...
}
```

由于负载均衡器不会对用户索赔加密，建议将目标组配置为使用 HTTPS。如果将目标组配置为使用 HTTP，请务必使用安全组限制至负载均衡器的流量。还建议在基于索赔执行任何授权之前验证签名。要获取公钥，请从 JWT 标头中获取密钥 ID 并使用它从以下区域终端节点查找公钥：

```
https://public-keys.auth.elb.region.amazonaws.com/key-id
```

对于 AWS GovCloud (US)，终端节点如下所示：

```
https://s3-us-gov-west-1.amazonaws.com/aws-elb-public-keys-prod-us-gov-west-1/key-id
```

以下示例演示如何在 Python 中获取公钥：

```
import jwt
import requests

# Step 1: Get the key id from JWT headers (the kid field)
encoded_jwt = headers.dict['x-amzn-oidc-data']
jwt_headers = encoded_jwt.split('.')[0]
decoded_jwt_headers = base64.b64decode(jwt_headers)
decoded_json = json.loads(decoded_jwt_headers)
kid = decoded_json['kid']

# Step 2: Get the public key from regional endpoint
url = 'https://public-keys.auth.elb' + region + '.amazonaws.com/' + kid
req = requests.get(url)
pub_key = req.text

# Step 3: Get the payload
payload = jwt.decode(encoded_jwt, pub_key, algorithms=['ES256'])
```

身份验证注销和会话超时

当应用程序需要注销经身份验证的用户时，应将身份验证会话 Cookie 的到期时间设置为 -1 并将客户端重定向到 IdP 注销终端节点（如果 IdP 支持一个终端节点）。为防止用户重复使用已删除的 Cookie，建议为访问令牌配置合理的短过期时间。如果客户端为负载均衡器提供了授权会话 Cookie（具有已到期的访问令牌和非 NULL 刷新令牌），负载均衡器将联系 IdP，以确定用户是否仍处于登录状态。

刷新令牌和会话超时将一起运行，如下所示：

- 如果会话超时短于访问令牌过期时间，则负载均衡器将遵守会话超时并在身份验证会话超时之后让用户再次登录。
- 如果会话超时长于访问令牌过期时间并且 IdP 不支持刷新令牌，则负载均衡器会将身份验证会话一直保留到其超时，之后让用户再次登录。
- 如果会话超时长于访问令牌过期时间并且 IdP 支持刷新令牌，则负载均衡器将在每次访问令牌到期时刷新用户会话。仅当身份验证会话超时或刷新流程失败之后，负载均衡器才会让用户再次登录。

删除 Application Load Balancer 的侦听器

可以随时删除侦听器。当您删除负载均衡器时，也会删除它的所有侦听器。

使用控制台删除侦听器

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择负载均衡器，然后选择 Listeners。
4. 选中 HTTPS 侦听器对应的复选框并选择 Delete (删除)。
5. 当系统提示进行确认时，选择 Yes, Delete。

使用 AWS CLI 删除侦听器

使用 delete-listener 命令。

应用程序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。在创建每个侦听器规则时，可以指定目标组和条件。满足规则条件时，流量会转发到相应的目标组。您可以为不同类型的请求创建不同的目标组。例如，为一般请求创建一个目标组，为应用程序的微服务请求创建其他目标组。有关更多信息，请参阅 应用程序负载均衡器 组件 (p. 1)。

您基于每个目标组定义负载均衡器的运行状况检查设置。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，负载均衡器将持续监控已注册到该目标组的所有目标 (这些目标位于已为负载均衡器启用的可用区中) 的运行状况。负载均衡器将请求路由到正常运行的已注册目标。

内容

路由配置

默认情况下，负载均衡器会使用您在创建目标组时指定的协议和端口号将请求路由到其目标。此外，您可以覆盖在将目标注册到目标组时用于将流量路由到目标的端口。

目标组支持以下协议和端口：

- 协议：HTTP、HTTPS
- 端口：1-65535

如果使用 HTTPS 协议配置目标组或使用 HTTPS 运行状况检查，SSL 和目标之间的连接将使用 `ELBSecurityPolicy2016-08` 策略中的安全设置。

目标类型

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)

凭借这些支持的 CIDR 块，您可以将以下内容注册到目标组：ClassicLink 实例、对等 VPC 中的实例、可通过 IP 地址和端口寻址的 AWS 资源 (例如数据库) 以及通过 AWS Direct Connect 或 VPN 连接链接到 AWS 的本地资源。

Important

不能指定可公开路由的 IP 地址。

如果使用实例 ID 指定目标，则使用实例的主网络接口中指定的主私有 IP 地址将流量路由到实例。如果使用 IP 地址指定目标，则可以使用来自一个或多个网络接口的任何私有 IP 地址将流量路由到实例。这使一个实例上的多个应用程序可以使用同一端口。每个网络接口都可以有自己的安全组。

已注册目标

您的负载均衡器充当客户端的单一接触点，并跨其正常运行的已注册目标分发传入流量。您可以将每个目标注册到一个或多个目标组中。您可以使用不同的端口多次向同一目标组注册每个 EC2 实例或 IP 地址，从而使负载均衡器能够将请求路由到微服务。

如果应用程序需求增加，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。只要注册过程完成且新注册的目标通过初始运行状况检查，负载均衡器就会开始将请求路由至此目标。

如果应用程序需求减少或者您需要为目标提供服务，您可以从目标组取消注册目标。取消注册目标将从目标组中删除目标，但不会影响目标。取消注册某个目标后，负载均衡器立即停止将请求路由到该目标。目标将进入 `draining` 状态，直至进行中请求完成。在您准备好目标以继续接收请求时，可以重新将目标注册到目标组。

如果要通过实例 ID 来注册目标，则可以将负载均衡器与 Auto Scaling 组一同使用。将一个目标组挂接到 Auto Scaling 组后，Auto Scaling 在启动目标时会为您向该目标组注册目标。有关更多信息，请参阅 Amazon EC2 Auto Scaling 用户指南 中的将负载均衡器附加到 Auto Scaling 组。

目标组属性

以下是目标组属性：

`deregistration_delay.timeout_seconds`

在取消注册目标前，Elastic Load Balancing 需等待的时间。范围为 0–3600 秒。默认值为 300 秒。

`slow_start.duration_seconds`

一个时间段 (秒)，在此期间，负载均衡器将进入目标组的流量的线性增加份额发送给新注册的目标。范围为 30–900 秒 (15 分钟)。默认值为 0 秒 (已禁用)。

`stickiness.enabled`

指示是否启用粘性会话。

`stickiness.lb_cookie.duration_seconds`

Cookie 有效期 (以秒为单位)。经过这个有效期后，Cookie 即过期。最小值为 1 秒，最大值为 7 天 (604800 秒)。默认值为 1 天 (86400 秒)。

`stickiness.type`

粘性的类型。可能的值为 `lb_cookie`。

取消注册延迟

Elastic Load Balancing 停止将请求发送到正在取消注册的目标。默认情况下，Elastic Load Balancing 在取消注册过程完成前会等待 300 秒，这有助于完成针对目标的进行中的请求。要更改 Elastic Load Balancing 的等待时间，请更新取消注册延迟值。

取消注册的目标的初始状态为 `draining`。取消注册延迟结束后，取消注册过程完成，目标状态变为 `unused`。如果目标是 Auto Scaling 组的一部分，便可以将其终止或替换。

如果取消注册的目标没有进行中的请求且没有活动连接，则 Elastic Load Balancing 将立即完成取消注册过程，而不等待取消注册延迟结束。但是，即使目标取消注册已完成，目标的状态也将显示为 `draining`，直至取消注册延迟结束。

如果正在取消注册的目标在取消注册延迟结束前终止连接，客户端将收到 500 级错误响应。

使用控制台更新取消注册延迟值

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。
3. 选择目标组。当前值在 Description (描述) 选项卡上显示为 Deregistration delay (取消注册延迟)。
4. 在 Description 选项卡上，选择 Edit attributes。
5. 在 Edit attributes 页面上，根据需要更改 Deregistration delay 的值，然后选择 Save。

使用 AWS CLI 更新取消注册延迟值

使用带 `deregistration_delay.timeout_seconds` 属性的 modify-target-group-attributes 命令。

慢启动模式

默认情况下，目标只要注册到目标组并通过了初始运行状况检查，就会开始接收其完整的请求份额。使用慢启动模式可给目标时间进行预热，然后负载均衡器向其发送完整的请求份额。为目标组启用慢启动之后，当目标注册到目标组时目标将进入慢启动模式，当配置的慢启动持续时间期限已过时目标退出慢启动模式。负

载均衡器线性增加它可以向慢启动模式下的目标发送的请求数量。当目标退出慢启动模式后，负载均衡器可以向它发送完整的请求份额。

注意事项

- 为目标组启用慢启动之后，已注册到目标组的目标不会进入慢启动模式。
- 当您为空的目标组启用慢启动，然后使用单一注册操作注册一个或多个目标时，这些目标不会进入慢启动模式。仅当至少有一个已注册的目标未处于慢启动模式时，新注册的目标才会进入慢启动模式。
- 如果您在慢启动模式下取消注册目标，目标将退出慢启动模式。如果您再次注册同一个目标，它会再次进入慢启动模式。
- 如果处于慢启动模式下的目标在持续时间期限过去之前变得运行状况不佳，然后再次变为运行状况良好状态，则目标将保持为慢启动模式，并在剩下的持续时间期限过去时退出慢启动模式。如果未处于慢启动模式的目标从运行状况不佳变为运行状况良好，则它不会进入慢启动模式。

使用控制台更新慢启动持续时间值

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。
3. 选择目标组。当前值在 Description (描述) 选项卡上显示为 Slow start duration (慢启动持续时间)。
4. 在 Description 选项卡上，选择 Edit attributes。
5. 在 Edit attributes (编辑属性) 页面上，根据需要更改 Slow start duration (慢启动持续时间) 的值，然后选择 Save (保存)。要禁用慢启动模式，请将持续时间设置为 0。

使用 AWS CLI 更新慢启动持续时间值

使用带 `slow_start.duration_seconds` 属性的 modify-target-group-attributes 命令。

粘性会话

粘性会话是用于将请求路由到目标组中的同一目标的机制。对于维护状态信息以便向客户端提供持续体验的服务器来说，这很有用。要使用粘性会话，客户端必须支持 Cookie。

当负载均衡器第一次收到来自客户端的请求时，它会将请求路由到目标并生成 Cookie 以包含在对客户端的响应中。来自客户端的下一个请求将包含 cookie。如果为目标组启用粘性会话，并且请求转至同一目标组，则负载均衡器将检测 cookie 并将请求路由到同一目标。

Application Load Balancer 仅支持负载均衡器生成的 Cookie。该 Cookie 的名称是 AWSALB。这些 Cookie 的内容使用轮换密钥进行加密。您无法解密或修改负载均衡器生成的 Cookie。

WebSockets 连接天生具有粘性。如果客户端请求 WebSockets 连接升级，则返回 HTTP 101 状态码以接受连接升级的目标将是在 WebSockets 连接中使用的目标。在 WebSockets 升级完成后，将不会使用基于 Cookie 的粘性。

您在目标组级别启用粘性会话。您还可以设置负载均衡器生成 Cookie 的粘性持续时间，以秒为单位。系统会随每个请求设置持续时间。因此，如果客户端在各个持续时间过期前发送请求，则粘性会话会继续。如果您对多个目标组启用粘性会话，则我们建议您为所有目标组配置相同的时段。

使用控制台启用粘性会话

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。

3. 选择目标组。
4. 在 Description 选项卡上，选择 Edit attributes。
5. 在 Edit attributes 页上，执行以下操作：
   a. 选择 Enable load balancer generated cookie stickiness。
   b. 对于 Stickiness duration，指定一个介于 1 秒和 7 天之间的值。
   c. 选择 Save。

使用 AWS CLI 启用粘性会话

使用带 `stickiness.enabled` 和 `stickiness.lb_cookie.duration_seconds` 属性的 modify-target-group-attributes 命令。

创建目标组

将目标注册到目标组。默认情况下，负载均衡器使用您为目标组指定的端口和协议将请求发送到已注册目标。在将每个目标注册到目标组时，可以覆盖此端口。

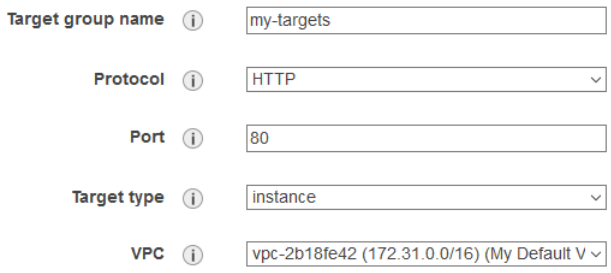
在创建目标组后，可以添加标签。

要将流量路由到目标组中的目标，请在创建侦听器或侦听器规则时，在操作中指定目标组。有关更多信息，请参阅 侦听器规则 (p. 22)。

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 向您的目标组注册目标 (p. 46)。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 修改目标组的运行状况检查设置 (p. 46)。

使用控制台创建目标组

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。
3. 选择 Create target group。
4. 对于 Target group name，键入目标组的名称。
5. (可选) 对于 Protocol 和 Port，根据需要修改默认值。
6. 对于 Target type，选择 `instance` 通过实例 ID 指定目标，或选择 `ip` 通过 IP 地址指定目标。
7. 对于 VPC，选择 Virtual Private Cloud (VPC)。



8. (可选) 对于 Health check settings 和 Advanced health check settings，根据需要修改默认设置。
9. 选择 Create。
10. (可选) 添加一个或多个标签，如下所示：
    a. 选择新创建的目标组。

b. 在 Tags 选项卡上，选择 Add/Edit Tags。
c. 在 Add/Edit Tags 页面上，对于添加的每个标签，选择 Create Tag，然后指定标签键和标签值。添加完标签后，选择 Save。

11. (可选) 要向目标组添加目标，请参阅向您的目标组注册目标 (p. 46)。

使用 AWS CLI 创建目标组

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

目标组的运行状况检查

您的应用程序负载均衡器会定期向其注册目标发送请求以测试其状态。这些测试称为运行状况检查。

每个负载均衡器节点仅将请求路由至负载均衡器的已启用可用区中的正常目标。每个负载均衡器节点均使用每个目标注册到的目标组的运行状况检查设置来检查该目标的运行状况。在注册目标后，目标必须通过一次运行状况检查才会被视为正常。在完成每次运行状况检查后，负载均衡器节点将关闭为运行状况检查而建立的连接。

如果任何可用区均不包含正常目标，则负载均衡器节点会将请求路由到所有目标。

运行状况检查不支持 WebSockets。

运行状况检查设置

可使用以下设置为目标组中的目标配置运行状况检查。负载均衡器使用指定的端口、协议和 ping 路径，每隔 HealthCheckIntervalSeconds 指定的秒数向每个已注册目标发送一次运行状况检查请求。每个运行状况检查请求都是独立的，并且在整个时间间隔内持续。目标响应所用时间不影响下一运行状况检查请求的时间间隔。如果运行状况检查超出了 UnhealthyThresholdCount 连续失败次数，则负载均衡器将使目标停止服务。如果运行状况检查超出了 HealthyThresholdCount 连续成功次数，则负载均衡器将使目标恢复服务。

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。可能的协议为 HTTP 和 HTTPS。默认值为 HTTP 协议。
HealthCheckPort	对目标执行运行状况检查时负载均衡器使用的端口。默认设置是使用每个目标用来从负载均衡器接收流量的端口。
HealthCheckPath	作为运行状况检查的目标上的目标 ping 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 2–60 秒。默认值为 5 秒。
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 5–300 秒。默认值为 30 秒。
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 2–10。默认值为 5。

设置	描述
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 2–10。默认值为 2。
Matcher	检查来自目标的成功响应时要使用的 HTTP 代码。您可以指定 200 到 499 之间的值或值的范围。默认值为 200。

目标运行状况

在负载均衡器向目标发送运行状况检查请求之前，您必须将目标注册到目标组，在侦听器规则中指定其目标组，并确保已为负载均衡器启用目标的可用区。目标必须先通过初始运行状况检查，然后才能接收来自负载均衡器的请求。在目标通过初始运行状况检查后，其状态为 `Healthy`。

下表描述已注册目标的正常状态的可能值。

值	描述
`initial`	负载均衡器正处于注册目标或对目标执行初始运行状况检查的过程中。
`healthy`	目标正常。
`unhealthy`	目标未响应运行状况检查或未通过运行状况检查。
`unused`	目标未注册到目标组，负载均衡器的侦听器规则中未使用目标组，或者目标在没有为负载均衡器启用的可用区中。
`draining`	目标正在取消注册，连接即将耗尽。

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，API 将返回问题的原因代码和描述，并且控制台将在工具提示中显示相同的描述。以 `Elb` 开头的原因代码源自负载均衡器端，以 `Target` 开头的原因代码源自目标端。

原因代码	描述
`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
`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

原因代码	描述
`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`	目标组没有被配置为接收来自负载均衡器的流量 目标处于没有为负载均衡器启用的可用区
`Target.NotRegistered`	目标未注册到目标组
`Target.ResponseCodeMismatch`	运行状况检查失败，显示以下代码：[code]
`Target.Timeout`	请求超时

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 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。
3. 选择目标组。
4. 在 Targets 选项卡上，Status 列指示每个目标的状态。
5. 如果状态是 `Healthy` 以外的任何值，请查看工具提示以了解更多信息。

使用 AWS CLI 检查目标的运行状况

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，则输出还包括原因代码。

修改目标组的运行状况检查设置

您可以随时修改目标组的运行状况检查设置。

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 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。
3. 选择目标组。
4. 在 Health checks 选项卡上，选择 Edit。
5. 在 Edit target group 页面上，根据需要修改设置，然后选择 Save。

使用 AWS CLI 修改目标组的运行状况检查设置

使用 modify-target-group 命令。

向您的目标组注册目标

将目标注册到目标组。您可以通过实例 ID 或 IP 地址注册目标。有关更多信息，请参阅 应用程序负载均衡器的目标组 (p. 39)。

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。在目标准备好处理请求后，将目标注册到您的目标组。只要注册过程完成且目标通过初始运行状况检查，负载均衡器就会开始将请求路由至目标。

如果已注册目标需求减少或者您需要为目标提供服务，您可以从目标组取消注册目标。取消注册某个目标后，负载均衡器立即停止将请求路由到该目标。在目标准备好接收请求时，您可以再次将目标注册到目标组。

在取消注册目标时，负载均衡器会一直等待，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`。

取消注册通过 IP 地址注册的目标后，必须等待取消注册延迟结束，然后才可以重新注册相同的 IP 地址。

如果要通过实例 ID 来注册目标，则可以将负载均衡器与 Auto Scaling 组一同使用。将目标组挂接到 Auto Scaling 组并且该组扩展后，由 Auto Scaling 组启动的实例将自动在目标组中注册。如果您将目标组与 Auto Scaling 组分离，则实例会自动从目标组中取消注册。有关更多信息，请参阅 Amazon EC2 Auto Scaling 用户指南 中的将负载均衡器附加到 Auto Scaling 组。

目标安全组

在将 EC2 实例注册为目标时，您必须确保实例的安全组允许负载均衡器在侦听器端口和运行状况检查端口上与您的实例进行通信。

推荐的规则

入站		
源	端口范围	评论
########	#####	在实例侦听器端口上允许来自负载均衡器的流量
########	######	在运行状况检查端口上允许来自负载均衡器的流量

我们还建议您允许入站 ICMP 流量以支持路径 MTU 发现。有关更多信息，请参阅 Amazon EC2 用户指南（适用于 Linux 实例）中的路径 MTU 发现。

注册或取消注册目标

在创建目标组时，指定是否必须通过实例 ID 或 IP 地址注册目标。

任务

- 通过实例 ID 注册或取消注册目标 (p. 47)
- 通过 IP 地址注册或取消注册目标 (p. 48)

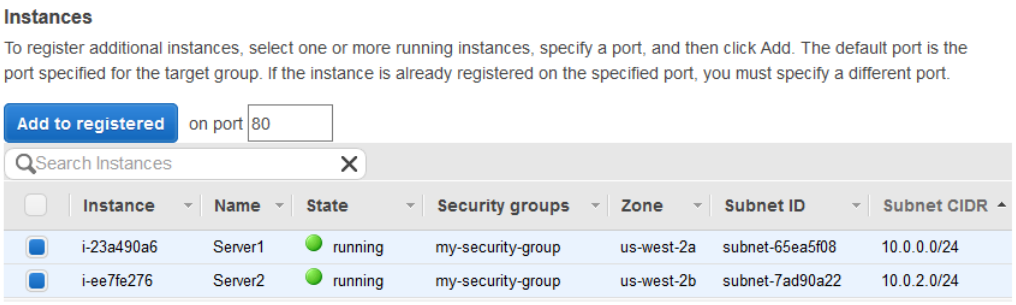
通过实例 ID 注册或取消注册目标

实例必须位于您为目标组指定的 Virtual Private Cloud (VPC) 中。当您注册实例时，实例还必须处于 `running` 状态。

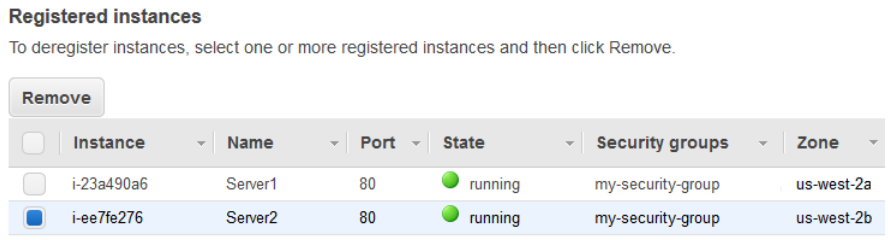
使用控制台按实例 ID 注册或取消注册目标

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。
3. 选择您的目标组。

4. 在 Targets 选项卡上，选择 Edit。
5. 要注册实例，请从 Instances 中选择实例，根据需要修改默认实例端口，然后选择 Add to registered。



6. 要取消注册实例，请从 Registered instances 中选择实例，然后选择 Remove。



7. 选择 Save。

使用 AWS CLI 注册或取消注册目标

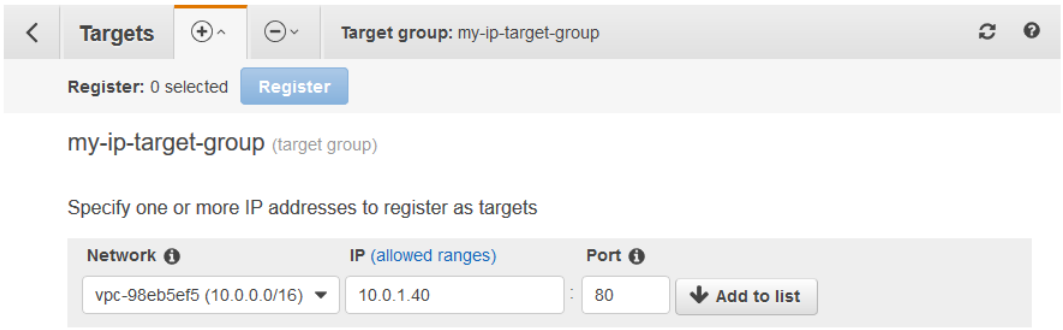
使用 register-targets 命令添加目标，并使用 deregister-targets 命令删除目标。

通过 IP 地址注册或取消注册目标

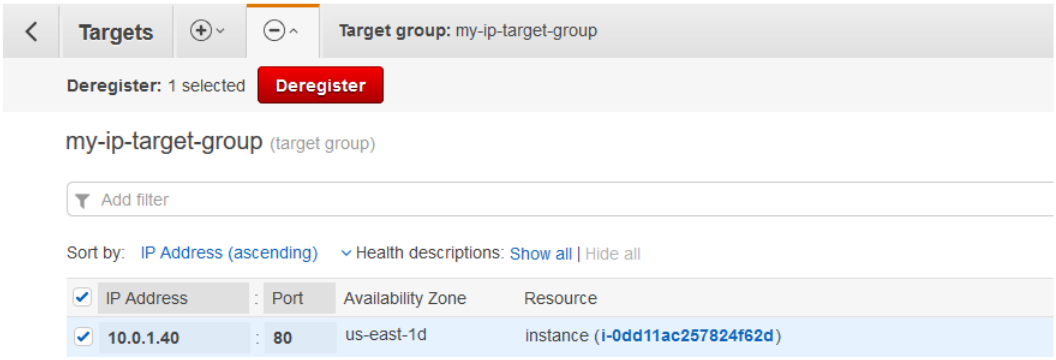
您注册的 IP 地址必须来自目标组的 VPC 的子网、RFC 1918 范围 (10.0.0.0/8、172.16.0.0/12 和 192.168.0.0/16) 和 RFC 6598 范围 (100.64.0.0/10)。不能注册可公开路由的 IP 地址。

使用控制台按 IP 地址注册或取消注册目标

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。
3. 选择您的目标组。
4. 在 Targets 选项卡上，选择 Edit。
5. 要注册 IP 地址，请在菜单栏中选择 Register targets 图标 (加号)。对于每个 IP 地址，请选择网络，键入 IP 地址和端口，然后选择 Add to list (添加到列表)。指定完地址后，选择 Register。



6. 要取消注册 IP 地址，请在菜单栏中选择 Deregister targets 图标 (减号)。如果您有多个注册的 IP 地址，则可能会发现添加筛选器或更改排序顺序很有帮助。选择 IP 地址并选择 Deregister。



7. 要离开此屏幕，请选择菜单栏中的 Back to target group 图标 (后退按钮)。

使用 AWS CLI 注册或取消注册目标

使用 register-targets 命令添加目标，并使用 deregister-targets 命令删除目标。

适用于目标组的标签

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，因为它专为 AWS 使用预留。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 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。
3. 选择目标组。
4. 在 Tags (标签) 选项卡上，选择 Add/Edit Tags (添加/编辑标签)，然后执行以下一项或多项操作：
   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 (创建标签) 并为 Key (密钥) 和 Value (值) 键入值。
   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 Save。

使用 AWS CLI 更新目标组的标签

使用 add-tags 和 remove-tags 命令。

删除目标组

如果目标组未由任何操作引用，则可删除目标组。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 https://console.amazonaws.cn/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Target Groups。
3. 选择目标组，然后依次选择 Actions、Delete。
4. 当系统提示您确认时，选择 Yes。

使用 AWS CLI 删除目标组

使用 delete-target-group 命令。

监控应用程序负载均衡器

您可使用以下功能监控负载均衡器，分析流量模式及解决与负载均衡器和目标相关的问题。

CloudWatch 指标

可以使用 Amazon CloudWatch 将有关负载均衡器和目标的数据点的统计数据作为一组有序时间序列数据 (称作指标) 进行检索。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 应用程序负载均衡器的 CloudWatch 指标 (p. 51)。

访问日志

您可以使用访问日志来捕获有关向负载均衡器发出的请求的详细信息，并将这些详细信息作为日志文件存储在 Amazon S3 中。您可以使用这些访问日志分析流量模式并解决与目标相关的问题。有关更多信息，请参阅 应用程序负载均衡器的访问日志 (p. 59)。

请求跟踪

您可以使用来请求跟踪来跟踪 HTTP 请求。负载均衡器会为它接收的每个请求添加一个包含跟踪标识符的标头。有关更多信息，请参阅 针对应用程序负载均衡器的请求跟踪 (p. 67)。

CloudTrail 日志

您可以使用 AWS CloudTrail 捕获有关向 Elastic Load Balancing API 发出的调用的详细信息，并将这些详细信息作为日志文件存储在 Amazon S3 中。可以使用这些 CloudTrail 日志确定已发出的调用、从中发出调用的源 IP 地址、调用的发出方、调用的发出时间等。有关更多信息，请参阅 应用程序负载均衡器的 AWS CloudTrail 日志记录 (p. 68)。

应用程序负载均衡器的 CloudWatch 指标

Elastic Load Balancing 为负载均衡器和目标向 Amazon CloudWatch 发布数据点。利用 CloudWatch，您可以按一组有序的时间序列数据 (称为指标) 来检索关于这些数据点的统计数据。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负载均衡器的正常目标的总数。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，并在指标超出您的可接受范围时启动某个操作 (如向电子邮件地址发送通知)。

只有当请求流经负载均衡器时，Elastic Load Balancing 才会向 CloudWatch 报告指标。如果有请求流经负载均衡器，则 Elastic Load Balancing 进行测量并以 60 秒的间隔发送其指标。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

应用程序负载均衡器 指标

AWS/ApplicationELB 命名空间包括负载均衡器的以下指标。

指标	描述
ActiveConnectionCount	从客户端到负载均衡器以及从负载均衡器到目标的并发活动 TCP 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
ClientTLSNegotiationErrorCount	由未与负载均衡器建立会话的客户端发起的 TLS 连接数。可能的原因包括密码或协议不匹配。 报告标准：有非零值。 Statistics：最有用的统计工具是 Sum。 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。您需要为每小时使用的 LCU 数量付费。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 Dimensions • LoadBalancer
HTTPCode_ELB_4XX_Count	源自负载均衡器的 HTTP 4XX 客户端错误代码的数量。如果请求格式错误或不完整，则会生成客户端错误。目标尚未收到这些请求。该计数不包含目标生成的任何响应代码。 报告标准：有非零值。 Statistics：最有用的统计工具是 Sum。请注意，Minimum、Maximum 和 Average 均返回 1。 Dimensions • LoadBalancer • AvailabilityZone、LoadBalancer
HTTPCode_ELB_5XX_Count	源自负载均衡器的 HTTP 5XX 服务器错误代码的数量。该计数不包含目标生成的任何响应代码。 报告标准：有非零值。

指标	描述
	Statistics：最有用的统计工具是 Sum。请注意，Minimum、Maximum 和 Average 均返回 1。 Dimensions • LoadBalancer • AvailabilityZone, LoadBalancer
IPv6ProcessedBytes	负载均衡器通过 IPv6 处理的总字节数。 报告标准：有非零值。 Statistics：最有用的统计工具是 Sum。 Dimensions • LoadBalancer
IPv6RequestCount	负载均衡器收到的 IPv6 请求的数量。 报告标准：有非零值。 Statistics：最有用的统计工具是 Sum。请注意，Minimum、Maximum 和 Average 均返回 1。 Dimensions • LoadBalancer • AvailabilityZone, LoadBalancer • TargetGroup、LoadBalancer • TargetGroup、AvailabilityZone、LoadBalancer
NewConnectionCount	从客户端到负载均衡器以及从负载均衡器到目标建立的新 TCP 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
ProcessedBytes	负载均衡器通过 IPv4 和 IPv6 处理的总字节数。 报告标准：有非零值。 Statistics：最有用的统计工具是 Sum。 Dimensions • LoadBalancer

指标	描述
RejectedConnectionCount	由于负载均衡器达到连接数上限被拒绝的链接的数量。 报告标准：有非零值。 Statistics：最有用的统计工具是 Sum。 Dimensions • LoadBalancer • AvailabilityZone、LoadBalancer
RequestCount	通过 IPv4 和 IPv6 处理的请求的数量。此计数仅包含具有由负载均衡器目标生成的响应的请求。 报告标准：始终报告 Statistics：最有用的统计工具是 Sum。 Dimensions • LoadBalancer • AvailabilityZone, LoadBalancer • TargetGroup、LoadBalancer • TargetGroup、AvailabilityZone、LoadBalancer
RuleEvaluations	在给定 1 小时的平均请求速率的情况下，负载均衡器处理的规则的数量。 报告标准：有非零值。 Statistics：最有用的统计工具是 Sum。 Dimensions • LoadBalancer

AWS/ApplicationELB 命名空间包括目标的以下指标。

指标	描述
HealthyHostCount	被视为正常运行的目标数量。 报告标准：始终报告 Statistics：最有用的统计工具是 Average、Minimum 和 Maximum。 Dimensions • TargetGroup、LoadBalancer • TargetGroup、AvailabilityZone、LoadBalancer
HTTPCode_Target_2XX_Count	目标生成的 HTTP 响应代码的数量。此不包括负载均衡器生成的任何响应代码。 报告标准：有非零值。

指标	描述
	Statistics：最有用的统计工具是 Sum。请注意，Minimum、Maximum 和 Average 均返回 1。 Dimensions • LoadBalancer • AvailabilityZone, LoadBalancer • TargetGroup、LoadBalancer • TargetGroup、AvailabilityZone、LoadBalancer
RequestCountPerTarget	目标组中每个目标收到的平均请求数量。您必须使用 TargetGroup 维度指定目标组。 报告标准：有非零值。 统计：唯一有效的统计数据是 Sum。请注意，这代表平均值，而不是总和。 Dimensions • TargetGroup • TargetGroup, LoadBalancer
TargetConnectionErrorCount	负载均衡器和目标之间连接建立不成功的次数。 报告标准：有非零值。 Statistics：最有用的统计工具是 Sum。 Dimensions • LoadBalancer • AvailabilityZone, LoadBalancer • TargetGroup、LoadBalancer • TargetGroup、AvailabilityZone、LoadBalancer
TargetResponseTime	请求离开负载均衡器直至收到来自目标的响应所用的时间（以秒为单位）。这与访问日志中的 target_processing_time 字段是等效的。 报告标准：有非零值。 Statistics：最有用的统计工具是 Average 和 pNN.NN（百分比）。 Dimensions • LoadBalancer • AvailabilityZone, LoadBalancer • TargetGroup、LoadBalancer • TargetGroup、AvailabilityZone、LoadBalancer

指标	描述
TargetTLSNegotiationErrorCount	由未与目标建立会话的负载均衡器发起的 TLS 连接数。可能的原因包括密码或协议不匹配。 报告标准：有非零值。 Statistics：最有用的统计工具是 Sum。 Dimensions • LoadBalancer • AvailabilityZone, LoadBalancer • TargetGroup、LoadBalancer • TargetGroup、AvailabilityZone、LoadBalancer
UnHealthyHostCount	被视为未正常运行的目标数量。 报告标准：始终报告 Statistics：最有用的统计工具是 Average、Minimum 和 Maximum。 Dimensions • TargetGroup、LoadBalancer • TargetGroup、AvailabilityZone、LoadBalancer

AWS/ApplicationELB 命名空间包括用户身份验证的以下指标。

指标	描述
ELBAuthError	由于身份验证操作配置错误、负载均衡器无法与 IdP 建立连接，或负载均衡器因内部错误无法完成身份验证流程，所导致的无法完成用户身份验证的次数。 报告标准：有非零值。 统计数据：唯一有意义的统计数据是 Sum。 维度: LoadBalancer
ELBAuthFailure	由于 IdP 拒绝用户访问或授权代码多次使用导致的无法完成用户身份验证的次数。 报告标准：有非零值。 统计数据：唯一有意义的统计数据是 Sum。 维度: LoadBalancer
ELBAuthLatency	向 IdP 查询 ID 令牌和用户信息所用的时间（毫秒）。如果这些操作中有一项或多项操作失败，这表示失败时间。 报告标准：有非零值。 统计数据：所有统计数据均有意义。

指标	描述
	维度: LoadBalancer
ELBAuthSuccess	成功的身份验证操作的次数。负载均衡器从 IdP 检索用户身份声明后，验证工作流结束时此指标会递增。 报告标准：有非零值。 Statistics：最有用的统计工具是 Sum。 维度: LoadBalancer

应用程序负载均衡器 的指标维度

要筛选您的应用程序负载均衡器的指标，可以使用以下维度。

维度	描述
AvailabilityZone	按照可用区筛选指标数据。
LoadBalancer	按负载均衡器筛选指标数据。按以下方式指定负载均衡器：app/load-balancer-name/1234567890123456（负载均衡器 ARN 的结尾部分）。
TargetGroup	按目标组筛选指标数据。按以下方式指定目标组：targetgroup/target-group-name/1234567890123456（目标组 ARN 的结尾部分）。

应用程序负载均衡器指标的统计数据

CloudWatch 提供基于 Elastic Load Balancing 发布的指标数据点的统计数据。统计数据是在指定的时间段内汇总的指标数据。当请求统计数据时，返回的数据流按指标名称和维度进行识别。维度是用于唯一标识指标的名称/值对。例如，您可以请求在特定可用区内启动的负载均衡器背后所有正常状态 EC2 实例的统计数据。

Minimum 和 Maximum 统计数据反映各个负载均衡器节点报告的最小值和最大值。例如，假定有两个负载均衡器节点。一个节点的 HealthyHostCount 的 Minimum 为 2，Maximum 为 10，Average 为 6，另一个节点的 HealthyHostCount 的 Minimum 为 1，Maximum 为 5，Average 为 3。因此，负载均衡器的 Minimum 为 1，Maximum 为 10，Average 大约为 4。

Sum 统计数据是所有负载均衡器节点的汇总值。由于这些指标在每个周期均包含多个报告，因此 Sum 仅适用于对所有负载均衡器节点进行汇总的指标。

SampleCount 统计数据是测量的样本数。由于这些指标是基于采样间隔和事件进行收集的，因此此统计信息一般没有用。例如，对于 HealthyHostCount，SampleCount 基于每个负载均衡器节点报告的样本数，而不是运行状况正常的主机数。

百分位数指示某个值在数据集中的相对位置。您可以指定任何百分位数，最多使用两位小数 (例如 p95.45)。例如，第 95 个百分位数表示 95% 的数据低于此值，5% 的数据高于此值。百分位数通常用于隔离异常值。例如，假设某个应用程序从缓存服务大多数请求的时间是 1-2 毫秒，但如果缓存是空的，则时间需要 100-200 毫秒。最大值反映了最慢的情况，也就是大约 200 毫秒。平均值不表示数据的分布。百分位提供了一个更有意义的应用程序性能视图。通过使用第 99 个百分位数作为 Auto Scaling 触发器或 CloudWatch 警报，您可以将目标设定为不超过 1% 的请求的处理时间会超过 2 毫秒。

查看负载均衡器的 CloudWatch 指标

您可以使用 Amazon EC2 控制台查看负载均衡器的 CloudWatch 指标。这些指标显示为监控图表。如果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 控制台查看负载均衡器的指标。

使用 Amazon EC2 控制台查看指标

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/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，然后选择 Monitoring 选项卡。
   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 Load Balancers。
   b. 选择您的负载均衡器，然后选择 Monitoring 选项卡。
   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. 通过以下网址打开 CloudWatch 控制台：https://console.amazonaws.cn/cloudwatch/。
2. 在导航窗格中，选择 Metrics。
3. 选择 ApplicationELB 命名空间。
4. (可选) 要跨所有维度查看某个指标，请在搜索字段中键入其名称。
5. (可选) 要按维度筛选，请选择下列选项之一：
   - 要仅显示为负载均衡器报告的指标，请选择 Per AppELB Metrics。要查看单个负载均衡器的指标，请在搜索字段中键入其名称。
   - 要仅显示为您的目标组报告的指标，请选择 Per AppELB, per TG Metrics。要查看单个目标组的指标，请在搜索字段中键入其名称。
   - 要仅按可用区显示为负载均衡器报告的指标，请选择 Per AppELB, per AZ Metrics。要查看单个负载均衡器的指标，请在搜索字段中键入其名称。要查看单个可用区的指标，请在搜索字段中键入其名称。
   - 要仅按可用区和目标组显示为负载均衡器报告的指标，请选择 Per AppELB, per AZ, per TG Metrics。要查看单个负载均衡器的指标，请在搜索字段中键入其名称。要查看单个目标组的指标，请在搜索字段中键入其名称。要查看单个可用区的指标，请在搜索字段中键入其名称。

使用 AWS CLI 查看指标

使用以下 list-metrics 命令列出可用指标：

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

使用 AWS CLI 获取指标的统计数据

使用以下 get-metric-statistics 命令获取指定指标和维度的统计数据。请注意 CloudWatch 将不同维度的每种唯一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/ApplicationELB \
--metric-name UnHealthyHostCount --statistics Average  --period 3600 \
--dimensions Name=LoadBalancer,Value=app/my-load-balancer/50dc6c495c0c9188 \
Name=TargetGroup,Value=targetgroup/my-targets/73e2d6bc24d8a067 \
--start-time 2016-04-18T00:00:00Z --end-time 2016-04-21T00:00:00Z
```

下面是示例输出：

```
{
    "Datapoints": [
        {
            "Timestamp": "2016-04-18T22:00:00Z",
            "Average": 0.0,
            "Unit": "Count"
        },
        {
            "Timestamp": "2016-04-18T04:00:00Z",
            "Average": 0.0,
            "Unit": "Count"
        },
        ...
    ],
    "Label": "UnHealthyHostCount"
}
```

应用程序负载均衡器的访问日志

Elastic Load Balancing 提供了访问日志，该访问日志可捕获有关发送到负载均衡器的请求的详细信息。每个日志都包含信息 (例如，收到请求的时间、客户端的 IP 地址、延迟、请求路径和服务器响应)。您可以使用这些访问日志分析流量模式并解决问题。

访问日志记录是 Elastic Load Balancing 的一项可选功能，默认情况下已禁用此功能。为负载均衡器启用访问日志记录之后，Elastic Load Balancing 捕获日志并将其作为压缩文件存储在您指定的 Amazon S3 存储桶中。您可以随时禁用访问日志记录。

Elastic Load Balancing 支持针对 应用程序负载均衡器 的访问日志的服务器端加密。这将保护存储在 S3 存储桶中的日志数据，并符合静态数据的合规性要求。每个访问日志文件在存储到 S3 存储桶中之前将自动加密，并在您访问它时进行解密。您不需要执行任何操作，因为这与您访问加密的日志文件或未加密的日志文件的方式基本相同。每个日志文件均利用多因素强加密来使用唯一密钥加密。作为额外的保护，密钥本身将使用定期轮换的主密钥进行加密。有关更多信息，请参阅 Amazon Simple Storage Service 开发人员指南 中的使用具有 Amazon S3 托管加密密钥的服务器端加密 (SSE-S3) 保护数据。

使用访问日志无需额外付费。您将需要支付 Amazon S3 的存储费用，但无需为 Elastic Load Balancing 用来将日志文件发送到 Amazon S3 的带宽付费。有关存储成本的更多信息，请参阅 Amazon S3 定价。

内容

访问日志文件

Elastic Load Balancing 每 5 分钟为每个负载均衡器节点发布一次日志文件。日志传输最终是一致的。负载均衡器可以传输相同时间段的多个日志。通常，如果站点具有高流量，会出现此情况。

访问日志的文件名采用以下格式：

```
bucket[/prefix]/AWSLogs/aws-account-id/elasticloadbalancing/region/yyyy/mm/dd/aws-account-id_elasticloadbalancing_region_load-balancer-id_end-time_ip-address_random-string.log.gz
```

存储桶

S3 存储桶的名称。

前缀

存储桶中的前缀 (逻辑层级结构)。如果您不指定前缀，则会将日志置于存储桶的根级。

aws-account-id

拥有者的 AWS 账户 ID。

区域

负载均衡器和 S3 存储桶所在的区域。

yyyy/mm/dd

传输日志的日期。

load-balancer-id

负载均衡器的资源 ID。如果资源 ID 包含任何正斜杠 (/)，这些正斜杠将替换为句点 (.)。

end-time

日志记录间隔结束的日期和时间。例如，结束时间 20140215T2340Z 包含在 23:35 和 23:40 之间发出的请求的条目。

ip-address

处理请求的负载均衡器节点的 IP 地址。对于内部负载均衡器，这是私有 IP 地址。

random-string

系统生成的随机字符串。

以下是示例日志文件名：

```
s3://my-bucket/prefix/AWSLogs/123456789012/elasticloadbalancing/us-west-2/2016/05/01/123456789012_elasticloadbalancing_us-west-2_my-loadbalancer_20140215T2340Z_172.160.001.192_20sg8hgm.log.gz
```

日志文件可以在存储桶中存储任意长时间，不过您也可以定义 Amazon S3 生命周期规则以自动存档或删除日志文件。有关更多信息，请参阅 Amazon Simple Storage Service 开发人员指南 中的对象生命周期管理。

访问日志条目

Elastic Load Balancing 会记录发送给负载均衡器的请求，包括从未到达目标的请求。例如，如果客户端发送格式错误的请求或者没有正常的目标响应请求，仍会记录请求。请注意，Elastic Load Balancing 不记录运行状况检查请求。

每个日志条目均包含向负载均衡器发出的一个请求 (如果使用 WebSockets，则为连接) 的详细信息。对于 WebSockets，仅在关闭连接后写入条目。如果无法建立升级后的连接，则 HTTP 或 HTTPS 请求的条目相同。

Important

Elastic Load Balancing 将尽力记录请求。我们建议您使用访问日志来了解请求性质，而不是作为所有请求的完整描述。

语法

下表按顺序描述了访问日志条目的字段。使用空格分隔所有字段。在引入新的字段时，会将这些字段添加到日志条目的末尾。您应忽略日志条目结尾的任何不需要的字段。

字段	描述
type	请求或连接的类型。可能的值如下 (忽略任何其他值)： • `http` - HTTP • `https` - HTTP over SSL/TLS • `h2` - HTTP/2 over SSL/TLS • `ws` - WebSockets • `wss` - WebSockets over SSL/TLS
timestamp	负载均衡器生成对客户端的响应的时间 (采用 ISO 8601 格式)。对于 WebSockets，这是关闭连接的时间。
elb	负载均衡器的资源 ID。如果您正在解析访问日志条目，请注意，资源 ID 可包含正斜杠 (/)。
client:port	请求客户端的 IP 地址和端口。
target:port	处理此请求的目标的 IP 地址和端口。 如果客户端没有发送完整请求，则负载均衡器无法将请求分派到目标，并且此值设置为 -。 如果 AWS WAF 拦截了此请求，该值会被设置为 -，且 elb_status_code 的值会被设置为 403。
request_processing_time	从负载均衡器收到请求到将请求发送到目标所用的总时间（以秒为单位，精度为毫秒）。 如果负载均衡器无法将请求分派到目标，则此值设置为 -1。如果目标在空闲超时前关闭连接，或客户端发送了格式错误的请求，则会发生这种情况。 如果注册目标在空闲超时之前未响应，则此值还可设置为 -1。
target_processing_time	从负载均衡器将请求发送到目标到该目标开始发送响应标头所用的总时间（以秒为单位，精度为毫秒）。 如果负载均衡器无法将请求分派到目标，则此值设置为 -1。如果目标在空闲超时前关闭连接，或客户端发送了格式错误的请求，则会发生这种情况。 如果注册目标在空闲超时之前未响应，则此值还可设置为 -1。
response_processing_time	从负载均衡器收到来自目标的响应标头到开始向客户端发送响应所用的总时间（以秒为单位，精度为毫秒）。此时间包括在负载均衡器上的排队时间以及从负载均衡器到客户端的连接获取时间。 如果负载均衡器无法将请求发送到目标，则此值设置为 -1。如果目标在空闲超时前关闭连接，或客户端发送了格式错误的请求，则会发生这种情况。

字段	描述
elb_status_code	来自负载均衡器的响应的状态代码。
target_status_code	来自目标的响应的状态代码。仅在已建立与目标的连接且目标已发送响应的情况下记录此值。否则，该值将设置为 -。
received_bytes	从客户端 (申请方) 接收的请求大小 (以字节为单位)。对于 HTTP 请求，这包括标头。对于 WebSockets，这是通过连接从客户端接收的字节总数。
sent_bytes	发送到客户端 (申请方) 的响应的大小 (以字节为单位)。对于 HTTP 请求，这包括标头。对于 WebSockets，这是通过连接发送到客户端的字节总数。
"请求"	来自客户端的请求行，包含在双引号内并采用以下格式进行记录：HTTP 方法 + 协议://主机:端口/uri + HTTP 版本。
"user_agent"	标识发出请求的客户端的用户代理字符串（用双引号括起来）。该字符串包含一个或多个产品标识符 (product[/version])。如果字符串长度超过 8 KB，则将被截断。
ssl_cipher	[HTTPS 侦听器] SSL 密码。仅当在成功协商后建立传入 连接时记录此值。否则，该值将设置为 -。
ssl_protocol	[HTTPS 侦听器] SSL 协议。仅当在成功协商后建立传入 连接时记录此值。否则，该值将设置为 -。
target_group_arn	目标组的 Amazon 资源名称 (ARN)。
"trace_id"	X-Amzn-Trace-Id 标头的内容（用双引号括起来）。
"domain_name"	[HTTPS 侦听器] TLS 握手期间客户端提供的 SNI 域（用双引号括起来）。如果客户端不支持 SNI 或此域与证书不匹配且将向客户端提供默认证书，则此值将设置为 -。
"chosen_cert_arn"	[HTTPS 侦听器] 向客户端提供的证书的 ARN（用双引号括起来）。如果重复使用会话，则将此值设置为 `session-reused`。
matched_rule_priority	匹配请求的规则的优先级值。如果匹配了某个规则，则此值的范围介于 1 和 50000 之间。如果未匹配任何规则并且已执行默认操作，则该值为 0。如果出错，则该值为 -1。
request_creation_time	负载均衡器从客户端收到请求的时间 (采用 ISO 8601 格式)。
"actions_executed"	处理请求时执行的操作（用双引号括起来）。此值是一个逗号分隔的列表，可以包含以下可能的值：`waf`、`authenticate` 和 `forward`。如果未执行任何操作（例如，针对格式错误的请求的操作），则此值设置为 - 。

示例

以下是示例日志条目。请注意，文本以多行形式显示只是为了更方便阅读。

示例 HTTP 条目

以下是 HTTP 侦听器 (端口 80 到端口 80) 的示例日志条目：

```
http 2018-07-02T22:23:00.186641Z app/my-loadbalancer/50dc6c495c0c9188
192.168.131.39:2817 10.0.0.1:80 0.000 0.001 0.000 200 200 34 366
"GET http://www.example.com:80/ HTTP/1.1" "curl/7.46.0" - -
arn:aws-cn:elasticloadbalancing:us-west-2:123456789012:targetgroup/my-
targets/73e2d6bc24d8a067
```

```
"Root=1-58337262-36d228ad5d99923122bbe354" "-" "-"
0 2018-07-02T22:22:48.364000Z "forward"
```

示例 HTTPS 条目

以下是 HTTPS 侦听器 (端口 443 到端口 80) 的示例日志条目：

```
https 2018-07-02T22:23:00.186641Z app/my-loadbalancer/50dc6c495c0c9188
192.168.131.39:2817 10.0.0.1:80 0.086 0.048 0.037 200 200 0 57
"GET https://www.example.com:443/ HTTP/1.1" "curl/7.46.0" ECDHE-RSA-AES128-GCM-SHA256
 TLSv1.2
arn:aws-cn:elasticloadbalancing:us-west-2:123456789012:targetgroup/my-
targets/73e2d6bc24d8a067
"Root=1-58337281-1d84f3d73c47ec4e58577259" "www.example.com" "arn:aws:acm:us-
west-2:123456789012:certificate/12345678-1234-1234-1234-123456789012"
1 2018-07-02T22:22:48.364000Z "authenticate,forward"
```

示例 HTTP/2 条目

以下是 HTTP/2 流的示例日志条目。

```
h2 2018-07-02T22:23:00.186641Z app/my-loadbalancer/50dc6c495c0c9188
10.0.1.252:48160 10.0.0.66:9000 0.000 0.002 0.000 200 200 5 257
"GET https://10.0.2.105:773/ HTTP/2.0" "curl/7.46.0" ECDHE-RSA-AES128-GCM-SHA256 TLSv1.2
arn:aws-cn:elasticloadbalancing:us-west-2:123456789012:targetgroup/my-
targets/73e2d6bc24d8a067
"Root=1-58337327-72bd00b0343d75b906739c42" "-" "-"
1 2018-07-02T22:22:48.364000Z "forward"
```

示例 WebSockets 条目

以下是 WebSockets 连接的示例日志条目。

```
ws 2018-07-02T22:23:00.186641Z app/my-loadbalancer/50dc6c495c0c9188
10.0.0.140:40914 10.0.1.192:8010 0.001 0.003 0.000 101 101 218 587
"GET http://10.0.0.30:80/ HTTP/1.1" "-" - -
arn:aws-cn:elasticloadbalancing:us-west-2:123456789012:targetgroup/my-
targets/73e2d6bc24d8a067
"Root=1-58337364-23a8c76965a2ef7629b185e3" "-" "-"
1 2018-07-02T22:22:48.364000Z "forward"
```

安全 WebSockets 条目示例

以下是安全 WebSockets 连接的示例日志条目。

```
wss 2018-07-02T22:23:00.186641Z app/my-loadbalancer/50dc6c495c0c9188
10.0.0.140:44244 10.0.0.171:8010 0.000 0.001 0.000 101 101 218 786
"GET https://10.0.0.30:443/ HTTP/1.1" "-" ECDHE-RSA-AES128-GCM-SHA256 TLSv1.2
arn:aws-cn:elasticloadbalancing:us-west-2:123456789012:targetgroup/my-
targets/73e2d6bc24d8a067
"Root=1-58337364-23a8c76965a2ef7629b185e3" "-" "-"
1 2018-07-02T22:22:48.364000Z "forward"
```

存储桶权限

在启用访问日志记录时，您必须为访问日志指定 S3 存储桶。此存储桶必须位于负载均衡器所在的区域，并且必须具有向 Elastic Load Balancing 授予向存储桶写入访问日志的权限的存储桶策略。存储桶策略是 JSON

语句的集合，这些语句以访问策略语言编写，用于为存储桶定义访问权限。每个语句都包括有关单个权限的信息并包含一系列元素。

> **Important**
>
> 如果您将使用控制台启用访问日志记录，则可跳至启用访问日志记录 (p. 66)。如果您将使用 AWS CLI 或 API 启用访问日志记录，则存储桶必须存在且必须具有所需的存储桶策略。

如果您需要为访问日志创建存储桶，请使用以下过程创建存储桶并添加所需的存储桶策略。如果您已拥有存储桶，请从步骤 4 开始以添加或更新存储桶的存储桶策略。

使用所需权限创建 Amazon S3 存储桶

1. 通过以下网址打开 Amazon S3 控制台：https://console.amazonaws.cn/s3/。
2. 选择 Create Bucket。
3. 在 Create a Bucket (创建存储桶) 对话框中，执行以下操作：
   a. 对于 Bucket Name，输入您的存储桶的名称 (例如 `my-loadbalancer-logs`)。此名称在 Amazon S3 内所有现有存储桶名称中必须唯一。在某些区域，可能对存储桶名称有其他限制。有关更多信息，请参阅 Amazon Simple Storage Service 开发人员指南 中的存储桶局限和限制。
   b. 对于 Region，选择在其中创建负载均衡器的区域。
   c. 选择 Create。
4. 选择存储桶，然后选择 Permissions。
5. 选择 Bucket Policy。如果存储桶具有已附加策略，您可以将所需的语句添加到现有策略中。
6. 选择 Policy generator。在 AWS Policy Generator 页面上，执行以下操作：
   a. 对于 Select Type of Policy，选择 S3 Bucket Policy。
   b. 对于 Effect，选择 Allow。
   c. 对于 Principal，指定下列 AWS 账户 ID 之一以向 Elastic Load Balancing 授予对 S3 存储桶的访问权限。使用与您的负载均衡器和存储桶所在区域对应的账户 ID。

区域	区域名称	Elastic Load Balancing 账户 ID
`us-east-1`	美国东部（弗吉尼亚北部）	127311923021
`us-east-2`	美国东部（俄亥俄州）	033677994240
`us-west-1`	美国西部（加利福尼亚北部）	027434742980
`us-west-2`	美国西部（俄勒冈）	797873946194
`ca-central-1`	加拿大 (中部)	985666609251
`eu-central-1`	欧洲（法兰克福）	054676820928
`eu-west-1`	欧洲（爱尔兰）	156460612806
`eu-west-2`	欧洲 (伦敦)	652711504416
`eu-west-3`	欧洲 (巴黎)	009996457667
`ap-northeast-1`	亚太区域（东京）	582318560864
`ap-northeast-2`	亚太区域（首尔）	600734575887

区域	区域名称	Elastic Load Balancing 账户 ID
ap-northeast-3	亚太区域 (大阪当地)	383597477331
ap-southeast-1	亚太区域（新加坡）	114774131450
ap-southeast-2	亚太区域（悉尼）	783225319266
ap-south-1	亚太地区（孟买）	718504428378
sa-east-1	南美洲（圣保罗）	507241528517
us-gov-west-1*	AWS GovCloud（美国）	048591011584
cn-north-1*	中国（北京）	638102146993
cn-northwest-1*	中国 (宁夏)	037604701340

* 这些区域需要单独的账户。有关更多信息，请参阅 AWS GovCloud（美国）和中国（北京）。

d. 对于 Actions，选择 PutObject 以允许 Elastic Load Balancing 将对象存储在 S3 存储桶中。

e. 对于 Amazon Resource Name (ARN)，采用以下格式输入 S3 存储桶的 ARN。对于 *aws-account-id*，指定拥有负载均衡器的 AWS 账户的 ID (例如 *123456789012*)。

```
arn:aws-cn:s3:::bucket/prefix/AWSLogs/aws-account-id/*
```

请注意，如果您使用的是 us-gov-west-1 区域，请在 ARN 中指定 arn:aws-us-gov 而非 arn:aws。

f. 依次选择 Add Statement 和 Generate Policy。该策略文档应该类似于以下内容：

```
{
  "Id": "Policy1429136655940",
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Stmt1429136633762",
      "Action": [
        "s3:PutObject"
      ],
      "Effect": "Allow",
      "Resource": "arn:aws-cn:s3:::my-loadbalancer-logs/my-app/
AWSLogs/123456789012/*",
      "Principal": {
        "AWS": [
          "797873946194"
        ]
      }
    }
  ]
}
```

g. 如果您创建新的存储桶策略，请复制整个策略文档，然后选择 Close。

如果您正在编辑现有存储桶策略，请复制策略文档中的新语句 (Statement 元素的 [和] 之间的文本)，然后选择 Close。

7. 返回 Amazon S3 控制台，并将策略粘贴到适当的文本区域中。

8. 选择 Save。

启用访问日志记录

在为负载均衡器启用访问日志记录时，您必须指定负载均衡器将在其中存储日志的 S3 存储桶的名称。存储桶必须位于负载均衡器所在的区域，并且必须具有向 Elastic Load Balancing 授予在存储桶中写入访问日志的权限的存储桶策略。该存储桶可由与拥有负载均衡器的账户不同的账户拥有。

使用控制台启用访问日志记录

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。
2. 在导航窗格中，选择 Load Balancers。
3. 选择您的负载均衡器。
4. 在 Description 选项卡上，选择 Edit attributes。
5. 在 Edit load balancer attributes 页面上，执行以下操作：
   a. 选择 Enable access logs。
   b. 对于 S3 location，键入 S3 存储桶的名称，包括任何前缀 (例如，`my-loadbalancer-logs/my-app`)。您可以指定现有存储桶的名称或新存储桶名称。如果您指定现有存储桶，请确保您拥有此存储桶，且配置了必要的存储桶策略。
   c. (可选) 如果存储桶不存在，请选择 Create this location for me。您必须指定在 Amazon S3 中的所有现有存储桶名称中唯一的名称，并遵循 DNS 命名约定。有关更多信息，请参阅 Amazon Simple Storage Service 开发人员指南 中的存储桶命名规则。
   d. 选择 Save。

使用 AWS CLI 启用访问日志记录

使用 modify-load-balancer-attributes 命令。

验证 Elastic Load Balancing 是否在 S3 存储桶中创建了测试文件

在为负载均衡器启用访问日志记录后，Elastic Load Balancing 将验证 S3 存储桶，并创建测试文件以确保存储桶策略指定所需权限。您可以使用 Amazon S3 控制台验证是否已创建测试文件。请注意，测试文件不是实际的访问日志文件；它不包含示例记录。

1. 通过以下网址打开 Amazon S3 控制台：https://console.amazonaws.cn/s3/。
2. 对于 All Buckets，选择您的 S3 存储桶。
3. 导航到测试日志文件。路径应如下所示：

```
my-bucket/prefix/AWSLogs/123456789012/ELBAccessLogTestFile
```

管理保存访问日志的 S3 存储桶

启用访问日志记录之后，要删除存有访问日志的存储桶，请确保首先禁用访问日志记录。否则，如果在一个不属于您的 AWS 账户中创建了具有相同名称和必要的存储桶策略的新存储桶，Elastic Load Balancing 会将您的负载均衡器的访问日志写入这个新存储桶。

禁用访问日志记录

您随时可为您的负载均衡器禁用访问日志记录。在禁用访问日志记录后，您的访问日志将在 S3 存储桶中保留，直至您将其删除。有关更多信息，请参阅 Amazon Simple Storage Service 控制台用户指南 中的使用存储桶。

使用控制台禁用访问日志记录

1. 打开 Amazon EC2 控制台 https://console.amazonaws.cn/ec2/。

2. 在导航窗格中，选择 Load Balancers。
3. 选择您的负载均衡器。
4. 在 Description 选项卡上，选择 Edit attributes。
5. 在 Edit load balancer attributes 页面上，清除 Enable access logs。
6. 选择 Save。

使用 AWS CLI 禁用访问日志记录

使用 modify-load-balancer-attributes 命令。

处理访问日志文件

访问日志文件是压缩文件。如果您使用 Amazon S3 控制台打开这些文件，则将对其进行解压缩，并且将显示信息。如果您下载这些文件，则必须对其进行解压才能查看信息。

如果您的网站上有大量需求，则负载均衡器可以生成包含大量数据的日志文件 (以 GB 为单位)。您可能无法通过逐行处理来处理数量如此庞大的数据。因此，您可能必须使用提供并行处理解决方案的分析工具。例如，您可以使用以下分析工具分析和处理访问日志：

- Amazon Athena 是一种交互式查询服务，让您能够使用标准 SQL 在 Amazon S3 中轻松分析数据。有关更多信息，请参阅 Amazon Athena 用户指南 中的查询应用程序负载均衡器日志。
- Loggly
- Splunk
- Sumo Logic

针对应用程序负载均衡器的请求跟踪

您可以使用请求跟踪来跟踪 HTTP 请求 (从客户端到目标或其他服务)。当负载均衡器从客户端接收到某个请求时，它将添加或更新 X-Amzn-Trace-Id 标头，然后再将该请求发送到目标。负载均衡器和目标之间的任何服务或应用程序也可以添加或更新此标头。

如果您启用访问日志，则将记录 X-Amzn-Trace-Id 标头的内容。有关更多信息，请参阅 应用程序负载均衡器的访问日志 (p. 59)。

语法

X-Amzn-Trace-Id 标头包含使用以下格式的字段：

```
Field=version-time-id
```

字段

字段的名称。支持的值是 `Root` 和 `Self`。

应用程序可以出于自身目的添加任意字段。负载均衡器将保留这些字段，但不会使用它们。

version

版本号。

time

新纪元时间 (用秒表示)。

id

跟踪标识符。

示例

如果传入的请求中不存在 X-Amzn-Trace-Id 标头，则负载均衡器会生成一个包含 `Root` 字段的标头，然后再转发该请求。例如：

```
X-Amzn-Trace-Id: Root=1-67891233-abcdef012345678912345678
```

如果 X-Amzn-Trace-Id 标头存在并且包含 `Root` 字段，则负载均衡器将插入 `Self` 字段，然后再转发该请求。例如：

```
X-Amzn-Trace-Id: Self=1-67891234-12456789abcdef012345678;Root=1-67891233-
abcdef012345678912345678
```

如果应用程序添加了包含一个 `Root` 字段和一个自定义字段的标头，则负载均衡器将保留这两个字段并插入一个 `Self` 字段，然后再转发该请求：

```
X-Amzn-Trace-Id: Self=1-67891234-12456789abcdef012345678;Root=1-67891233-
abcdef012345678912345678;CalledFrom=app
```

如果 X-Amzn-Trace-Id 标头存在并包含 `Self` 字段，则负载均衡器将更新 `Self` 字段的值。

限制

- 负载均衡器在接收到传入的请求时将更新标头，而在接收到响应时不进行更新。
- 如果 HTTP 标头大于 7 KB，则负载均衡器将重写编写包含 `Root` 字段的 X-Amzn-Trace-Id 标头。
- 利用 WebSockets，您只可以跟踪到升级请求成功。

应用程序负载均衡器的 AWS CloudTrail 日志记录

Elastic Load Balancing 与 AWS CloudTrail 相集成，后者可捕获由您的 AWS 账户或代表该账户向 AWS 发出的 API 调用，并将日志文件传输到您指定的 Amazon S3 存储桶。使用 CloudTrail 不产生任何费用。但是，将按照 Amazon S3 的标准费率收费。

无论您是直接使用还是通过 AWS 管理控制台间接使用 AWS API (包括 Elastic Load Balancing API)，CloudTrail 都将记录对这些 API 的调用。您可以使用 CloudTrail 收集的信息，来确定发出了哪些 API 调用、使用了哪个源 IP 地址、发出调用的用户以及发出调用的时间等。

要了解有关 CloudTrail 的更多信息，包括如何对其进行配置和启用，请参阅 AWS CloudTrail User Guide有关 Elastic Load Balancing API 操作的完整列表，请参阅 Elastic Load Balancing API 参考第 2015-12-01 版。

要监控负载均衡器的其他操作 (例如，当客户端向负载均衡器发出请求时)，请使用访问日志。有关更多信息，请参阅 应用程序负载均衡器的访问日志 (p. 59)。

内容

启用 CloudTrail 事件日志记录

如果您还没有为您的账户启用 CloudTrail 事件日志记录，请按照下面的步骤操作。

启用 CloudTrail 事件日志记录

1. 在 https://console.amazonaws.cn/cloudtrail/ 打开 CloudTrail 控制台。
2. 选择 Get Started Now。
3. 对于 Trail name，键入跟踪的名称。
4. 将 Apply trail to all regions 保留为 Yes。
5. 为您的 CloudTrail 日志文件选择现有的 S3 存储桶，或者创建新存储桶。要创建新的存储桶，请为 S3 bucket 键入唯一名称。要使用现有存储桶，请将 Create a new S3 bucket 更改为 No，然后从 S3 bucket 中选择您的存储桶。
6. 选择 Turn on。

日志文件写入到 S3 存储桶中的以下位置：

```
my-bucket/AWSLogs/123456789012/CloudTrail/region/yyyy/mm/dd/
```

有关更多信息，请参阅 AWS CloudTrail User Guide。

CloudTrail 日志文件中的 Elastic Load Balancing 事件记录

CloudTrail 中的日志文件包含以 JSON 格式表示的事件信息。一条事件记录表示一次 AWS API 调用，并包含有关所请求操作的信息，例如请求操作的用户、请求的日期和时间、请求参数以及响应元素等。

这些日志文件包含 AWS 账户的所有 AWS API 调用 (而不只是 Elastic Load Balancing API 调用) 的相关事件。您可通过检查是否有包含值 `elasticloadbalancing.amazonaws.com` 的 `eventSource` 元素来查找对 Elastic Load Balancing API 的调用。要查看特定操作 (如 `CreateLoadBalancer`) 的记录，请检查是否有具有操作名称的 `eventName` 元素。要查看由 Elastic Load Balancing 代表您对 Amazon EC2 API 进行的调用的记录，请检查是否有包含值 `ec2.amazonaws.com` 的 `eventSource` 元素和包含值 `elasticloadbalancing.amazonaws.com` 的 `invokedBy` 元素的记录。例如，当您的负载均衡器扩展以应对流量的变化时，它将调用 Amazon EC2 API 创建和删除网络接口。

以下示例显示了一位用户的 CloudTrail 日志记录，该用户使用 AWS CLI 创建了一个负载均衡器，然后又删除了该负载均衡器。您可以使用 `userAgent` 元素标识 CLI。可使用 `eventName` 元素标识请求的 API 调用。有关用户 (`Alice`) 的信息可在 `userIdentity` 元素中找到。有关 CloudTrail 日志文件中不同元素和值的更多信息，请参阅 AWS CloudTrail User Guide 中的 CloudTrail 事件参考。

```
{
   "Records": [
   . . .
   {
      "eventVersion: "1.03",
      "userIdentity": {
         "type": "IAMUser",
         "principalId": "123456789012",
         "arn": "arn:aws:iam::123456789012:user/Alice",
         "accountId": "123456789012",
         "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
         "userName": "Alice"
      },
```

```
      "eventTime": "2016-04-01T15:31:48Z",
      "eventSource": "elasticloadbalancing.amazonaws.com",
      "eventName": "CreateLoadBalancer",
      "awsRegion": "us-west-2",
      "sourceIPAddress": "198.51.100.1",
      "userAgent": "aws-cli/1.10.10 Python/2.7.9 Windows/7 botocore/1.4.1",
      "requestParameters": {
         "subnets": ["subnet-8360a9e7","subnet-b7d581c0"],
         "securityGroups": ["sg-5943793c"],
         "name": "my-load-balancer",
         "scheme": "internet-facing"
      },
      "responseElements": {
         "loadBalancers":[{
            "type": "application",
            "loadBalancerName": "my-load-balancer",
            "vpcId": "vpc-3ac0fb5f",
            "securityGroups": ["sg-5943793c"],
            "state": {"code":"provisioning"},
            "availabilityZones": [
               {"subnetId":"subnet-8360a9e7","zoneName":"us-west-2a"},
               {"subnetId":"subnet-b7d581c0","zoneName":"us-west-2b"}
            ],
            "dNSName": "my-load-balancer-1836718677.us-west-2.elb.amazonaws.com",
            "canonicalHostedZoneId": "Z2P70J7HTTTPLU",
            "createdTime": "Apr 11, 2016 5:23:50 PM",
            "loadBalancerArn": "arn:aws:elasticloadbalancing:us-
west-2:123456789012:loadbalancer/app/my-load-balancer/ffcddace1759e1d0",
            "scheme": "internet-facing"
         }]
      },
      "requestID": "b9960276-b9b2-11e3-8a13-f1ef1EXAMPLE",
      "eventID": "6f4ab5bd-2daa-4d00-be14-d92efEXAMPLE",
      "eventType": "AwsApiCall",
      "apiVersion": "2015-12-01",
      "recipientAccountId": "123456789012"
   },
   . . .
   {
      "eventVersion: "1.03",
      "userIdentity": {
         "type": "IAMUser",
         "principalId": "123456789012",
         "arn": "arn:aws:iam::123456789012:user/Alice",
         "accountId": "123456789012",
         "accessKeyId": "AKIAIOSFODNN7EXAMPLE",
         "userName": "Alice"
      },
      "eventTime": "2016-04-01T15:31:48Z",
      "eventSource": "elasticloadbalancing.amazonaws.com",
      "eventName": "DeleteLoadBalancer",
      "awsRegion": "us-west-2",
      "sourceIPAddress": "198.51.100.1",
      "userAgent": "aws-cli/1.10.10 Python/2.7.9 Windows/7 botocore/1.4.1",
      "requestParameters": {
         "loadBalancerArn": "arn:aws:elasticloadbalancing:us-
west-2:123456789012:loadbalancer/app/my-load-balancer/ffcddace1759e1d0"
      },
      "responseElements": null,
      "requestID": "349598b3-000e-11e6-a82b-298133eEXAMPLE",
      "eventID": "75e81c95-4012-421f-a0cf-babdaEXAMPLE",
      "eventType": "AwsApiCall",
      "apiVersion": "2015-12-01",
      "recipientAccountId": "123456789012"
   },
   . . .
```

```
]}
```

您还可以使用与 CloudTrail 集成的 Amazon 合作伙伴解决方案之一来读取和分析 CloudTrail 日志文件。有关更多信息，请参阅 AWS CloudTrail 合作伙伴页面。

对应用程序负载均衡器进行故障排除

以下信息可帮助您解决与 应用程序负载均衡器 相关的问题。

问题

已注册目标未处于可用状态

如果目标进入 `InService` 状态所花费的时间超过预期，则该目标可能无法通过运行状况检查。您的目标未处于可用状态，除非通过一次运行状况检查。有关更多信息，请参阅 目标组的运行状况检查 (p. 44)。

验证您的实例是否通过运行状况检查，然后检查以下各项：

安全组不允许流量

与实例关联的安全组必须允许来自负载均衡器的使用运行状况检查端口和运行状况检查协议的流量。您可以向实例安全组添加一个规则以允许来自负载均衡器安全组的所有流量。负载均衡器的安全组也必须允许流入实例的流量。

网络访问控制列表 (ACL) 不允许流量

与实例的子网关联的网络 ACL 必须允许运行状况检查端口上的入站流量，以及临时端口 (1024-65535) 上的出站流量。与您负载均衡器节点的子网关联的网络 ACL 必须允许临时端口上的入站流量，以及运行状况检查端口和临时端口上的出站流量。

ping 路径不存在

创建运行状况检查的目标页并将其路径指定为 ping 路径。

连接超时

首先，验证您是否能使用目标的私有 IP 地址和运行状况检查协议直接从网络中连接到目标。如果您无法连接，请检查实例是否被过度使用，并将多个目标添加到目标组 (如果它太忙而无法响应)。如果您可以连接，则在运行状况检查超时期限之前，目标页可能不会响应。为运行状况检查选择更简单的目标页，或者调整运行状况检查设置。

目标未返回成功响应代码

默认情况下，成功代码为 200，但您可以选择在配置运行状况检查时指定其他成功代码。确认负载均衡器所需的成功代码，并且应用程序已配置为在成功时返回这些代码。

客户端无法连接到面向 Internet 的负载均衡器

如果负载均衡器未响应请求，请检查以下各项：

您的面向 Internet 的负载均衡器已连接到私有子网

确认您已为负载均衡器指定公有子网。公有子网有一个指向 Virtual Private Cloud (VPC) 的 Internet 网关的路由。

安全组或网络 ACL 不允许流量

负载均衡器的安全组和负载均衡器子网的任何网络 ACL 必须允许客户端侦听器端口上的入站流量和出站流量。

负载均衡器将请求发送到运行状况不佳的目标

如果您的负载均衡器至少有一个运行正常的已注册目标，则负载均衡器仅将请求路由到运行正常的已注册目标。如果只有运行状况不佳的已注册目标，则负载均衡器将请求路由到所有已注册目标。

负载均衡器生成 HTTP 错误

以下 HTTP 错误由负载均衡器生成。负载均衡器将 HTTP 代码发送到客户端，将请求保存到访问日志并增加 `HTTPCode_ELB_4XX_Count` 或 `HTTPCode_ELB_5XX_Count` 指标。

错误

HTTP 400：错误请求

可能的原因：

- 客户端发送的请求格式错误，不符合 HTTP 规范。
- 请求标头超出了每个请求行 16K、每单个标头 16K 或整个标头 64K 的限制。

HTTP 401: 未授权

配置了侦听器规则以验证用户的身份。已将 `OnUnauthenticatedRequest` 配置为拒绝未经身份验证的用户或 IdP 拒绝访问。

HTTP 403：禁止访问

您已配置 AWS WAF Web 访问控制列表 (Web ACL) 以监控发往您的应用程序负载均衡器的请求，并且它拦截了请求。

HTTP 460

负载均衡器收到了来自客户端的请求，但客户端在空闲超时期限结束前就关闭了与负载均衡器的连接。

检查客户端超时期限是否超过负载均衡器的空闲超时期限。确保目标在客户端超时期限之前向客户端提供响应，或增加客户端超时期限以匹配负载均衡器空闲超时 (如果客户端支持这样做)。

HTTP 463

负载均衡器收到一个包含超过 30 个 IP 地址的 X-Forwarded-For 请求标头。

HTTP 500：内部服务器错误

可能的原因：

- 您配置了一个 AWS WAF Web 访问控制列表 (web ACL)，并且在执行 Web ACL 规则时出现了错误。
- 您已配置侦听器规则以验证用户身份，但 IdP 返回的声明大小超出了负载均衡器支持的最大大小。
- 您已配置侦听器规则以验证用户身份，客户端提交了一个不带主机标头的 HTTP/1.0 请求，并且负载均衡器未能生成重定向 URL。

HTTP 502：无效网关

可能的原因：

- 负载均衡器在尝试建立连接时从目标收到了 TCP RST。
- 当负载均衡器具有目标的未完成请求时，目标关闭了具有 TCP RST 或 TCP FIN 的连接。检查目标的保持活动状态持续时间是否短于负载均衡器的空闲超时值。
- 目标响应格式错误，或者包含无效的 HTTP 标头。
- 使用了新目标组，但还没有任何目标通过初始运行状况检查。目标必须通过一次运行状况检查才能被视为正常运行。
- 负载均衡器在连接到目标时遇到 SSL 握手错误或 SSL 握手超时 (10 秒)。
- 对于已取消注册的目标正在处理的请求，取消注册延迟期已结束。增加延迟期，以便较长的操作能够完成。

HTTP 503：服务不可用

负载均衡器的目标组没有已注册目标。

HTTP 504：网关超时

可能的原因：

- 负载均衡器未能在连接超时到期 (10 秒) 之前建立与目标的连接。
- 负载均衡器与目标建立了连接，但在空闲超时周期到期之前未响应。
- 子网的网络 ACL 不允许临时端口 (1024-65535) 上从目标到负载均衡器节点的流量。
- 目标返回的 content-length 标头大于整个正文。负载均衡器因等待缺少的字节而超时。

HTTP 561: 未授权

已配置侦听器规则以验证用户的身份，但在验证用户身份时，IdP 返回错误代码。

目标生成 HTTP 错误

负载均衡器将有效的 HTTP 响应从目标转发到客户端，包括 HTTP 错误。目标生成的 HTTP 错误记录在 `HTTPCode_Target_4XX_Count` 和 `HTTPCode_Target_5XX_Count` 指标中。

应用程序负载均衡器的限制

要查看 Application Load Balancer 的当前限制，请使用 Amazon EC2 控制台的限制页面，或者使用 describe-account-limits (AWS CLI) 命令。要请求提高限制，请使用 Elastic Load Balancing Limits form。

您的 AWS 账户存在以下与 Application Load Balancer 相关的限制。

区域限制

- 每个区域的负载均衡器：20 *
- 每个区域的目标组：3000

负载均衡器限制

- 每个负载均衡器的侦听器：50
- 每个负载均衡器的目标：1000
- 每个负载均衡器每个可用区的子网数：1
- 每个负载均衡器的安全组数：5
- 每个负载均衡器的规则 (不计入默认规则)：100
- 每个负载均衡器的证书 (不计入默认证书)：25
- 每个负载均衡器可注册目标的次数：100

目标组限制

- 每个目标组的负载均衡器数：1
- 每个目标组的目标数：1000

规则限制

- 每个规则的条件数：2 (一个主机条件，一个路径条件)
- 每个规则的操作数：1
- 每个操作的目标组数：1

* 此限制包括您的Application Load Balancer和 Classic Load Balancer。

应用程序负载均衡器的文件历史记录

下表描述 Application Load Balancer 文档的重要补充部分。

- API 版本：2015-01-12

更改	描述	日期
用于 FS 和 TLS 1.2 的安全策略	此版本增加了用于向前保密 (FS) 和 TLS 1.2 的安全策略。有关更多信息，请参阅安全策略 (p. 26)。	2018 年 6 月 6 日
身份验证支持	此版本支持负载均衡器在路由请求之前使用应用程序用户的企业或社交身份对这些用户进行身份验证。有关更多信息，请参阅 使用 应用程序负载均衡器 验证用户身份 (p. 33)。	2018 年 5 月 30 日
慢启动模式	此版本增加了对慢启动模式的支持，这种模式会在新注册的目标预热时，逐渐增加负载均衡器向此目标发送的请求份额。有关更多信息，请参阅 慢启动模式 (p. 41)。	2018 年 3 月 24 日
资源级权限	此版本支持资源级权限和标记条件键。有关更多信息，请参阅 Elastic Load Balancing 用户指南 中的身份验证和访问控制。	2018 年 5 月 10 日
SNI 支持	此版本增加了对服务器名称指示 (SNI) 的支持。有关更多信息，请参阅 SSL 证书 (p. 26)。	2017 年 10 月 10 日
IP 地址即目标	此版本增加了将 IP 地址注册为目标的支持。有关更多信息，请参阅 目标类型 (p. 39)。	2017 年 8 月 31 日
基于主机的路由	此版本支持根据主机标头中的主机名路由请求。有关更多信息，请参阅 主机条件 (p. 23)。	2017 年 4 月 5 日
用于 TLS 1.1 和 TLS 1.2 的安全策略	此版本增加了用于 TLS 1.1 和 TLS 1.2 的安全策略。有关更多信息，请参阅安全策略 (p. 26)。	2017 年 2 月 6 日
IPv6 支持	此版本增加了对 IPv6 地址的支持。有关更多信息，请参阅 IP 地址类型 (p. 14)。	2017 年 1 月 25 日
请求跟踪	此版本增加了对请求跟踪的支持。有关更多信息，请参阅 针对应用程序负载均衡器的请求跟踪 (p. 67)。	2016 年 11 月 22 日

更改	描述	日期
对 TargetResponseTime 指标的百分位支持	此版本增加了对 Amazon CloudWatch 支持的新的百分位统计信息的支持。有关更多信息，请参阅 应用程序负载均衡器指标的统计数据 (p. 57)。	2016 年 11 月 17 日
新负载均衡器类型	此版本的 Elastic Load Balancing 引入了 Application Load Balancer。	2016 年 8 月 11 日