弹性负载均衡

用户指南

弹性负载均衡: 用户指南

Table of Contents

什么是 Elastic Load Balancing？

Elastic Load Balancing 在一个或多个可用区中的多个目标（如 EC2 实例、容器和 IP 地址）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。弹性负载均衡将会扩展负载均衡器容量，以响应传入流量中的变化。

负载均衡器优势

负载均衡器跨多个计算资源 (如虚拟服务器) 分布工作负载。使用负载均衡器可提高您的应用程序的可用性和容错性。

可以根据需求变化在负载均衡器中添加和删除计算资源，而不会中断应用程序的整体请求流。

您可以配置运行状况检查，这些检查监控计算资源的运行状况，以便负载均衡器只将请求发送到正常运行的目标。此外，您可以将加密和解密的工作交给负载均衡器完成，以使您的计算资源能够专注于完成主要工作。

Elastic Load Balancing 的功能

Elastic Load Balancing 支持以下负载均衡器：应用程序负载均衡器、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。有关更多信息，请参阅产品对比。

有关使用每个负载均衡器的详细信息，请参阅以下文档：

- 适用于应用程序负载均衡器的用户指南
- 适用于网络负载均衡器的用户指南
- 网关负载均衡器用户指南
- 经典负载均衡器用户指南

访问 Elastic Load Balancing

可以使用以下任意接口创建、访问和管理负载均衡器：

- AWS Management Console –提供可用于访问 Elastic Load Balancing 的 Web 界面。
- AWS Command Line Interface (AWS CLI) — 为众多 AWS 服务（包括 Elastic Load Balancing）提供命令。AWS CLI 在 Windows、macOS 和 Linux 上受支持。有关更多信息，请参阅 AWS Command Line Interface。
- AWS 开发工具包 — 提供特定于语言的 API，并关注许多连接详细信息，例如计算签名、处理请求重试和错误处理。有关更多信息，请参阅 AWS 开发工具包。
- 查询 API — 提供您使用 HTTPS 请求调用的低级别 API 操作。使用查询 API 是访问 Elastic Load Balancing 的最直接方式。但是，查询 API 需要您的应用程序处理低级别的详细信息，例如生成哈希值以签署请求以及进行错误处理。有关更多信息，请参阅下列内容：
  - 应用程序负载均衡器 和 Network Load Balancer — API 版本 2015-12-01
  - 经典负载均衡器— API 版本 2012-06-01

相关服务

弹性负载均衡 可与以下服务一起使用，以提高应用程序的可用性和可扩展性。

- Amazon EC2 — 在云中运行应用程序的虚拟服务器。您可以将负载均衡器配置为将流量路由到您的 EC2 实例。有关更多信息，请参阅适用于 Linux 实例的 Amazon EC2 用户指南或适用于 Windows 实例的 Amazon EC2 用户指南。
- Amazon EC2 Auto Scaling — 确保运行所需数量的实例，即使实例失败也是如此。您还可以利用 Amazon EC2 Auto Scaling 在实例需求变化时自动增加或减少实例数量。如果通过 Elastic Load Balancing 启用 Auto Scaling，则由 Auto Scaling 启动的实例将自动注册到负载均衡器。同样，由 Auto Scaling 终止的实例将自动从负载均衡器取消注册。有关更多信息，请参阅 Amazon EC2 Auto Scaling 用户指南。
- AWS Certificate Manager – 在创建 HTTPS 侦听器时，您必须指定由 ACM 提供的证书。负载均衡器使用证书终止连接并解密来自客户端的请求。
- Amazon CloudWatch — 使您能够监控负载均衡器并执行所需操作。有关更多信息，请参阅 Amazon CloudWatch 用户指南。
- Amazon ECS — 使您能够在 EC2 实例集群上运行、停止和管理 Docker 容器。您可以将负载均衡器配置为将流量路由到您的容器。有关更多信息，请参阅 Amazon Elastic Container Service 开发人员指南。
- AWS Global Accelerator — 提高应用程序的可用性和性能。使用加速器在一个或多个 AWS 区域的多个负载均衡器之间分配流量。有关更多信息，请参阅 AWS Global Accelerator 开发人员指南。
- Route 53 — 通过将域名转换为计算机相互连接所用的数字 IP 地址，以一种可靠且经济的方式将访问者路由至网站。例如，它会将 `www.example.com` 转换为数字 IP 地址 `192.0.2.1`。AWS 会向您的资源（例如负载均衡器）分配 URL。不过，您可能希望使用方便用户记忆的 URL。例如，您可以将域名映射到负载均衡器。有关更多信息，请参阅 Amazon Route 53 开发人员指南。
- 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。

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。

Elastic Load Balancing 的工作原理

负载均衡器接受来自客户端的传入流量并将请求路由到一个或多个可用区中的已注册目标 (例如 EC2 实例)。负载均衡器还会监控已注册目标的运行状况，并确保它只将流量路由到正常运行的目标。当负载均衡器检测到不正常目标时，它会停止将流量路由到该目标。然后，当它检测到目标再次正常时，它会恢复将流量路由到该目标。

您可通过指定一个或多个侦听器将您的负载均衡器配置为接受传入流量。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号。同样，它配置了用于从负载均衡器连接到目标的协议和端口号。

Elastic Load Balancing 支持以下类型的负载均衡器：

- Application Load Balancer
- Network Load Balancer
- 网关负载均衡器
- 经典负载均衡器

负载均衡器类型的配置方式具有一个关键区别。对于 Application Load Balancer、Network Load Balancer 和 Gateway Load Balancer，可以在目标组中注册目标，并将流量路由到目标组。通过经典负载均衡器，可以在负载均衡器中注册实例。

可用区与负载均衡器节点

当您为负载均衡器启用可用区时，Elastic Load Balancing 会在该可用区中创建一个负载均衡器节点。如果您在可用区中注册目标但不启用可用区，这些已注册目标将无法接收流量。当您确保每个启用的可用区均具有至少一个已注册目标时，负载均衡器将具有最高效率。

我们建议为所有负载均衡器启用多个可用区。但对于 Application Load Balancer，要求您至少启用两个或更多可用区。此配置有助于确保负载均衡器可以继续路由流量。如果一个可用区变得不可用或没有正常目标，则负载均衡器会将流量路由到其他可用区中的正常目标。

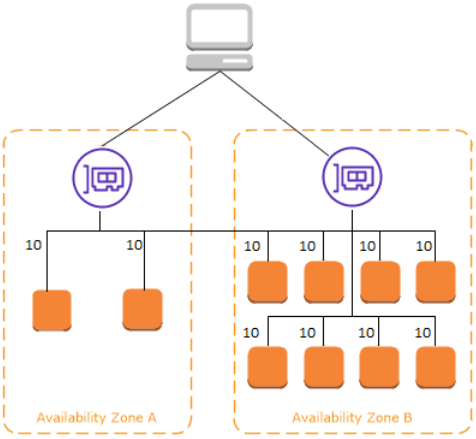
在禁用一个可用区后，该可用区中的目标将保持已注册到负载均衡器的状态。但是，即使它们保持已注册状态，负载均衡器也不会将流量路由到它们。

跨区域负载均衡

负载均衡器的节点将来自客户端的请求分配给已注册目标。启用了跨区域负载均衡后，每个负载均衡器节点会在所有启用的可用区中的已注册目标之间分配流量。禁用了跨区域负载均衡后，每个负载均衡器节点会仅在其可用区中的已注册目标之间分配流量。

下图演示了以轮询为默认路由算法的跨可用区负载均衡效果。有 2 个已启用的可用区，其中可用区 A 中有 2 个目标，可用区 B 中有 8 个目标。客户端发送请求，Amazon Route 53 使用负载均衡器节点之一的 IP 地址响应每个请求。基于轮询路由算法，系统会分配流量，以便每个负载均衡器节点接收来自客户端 50% 的流量。每个负载均衡器节点会在其范围中的已注册目标之间分配其流量份额。

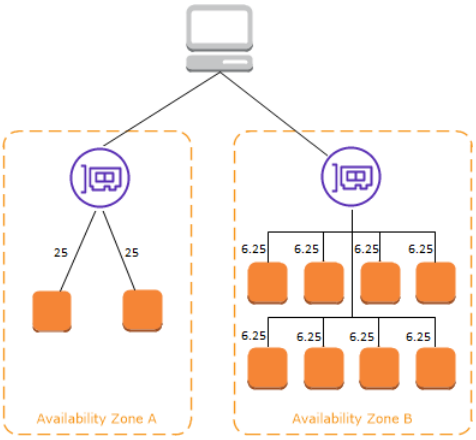
如果启用了跨区域负载均衡，则 10 个目标中的每个目标接收 10% 的流量。这是因为每个负载均衡器节点可将其 50% 的客户端流量路由到所有 10 个目标。



如果禁用了跨区域负载均衡：

- 可用区 A 中的两个目标中的每个目标接收 25% 的流量。
- 可用区 B 中的八个目标中的每个目标接收 6.25% 的流量。

这是因为每个负载均衡器节点只能将其 50% 的客户端流量路由到其可用区中的目标。



对于 Application Load Balancer，跨区域负载均衡始终处于启用状态。

对于 Network Load Balancer 和 Gateway Load Balancer，默认情况下会禁用跨区域负载均衡。创建负载均衡器后，您随时可以启用或禁用跨区域负载均衡。

在创建经典负载均衡器时，跨区域负载均衡的默认值取决于创建负载均衡器的方式。默认情况下，使用 API 或 CLI 时将禁用跨区域负载均衡。默认情况下，使用 AWS Management Console时启用跨区域负载均衡的选项处于选中状态。创建经典负载均衡器后，您随时可以启用或禁用跨区域负载均衡。有关更多信息，请参阅《经典负载均衡器用户指南》中的启用跨区域负载均衡。

请求路由选择

在客户端将请求发送到负载均衡器之前，它会利用域名系统 (DNS) 服务器解析负载均衡器的域名。DNS 条目由 Amazon 控制，因为您的负载均衡器位于 amazonaws.com 域中。Amazon DNS 服务器会将一个或多个 IP 地址返回到客户端。这些是您的负载均衡器的负载均衡器节点的 IP 地址。对于 Network Load Balancer，Elastic Load Balancing 将为启用的每个可用区创建一个网络接口。可用区内的每个负载均衡器节

点使用该网络接口来获取一个静态 IP 地址。在您创建负载均衡器时，可以选择将一个弹性 IP 地址与每个网络接口关联。

当流向应用程序的流量随时间变化时，Elastic Load Balancing 会扩展负载均衡器并更新 DNS 条目。DNS 条目还指定生存时间 (TTL) 为 60 秒。这有助于确保可以快速重新映射 IP 地址以响应不断变化的流量。

客户端可以确定使用哪个 IP 地址将请求发送到负载均衡器。用于接收请求的负载均衡器节点会选择一个正常运行的已注册目标，并使用其私有 IP 地址将请求发送到该目标。

路由算法

借助 Application Load Balancer，接收请求的负载均衡器节点使用以下过程：

1. 按优先级顺序评估侦听器规则以确定要应用的规则。
2. 使用为目标组配置的路由算法，从目标组中为规则操作选择目标。默认路由算法是轮询。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。

借助 Network Load Balancer，接收连接的负载均衡器节点使用以下过程：

1. 使用流哈希算法从目标组中为默认规则选择目标。它使算法基于：
   - 协议
   - 源 IP 地址和源端口
   - 目标 IP 地址和目标端口
   - TCP 序列号
2. 将每个单独的 TCP 连接在连接的有效期内路由到单个目标。来自客户端的 TCP 连接具有不同的源端口和序列号，可以路由到不同的目标。

借助 经典负载均衡器，接收请求的负载均衡器节点按照以下方式选择注册实例：

- 使用适用于 TCP 侦听器的轮询路由算法
- 使用适用于 HTTP 和 HTTPS 侦听器的最少未完成请求路由算法

HTTP 连接

经典负载均衡器会使用预打开连接，但 Application Load Balancer 不会使用预打开连接。经典负载均衡器和 Application Load Balancer 均使用多路复用连接。也就是说，来自多个前端连接上的多个客户端的请求可通过单一的后端连接路由到指定目标。多路复用连接可缩短延迟并减少您的应用程序上的负载。要禁止多路复用连接，请在您的 HTTP 响应中设置 `Connection: close` 标头来禁用 HTTP `keep-alive` 标头。

对于前端连接，Application Load Balancer 和经典负载均衡器支持管道化 HTTP。对于后端连接它们均不支持管道化 HTTP。

对于前端连接，Application Load Balancer 支持以下协议：HTTP/0.9、HTTP/1.0、HTTP/1.1 和 HTTP/2。HTTP/2 仅适用于 HTTPS 侦听器，使用一个 HTTP/2 连接最多可并行发送 128 个请求。Application Load Balancer 还支持将连接从 HTTP 升级到 WebSocket。但是，如果进行连接升级，Application Load Balancer 侦听器路由规则和 AWS WAF 集成将不再适用。

默认情况下，Application Load Balancer 在后端连接上使用 HTTP/1.1（负载均衡器连接到已注册的目标）。但是，您可以通过协议版本使用 HTTP/2 或 gRPC 将请求发送到目标。有关更多信息，请参阅协议版本。默认情况下，`keep-alive` 标头在后端连接上受支持。如果 HTTP/1.0 请求来自没有主机标头的客户端，负载均衡器会对后端连接发送的 HTTP/1.1 请求生成一个主机标头。主机标头包含负载均衡器的 DNS 名称。

对于前端连接（客户端到负载均衡器），经典负载均衡器支持以下协议：HTTP/0.9、HTTP/1.0 和 HTTP/1.1。默认情况下，它们在后端连接（已注册目标的负载均衡器）上使用 HTTP/1.1。默认情况

下，keep-alive 标头在后端连接上受支持。如果 HTTP/1.0 请求来自没有主机标头的客户端，负载均衡器会对后端连接发送的 HTTP/1.1 请求生成一个主机标头。主机标头包含负载均衡器节点的 IP 地址。

HTTP 标头

Application Load Balancer 和经典负载均衡器会将 X-Forwarded-For、X-Forwarded-Proto 和 X-Forwarded-Port 标头自动添加到请求。

应用程序负载均衡器将 HTTP 主机标头中的主机名转换为小写，然后再将其发送到目标。

对于使用 HTTP/2 的前端连接，标头名称是小写的。使用 HTTP/1.1 将请求发送到目标之前，以下标头名称将转换为混合大小写：X-Forwarded-For、X-Forwarded-Proto、X-Forwarded-Port、Host、X-Amzn-Trace-Id、Upgrade 和 Connection。所有其他标头名称是小写的。

Application Load Balancer 和经典负载均衡器将响应代理返回客户端后，遵守来自传入客户端请求的连接标头。

当 Application Load Balancer 和经典负载均衡器收到 Expect 标头时，它们会立即使用 HTTP 100 Continue 响应客户端而不测试内容长度标头，然后会删除 Expect 标头，再路由请求。

HTTP 标头限制

应用程序负载均衡器的以下大小限制是无法更改的硬限制：

- 请求行：16K
- 单个标头：16K
- 整个响应标头：32 K
- 整个请求标头：64 K

负载均衡器模式

在创建负载均衡器时，您必须选择使其成为内部负载均衡器还是面向 Internet 的负载均衡器。请注意，当您在 EC2-Classic 中创建经典负载均衡器时，它必须是面向 Internet 的负载均衡器。

> 我们将于 2022 年 8 月 15 日停用 EC2-Classic 网络。我们建议您将经典负载均衡器从 EC2-Classic 网络迁移到 VPC。有关更多信息，请参阅《Amazon EC2 用户指南》中的从 EC2-Classic 迁移到 VPC 和博客 EC2-Classic Networking 即将停用 – 以下是准备方法。

面向 Internet 的负载均衡器的节点具有公共 IP 地址。面向 Internet 的负载均衡器的 DNS 名称可公开解析为节点的公共 IP 地址。因此，面向 Internet 的负载均衡器可以通过 Internet 路由来自客户端的请求。

内部负载均衡器的节点只有私有 IP 地址。内部负载均衡器的 DNS 名称可公开解析为节点的私有 IP 地址。因此，内部负载均衡器可路由的请求只能来自对负载均衡器的 VPC 具有访问权限的客户端。

面向 Internet 的负载均衡器和内部负载均衡器均使用私有 IP 地址将请求路由到您的目标。因此，您的目标无需使用公有 IP 地址从内部负载均衡器或面向 Internet 的负载均衡器接收请求。

如果您的应用程序具有多个层，则可以设计一个同时使用内部负载均衡器和面向 Internet 的负载均衡器的架构。例如，如果您的应用程序使用必须连接到 Internet 的 Web 服务器，以及仅连接到 Web 服务器的应用程序服务器，则可以如此。创建一个面向 Internet 的负载均衡器并向其注册 Web 服务器。创建一个内部负载均衡器并向它注册应用程序服务器。Web 服务器从面向 Internet 的负载均衡器接收请求，并将对应用程序服务器的请求发送到内部负载均衡器。应用程序服务器从内部负载均衡器接收请求。

您的负载均衡器的网络 MTU

网络连接的最大传输单位 (MTU) 是能够通过该连接传递的最大可允许数据包的大小 (以字节为单位)。连接的 MTU 越大，可在单个数据包中传递的数据越多。以太网数据包由帧 (或您发送的实际数据) 和围绕它的网络开销信息组成。通过互联网网关发送的流量限制为 1500 MTU。这就意味着，如果数据包大于 1500 字节，则对数据包进行分段；如果在 IP 标头中设置了 `Don't Fragment` 标记，则丢弃数据包。

负载均衡器节点上的 MTU 大小不可配置。Jumbo 帧 (9001 MTU) 在应用程序负载均衡器、网络负载均衡器和经典负载均衡器的负载均衡器节点中是标准的。网关负载均衡器支持 8500 MTU。有关更多信息，请参阅网关负载均衡器用户指南中的最大传输单位 (MTU)。

路径 MTU 是原始主机和接收主机之间的路径所支持的最大数据包大小。路径 MTU 发现 (PMTUD) 用于确定两台设备之间的路径 MTU。如果客户端或目标不支持巨型帧，路径 MTU 发现特别重要。

如果主机发送一个大于接收主机的 MTU 或大于路径上某台设备的 MTU 的数据包，则接收主机或设备将丢弃此数据包，然后返回以下 ICMP 消息：`Destination Unreachable: Fragmentation Needed and Don't Fragment was Set (Type 3, Code 4)`。这将指示传输主机将有效负载拆分为多个较小的数据包，并重新传输。

如果继续丢弃大于客户端或目标接口 MTU 大小的数据包，则可能是路径 MTU 发现 (PMTUD) 不起作用。为了避免这种情况，请确保路径 MTU 发现端到端工作，并且您已在客户端和目标上启用了巨型帧。有关路径 MTU 发现和启用巨型帧的详细信息，请参阅 Amazon EC2 用户指南中的路径 MTU 发现。

Elastic Load Balancing 入门

Elastic Load Balancing 支持以下负载均衡器：Application Load Balancer、Network Load Balancer、Gateway Load Balancer 和 Classic Load Balancer。您可以选择最适合自己需求的负载均衡器类型。有关更多信息，请参阅产品对比。

有关常见负载均衡器配置的演示，请参阅 Elastic Load Balancing 演示。

如果您有现有的 Classic Load Balancer，则可以迁移到 Application Load Balancer 或 Network Load Balancer。有关更多信息，请参阅 迁移您的经典负载均衡器 (p. 24)。

目录

创建 Application Load Balancer

要使用 AWS Management Console 创建 Application Load Balancer，请参阅 Application Load Balancers 用户指南中的 Application Load Balancer 入门。

要使用 AWS CLI 创建 Application Load Balancer，请参阅 Application Load Balancers 用户指南中的使用 AWS CLI 创建 Application Load Balancer。

创建网络负载均衡器

要使用 AWS Management Console 创建 Network Load Balancer，请参阅 Network Load Balancers 用户指南中的Network Load Balancers 入门。

要使用 AWS CLI 创建 Network Load Balancer，请参阅 Network Load Balancers 用户指南中的使用 AWS CLI 创建 Network Load Balancer。

创建网关负载均衡器

要使用 AWS Management Console 创建网关负载均衡器，请参阅网关负载均衡器用户指南中的网关负载均衡器入门。

要使用 AWS CLI 创建网关负载均衡器，请参阅网关负载均衡器用户指南中的使用 AWS CLI 网关负载均衡器入门。

创建 Classic Load Balancer

要使用 AWS Management Console 创建 Classic Load Balancer，请参阅 Classic Load Balancers 用户指南中的创建 Classic Load Balancer。

Elastic Load Balancing 中的安全性

AWS的云安全性的优先级最高。作为 AWS 客户，您将从专为满足大多数安全敏感型组织的要求而打造的数据中心和网络架构中受益。

安全性是 AWS 和您的共同责任。责任共担模式将其描述为云的安全性和云中的安全性：

- 云的安全性 – AWS 负责保护在 AWS 云中运行 AWS 服务的基础设施。AWS 还向您提供可安全使用的服务。作为 AWS 合规性计划的一部分，第三方审计人员将定期测试和验证安全措施的有效性。要了解适用于 Elastic Load Balancing 的合规性计划，请参阅合规性计划范围内的 AWS 服务。
- 云中的安全性 - 您的责任由您使用的 AWS 服务决定。您还需要对其他因素负责，包括您的数据的敏感性、您公司的要求以及适用的法律法规。

此文档将帮助您了解如何在使用 Elastic Load Balancing 时应用责任共担模式。其中说明了如何配置 Elastic Load Balancing 以实现您的安全性和合规性目标。您还会了解如何使用其他 AWS 服务以帮助您监控和保护 Elastic Load Balancing 资源。

对于网关负载均衡器，您要负责从设备供应商那里选择和鉴定软件。您必须信任设备软件才能检查或修改来自负载均衡器的流量，负载均衡器在开放系统互连 (OSI) 模型的第 3 层（网络层）运行。列为 Elastic Load Balancing 合作伙伴的设备供应商已与 AWS 集成并已鉴定其提供的设备软件。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。

目录

Elastic Load Balancing 中的数据保护

AWS 责任共担模式适用于 Elastic Load Balancing 中的数据保护。如该模式中所述，AWS 负责保护运行所有 AWS Cloud 的全球基础设施。您负责维护对托管在此基础设施上的内容的控制。此内容包括您所使用的 AWS 服务的安全配置和管理任务。有关数据隐私的更多信息，请参阅数据隐私常见问题。有关欧洲数据保护的信息，请参阅AWS安全性博客上的AWS责任共担模式和 GDPR 博客文章。

出于数据保护目的，我们建议您保护 AWS 账户 凭证并使用 AWS Identity and Access Management (IAM) 设置单独的用户账户。这仅向每个用户授予履行其工作职责所需的权限。我们还建议您通过以下方式保护您的数据：

- 对每个账户使用 Multi-Factor Authentication (MFA)。
- 使用 SSL/TLS 与 AWS 资源进行通信。建议使用 TLS 1.2 或更高版本。
- 使用 AWS CloudTrail 设置 API 和用户活动日志记录。
- 使用 AWS 加密解决方案以及 AWS 服务中的所有默认安全控制。
- 使用高级托管安全服务（例如 Amazon Macie），它有助于发现和保护存储在 Simple Storage Service（Amazon S3）中的个人数据。
- 如果在通过命令行界面或 API 访问 AWS 时需要经过 FIPS 140-2 验证的加密模块，请使用 FIPS 端点。有关可用的 FIPS 端点的更多信息，请参阅《美国联邦信息处理标准 (FIPS) 第 140-2 版》。

我们强烈建议您切勿将机密信息或敏感信息（例如您客户的电子邮件地址）放入标签或自由格式字段（例如名称字段）。这包括通过控制台、API、AWS CLI CLI 或 AWS 开发工具包使用 Elastic Load Balancing 或其他 AWS 服务时。您在用于名称的标签或自由格式字段中输入的任何数据都可能会用于计费或诊断日志。当您向外部服务器提供 URL 时，强烈建议您不要在 URL 中包含凭证信息来验证您对该服务器的请求。

静态加密

如果您为用于 Elastic Load Balancing 访问日志的 S3 存储桶启用了使用 Amazon S3 托管加密密钥 (SSE-S3) 的服务器端加密，则 Elastic Load Balancing 会先自动加密每个访问日志文件，然后再存储到 S3 存储桶中。Elastic Load Balancing 还会在您对访问日志文件进行访问时对其进行解密。每个日志文件都使用一个唯一密钥进行加密，此密钥本身将使用定期轮换的主密钥进行加密。

传输中加密

Elastic Load Balancing 通过在负载均衡器上终止来自客户端的 HTTPS 和 TLS 流量，从而简化了构建安全 Web 应用程序的过程。负载均衡器会执行加密和解密流量的工作，而不要求每个 EC2 实例来处理 TLS 终止工作。在配置安全侦听器时，您可以指定应用程序支持的密码套件和协议版本，以及要在您的负载均衡器上安装的服务器证书。您可以使用 AWS Certificate Manager (ACM) 或者 AWS Identity and Access Management (IAM) 来管理您的服务器证书。Application Load Balancer 支持 HTTPS 侦听器。Network Load Balancer 支持 TLS 侦听器。Classic Load Balancer 同时支持 HTTPS 和 TLS 侦听器。

适用于 Elastic Load Balancing 的 Identity and Access Management

AWS 使用安全凭证来识别您的身份并向您授予对 AWS 资源的访问权限。AWS Identity and Access Management (IAM) 的功能允许其他用户、服务和应用程序完全使用或受限使用您的 AWS 资源。您可以在不共享安全凭证的情况下执行此操作。

默认情况下，IAM 用户没有创建、查看或修改 AWS 资源的权限。要允许 IAM 用户访问资源（如负载均衡器）并执行任务，您可以：

1. 创建授予 IAM 用户使用所需特定资源和 API 操作的权限的 IAM 策略。
2. 将该策略附加到 IAM 用户或 IAM 用户所属的组。

在将策略附加到一个用户或一组用户时，它会授权或拒绝用户使用指定资源执行指定任务。

例如，您可以使用 IAM 在您的 AWS 账户下创建用户和组。IAM 用户可以是人员、系统或应用程序。然后，使用 IAM 策略向用户和组授予对指定资源执行特定操作的权限。

使用 IAM 策略授予权限

在将策略附加到一个用户或一组用户时，它会授权或拒绝用户使用指定资源执行指定任务。

IAM 策略是包含一个或多个语句的 JSON 文档。每个语句的结构如下例所示。

```
{
  "Version": "2012-10-17",
  "Statement":[{
    "Effect": "effect",
    "Action": "action",
    "Resource": "resource-arn",
    "Condition": {
      "condition": {
        "key":"value"
```

```
      }
    }
  }]
}
```

- Effect（效果）：effect（效果）可以是 `Allow` 或 `Deny`。在默认情况下，IAM 用户没有使用资源和 API 操作的许可，因此，所有请求均会被拒绝。显式允许将覆盖默认规则。显式拒绝将覆盖任何允许。
- Action（操作）：action（操作）是对其授予或拒绝权限的特定 API 操作。有关指定 action（操作）的更多信息，请参阅Elastic Load Balancing 的 API 操作 (p. 11)。
- Resource（资源）：受操作影响的资源。利用许多 Elastic Load Balancing API 操作，您可以限制对特定负载均衡器授予或拒绝的权限。为此，请在此语句中指定其 Amazon Resource Name (ARN)。否则，您可以使用 * 通配符指定所有负载均衡器。有关更多信息，请参阅 Elastic Load Balancing 资源 (p. 11)。
- Condition（条件）：您可以选择性地使用条件来控制策略的生效时间。有关更多信息，请参阅 Elastic Load Balancing 的条件键 (p. 14)。

有关更多信息，请参阅 IAM 用户指南。

Elastic Load Balancing 的 API 操作

在 IAM 策略语句的 Action（操作）元素中，您可以指定 Elastic Load Balancing 所提供的任意 API 操作。如以下示例所示，您必须使用小写形式的字符串 `elasticloadbalancing:` 作为操作名称的前缀。

```
"Action": "elasticloadbalancing:DescribeLoadBalancers"
```

要在单个语句中指定多项操作，请使用方括号将操作括起来并以逗号分隔，如以下示例所示。

```
"Action": [
    "elasticloadbalancing:DescribeLoadBalancers",
    "elasticloadbalancing:DeleteLoadBalancer"
]
```

您也可以使用 * 通配符指定多项操作。以下示例指定以 `Describe` 开头的所有 Elastic Load Balancing API 操作名称。

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

要指定所有 Elastic Load Balancing API 操作，请按以下示例所示使用 * 通配符。

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

有关 Elastic Load Balancing API 操作的完整列表，请参阅以下文档：

- Application Load Balancer 和 Network Load Balancer — API 参考版本 2015-12-01
- Classic Load Balancer — API 参考版本 2012-06-01

Elastic Load Balancing 资源

资源级权限指的是能够指定允许用户对哪些资源执行操作的能力。Elastic Load Balancing 部分支持资源级权限。对于支持资源级权限的 API 操作，您可以控制用户可与操作结合使用的资源。要在策略中指定资源，您必须使用其 Amazon Resource Name (ARN)。指定 ARN 时，您可以在路径中使用 * 通配符。例如，当您不想指定确切的负载均衡器名称时，可以使用 * 通配符。

Application Load Balancer 的 ARN 具有以下示例中显示的格式。

```
arn:aws:elasticloadbalancing:region-code:account-id:loadbalancer/app/load-balancer-name/load-balancer-id
```

Network Load Balancer 的 ARN 具有以下示例中显示的格式。

```
arn:aws:elasticloadbalancing:region-code:account-id:loadbalancer/net/load-balancer-name/load-balancer-id
```

Classic Load Balancer 的 ARN 具有以下示例中显示的格式。

```
arn:aws:elasticloadbalancing:region-code:account-id:loadbalancer/load-balancer-name
```

侦听器的 ARN 和 Application Load Balancer 的侦听器规则具有以下示例中显示的格式。

```
arn:aws:elasticloadbalancing:region-code:account-id:listener/app/load-balancer-name/load-balancer-id/listener-id
arn:aws:elasticloadbalancing:region-code:account-id:listener-rule/app/load-balancer-name/load-balancer-id/listener-id/rule-id
```

Network Load Balancer 的侦听器的 ARN 具有以下示例中显示的格式。

```
arn:aws:elasticloadbalancing:region-code:account-id:listener/net/load-balancer-name/load-balancer-id/listener-id
```

目标组的 ARN 具有以下示例中显示的格式。

```
arn:aws:elasticloadbalancing:region-code:account-id:targetgroup/target-group-name/target-group-id
```

不支持资源级权限的 API 操作

以下 Elastic Load Balancing 操作不支持资源级权限：

- API 版本 2015-12-01：
  - DescribeAccountLimits
  - DescribeListenerCertificates
  - DescribeListeners
  - DescribeLoadBalancerAttributes
  - DescribeLoadBalancers
  - DescribeRules
  - DescribeSSLPolicies
  - DescribeTags
  - DescribeTargetGroupAttributes
  - DescribeTargetGroups
  - DescribeTargetHealth
- API 版本 2012-06-01：
  - DescribeInstanceHealth
  - DescribeLoadBalancerAttributes
  - DescribeLoadBalancerPolicyTypes
  - DescribeLoadBalancers

- DescribeLoadBalancerPolicies
- DescribeTags

对于不支持资源级权限的 API 操作，必须指定以下示例中显示的资源语句。

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

Elastic Load Balancing 的资源级权限

下列各表介绍了支持资源级权限的 Elastic Load Balancing 操作，以及每个操作支持的资源。

API 版本 2015-12-01

API 操作	资源 ARN
AddListenerCertificates	侦听器
AddTags	负载均衡器，目标组
CreateListener	负载均衡器
CreateLoadBalancer	负载均衡器
CreateRule	侦听器
CreateTargetGroup	目标组
DeleteListener	侦听器
DeleteLoadBalancer	负载均衡器
DeleteRule	侦听器规则
DeleteTargetGroup	目标组
DeregisterTargets	目标组
ModifyListener	侦听器
ModifyLoadBalancerAttributes	负载均衡器
ModifyRule	侦听器规则
ModifyTargetGroup	目标组
ModifyTargetGroupAttributes	目标组
RegisterTargets	目标组
RemoveListenerCertificates	侦听器
RemoveTags	负载均衡器，目标组
SetIpAddressType	负载均衡器
SetRulePriorities	侦听器规则
SetSecurityGroups	负载均衡器
SetSubnets	负载均衡器

API 版本 2012-06-01

API 操作	资源 ARN
AddTags	负载均衡器
ApplySecurityGroupsToLoadBalancer	负载均衡器
AttachLoadBalancerToSubnets	负载均衡器
ConfigureHealthCheck	负载均衡器
CreateAppCookieStickinessPolicy	负载均衡器
CreateLBCookieStickinessPolicy	负载均衡器
CreateLoadBalancer	负载均衡器
CreateLoadBalancerListeners	负载均衡器
CreateLoadBalancerPolicy	负载均衡器
DeleteLoadBalancer	负载均衡器
DeleteLoadBalancerListeners	负载均衡器
DeleteLoadBalancerPolicy	负载均衡器
DeregisterInstancesFromLoadBalancer	负载均衡器
DetachLoadBalancerFromSubnets	负载均衡器
DisableAvailabilityZonesForLoadBalancer	负载均衡器
EnableAvailabilityZonesForLoadBalancer	负载均衡器
ModifyLoadBalancerAttributes	负载均衡器
RegisterInstancesWithLoadBalancer	负载均衡器
RemoveTags	负载均衡器
SetLoadBalancerListenerSSLCertificate	负载均衡器
SetLoadBalancerPoliciesForBackendServer	负载均衡器
SetLoadBalancerPoliciesOfListener	负载均衡器

Elastic Load Balancing 的条件键

在创建策略时，您可指定控制策略生效时间的条件。每个条件都包含一个或多个键值对。有全局条件键和特定于服务的条件键。

不能将 aws:SourceIp 条件键与 Elastic Load Balancing 一起使用。

elasticloadbalancing:ResourceTag/*key* 条件键特定于 Elastic Load Balancing。以下操作支持此条件键：

API 版本 2015-12-01

- AddTags

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

API 版本 2012-06-01

- AddTags
- ApplySecurityGroupsToLoadBalancer
- AttachLoadBalancersToSubnets
- ConfigureHealthCheck
- CreateAppCookieStickinessPolicy
- CreateLBCookieStickinessPolicy
- CreateLoadBalancer
- CreateLoadBalancerListeners
- CreateLoadBalancerPolicy
- DeleteLoadBalancer
- DeleteLoadBalancerListeners
- DeleteLoadBalancerPolicy
- DeregisterInstancesFromLoadBalancer
- DetachLoadBalancersFromSubnets
- DisableAvailabilityZonesForLoadBalancer
- EnableAvailabilityZonesForLoadBalancer
- ModifyLoadBalancerAttributes
- RegisterInstancesWithLoadBalancer
- RemoveTags
- SetLoadBalancerListenerSSLCertificate
- SetLoadBalancerPoliciesForBackendServer
- SetLoadBalancerPoliciesOfListener

有关全局条件键的更多信息，请参阅 IAM 用户指南中的 AWS 全球条件上下文键。

以下操作支持 aws:RequestTag/*key* 和 aws:TagKeys 条件键：

- AddTags
- CreateLoadBalancer

- RemoveTags

预定义的 AWS 托管策略

AWS 创建的托管策略将授予针对常用案例的必要权限。您可以根据您的 IAM 用户对 Elastic Load Balancing 所需的访问权限将这些策略附加到这些用户：

- ElasticLoadBalancingFullAccess — 授予使用 Elastic Load Balancing 功能所需的完整访问权限。
- ElasticLoadBalancingReadOnly — 授予对 Elastic Load Balancing 功能的只读访问权限。

有关每个 Elastic Load Balancing 操作所需的权限的更多信息，请参阅Elastic Load Balancing API 权限 (p. 16)。

Elastic Load Balancing API 权限

您必须为 IAM 用户授予调用所需 Elastic Load Balancing API 操作的权限，如Elastic Load Balancing 的 API 操作 (p. 11)中所述。此外，对于某些 Elastic Load Balancing 操作，您必须授予 IAM 用户从 Amazon EC2 API 调用特定操作的权限。

2015-12-01 API 所需的权限

从 2015-12-01 API 调用以下操作时，您必须授予 IAM 用户调用指定操作的权限。

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

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

RegisterTargets
- elasticloadbalancing:RegisterTargets
- ec2:DescribeInstances
- ec2:DescribeInternetGateways
- ec2:DescribeSubnets
- ec2:DescribeVpcs

SetIpAddressType
- elasticloadbalancing:SetIpAddressType
- ec2:DescribeSubnets

SetSubnets
- elasticloadbalancing:SetSubnets
- ec2:DescribeSubnets

2012-06-01 API 所需的权限

从 2012-06-01 API 调用以下操作时，您必须授予 IAM 用户调用指定操作的权限。

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

AttachLoadBalancerToSubnets
- elasticloadbalancing:AttachLoadBalancerToSubnets
- ec2:DescribeSubnets

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

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

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

DescribeLoadBalancers
- elasticloadbalancing:DescribeLoadBalancers
- ec2:DescribeSecurityGroups

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

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

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

- ec2:DescribeInstances
- ec2:DescribeVpcClassicLink

Elastic Load Balancing 服务相关角色

Elastic Load Balancing 使用服务相关角色来获取它代表您调用其他 AWS 服务所需的权限。有关更多信息，请参阅 IAM 用户指南 中的使用服务相关角色。

服务相关角色授予的权限

Elastic Load Balancing 使用名为 AWSServiceRoleForElasticLoadBalancing 的服务相关角色代表您调用以下操作：

- ec2:DescribeAddresses
- ec2:DescribeInstances
- ec2:DescribeNetworkInterfaces
- ec2:DescribeSubnets
- ec2:DescribeSecurityGroups
- ec2:DescribeVpcs
- ec2:DescribeInternetGateways
- ec2:DescribeAccountAttributes
- ec2:DescribeClassicLinkInstances
- ec2:DescribeVpcClassicLink
- ec2:CreateSecurityGroup
- ec2:CreateNetworkInterface
- ec2:DeleteNetworkInterface
- ec2:ModifyNetworkInterfaceAttribute
- ec2:AuthorizeSecurityGroupIngress
- ec2:AssociateAddress
- ec2:DisassociateAddress
- ec2:AttachNetworkInterface
- ec2:DetachNetworkInterface
- ec2:AssignPrivateIpAddresses
- ec2:AssignIpv6Addresses
- ec2:UnassignIpv6Addresses
- logs:CreateLogDelivery
- logs:GetLogDelivery
- logs:UpdateLogDelivery
- logs:DeleteLogDelivery
- logs:ListLogDeliveries

AWSServiceRoleForElasticLoadBalancing 信任 elasticloadbalancing.amazonaws.com 服务来代入该角色。

创建服务相关角色

您无需手动创建 AWSServiceRoleForElasticLoadBalancing 角色。Elastic Load Balancing 将在您创建负载均衡器或目标组时为您创建此角色。

要让 Elastic Load Balancing 用户代表您创建服务相关角色，您必须具有所需权限。有关更多信息，请参阅《IAM 用户指南》中的服务相关角色权限。

如果您在 2018 年 1 月 11 日之前创建了负载均衡器，则 Elastic Load Balancing 已在您的 AWS 账户中创建 AWSServiceRoleForElasticLoadBalancing。有关更多信息，请参阅 IAM 用户指南中的我的 AWS 帐户中出现新角色。

编辑服务相关角色

您可以使用 IAM 编辑 AWSServiceRoleForElasticLoadBalancing 的描述。有关更多信息，请参阅 IAM 用户指南 中的编辑服务相关角色。

删除服务相关角色

如果您不再需要使用 Elastic Load Balancing，我们建议您删除 AWSServiceRoleForElasticLoadBalancing。

只有在删除 AWS 账户中的所有负载均衡器后，才能删除此服务相关角色。这可确保您不会无意中删除访问您的负载均衡器的权限。有关更多信息，请参阅删除 Application Load Balancer、删除 Network Load Balancer 和删除 Classic Load Balancer。

您可以使用 IAM 控制台、IAM CLI 或 IAM API 删除服务相关角色。有关更多信息，请参阅 IAM 用户指南 中的删除服务相关角色。

在您删除 AWSServiceRoleForElasticLoadBalancing 之后，Elastic Load Balancing 将在您创建负载均衡器时再次为您创建该角色。

Elastic Load Balancing 的 AWS 托管策略

要向用户、组和角色添加权限，与自己编写策略相比，使用 AWS 托管策略更简单。创建仅为团队提供所需权限的 IAM 客户托管式策略需要时间和专业知识。要快速入门，您可以使用我们的 AWS 托管策略。这些策略涵盖常见使用案例，可在您的 AWS 账户 中使用。有关 AWS 托管策略的更多信息，请参阅 IAM 用户指南中的AWS 托管策略。

AWS 服务负责维护和更新 AWS 托管策略。您无法更改 AWS 托管策略中的权限。服务会向 AWS 托管式策略添加额外权限以支持新功能。此类更新会影响附加策略的所有身份（用户、组和角色）。当启动新功能或新操作可用时，服务最有可能会更新 AWS 托管策略。服务不会从 AWS 托管策略中删除权限，因此策略更新不会破坏您的现有权限。

此外，AWS 还支持跨多种服务的工作职能的托管策略。例如，ReadOnlyAccess AWS 托管策略提供对所有 AWS 服务和资源的只读访问权限。当服务启动新功能时，AWS 会为新操作和资源添加只读权限。有关工作职能策略的列表及其说明，请参阅 IAM 用户指南中的适用于工作职能的 AWS 托管策略。

AWS 托管策略：AWSElasticLoadBalancingClassicServiceRolePolicy

此策略包括 Elastic Load Balancing (Classic Load Balancer) 代表您调用其他 AWS 服务所需的所有权限。服务相关角色已预先定义。使用预定义角色，您不必手动添加 Elastic Load Balancing 代表您完成操作所需的权限。您不能附加、分离、修改或删除此策略。

要查看此策略的权限，请参阅 AWS 管理控制台中的 AWSElasticLoadBalancingClassicServiceRolePolicy。

AWS 托管策略：AWSElasticLoadBalancingServiceRolePolicy

此策略包含 Elastic Load Balancing 代表您调用其他 AWS 服务所需的所有权限。服务相关角色已预先定义。使用预定义角色，您不必手动添加 Elastic Load Balancing 代表您完成操作所需的权限。您不能附加、分离、修改或删除此策略。

要查看此策略的权限，请参阅 AWS 管理控制台中的 AWSElasticLoadBalancingServiceRolePolicy。

AWS 托管策略：ElasticLoadBalancingFullAccess

此策略授予对 Elastic Load Balancing 服务的完全访问权限以及通过 AWS 管理控制台有限访问其他服务的权限。

要查看此策略的权限，请参阅 AWS 管理控制台中的 ElasticLoadBalancingFullAccess。

AWS 托管策略：ElasticLoadBalancingReadOnly

此策略提供对 Elastic Load Balancing 和相关服务的只读访问权限

要查看此策略的权限，请参阅 AWS 管理控制台中的 ElasticLoadBalancingReadOnly。

Elastic Load Balancing 的 AWS 托管策略更新

查看有关 Elastic Load Balancing 的 AWS 托管策略更新的详细信息（自该服务开始跟踪这些更改以来）。

更改	描述	日期
AWS 托管策略：AWSElasticLoadBalancingServiceRolePolicy (p. 20) – 对现有策略的更新	Elastic Load Balancing 添加了一个新的操作来授予使用对等连接的权限。此操作已添加到适用于 Elastic Load Balancing 控制面板的服务相关角色策略中。它与 `ec2:DescribeVpcPeeringConnections` API 操作关联。	2021 年 10 月 11 日
AWS 托管策略：ElasticLoadBalancingFullAccess (p. 20) – 对现有策略的更新	Elastic Load Balancing 添加了一个新的操作来授予使用对等连接的权限。此操作已添加到 Elastic Load Balancing 完全访问策略中。它与 `ec2:DescribeVpcPeeringConnections` API 操作关联。	2021 年 10 月 11 日
AWS 托管策略：AWSElasticLoadBalancingClassicServiceRolePolicy (p. 20) – 对现有策略的更新	Elastic Load Balancing 为 Classic Load Balancer 添加了服务相关角色策略（用于控制面板）。此更新适用于版本 2（默认模式）。	2019 年 10 月 7 日
AWS 托管策略：ElasticLoadBalancingReadOnly (p. 20)	提供对 Elastic Load Balancing 和相关服务的只读访问。这是版本 1（默认模式）。	2018 年 9 月 20 日
Elastic Load Balancing 开始跟踪更改	Elastic Load Balancing 开始跟踪其 AWS 托管策略的更改。	2021 年 7 月 23 日

Elastic Load Balancing 的合规性验证

作为多项 AWS 合规性计划的一部分，第三方审计员将评估 Elastic Load Balancing 的安全性和合规性。其中包括 SOC、PCI、FedRAMP、HIPAA 及其他。

有关特定合规性计划范围内的 AWS 服务列表，请参阅合规性计划范围内的 AWS 服务。有关常规信息，请参阅 AWS 合规性计划。

您可以使用 AWS Artifact 下载第三方审计报告。有关更多信息，请参阅在 AWS Artifact 中下载报告。

您在使用 Elastic Load Balancing 时的合规性责任由您的数据的敏感性、您公司的合规性目标以及适用的法律法规决定。AWS 提供以下资源以帮助满足合规性要求：

- 安全性与合规性 Quick Start 指南 – 这些部署指南讨论了架构注意事项，并提供了在 AWS 上部署关注安全性和合规性的基准环境的步骤。
- 《设计符合 HIPAA 安全性和合规性要求的架构》白皮书 – 此白皮书介绍公司如何使用 AWS 创建符合 HIPAA 标准的应用程序。
- AWS 合规性资源 – 此业务手册和指南集合可能适用于您的行业和位置。
- AWS Config 开发人员指南中的使用规则评估资源 – AWS Config；评估您的资源配置对内部实践、行业指南和法规的遵循情况。
- AWS Security Hub – 此AWS服务提供了AWS中安全状态的全面视图，可帮助您检查是否符合安全行业标准和最佳实践。

Elastic Load Balancing 中的故障恢复能力

AWS 全球基础设施围绕 AWS 区域和可用区构建。区域提供多个在物理上独立且隔离的可用区，这些可用区通过延迟低、吞吐量高且冗余性高的网络连接在一起。利用可用区，您可以设计和操作在可用区之间无中断地自动实现故障转移的应用程序和数据库。与传统的单个或多个数据中心基础设施相比，可用区具有更高的可用性、容错性和可扩展性。

有关 AWS 区域和可用区的更多信息，请参阅 AWS 全球基础设施。

除了 AWS 全球基础设施以外，Elastic Load Balancing 还提供以下功能以支持数据恢复：

- 在一个或多个可用区中的多个实例之间分配传入流量。
- 您可以将 AWS Global Accelerator 与 Application Load Balancer 结合使用，以在一个或多个 AWS 区域的多个负载均衡器之间分配传入流量。有关更多信息，请参阅 AWS Global Accelerator 开发人员指南。
- Amazon ECS 使您能够在 EC2 实例集群上运行、停止和管理 Docker 容器。您可以将 Amazon ECS 服务配置为使用负载均衡器在集群中的服务之间分配传入流量。有关更多信息，请参阅 Amazon Elastic Container Service 开发人员指南。

Elastic Load Balancing 中的基础设施安全性

作为一项托管服务，Elastic Load Balancing 由 Amazon Web Services：安全流程概述白皮书中所述的 AWS 全球网络安全程序提供保护。

您可以使用 AWS 发布的 API 调用，以通过网络访问 Elastic Load Balancing。客户端必须支持传输层安全性 (TLS) 1.0 或更高版本。建议使用 TLS 1.2 或更高版本。客户端还必须支持具有完全向前保密 (PFS) 的密码套件，例如 Ephemeral Diffie-Hellman (DHE) 或 Elliptic Curve Ephemeral Diffie-Hellman (ECDHE)。大多数现代系统（如 Java 7 及更高版本）都支持这些模式。

此外，必须使用访问密钥 ID 和与 IAM 主体关联的秘密访问密钥来对请求进行签名。或者，您可以使用 AWS Security Token Service (AWS STS) 生成临时安全凭证来对请求进行签名。

网络隔离

Virtual Private Cloud (VPC) 是 AWS 云内您自己的逻辑隔离区域中的虚拟网络。子网是 VPC 中的 IP 地址范围。当您创建负载均衡器时，可以为负载均衡器节点指定一个或多个子网。您可以在您的 VPC 的子网中部署 EC2 实例，并将这些实例注册到您的负载均衡器。有关 VPC 和子网的更多信息，请参阅 Amazon VPC 用户指南。

当您在 VPC 中创建负载均衡器时，它可以面向 Internet，也可以面向内部。内部负载均衡器可路由的请求只能来自对负载均衡器的 VPC 具有访问权限的客户端。

您的负载均衡器会使用私有 IP 地址向已注册目标发送请求。因此，您的目标无需使用公有 IP 地址，即可接收来自负载均衡器的请求。

要从 VPC 调用 Elastic Load Balancing API，而不通过公有 Internet 发送流量，请使用 AWS PrivateLink。有关更多信息，请参阅 使用接口端点访问 Elastic Load Balancing（AWS PrivateLink）(p. 22)。

控制网络流量

当您使用负载均衡器时，请考虑使用以下选项来保护网络流量：

- 使用安全侦听器支持客户端和负载均衡器之间的加密通信。Application Load Balancer 支持 HTTPS 侦听器。Network Load Balancer 支持 TLS 侦听器。经典负载均衡器同时支持 HTTPS 和 TLS 侦听器。您可以从您的负载均衡器的预定义安全策略中选择，指定您的应用程序支持的密码套件和协议版本。可以使用 AWS Certificate Manager (ACM) 或者 AWS Identity and Access Management (IAM) 管理安装在您的负载均衡器上的服务器证书。您可以利用服务器名称指示 (SNI) 协议，使用单个安全侦听器为多个安全网站提供服务。当您将多个服务器证书与安全侦听器关联时，会自动为您的负载均衡器启用 SNI。
- 配置 Application Load Balancer 和经典负载均衡器的安全组，以仅接受来自特定客户端的流量。这些安全组必须在侦听器端口上允许来自客户端的入站流量以及流向客户端的出站流量。
- 为您的 Amazon EC2 实例配置安全组，以仅接受来自负载均衡器的流量。这些安全组必须在侦听器端口和运行状况检查端口上允许来自负载均衡器的入站流量。
- 配置您的 Application Load Balancer，以通过身份提供商或使用公司身份安全地对用户进行身份验证。有关更多信息，请参阅使用 Application Load Balancer 对用户进行身份验证。
- 将 AWS WAF 与 Application Load Balancer 结合使用，根据 Web 访问控制列表 (Web ACL) 中的规则允许或阻止请求。

使用接口端点访问 Elastic Load Balancing（AWS PrivateLink）

您可以通过创建接口 VPC 终端节点在 Virtual Private Cloud (VPC) 与 Elastic Load Balancing API 之间建立私有连接。您可以使用此连接从 VPC 调用 Elastic Load Balancing API，而无需将互联网网关、NAT 实例或 VPN 连接附加到您的 VPC。终端节点提供了与用于创建和管理负载均衡器的 2015-12-01 版和 2012-06-01 版 Elastic Load Balancing API 的可靠、可扩展连接。

接口 VPC 端点由 AWS PrivateLink 提供支持，此功能使用私有 IP 地址在应用程序与 AWS 服务 之间进行通信。有关更多信息，请参阅AWS PrivateLink。

限制

AWS PrivateLink 不支持包含超过 50 个侦听器的 Network Load Balancer。

为 Elastic Load Balancing 创建接口终端节点

使用以下服务名称为 Elastic Load Balancing 创建终端节点：

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

有关更多信息，请参阅 AWS PrivateLink 指南中的创建接口端点。

为 Elastic Load Balancing 创建 VPC 终端节点策略

您可以向 VPC 终端节点附加策略，以控制对 Elastic Load Balancing API 的访问。该策略指定：

- 可执行操作的委托人。
- 可执行的操作。

- 可对其执行操作的资源。

以下示例显示了一个 VPC 终端节点策略，该策略拒绝所有人通过终端节点创建负载均衡器的权限。示例策略还授予所有人执行所有其他操作的权限。

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

有关更多信息，请参阅 AWS PrivateLink 指南中的使用端点策略控制对服务的访问权限。

迁移您的经典负载均衡器

我们即将停用 EC2-Classic 网络。我们建议您将经典负载均衡器从 EC2-Classic 网络迁移到 VPC。有关更多信息，请参阅 Amazon EC2 用户指南中的从 EC2-Classic 迁移到 VPC 和博客 EC2-Classic Networking 即将停用 – 以下是准备方法。

Elastic Load Balancing 支持以下类型的负载均衡器：应用程序负载均衡器、网络负载均衡器、网关负载均衡器和经典负载均衡器。有关每种负载均衡器类型的不同功能的信息，请参阅 Elastic Load Balancing 产品对比。

如果您在 EC2-Classic 网络中有经典负载均衡器，则必须将其迁移到 VPC。AWS 将于 2022 年 8 月 15 日停用 EC2-Classic 网络。在此日期之后，经典负载均衡器将仅在 VPC 中提供。为避免中断，您必须在 2022 年 8 月 15 日之前将您在 EC2-Classic 中的任何经典负载均衡器迁移到 VPC。有关更多信息，请参阅 迁移到 VPC 中的经典负载均衡器 (p. 26)。

您也可以选择使用 Application Load Balancer 或 Network Load Balancer 替换 VPC 中的经典负载均衡器。有关更多信息，请参阅 迁移到 Application Load Balancer 或 Network Load Balancer (p. 24)。

迁移过程

- 步骤 1：创建新负载均衡器 (p. 24)
- 步骤 2：逐步将流量重定向到您的新负载均衡器 (p. 26)
- 步骤 3：更新策略、脚本和代码 (p. 26)
- 步骤 4：删除旧负载均衡器 (p. 27)

步骤 1：创建新负载均衡器

创建配置等效于经典负载均衡器的负载均衡器以进行迁移。在迁移过程完成后，您就可以利用新负载均衡器的功能了。

要创建 Application Load Balancer 或 Network Load Balancer 以替换 VPC 中的经典负载均衡器，请使用以下选项之一：

- 选项 1：在控制台中使用迁移向导 (p. 24)
- 选项 2：使用 github 中的负载均衡器复制实用程序 (p. 25)
- 选项 3：手动迁移到 Application Load Balancer 或 Network Load Balancer (p. 25)

要在 VPC 中创建经典负载均衡器以替换 EC2-Classic 中的经典负载均衡器，请使用以下选项：

- 选项 4：手动迁移到 VPC 中的经典负载均衡器 (p. 26)

选项 1：在控制台中使用迁移向导

迁移向导会根据 VPC 中经典负载均衡器的配置创建 Application Load Balancer 或 Network Load Balancer。所创建的负载均衡器的类型取决于经典负载均衡器的配置。

迁移向导发布说明

- 经典负载均衡器必须位于 VPC 中。

- 如果经典负载均衡器具有 HTTP 或 HTTPS 侦听器，则向导将创建 Application Load Balancer。如果经典负载均衡器具有 TCP 侦听器，则向导将创建 Network Load Balancer.。
- 如果经典负载均衡器的名称与现有 Application Load Balancer 或 Network Load Balancer 的名称匹配，则向导将要求您在迁移过程中指定其他名称。
- 如果经典负载均衡器有一个子网，则向导将要求您在创建 Application Load Balancer 时指定第二个子网。
- 如果经典负载均衡器已在 EC2-Classic 中注册实例，则这些实例不会注册到新负载均衡器的目标组。
- 如果经典负载均衡器具有以下类型的已注册实例，则它们不会注册到 Network Load Balancer 的目标组：C1、CC1、CC2、CG1、CG2、CR1、CS1、G1、G2、HI1、HS1、M1、M2、M3 和 T1。
- 如果经典负载均衡器具有 HTTP/HTTPS 侦听器，但使用 TCP 运行状况检查，则向导将更改为 HTTP 运行状况检查。然后向导在创建 Application Load Balancer 时会将路径默认设置为“/”。
- 如果将经典负载均衡器迁移到 Network Load Balancer，运行状况检查设置则将更改，以满足 Network Load Balancer 的要求。
- 如果经典负载均衡器具有多个 HTTPS 侦听器，则向导将选择一个侦听器并使用其证书和策略。如果端口 443 上有一个 HTTPS 侦听器，向导将选择此侦听器。如果所选侦听器使用自定义策略或 Application Load Balancer 不支持的策略，则向导将更改为默认安全策略。
- 如果经典负载均衡器具有安全的 TCP 侦听器，则 Network Load Balancer 将使用 TCP 侦听器。但它不使用证书或安全策略。
- 如果经典负载均衡器具有多个侦听器，则向导将使用端口值最低的侦听器端口作为目标组端口。注册到这些侦听器的每个实例都会在所有侦听器的侦听器端口上注册到目标组。
- 如果经典负载均衡器的一些标签在标签名称中具有 `aws` 前缀，则这些标签不会添加到新的负载均衡器。

使用迁移向导迁移经典负载均衡器

1. 通过以下网址打开 Amazon EC2 控制台：https://console.aws.amazon.com/ec2/。
2. 在导航窗格上的 LOAD BALANCING 下，选择 Load Balancers。
3. 选择您的经典负载均衡器。
4. 在 Migration 选项卡上，选择 Launch ALB Migration Wizard 或 Launch NLB Migration Wizard。显示的按钮取决于在检查经典负载均衡器后由向导选择的负载均衡器类型。
5. 在 Review 页面上，验证向导选择的配置选项。要更改某个选项，请选择 Edit。
6. 当您完成配置新的负载均衡器时，选择 Create。

选项 2：使用 github 中的负载均衡器复制实用程序

此负载均衡器复制实用程序在 GitHub 上可用。有关更多信息，请参阅 Elastic Load Balancing 工具。

选项 3：手动迁移到 Application Load Balancer 或 Network Load Balancer

以下信息提供了基于 VPC 中的现有经典负载均衡器手动创建新的 Application Load Balancer 或 Network Load Balancer 的常规说明。您可以使用 AWS Management Console、AWS CLI 或 AWS 开发工具包进行迁移。有关更多信息，请参阅 Elastic Load Balancing 入门 (p. 8)。

1. 创建具有与经典负载均衡器相同的模式 (面向 Internet 或内部)、子网和安全组的新负载均衡器。
2. 使用与经典负载均衡器相同的运行状况检查设置为负载均衡器创建一个目标组。
3. 请执行下列操作之一：
   - 如果您的经典负载均衡器已附加到 Auto Scaling 组，请将目标组附加到 Auto Scaling 组。这样还可以向目标组注册 Auto Scaling 实例。
   - 向目标组注册您的 EC2 实例。

4. 创建一个或多个侦听器，每个都具有将请求转发到目标组的默认规则。如果创建 HTTPS 侦听器，则可指定您为经典负载均衡器所指定的同一证书。建议您使用默认安全策略。
5. 如果您的经典负载均衡器具有标签，请进行检查并将相关标签添加到新负载均衡器。

选项 4：手动迁移到 VPC 中的经典负载均衡器

以下信息提供了基于 EC2-Classic 中的经典负载均衡器在 VPC 中手动创建新经典负载均衡器的常规说明。您可以使用 AWS Management Console、AWS CLI 或 AWS 开发工具包进行迁移。有关更多信息，请参阅经典负载均衡器用户指南中的教程：创建经典负载均衡器。

1. 将 EC2 资源（例如实例和安全组）从 EC2-Classic 迁移到 VPC。有关更多信息，请参阅 Amazon EC2 用户指南中的将资源迁移到 VPC。
2. 在 VPC 中创建新的经典负载均衡器。
3. 创建负载均衡器时，请选择您从 EC2-Classic 迁移实例时使用的 VPC。从您计划向新负载均衡器注册的实例所在的每个可用区中选择一个子网。
4. 出现提示时，选择要向负载均衡器注册的实例。
5. 如果旧的经典负载均衡器具有标签，请进行检查并将相关标签添加到新的经典负载均衡器。

步骤 2：逐步将流量重定向到您的新负载均衡器

在向新负载均衡器注册您的实例后，您可以开始将流量从旧负载均衡器重定向到新负载均衡器的过程。这使您能够测试新的负载均衡器，同时将应用程序可用性风险降至最低。

逐步将流量重定向到您的新负载均衡器

1. 将新负载均衡器的 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，浏览器会显示您应用程序的默认页面。
2. 创建一个用于将域名与您的新负载均衡器关联的新 DNS 记录。如果您的 DNS 服务支持权重，则在新 DNS 记录中指定权重为 1；对于您的旧负载均衡器的现有 DNS 记录，指定权重为 9。这样可以将 10% 的流量定向到新负载均衡器，而将 90% 的流量定向到旧负载均衡器。
3. 监控您的新负载均衡器，验证它能否接收流量并将请求路由到您的实例。

   Important

   DNS 记录中的生存时间 (TTL) 为 60 秒。这意味着，解析域名的任何 DNS 服务器在其缓存中保留记录信息的时间为 60 秒，同时更改会传播。因此，在您完成上一步后，这些 DNS 服务器仍然可以在 60 秒内将流量路由到旧负载均衡器。在传输过程中，流量可以定向到任一负载均衡器。
4. 继续更新您的 DNS 记录的权重，直到所有流量都定向到您的新负载均衡器。完成后，您可以删除旧负载均衡器的 DNS 记录。

步骤 3：更新策略、脚本和代码

如果要将经典负载均衡器迁移到 Application Load Balancer 或 Network Load Balancer，请务必执行以下操作：

- 将使用 API 版本 2012-06-01 的 IAM 策略更新为使用版本 2015-12-01。
- 将使用 `AWS/ELB` 命名空间中的 CloudWatch 指标的进程更新为使用 `AWS/ApplicationELB` 或 `AWS/NetworkELB` 命名空间中的指标。
- 将使用 aws elb AWS CLI 命令的脚本更新为使用 aws elbv2 AWS CLI 命令。

- 将使用 AWS CloudFormation 资源的 `AWS::ElasticLoadBalancing::LoadBalancer` 模板更新为使用 `AWS::ElasticLoadBalancingV2` 资源。
- 将使用 Elastic Load Balancing API 版本 2012-06-01 的代码更新为使用版本 2015-12-01。

资源

- AWS CLI 命令参考中的 elbv2
- Elastic Load Balancing API 参考（2015 年 12 月 1 日版）
- 适用于 Elastic Load Balancing 的 Identity and Access Management (p. 10)
- Application Load Balancer 用户指南中的 Application Load Balancer 指标
- Network Load Balancer 用户指南中的 Network Load Balancer 指标
- 《AWS CloudFormation 用户指南》中的 AWS::ElasticLoadBalancingV2::LoadBalancer

步骤 4：删除旧负载均衡器

您可以在完成以下步骤后删除旧经典负载均衡器：

- 您已将旧负载均衡器的所有流量重定向到新负载均衡器。
- 已完成路由到旧负载均衡器的所有现有请求。