天翼云 3.0• 弹性负载均衡

（性能保障型）

用户使用指南

中国电信股份有限公司云计算分公司

修订记录

修改时间	修改说明
1　2019-12-19	● 发布 v1.0 版本

目 录

1 产品概述 5

1.1 概念 5

1.1.1 弹性负载均衡 5

1.1.2 监听器 6

1.1.3 健康检查 7

1.2 弹性负载均衡是如何工作的 7

1.3 公网和私网负载均衡器 8

1.4 功能介绍 9

1.5 应用场景 10

1.6 关联业务 13

2 快速入门 13

2.1 典型场景说明 13

2.2 创建负载均衡器 14

2.2.1 创建增强型负载均衡器 14

2.2.2 添加监听器 1

2.2.3 添加后端云主机 4

3 负载均衡管理 6

3.1 负载均衡器管理 6

3.1.1 查询负载均衡器 6

3.1.2 删除负载均衡器 7

3.1.3 停用负载均衡器 8

3.2 监听器管理 9

3.2.1 添加监听器 9

3.2.2 修改监听器 12

3.2.3 删除监听器 13

3.3 后端云主机管理 13

3.3.1 添加后端云主机 13

3.3.2 移除后端云主机 14

3.4 白名单 14

3.4.1 添加白名单 14

3.5 证书 15

3.5.1 创建证书 16

3.5.2 删除证书 17

3.5.3 绑定证书 17

3.6 URL 转发策略 18

3.6.1 添加转发策略 18

3.6.2 删除转发策略 20

4 监控 29

4.1 支持的监控指标 29

4.2 告警规则 31

4.2.1 添加告警规则 31

4.2.2 修改告警规则 31

5 常见问题 37

5.1 弹性负载均衡（增强型）是什么？ 37

5.2 弹性负载均衡服务是否收费？ 37

5.3 弹性负载均衡支持哪些转发方式？ 37

5.4 弹性负载均衡是否可以添加不同操作系统的云主机？ 38

5.5 单个用户支持保有多少个弹性负载均衡？ 39

5.6 什么是配额？ 39

5.7 弹性负载均衡如何支持多证书？ 39

5.8 如何配置私网或公网负载均衡？ 39

5.9 监听器是什么？ 39

5.10 监听器中分配算法和会话保持算法是什么关系？ 39

5.11 什么是负载均衡协议（端口）？ 40

5.12 什么是云主机协议（端口）？ 40

5.13 弹性负载均衡分配的弹性 IP 是否为独占？ 41

5.14 删除弹性负载均衡有什么影响？ 41

5.15 健康检查异常如何排查？ 41

5.16 为什么很多访问 ELB 实例后端云主机的 IP 是 100.125 开头？ 42

5.17 如何获得来访者的真实 IP？ 42

5.18 ELB 支持什么类型的会话保持？ 42

5.19 使用 UDP 协议有什么注意事项？ 42

5.20 什么是 UDP 健康检查？ 43

5.21 如何检查弹性负载均衡会话保持不生效问题？ 44

5.22 如何检查弹性负载均衡业务访问延时大？ 45

5.23 如何检查弹性负载均衡服务不通或异常中断？ 45

5.24 如何检查弹性负载均衡前后端流量不一致？ 45

5.25 如何检查请求不均衡？ 46

5.26 如何检查弹性负载均衡健康检查异常？ 46

5.27 如何检测压测性能上不去？ 47

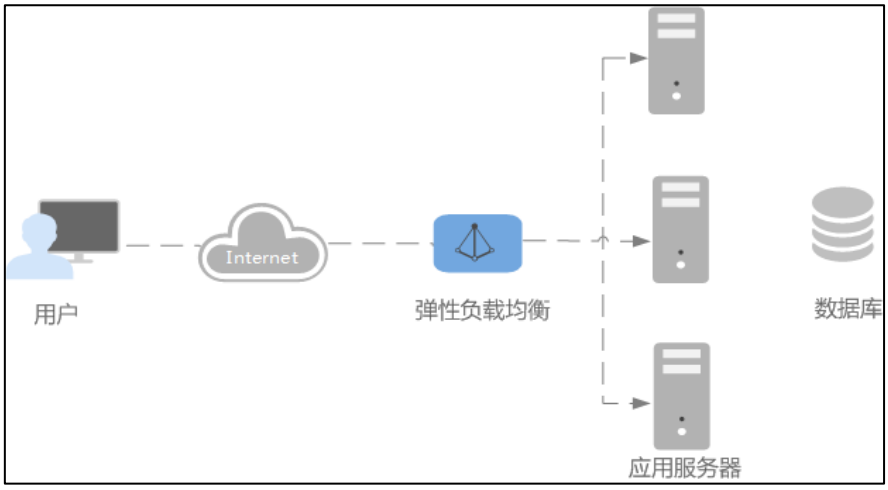
1 产品概述

1.1 概念

1.1.1 弹性负载均衡

弹性负载均衡（Elastic Load Balancing ，简称 ELB）是将访问流量根据转发策略分发到后端多台弹性云主机的流量分发控制服务。弹性负载均衡可以通过流量分发扩展应用系统对外的服务能力，实现更高水平的应用程序容错性能。

用户通过基于浏览器、统一化视图的云计算管理图形化界面，可以创建 ELB，为服务配置需要监听的端口，配置云主机，消除单点故障，提高整个系统的可用性。



- 弹性负载均衡的类型

弹性负载均衡增强型支持两种形式：

- 性能共享型：低成本，提供较强大的负载均衡能力，适用于常规用户。
- 性能保障型：为用户提供独立的 ELB 底层实例资源，实例性能不受其它实例影响。按需实例会根据业务量进行即时调整，包周期实例用户可自选实例规格（不同规格的性能指标不同）。适用于对性能有较高要求的用户。

- 弹性负载均衡的组件

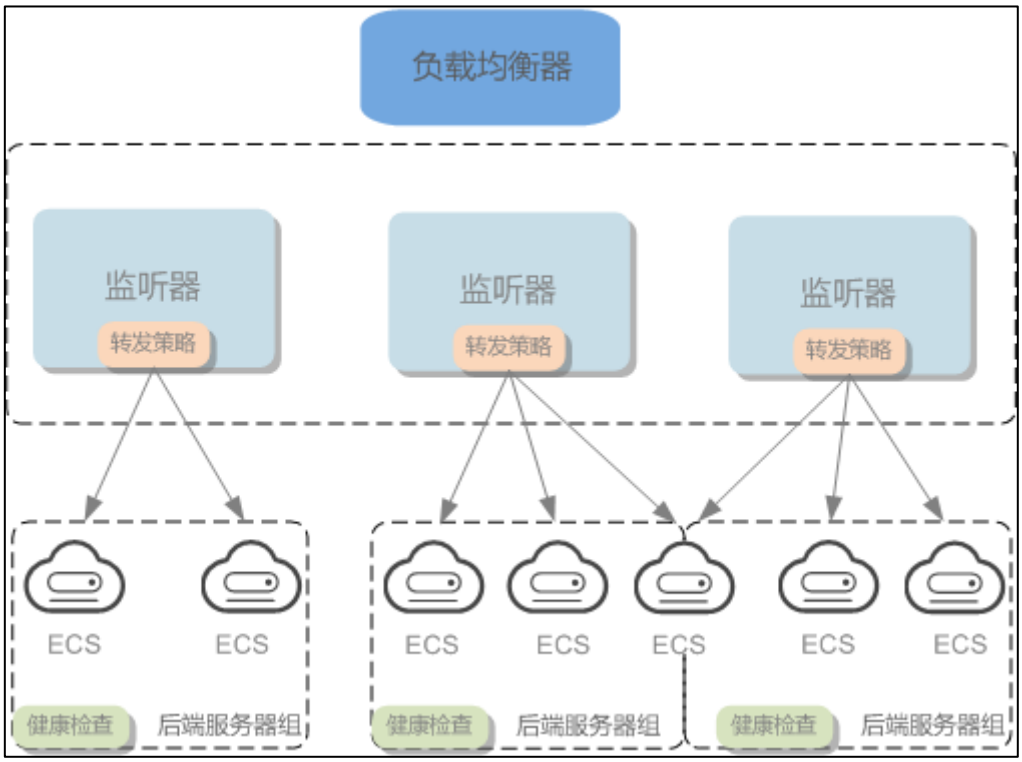
弹性负载均衡器接受来自客户端的传入流量并将请求转发到一个或多个可用区中的后端云主机。

您可以向您的弹性负载均衡器添加一个或多个监听器。监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，并根据您定义的转发策略将请求转发到一个后端云主机组里的后端云主机。

每个后端云主机组使用您指定的协议和端口号将请求转发到一个或多个后端云主机。

您可以开启健康检查功能，对每个后端云主机组配置运行状况检查。当后端某台云主机健康检查出现异常时，弹性负载均衡会自动将新的请求分发到其它健康检查正常的后端云主机上；而当该后端云主机恢复正常运行时，弹性负载均衡会将其自动恢复到弹性负载均衡服务中。



1.1.2 监听器

用户定制的监听器，定义了负载均衡策略和转发规则。负载均衡策略和转发规则的相关概念如下：

监听器使用前端（客户端到负载均衡器）连接的协议以及端口和后端（负载均衡器到后端弹性云主机）连接的协议以及端口配置负载均衡策略。负载均衡器支持协议 HTTP、TCP。负载均衡器可以监听端口 1-65535。

ELB（性能保障型）支持三种转发规则，用户可以根据自身需求选择相应的算法来分配用户访问流量，提升负载均衡能力。增强型负载均衡的分配策略，支持以下三种调度算法：

- 加权轮询算法：按顺序依次将请求分发给不同的云主机。它用相应的权重表示云主机的处理性能，按照权重的高低以及轮询方式将请求分配给各云主机，相同权重的云主机处理相同数目的连接数。

- 加权最少连接：通过当前活跃的连接数来估计云主机负载情况的一种动态调度算法。
- 源 IP 算法：将请求的源 IP 地址作为散列键（HashKey），从静态分配的散列表找出对应的云主机；

ELB（性能保障型）支持如下三种会话保持方式：

- 源 IP 地址：将请求的源 IP 地址作为散列键（HashKey），从静态分配的散列表找出对应的云主机；
- HTTP cookie：负载均衡器会根据客户端第一个请求生成一个 cookie，后续所有包含这个 cookie 值的请求都会由同一个后端云主机处理；
- 应用程序 cookie：该选项依赖于后端应用，后端应用生成一个 cookie 值，后续所有包含这个 cookie 值的请求都会由同一个后端云主机处理；

1.1.3 健康检查

用户可以配置运行状况检查，这些检查可用来监控后端云主机的运行状况，以便负载均衡器只将请求发送到正常运行的后端云主机。而当该故障云主机恢复正常运行时，负载均衡会将其自动恢复到对外或对内的服务中。健康检查支持协议 TCP 和 HTTP。

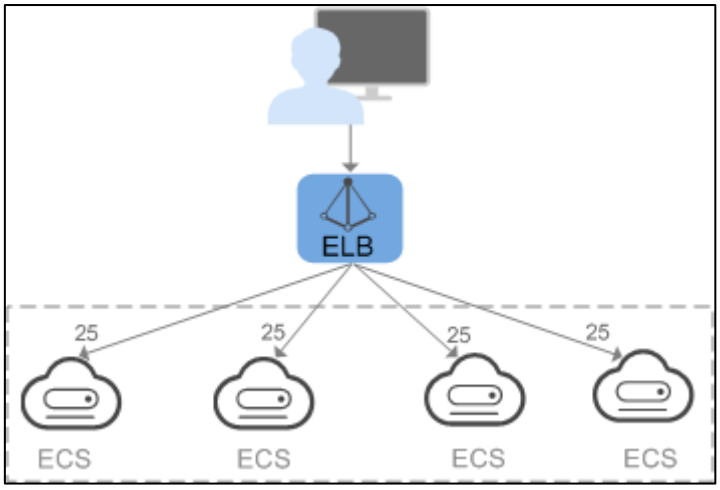
1.2 弹性负载均衡是如何工作的

您可以在弹性负载均衡服务中创建一个负载均衡器。该负载均衡器会接收来自客户端的请求，并将请求转发到一个或多个可用区的后端云主机中进行处理。请求的流量分发与负载均衡器配置的分配策略类型相关。

ELB（性能保障型）负载均衡的分配策略，支持以下三种调度算法：

- 加权轮询算法：按顺序依次将请求分发给不同的云主机。它用相应的权重表示云主机的处理性能，按照权重的高低以及轮询方式将请求分配给各云主机，相同权重的云主机处理相同数目的连接数。
- 加权最少连接：通过当前活跃的连接数来估计云主机负载情况的一种动态调度算法。
- 源 IP 算法：将请求的源 IP 地址作为散列键（HashKey），从静态分配的散列表找出对应的云主机。

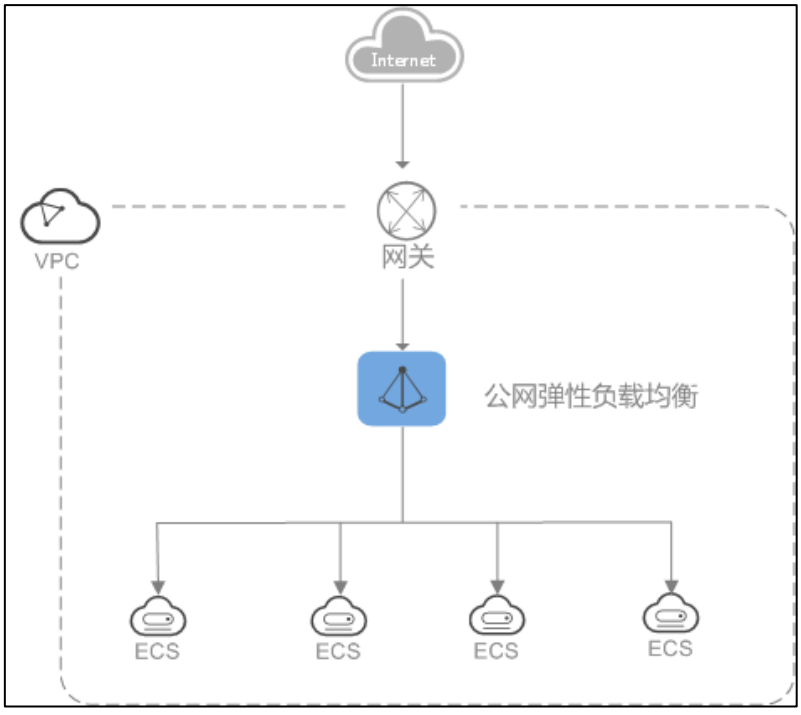
下图展示弹性负载均衡器使用加权轮询算法的流量分发流程。假设可用内有 4 台权重相同的后端云主机，负载均衡器节点会将 25%的客户端流量分发到其可用区中的每一台后端云云主机。



1.3 公网和私网负载均衡器

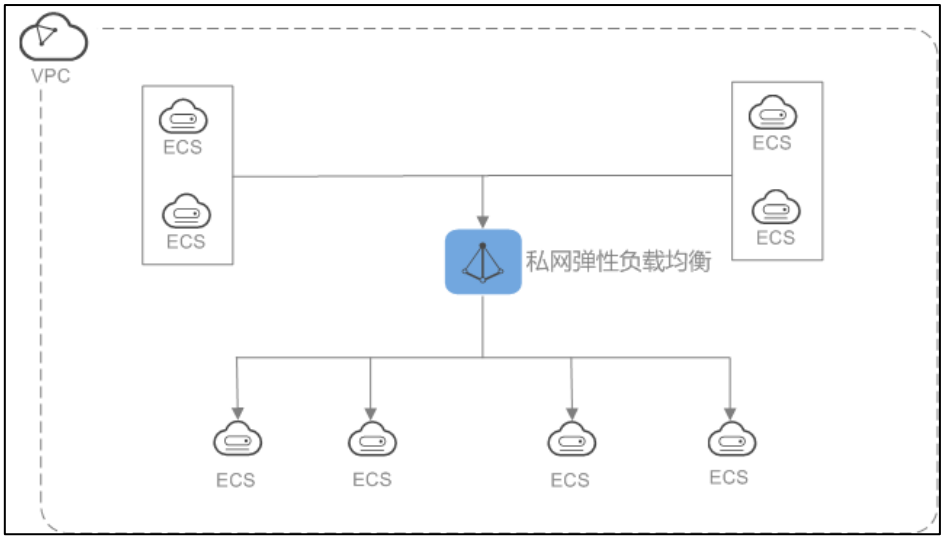
- 公网负载均衡器

公网负载均衡器通过公网 IP 对外提供服务，将来自公网的客户端请求按照指定的负载均衡策略分发到后端云主机进行处理。



- 私网负载均衡器

私网负载均衡器通过私网 IP 对外提供服务，将来自同一个 VPC 的客户端请求按照指定的负载均衡策略分发到后端云主机进行处理。



1.4 功能介绍

弹性负载均衡有两种不同的负载均衡，分别是增强型负载均衡和经典型负载均衡，便于用户根据不同的应用场景和功能需求选择合适的负载均衡器类型。

经典型负载均衡：适用于访问量较小，应用模型简单的 web 业务。

增强型负载均衡：适用于访问量较大的 web 业务，不仅提供基于域名和 URL 的路由均衡能力，还提供多可用区和 IPv6 功能，实现更加灵活的业务需求。增强型负载均衡根据性能又进一步分为性能保障型和性能共享型。

- 性能保障型：性能保障型负载均衡实例资源独享，实例的性能不受其它实例的影响，您可根据业务需要选择不同规格的实例。

- 性能共享型：性能共享型负载均衡实例资源共享，实例的性能会受其它实例的影响。

增强型负载均衡对比经典型负载均衡，不仅支持多可用区和 IPv6 功能，同时提供了更丰富的 HTTP 和 HTTPS 转发能力，在转发性能和稳定性上也有较大提升。具体的功能差异如表 1-1 所示。（“√”表示支持，“—”表示不支持。）。

功能	经典型负载均衡	增强型负载均衡（性能共享型）	增强型负载均衡（性能保障型）
支持公网和私网负载均衡	√	√	√
支持四层（TCP/UDP）和七层负载均衡	√（私网类型不支持 UDP 协议）	√	√

支持轮询 /最少连接/源 IP	√	√	√
支持会话保持	√	√	√
支持 WebSocket 协议	√	√	√
支持按域名和 URL 转发	—	√	√
支持 HTTP/2	—	√	√
支持后端服务器为 ECS	√	√	√
支持访问控制（白名单）	—	√	√
支持标准 OpenStack API	—	√	√
支持后端服务器为裸机	—	√	√
支持 SNI 多证书特性	—	√	√
支持 SSL 协议/加密算法可选	√	—	—
支持访问日志	—	√	√
支持权重	—	√	√
支持修改证书内容	—	√	√
支持双向认证	—	√	√
支持 HTTP 重定向	—	√	√
支持获取弹性公网 IP	—	√	√
支持 IPv6	—	—	√
支持多可用区	—	—	√
支持选择规格	—	—	√

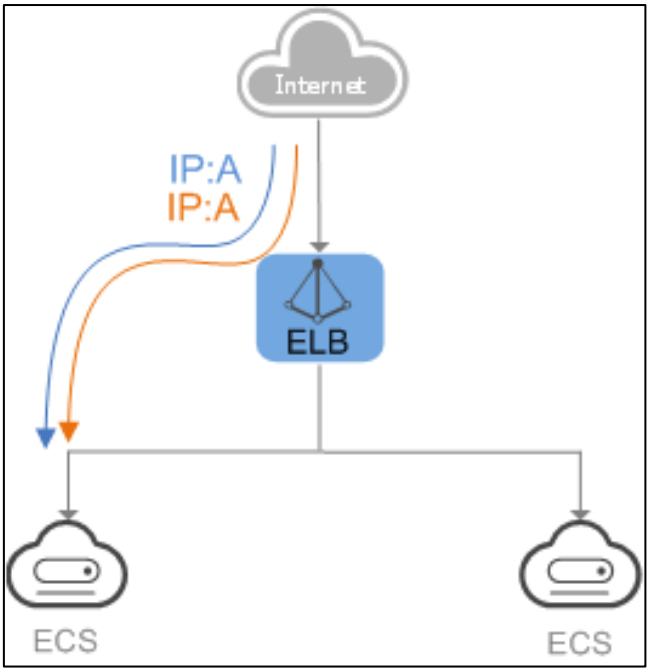
ELB（增强型）服务为用户提供了自助控制负载均衡的能力，并配套提供一个高度管控、灵活使用的管理平台，达到配置简单、服务资源快速添加的目标。

1.5 应用场景

- 使用 ELB 为高访问量业务进行流量分发

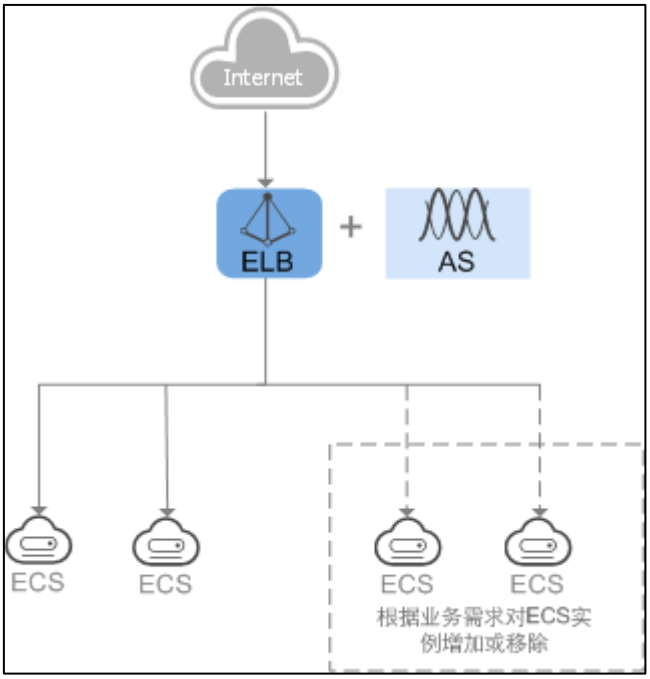
对于业务量访问较大的业务，可以通过 ELB 设置相应的转发策略，将访问量均匀的分到多个后端云主机处理。例如大型门户网站，移动应用市场等。

同时您还可以开启会话保持功能，保证同一个客户请求转发到同一个后端云主机，从而提升访问效率，如下图所示。



- 使用 ELB 和 AS 为潮汐业务弹性分发流量

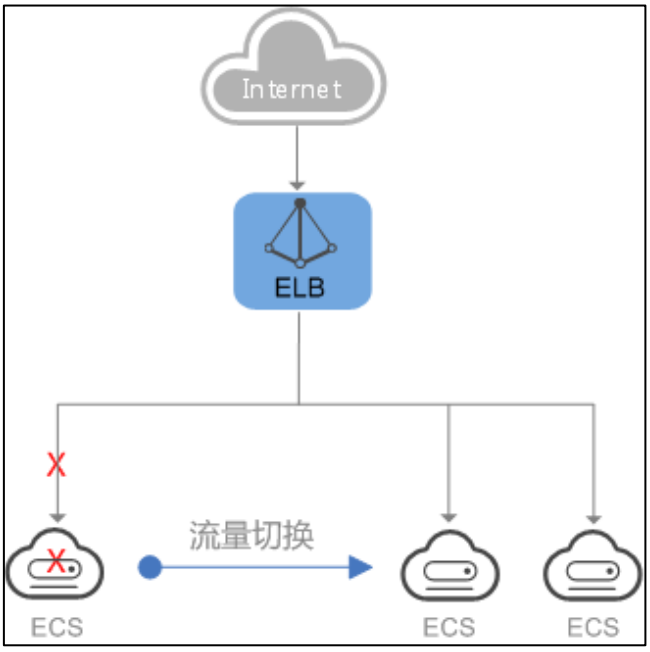
对于存在潮汐效应的业务，结合弹性伸缩服务，可以随时在 ELB 上添加和移除后端云主机，更好的提升业务的灵活扩展能力，如下图所示。例如电商，手游，直播网站等。



- 使用 ELB 消除单点故障

对于可靠性有较高要求的业务，可以在负载均衡器上添加多个后端云主机。负载均衡器会通过健康检查及时发现并屏蔽有故障的云主机，并将流量转发到其他正常运行的后端云主机，确保业务不中断，如下图所示。

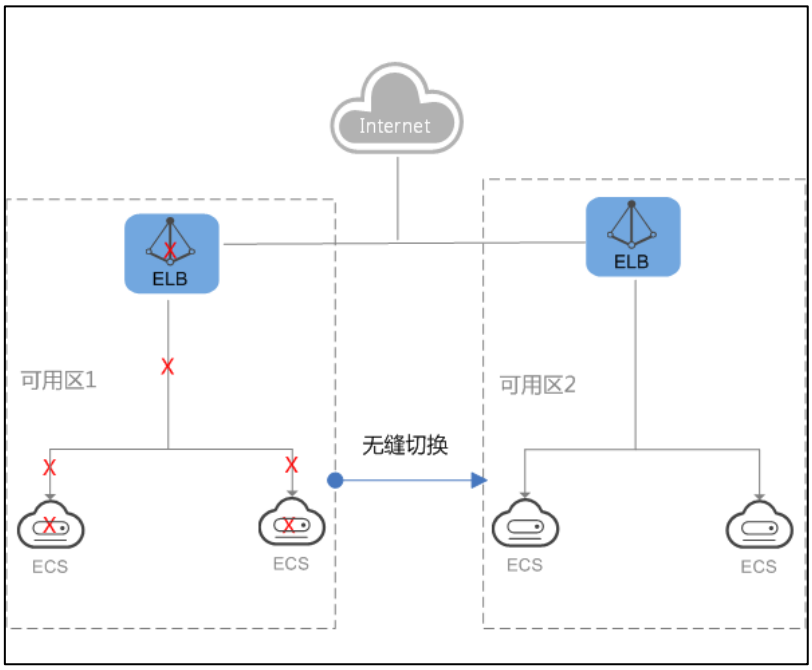
例如官网，计费业务，Web 业务等。



- 使用 ELB 跨可用区特性实现业务容灾部署

对于可靠性和容灾有很高要求的业务，弹性负载均衡可将流量跨可用区进行分发，建立实时的业务容灾部署。即使出现某个可用区网络故障，负载均衡器仍可将流量转发到其他可用区的后端云主机进行处理，如下图所示。

例如银行业务，警务业务，大型应用系统等。



1.6 关联业务

- 虚拟私有云

创建 ELB 时需要使用虚拟私有云服务创建的弹性 IP、带宽。

- 弹性伸缩

当配置了负载均衡服务后，弹性伸缩在添加和移除云云主机时，自动在负载均衡服务中添加和移除云云主机。

- 云监控

当用户开通了弹性负载均衡服务后，无需额外安装其他插件，即可在云监控查看对应服务的实例状态。

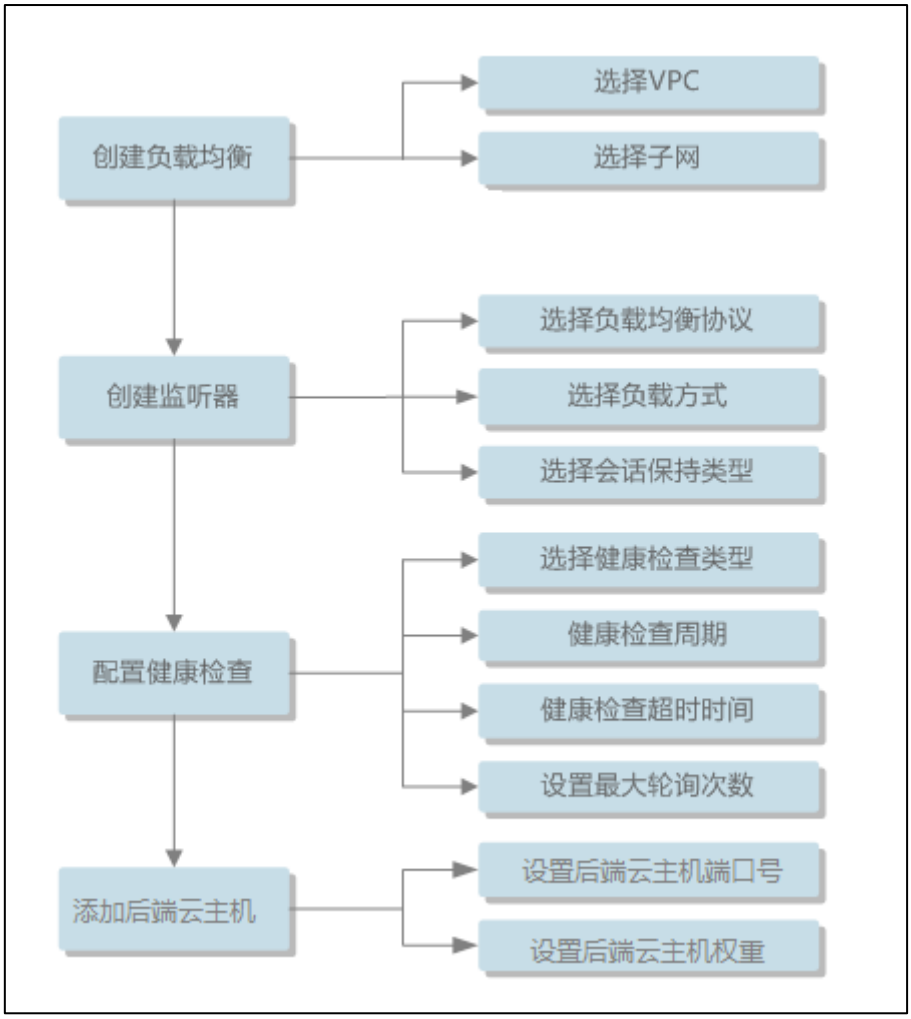
2 快速入门

2.1 典型场景说明

使用负载均衡对后端多台云主机进行流量分发时，需要创建负载均衡，在负载均衡下添加监听器并设置健康检查，最后将后端云主机添加至监听器。

当对用户提供增强型负载均衡服务时，将来自同一个 VPC 下的访问流量自动分发到多台云主机。

配置流程如图所示：



2.2 创建负载均衡器

【注】经典型负载均衡不再支持新创建，建议用户使用增强型负载均衡，可满足经典型 ELB 的全部功能。

本节将说明通过控制中心创建增强型负载均衡器的方法。在这一章节，您将创建一个增强型负载均衡器。

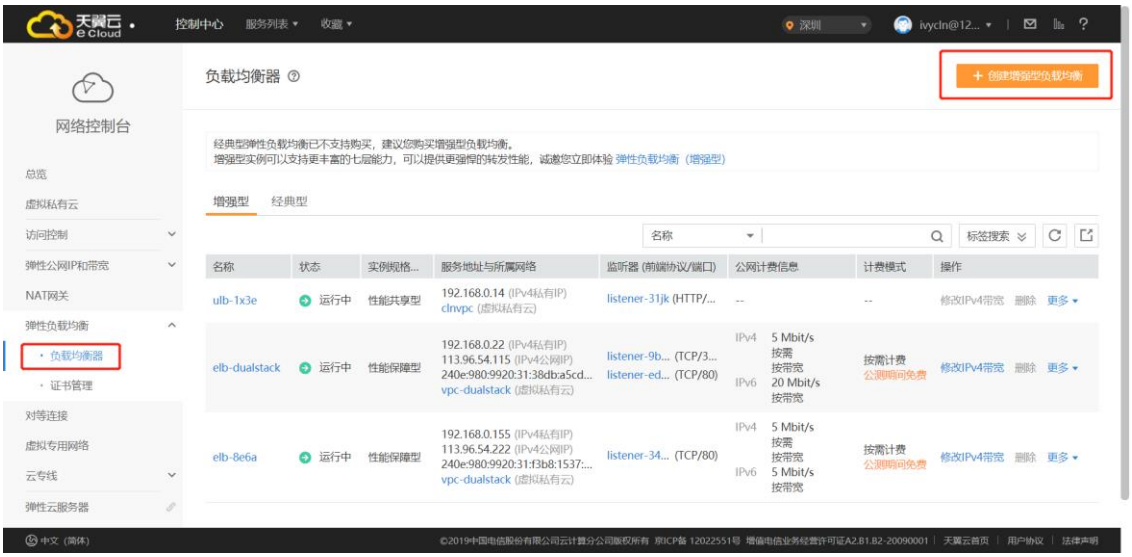
在创建前，请启动您计划添加到负载均衡器的后端云主机，并确保这些云主机的安全组允许端口 22 上的 TCP 访问。

操作步骤见下文

2.2.1 创建增强型负载均衡器

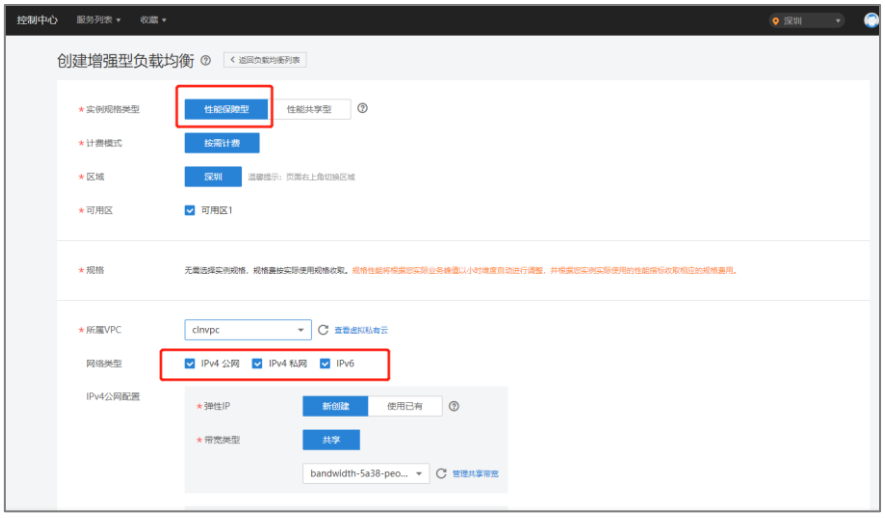
- 登录天翼云控制中心；
- 在【所有服务】标签下，选择【网络】【弹性负载均衡】;

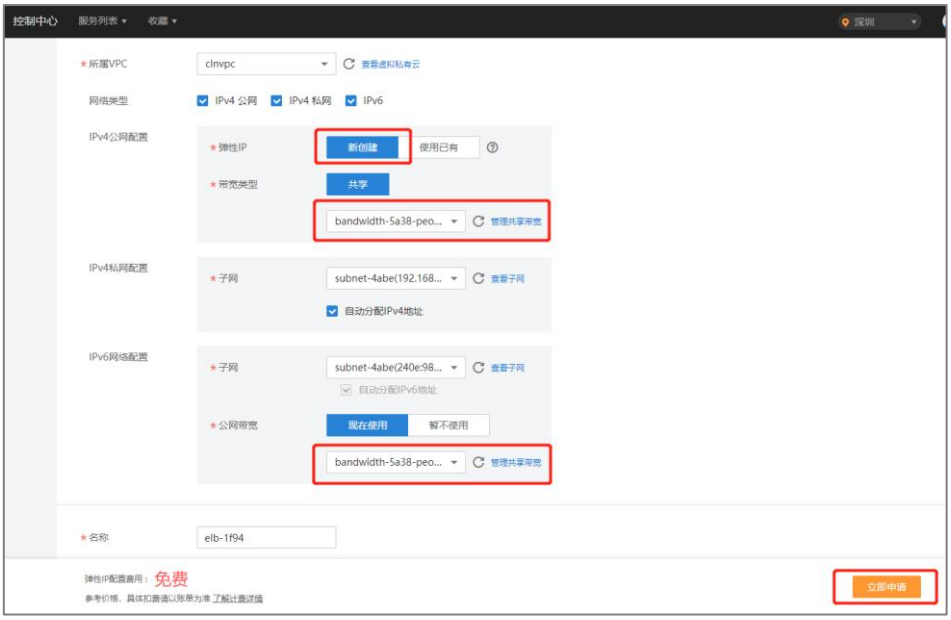
- 在左侧导航栏单击【弹性负载均衡】;



- 在【负载均衡器】界面单击【创建增强型负载均衡】;

- 在【创建增强型负载均衡】界面，选择【性能保障型】实例，根据界面提示配置参数;





注意：如果要分配 IPv6 地址，必须确保 ELB 实例分配在一个已有 IPv6 子网的 VPC 中，并在 IPv6 地址配置处选择 IPv6 子网。

实例配置指南：

参数	说明	取值样例
实例规格类型	选择性能保障型。 - 性能保障型：性能保障型负载均衡 实例资源独享，实例的性能不受其 它实例的影响，您可根据业务需要 选择不同规格的实例。 - 性能共享型：性能共享型负载均衡 实例资源共享，实例的性能会受其 它实例的影响。	性能保障型
计费模式	性能保障型负载均衡器的收费类型。 按需计费（公测期间）	包年/包月
区域	不同区域的资源之间内网不互通。请选择靠近您客户的区域，可以降低网络时延、提高访问速度。	-

可用区	可以选择在多个可用区创建负载均衡实例，提高服务的可用性。如果业务需要考虑容灾能力，建议选择多个可用区。当一个可用区出现故障或不可用时，业务可以快速切换到另一个可用区的负载均衡继续提供服务。 说明： 针对已有实例，如果进行可用区配置修改，可能会导致该实例的业务闪断数秒， 请在创建时做好规划，确实要修改的话建 议选择闲时操作。	-
规格	只有当计费模式为“包年/包月”时，才支持选择规格。当计费模式为“按需计费”时，无需选择实例规格，规格费按实际使用规格收取。规格性能将根据您实际业务峰值以小时维度自动进行调整，并根据您实例实际使用的性能指标收取相应的规格费用。 “支持七层能力”和“支持四层能力”请至少勾选一种，勾选后可选择相应能力的规格。 不同的实例规格在性能上存在差异。可以从最大连接数，新建连接数（CPS）和每秒查询数（QPS）三个维度对实际业务进行评估，然后选择适合的规格。	中型II
所属 VPC	所属虚拟私有云。您可以选择使用已有的虚拟私有云网络，或者单击“查看虚拟私有云”创 建新的虚拟私有云。	vpc-4536

网络配置指南：

参数		说明	取值样例
IPv4 公网配置	弹性公网 IP	当网络类型勾选“IPv4 公网”时，需要指定弹性公网 IP。弹性公网 IP 可以使用已有的 IP 地址，也可以新创建。弹性公网 IP 选择使用已有时，需要选择已有的弹性公网 IP 地址。 新创建：系统为弹性负载均衡实例新创建一个弹性公网 IP。 使用已有：为弹性负载均衡实例选择一个已有的弹性公网 IP 地址。	-
	带宽类型	当选择新创建弹性公网 IP 时，需要指定带宽类型，支持以下两种带宽类型: 独享：一个带宽只能被一个弹性公网 IP 使用。 共享：一个带宽中可以加入多个弹性公网 IP，多个弹性公网 IP 共用一个带宽。	独享
	计费方式	当选择新创建弹性公网 IP 且带宽类型 为独享时，需要指定计费方式。可选 择按带宽计费或者按流量计费。 按带宽计费：按照购买的带宽大小计费。 按流量计费：按照实际使用的流量来计费。	按带宽计费
	带宽	当计费方式选择按带宽计费时，需要指定带宽大小。	100 Mbit/s
IPv4 私网配置	子网	当网络类型勾选“IPv4 私网”时，需要选择创建负载均衡实例的子网。默认勾选“自动分配 IPv4 地址”，系统会自动为负载均衡实例分配一个 IPv4 私网地址。	subnet-4536
	私有 IP 地址	当去勾选“自动分配 IPv4 地址”时，需要为负载均衡实例手动填写一个未 使用的 IPv4 私有 IP 地址。	-
IPv6 网络配置	子网	当网络类型勾选“IPv6”时，需要选择创建负载均衡实例且已开启 IPv6 的子网。当前 IPv6 网络只支持自动获取 IP 地址。	subnet-4537

	公网带宽	IPv6 网络支持使用公网共享带宽，可 以选择是否使用公网带宽。 现在使用：选择已创建的共享带宽。 暂不使用：暂不使用公网带宽，IPv6 只有私网能力，不具备公网能力。	现在使用
名称		负载均衡器的名称。	elb93wd
企业项目		企业项目是一种云资源管理方式，企 业项目管理服务提供统一的云资源按 项目管理，以及项目内的资源管理、 成员管理。	default
描述		可添加负载均衡器相关描述。	-
标签		标签用于标识云资源，可对云资源进 行分类和搜索。标签由标签“键”和 标签“值”组成，标签键用于标记标签，标签值用于表示具体的标签内容。	键： elb_key1 值： elb-01
购买时长		当计费模式为“包年/包月”时，需要指定购买实例的时长。	2 个月

- 单击【立即申请】按钮；
- 确认配置无误后，单击【提交】，任务下发成功后，关闭创建界面；

2.2.2 添加监听器

- 登录天翼云控制中心；
- 选择【弹性负载均衡】【负载均衡器】；
- 单击已创建的负载均衡器实例名称；
- 在该负载均衡界面的【监听器】区域，单击【添加监听器】按钮；

- 在【添加监听器】界面，根据提示配置参数；

各参数说明如下：

参数	说明	取值样例
名称	监听器名称。	listener01
协议/前端端口	负载分发的协议和端口。 支持以下协议，端口取值范围[1-65535]。 • HTTP • TCP • HTTPS（Termination） • UDP	HTTP/80
分配策略类型	负载均衡采用的算法，用户可以根据自身需求选择相应的算法来分配用户访问流量： • 加权轮询算法：按顺序依次将请求分发给不同的云主机。它用相应的权重表示云主机的处理性能，按照权重的高低以及轮询方式将请求分配给各云主机，相同权重的云主机处理相同数目的连接数； • 加权最少连接：通过当前活跃的连接数来估计云主机负载情况的一种动态调度算法； • 源 IP 算法：将请求的源 IP 地址作为散列键（HashKey），从静态分配的散列表找出对应的云主机。	加权轮询算法
会话保持类型	会话保持的方式，用户可以根据自身需求选择相应的会话保	HTTP_COOKIE

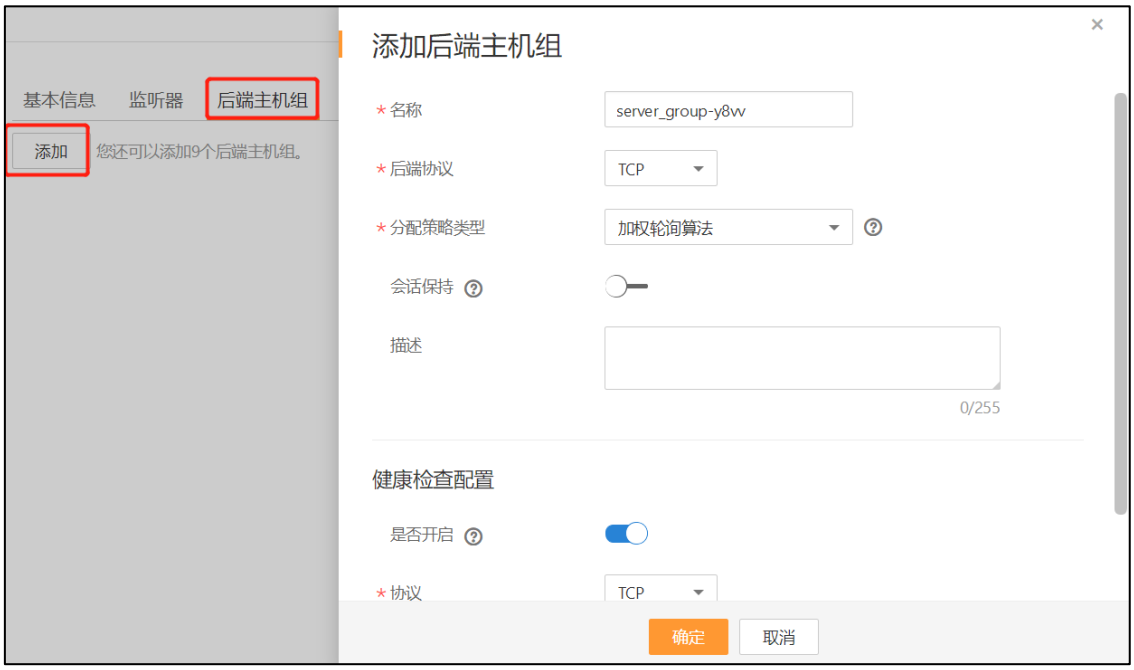
参数	说明	取值样例
	持方式来分配用户访问流量： • 源 IP：将请求的源 IP 地址作为散列键（HashKey），从静态分配的散列表找出对应的云主机； • HTTP COOKIE：负载均衡器会根据客户端第一个请求生成一个 cookie，后续所有包含这个 cookie 值的请求都会由同一个后端云主机处理； • 应用程序 COOKIE：该选项依赖于后端应用。后端应用生成一个 cookie 值，后续所有包含这个 cookie 值的请求都会由同一个后端云主机处理；	
描述	对于监听器描述；	-
健康检查协议	当负载分发协议选择“HTTP”时，健康检查支持的三种类型，设置后不可修改： • TCP • HTTP 当负载分发协议选择“TCP”时，健康检查支持的三种类型，设置后不可修改： • HTTPS • TCP • HTTP 当负载分发协议选择“TCP”时，健康检查支持的三种类型，设置后不可修改： • TCP • PING	HTTP
检查周期(秒)	每次健康检查响应的最大间隔时间；	5
超时时间（秒）	每次健康检查响应的最大超时时间；	10
检查路径	当健康检查方式为 HTTP 时需要配置的选项。为需要被请求的 URL 地址；	/index.html
最大重试次数	健康检查最大的重试次数，范围[1-10]；	3
HTTP 方法	当健康检查方式为 HTTP 或 HTTPS 时需要配置的选项，为 HTTP 或 HTTPS 请求的方法；	GET
HTTP 状态码	当健康检查方式为 HTTP 或 HTTPS 时需要配置的选项，为 HTTP 或 HTTPS 请求后表示请求返回的状态码；	201

- 单击【确定】按钮完成监听器创建；

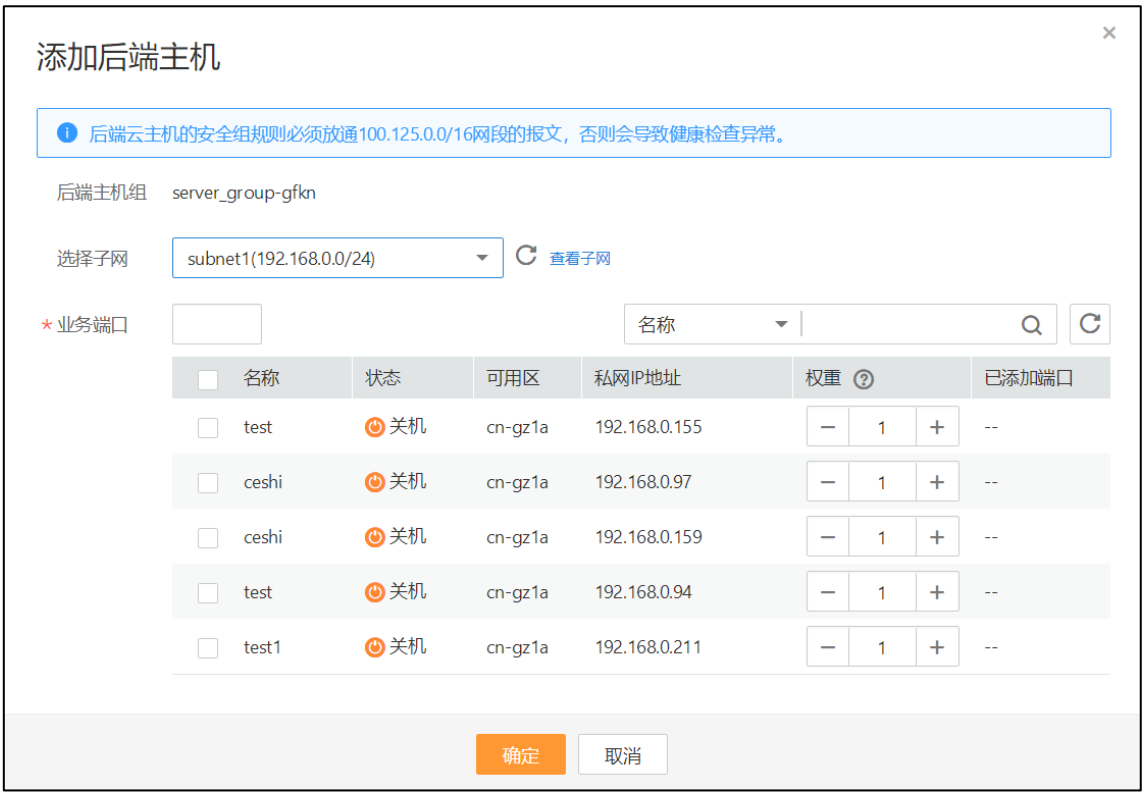
2.2.3 添加后端云主机

您必须将在运行中的云主机添加至您的的负载均衡器中，才能实现负载均衡器对云主机流量分发的功能。

- 登录天翼云控制中心；
- 选择【弹性负载均衡】【负载均衡器】;
- 单击已创建的负载均衡器实例名称，选择【后端主机组】标签，点击添加，来添加后端主机组；



- 创建云主机组后，点击【添加】添加需要和负载均衡关联的云主机，并配置参数；



各参数说明如下：

参数	说明	取值样例
业务端口	后端云主机的服务监听端口，取值范围[1-65535]；	123
权重	后端虚拟机权重。权重值决定了后端云主机处理的请求的比例。例如，一个权重为 2 的云主机处理的请求数是权重为 1 的两倍。默认情况下，权重为 1；	10

- 单击【确定】按钮完成后端云主机添加；

3 负载均衡管理

3.1 负载均衡器管理

本章节提供查询和删除负载均衡器的操作步骤。当您需要查看某负载均衡器详情或不再使用该负载均衡器时可参考本章节。

3.1.1 查询负载均衡器

在控制中心的【弹性负载均衡（增强型）】界面的信息列表，可以查看已创建负载均衡器的状态、子网等详细信息。

- 登录天翼云控制中心；
- 选择【网络】【弹性负载均衡】【负载均衡器】;
- 在负载均衡器信息列表右上角的下拉框中，可设置通过名称、子网等参数搜索负载均衡器；

- 单击负载均衡器名称，进入负载均衡器详情页面，查看负载均衡器的详细信息；

3.1.2 删除负载均衡器

当您不需要再使用某个负载均衡器时，可删除该负载均衡器。

- 登录天翼云控制中心；
- 选择【网络】【弹性负载均衡】【负载均衡器】;
- 在【负载均衡器】界面，单击负载均衡器所在行的【删除】按钮；
- 在确认对话框单击【确定】;

【注 1】如果该负载均衡器下有监听器，不能删除，需先删除监听器后才可删除负载均衡器。

【注 2】当开启短信验证后，删除负载均衡器需要进行短信身份验证。操作如下：

1、进行负载均衡删除操作。

2、弹出删除提醒框，并带有验证提示栏如下图，

3、点击去验证，跳转操作保护页面。如下图，

4、点击【免费获取验证码】，验证码会发送到天翼云账号预留手机号，填写正确的验证码，点击【认证】，弹出认证成功（失败）提示框。

5、认证成功自动跳转回负载均衡删除弹窗界面，继续进行删除操作。

3.1.3 停用负载均衡器

【注】当开启短信验证后，停用负载均衡器需要进行短信身份验证。操作如下：

1、进行负载均衡停用操作。

2、弹出停用提醒框，并带有验证提示栏如下图，

3、点击去验证，跳转操作保护页面。如下图，

4、点击【免费获取验证码】，验证码会发送到天翼云账号预留手机号，填写正确的验证码，点击【认证】，弹出认证成功（失败）提示框。

5、认证成功自动跳转回负载均衡停用弹窗界面，继续进行停用操作。

3.2 监听器管理

本章节提供添加、修改监听器和删除监听器的操作步骤。当您需要向负载均衡器添加监听器，修改监听器的负载均衡模式、会话保持类型、健康检查配置或不再使用该监听器时可参考本章节。

3.2.1 添加监听器

- 登录天翼云控制中心；
- 选择【弹性负载均衡】【负载均衡器】；
- 单击已创建的负载均衡器实例名称；
- 在该负载均衡界面的【监听器】区域，单击【添加监听器】按钮；

- 在【添加监听器】界面，根据提示配置参数；

各参数说明如下：

参数	说明	取值样例
名称	监听器名称。	listener01
协议/前端端口	负载分发的协议和端口。 支持以下协议，端口取值范围	HTTP/80

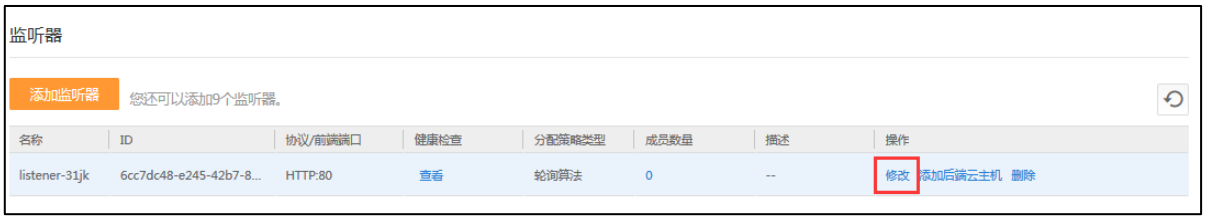
参数	说明	取值样例
	[1-65535]。 • HTTP • TCP • HTTPS(Termination) • UDP	
负载方式	负载均衡采用的算法，用户可以根据自身需求选择相应的算法来分配用户访问流量： 加权轮询算法：按顺序依次将请求分发给不同的云主机。它用相应的权重表示云主机的处理性能，按照权重的高低以及轮询方式将请求分配给各云主机，相同权重的云主机处理相同数目的连接数； 加权最少连接：通过当前活跃的连接数来估计云主机负载情况的一种动态调度算法； 源 IP 算法：将请求的源 IP 地址作为散列键（HashKey），从静态分配的散列表找出对应的云主机。	加权轮询算法
会话保持类型	会话保持的方式，用户可以根据自身需求选择相应的会话保持方式来分配用户访问流量： • 源 IP：将请求的源 IP 地址作为散列键（HashKey），从静态分配的散列表找出对应的云主机； • HTTP COOKIE：负载均衡器会根据客户端第一个请求生成一个 cookie，后续所有包含这个 cookie 值的请求都会由同一个后端云主机处理； • 应用程序 COOKIE：该选项依赖于后端应用。后端应用生成一个 cookie 值，后续所有包含这个 cookie 值的请求都会由同一个后端云主机处理；	HTTP_COOKIE
描述	对于监听器描述；	-
健康检查方式	当负载分发协议选择“HTTP”时，健康检查支持的三种类型，设置后不可修改： • TCP • HTTP 当负载分发协议选择“TCP”时，健康检查支持的三种类型，设置后不可修改： • HTTPS	HTTP

参数	说明	取值样例
	• TCP • HTTP 当负载分发协议选择“TCP”时，健康检查支持的三种类型，设置后不可修改： • TCP • PING	
检查周期(秒)	每次健康检查响应的最大间隔时间；	5
超时时间（秒）	每次健康检查响应的最大超时时间；	10
检查路径	当健康检查方式为 HTTP 时需要配置的选项。为需要被请求的 URL 地址；	/index.html
最大轮询次数	健康检查最大的重试次数，范围[1-10]；	3
HTTP 方法	当健康检查方式为 HTTP 或 HTTPS 时需要配置的选项，为 HTTP 或 HTTPS 请求的方法；	GET
HTTP 状态码	当健康检查方式为 HTTP 或 HTTPS 时需要配置的选项，为 HTTP 或 HTTPS 请求后表示请求返回的状态码；	201

- 单击【确定】按钮完成监听器创建；

3.2.2 修改监听器

- 登录天翼云控制中心；
- 选择【弹性负载均衡】【负载均衡器】；
- 单击已创建的负载均衡器实例名称；
- 在该负载均衡界面的【监听器】区域，单击监听器所在行的【修改】选项；



- 在【修改监听器】界面，根据页面提示配置参数，参数说明请见“3.2.1 添加监听器”；

- 单击【确定】按钮；

3.2.3 删除监听器

- 登录天翼云控制中心；
- 选择【弹性负载均衡】【负载均衡器】；
- 单击已创建的负载均衡器实例名称；
- 在该负载均衡界面的【监听器】区域，单击监听器所在行的【删除】选项；

  说明：如果该负载均衡器下有监听器，不能删除，需先删除监听器后才可删除负载均衡器。

3.3 后端云主机管理

本章节提供添加、修改监听器和删除监听器的操作步骤。当您需要向负载均衡器添加监听器，修改监听器的负载均衡模式、会话保持类型、健康检查配置或不再使用该监听器时可参考本章节。

3.3.1 添加后端云主机

本章节提供添加和移除后端云主机的操作步骤。当您需要将云主机添加至负载均衡器或将云主机从负载均衡器下移除时可参考本章节。

- 登录天翼云控制中心；
- 选择【弹性负载均衡】【负载均衡器】；
- 单击已创建的负载均衡器实例名称；
- 在该负载均衡详情界面的【监听器】所在行的操作栏中，单击【添加后端云主机】选项；
- 选择需要和负载均衡关联的云主机，并配置参数；

  各参数说明如下：

参数	说明	取值样例
业务端口	后端云主机的服务监听端口，取值范围[1-65535]；	123
权重	后端虚拟机权重。权重值决定了后端云主机处理的请	10

参数	说明	取值样例
	求的比例。例如，一个权重为 2 的云主机处理的请求数是权重为 1 的两倍。默认情况下，权重为 1；	

- 单击【确定】按钮完成后端云主机添加；

3.3.2 移除后端云主机

- 登录天翼云控制中心；
- 选择【弹性负载均衡】【负载均衡器】;
- 单击已创建的负载均衡器实例名称；
- 在该负载均衡详情界面，选择【后端主机组】标签；
- 需要移除多个后端云主机时，可勾选云主机并单击列表上方的【移除】按钮；需要移除单个后端云主机，可单击列表中云主机所在行的【移除】按钮或勾选云主机并单击列表上方的【移除】按钮；
- 单击【确定】按钮；

3.4 白名单

增强型负载均衡器用户可以通过添加白名单的方式控制访问负载均衡的监听器的 IP，能够设置允许特定 IP 访问，而其他 IP 不许访问。

【注】当开启短信验证后，删除负载均衡器需要进行短信身份验证。操作如下：

设置白名单是增强型负载均衡的功能，设置白名单存在一定业务风险。一旦设置白名单，就只有白名单中的 IP 可以访问负载均衡监听。

如开启访问控制而不设置白名单列表，则这个负载均衡监听就无人可以访问。

3.4.1 添加白名单

- 登录管理控制台。
- 单击“服务列表 > 弹性负载均衡”。

- 在“弹性负载均衡”界面，单击负载均衡名称，进入监听器管理界面。
- 在需要添加白名单的监听器的基本信息页面，单击访问控制右侧“设置”按钮，如下表所示配置白名单。

- 点击“确定”。

参数	说明	样例
访问控制开关	开启 开启访问控制开关而不设置白名单列表，表示不允许任何 IP 访问负载均衡监听器。 开启访问控制开关且在白名单列表中设置了 IP，表示允许该 IP 访问负载均衡监听器。 关闭 关闭访问控制开关，表示允许任何 IP 访问负载均衡监听器。	-
白名单	允许能够访问负载均衡的监听器的 IP 或网段。 说明 多个 IP 或网段间用逗号隔开，最多可以输入 300 个 IP 或网段。	10.168.2.24， 10.168.16.0/24

3.5 证书

仅用于 HTTPS。用户将证书上传到负载均衡中，在创建 HTTPS 协议监听的时候需绑定证书，提供

HTTPS 服务。

3.5.1 创建证书

- 登录管理控制台。
- 单击“服务列表 > 弹性负载均衡”。
- 在“弹性负载均衡”界面选择“证书管理”页签。
- 单击“创建证书”，配置证书内容。
  - 负载均衡类型：增强型、经典型
  - 证书名称
  - 证书类型：

服务器证书：在使用 HTTPS 协议时，服务器证书用于 SSL 握手协商，需提供证书内容和私钥。

CA 证书：又称客户端 CA 公钥证书，用于验证客户端证书的签发者；在 HTTPS 双向认证功能时，只有当客户端能够出具指定 CA 签发的证书时，HTTPS 连接才能成功。

  - 描述
  - 证书内容：证书内容必须为 PEM 格式
  - 私钥

需注意必须是无密码的私钥。私钥格式如下：

```
-----BEGIN PRIVATE KEY-----
[key]
-----END PRIVATE KEY-----
```

【注】若是证书链，需要配置从子证书到根证书的所有证书内容和私钥，且证书内容和私钥的配置顺序必须保持一致。

例如：某用户有三个证书，它们的关系是子证书>中级证书>根证书，则正确的配置顺序是子证

书>中级证书>根证书。

- 填写完成后，单击“确定”。

3.5.2 删除证书

删除证书时，只能删除未使用的证书，在使用中的证书无法删除。

- 登录管理控制台。
- 单击“服务列表 > 弹性负载均衡”。
- 在“弹性负载均衡”界面选择“证书”页签。
- 在证书列表中，在需要修改的证书所在行，单击“删除”。
- 在确认对话框中单击“确定”，完成删除。

3.5.3 绑定证书

- 登录管理控制台。
- 单击“服务列表 > 弹性负载均衡”。
- 在“弹性负载均衡”界面，选中需要添加 HTTPS 协议的弹性负载均衡，单击负载均衡器名称。
- 单击“添加监听器”。
- 在弹出的“添加监听器”对话框中，完成参数配置。
- 填写完成后，单击“确定”。

– 经典型负载均衡当“前端协议/端口”选择 HTTPS 协议时，监听器需绑定证书，即设置“默认证书”参数。

– 增强型负载均衡当“前端协议/端口”选择 HTTPS 协议时，监听器需绑定证书，即设置“服务器证书”参数。

3.6 HTTP/HTTPS 高级配置

3.6.1 URL 转发策略

增强型负载均衡用户可以通过添加转发策略支持自行设定的域名和 URL，将来自不同域名或者不同 URL 的请求转发到不同的后端云主机组处理。此功能目前仅支持协议类型为 HTTP、HTTPS 的监听器。

一个监听器最多可添加 20 条转发策略，您可以将视频、图片、音频、文本等请求分别转发到不同的后端云主机组上去处理，便于灵活的分流业务，合理的分配资源。

在添加了转发策略后，负载均衡器将按以下规则转发前端请求：

- 如果能匹配到监听器的转发策略，则按该转发策略将请求转发到对应的后端云主机组。
- 如果不能匹配到监听器的转发策略，则将请求转发到监听器对应的后端云主机组。

3.6.1.1 添加转发策略

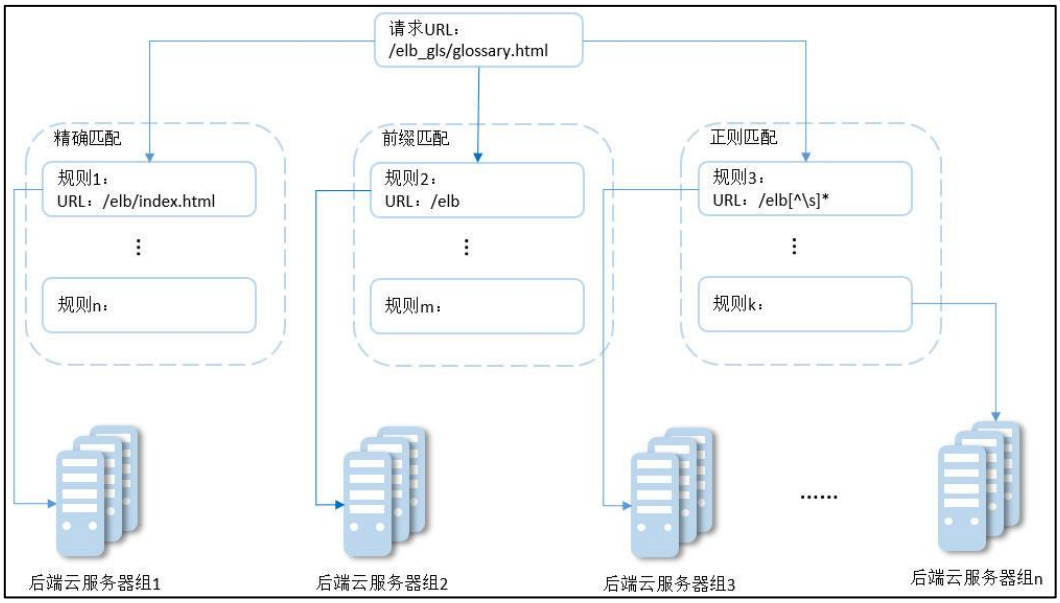
- 登录管理控制台。
- 单击“服务列表 > 弹性负载均衡”。
- 在“弹性负载均衡”界面，单击需要添加转发策略的负载均衡器名称。
- 选择需要添加转发策略的监听器，单击“更多 > 添加转发策略”。或者在监听器页面直接单击监听器的名称，跳转至添加转发策略页面，添加转发策略。
- 单击“添加转发策略”，参数配置如下表所示。
- 配置完成，单击“确定”。

参数	说明	样例
名称	转发策略的名称。	l7policy-l2dc
域名	触发转发的域名，仅支持精确域名。注意，域名或者 URL 至少要指定一个。	www.test.com
URL	触发转发的 URL。	/login.php
URL 匹配规则	精确匹配	-

参数	说明	样例
	请求的 URL 和设定 URL 完全一致。 前缀匹配 请求的 URL 匹配以设定 URL 开头的 URL。 正则匹配 请求的 URL 和设定的 URL 正则表达式匹配。 注 匹配的优先级为： 精确匹配>前缀匹配>正则匹配	
转发对象	用于处理请求的后端云主机组。后端云主机组是弹性负载均衡器的必要组成部分，用于接收并处理监听器转发的请求。	pool-nvhc
描述	转发策略的描述。	-

URL 匹配示例

模式	请求 UR	设定 URL			
-	-	/elb/index.html	/elb	/elb[^\s]*	/index.html
精确匹配	/elb/index.html	√	-	-	-
前缀匹配		√	√	-	-
正则匹配		√	-	√	-



以上图为例

请求的 URL：/elb_gls/glossary.html 先在精确匹配规则中查找，如果没有找到精确匹配的规则，则继续在前缀匹配规则中查找，找到匹配的规则 2，将该请求转发到规则 2 对应的后端云主机组 2。此时虽然请求 URL 和正则匹配规则中的规则 3 相匹配，但由于前缀匹配的优先级比较高，所以最终将请求转发至后端云主机组 2。

3.6.1.2 删除转发策略

- 登录管理控制台。
- 单击“服务列表 > 弹性负载均衡”。
- 在“弹性负载均衡”界面，单击需要删除转发策略的增强型负载均衡器名称。
- 单击需要删除转发策略的监听器名称，进入“添加转发策略”界面。
- 选择要删除转发策略，单击“删除”。
- 在弹出的“删除转发策略”对话框中，单击“是”，删除转发策略。

3.6.2 HTTPS 双向认证

使用场景

一般的 HTTPS 业务场景只对服务器做认证，因此只需要配置服务器的证书即可，某些关键业务（如银行支付），需要对通信双方的身份都要做认证，即双向认证，以确保业务的安全性，此时，除了配置服务器的证书之外，还需要配置客户端的证书，以实现通信双方的双向认证功能。以下说明如何在 ELB 上配置双向认证功能。

3.6.2.1 证书准备

服务器证书准备

用户可以用权威 CA 签发的证书或者自签名的证书，这里以自签名证书为例说明如何创建服务器证书。

- 登录到任意一台安装有 openssl 工具的 linux 机器。
- 创建工作目录并进入该目录。

  ```
  mkdir server
  cd server
  ```

- 创建 CA 证书的 openssl 配置文件 ca_cert.conf，内容如下：

  ```
  [ req ]
  distinguished_name     = req_distinguished_name
  prompt                 = no

  [ req_distinguished_name ]
   O                      = ELB
  ```

- 创建服务器证书的 openssl 配置文件 server_cert.conf，内容如下：

  ```
  [ req ]
  distinguished_name     = req_distinguished_name
  prompt                 = no

  [ req_distinguished_name ]
   O                      = ELB
   CN                     = www.test.com
  ```

  说明：

  CN 字段可以根据需求改为服务器对应的域名、IP 地址。

- 创建 CA 证书私钥文件 ca.key 以及服务器证书私钥文件 server.key。

  ```
  openssl genrsa -out ca.key 2048
  openssl genrsa -out server.key 2048
  ```

- 创建 CA 证书以及服务器证书的 csr 请求文件 ca.csr/server.csr。

  openssl req -out ca.csr -key ca.key -new -config ./ca_cert.conf

  openssl req -out server.csr -key server.key -new -config ./server_cert.conf

- 创建自签名的 CA 证书 ca.crt 以及用该 CA 证书签名的服务器证书 server.crt。

  openssl x509 -req -in ca.csr -out ca.crt -sha1 -days 5000 -signkey ca.key

  openssl x509 -req -in server.csr -out server.crt -sha1 -CAcreateserial -days 5000 -CA ca.crt -CAkey ca.key

3.6.2.2 客户端证书准备

登录到任意一台安装有 openssl 工具的 linux 机器。

- 创建工作目录并进入该目录。

  mkdir client

  cd client

- 创建 CA 证书的 openssl 配置文件 ca_cert.conf，内容如下：

```
[ req ]
distinguished_name     = req_distinguished_name
prompt                 = no

[ req_distinguished_name ]
 O                      = ELB
```

- 创建客户端证书的 openssl 配置文件 client_cert.conf，内容如下：

```
[ req ]
distinguished_name     = req_distinguished_name
prompt                 = no

[ req_distinguished_name ]
 O                      = ELB
 CN                     = www.test.com
```

  说明：

  CN 字段可以根据需求改为服务器对应的域名、IP 地址。

- 创建 CA 证书私钥文件 ca.key 以及客户端证书私钥文件 client.key。

  openssl genrsa -out ca.key 2048

  openssl genrsa -out client.key 2048

- 创建 CA 证书以及客户端证书的 csr 请求文件 ca.csr/client.csr。

```
openssl req -out ca.csr -key ca.key -new -config ./ca_cert.conf

openssl req -out client.csr -key client.key -new -config ./client_cert.conf
```

- 创建自签名的 CA 证书 ca.crt 以及用该 CA 证书签名的客户端证书 client.crt。

```
openssl x509 -req -in ca.csr -out ca.crt -sha1 -days 5000 -signkey ca.key

openssl x509 -req -in client.csr -out client.crt -sha1 -CAcreateserial -days
5000                -CA ca.crt -CAkey ca.key
```

- 把客户端证书格式转为浏览器可识别的 p12 格式。

```
openssl pkcs12 -export -clcerts -in client.crt -inkey client.key -out clie
```

说明：

该命令执行时需要输入导出密码，请输入并记住该密码，在证书导入浏览器时需要

3.6.2.3 配置证书

配置服务器证书和私钥

- 登录负载均衡控制台页面。

在创建证书页面，证书类型选择“服务器证书”，同时把上文证书准备部分创建的服务器证书 server.crt 以及私钥 server.key 的内容复制到对应的区域，点击“确定”按钮。

说明：

服务器证书和私钥内容只支持 pem 格式。

配置 CA 证书

- 登录负载均衡控制台页面。

在创建证书页面，证书类型选择“CA 证书”，同时把上文证书准备部分创建的 ca.crt 的内容复制到证书内容区域，点击“确定”按钮。

说明：

CA 证书内容只支持 pem 格式。

配置监听器

绑定服务器证书和 CA 证书

- 登录负载均衡控制台页面。

- 在添加监听器页面，协议类型选择“HTTPS(Termination)”，打开双向认证开关，并且在证书和 CA 证书两个配置项中选择所添加的服务器证书和 CA 证书对应的 ID。

  说明

  只有增强型监听器才支持双向认证功能。

添加后端服务器

此步骤请参考 3.3.1 添加后端服务器。

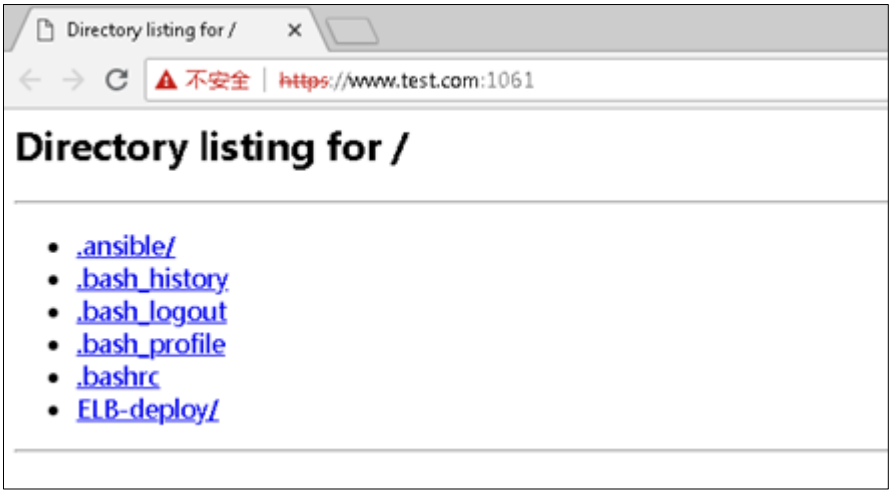
3.6.2.4 功能测试

浏览器导入客户端证书

- 把客户端证书从 linux 机器导出来，即在“客户端证书准备”中生成的 client.p12 文件。
- 双击 client.p12 文件，按提示导入证书文件，这个过程中系统会提示输入密码，该密码为“客户端证书准备”导出该文件时所输入的密码。

测试验证

在浏览器中输入地址，浏览器会弹出证书选择窗口，如下，选择客户端证书，然后点确定按钮，可以正常访问网站，如下图。



3.6.3 HTTP 重定向至 HTTPS

操作场景

对于需要保证业务建立安全连接 的用户，如果您已创建了 HTTPS 监听器和 HTTP 监听器，可以通过

负载均衡 HTTP 重定向功能，将 HTTP 访问重定向至 HTTPS。HTTPS 是加密数据传输协议，安全性高。

说明：

目前只有增强型负载均衡支持此功能，经典型负载均衡不支持。

3.6.3.1 添加重定向

- 登录管理控制台。
- 选择“服务列表 > 网络 > 弹性负载均衡”。
- 在“负载均衡器”界面，单击需要重定向的 HTTP 监听器的负载均衡名称。
- 在该负载均衡界面的“监听器”页签，单击需要重定向的 HTTP 监听器名称。
- 选择“重定向 > 添加”，选择需要重定向至 HTTPS 监听器的名称。

参数	说明	取值样例
名称	重定向的名称。	
重定向至	选择需要重定向 HTTPS 监听器的名称。	

- 在确认对话框单击“确定”。

说明：

HTTP 监听器被重定向后，原有监听器转发与转发策略转发都会失效。

3.6.3.2 修改重定向

- 登录管理控制台。
- 选择“服务列表 > 网络 > 弹性负载均衡”。
- 在“负载均衡器”界面，单击需要重定向的 HTTP 监听器的负载均衡名称。
- 在该负载均衡界面的“监听器”页签，单击需要重定向的 HTTP 监听器名称。
- 选择“重定向 > 修改”，选择需要重定向至 HTTPS 监听器的名称。
- 在确认对话框单击“确定”。

3.6.3.3 删除重定向

- 登录管理控制台。
- 选择“服务列表 > 网络 > 弹性负载均衡”。
- 在“负载均衡器”界面，单击已经重定向的 HTTP 监听器的负载均衡名称。
- 在该负载均衡界面的“监听器”页签，单击已经重定向的 HTTP 监听器名称。
- 选择“重定向 > 删除”。
- 在确认对话框单击“是”。

3.7 标签管理

操作场景

对于拥有大量云资源的用户，可以通过给云资源打标签，快速查找具有某标签的云资源，可对这些资源标签统一进行检视、修改、删除等操作，方便用户对云资源的管理。

目前只有增强型负载均衡支持此功能，经典型负载均衡不支持。

3.7.1 为负载均衡器添加标签

给负载均衡器添加标签有以下两种方法。

——在创建负载均衡器的时候，输入标签的“键”和“值”。

操作步骤和配置参数，请参见 2.2 创建增强型负载均衡器。

——给已创建的负载均衡器添加标签。

- 登录管理控制台。
- 单击“服务列表 > 弹性负载均衡”。
- 在“弹性负载均衡”界面，单击已创建的负载均衡器名称。
- 在“标签”页签下，单击“添加标签”，输入“键”和“值”。

- 确认正确，单击“确认”。

说明：

一个负载均衡器最多可以增加 10 个标签。

标签的“键”和“值”是一一对应的，其中“键”值是唯一的。

3.7.2 监听器添加标签

给已创建的监听器添加标签的方法如下：

- 登录管理控制台。
- 单击“服务列表 > 弹性负载均衡”。
- 在“弹性负载均衡”界面，单击已创建的负载均衡器名称。
- 切换到监听器页签，单击需要添加标签的监听器名称。
- 切换到监听器子页面的标签页签，单击“添加标签”，输入“键”和“值”。
- 确认正确，单击“确认”。

说明：

一个监听器最多可以增加 10 个标签。

标签的“键”和“值”是一一对应的，其中“键”值是唯一的。

3.7.3 修改标签

- 登录管理控制台。
- 单击“服务列表 > 弹性负载均衡”。
- 在“弹性负载均衡”界面，单击需要修改标签的负载均衡器名称。
- 在“标签”页签下，在需要修改的标签所在行，单击 “编辑”，输入修改的“值”。

说明：

“键” 值不支持修改。

- 确认正确，单击“确认”。

以上步骤描述的是修改负载均衡器的标签，修改监听器的标签可参考上面步骤进行，仅有操作入口

不同。

3.7.4 删除标签

- 登录管理控制台。
- 单击“服务列表 > 弹性负载均衡”。
- 在“弹性负载均衡”界面，单击需要删除标签的负载均衡器名称。
- 在“标签”页签下，在需要删除的标签所在行，单击 "删除"。
- 确认正确，单击“确认”。

以上步骤描述的是删除负载均衡器的标签，删除监听器的标签可参考上面步骤进行，仅有操作入口不同。

4 监控

4.1 支持的监控指标

当用户开通了弹性负载均衡服务后，无需额外安装其他插件，即可在云监控查看对应服务的实例状态，云监控支持监控经典型负载均衡器的相关指标如表一所示，支持增强型负载均衡器和增强型负载均衡监听器的相关指标如表二所示。

表一　经典型负载均衡指标

指标名称	指标含义	监控周期（原始指标）	备注
并发连接数	统计监控对象当前处理的并发连接数。 单位：个	1 分钟	测量对象：经典型负载均衡器。
活跃连接数	统计监控对象当前处理的活跃连接数量。 单位：个	1 分钟	测量对象：经典型负载均衡器。
非活跃连接数	统计监控对象当前处理的非活跃连接数量。 单位：个	1 分钟	测量对象：经典型负载均衡器。
新建连接数	统计监控对象当前处理的新建连接数量。 单位：个	1 分钟	测量对象：经典型负载均衡器。
流入数据包数	统计当前流入监控对象的数据包。 单位：个	1 分钟	测量对象：经典型负载均衡器。
流出数据包数	统计当前流出监控对象的数据包。 单位：个	1 分钟	测量对象：经典型负载均衡器。
网络流入速率	统计每秒流入监控对象的网络流量。 单位：字节/秒	1 分钟	测量对象：经典型负载均衡器。

指标名称	指标含义	监控周期（原始指标）	备注
网络流出速率	统计每秒流出监控对象的网络流量。 单位：字节/秒	1 分钟	测量对象：经典型负载均衡器。
异常主机数	统计监控对象后端异常的主机个数。 单位：个	1 分钟	测量对象：经典型负载均衡器。
正常主机数	统计监控对象后端正常的主机个数。 单位：个	1 分钟	测量对象：经典型负载均衡器。

表二　增强型负载均衡指标

指标名称	指标含义	监控周期（原始指标）	备注
并发连接数	统计监控对象当前处理的并发连接数。 单位：个。	1 分钟	测量对象：增强型负载均衡器和增强型负载均衡监听器。
活跃连接数	统计监控对象当前处理的活跃连接数量。 单位：个。	1 分钟	测量对象：增强型负载均衡器和增强型负载均衡监听器。
非活跃连接数	统计监控对象当前处理的非活跃连接数量。 单位：个。	1 分钟	测量对象：增强型负载均衡器和增强型负载均衡监听器。
新建连接数	统计监控对象当前处理的新建连接数量。 单位：个。	1 分钟	测量对象：增强型负载均衡器和增强型负载均衡监听器。
流入数据包数	统计当前流入监控对象的数据包。 单位：个。	1 分钟	测量对象：增强型负载均衡器和增强型负载均衡监听器。

指标名称	指标含义	监控周期（原始指标）	备注
流出数据包数	统计当前流出监控对象的数据包。 单位：个。	1 分钟	测量对象：增强型负载均衡器和增强型负载均衡监听器。
网络流入速率	统计每秒流入监控对象的网络流量。 单位：字节/秒。	1 分钟	测量对象：增强型负载均衡器和增强型负载均衡监听器。
网络流出速率	统计每秒流出监控对象的网络流量。 单位：字节/秒。	1 分钟	测量对象：增强型负载均衡器和增强型负载均衡监听器。
异常主机数	统计监控对象后端异常的主机个数。 单位：个	1 分钟	测量对象：增强型负载均衡器。
正常主机数	统计监控对象后端正常的主机个数。 单位：个	1 分钟	测量对象：增强型负载均衡器。

4.2 告警规则

4.2.1 添加告警规则

- 登录管理控制台。
- 选择“管理与部署 > 云监控”。
- 在左侧导航树栏，选择“告警 > 告警规则”。
- 在“告警规则”界面，单击“创建告警规则”进行添加，设置弹性负载均衡器的告警规则。

更多关于弹性负载均衡器监控规则的信息，请参见《云监控用户指南》。

4.2.2 修改告警规则

- 登录管理控制台。
- 选择“管理与部署 > 云监控”。

- 在左侧导航树栏，选择“告警 > 告警规则”。
- 在“告警规则”界面，选择已有的告警规则进行修改，设置弹性负载均衡器的告警规则。

更多关于弹性负载均衡器监控规则的信息，请参见《云监控用户指南》。

5 审计

5.1 支持审计的关键操作列表

通过云审计服务，您可以记录与弹性负载均衡相关的操作事件，便于日后的查询、审计和回溯。

云审计支持的弹性负载均衡列表操作事件如下表所示。

操作名称	资源类型	事件名称
配置访问日志	accesslog	create access log
删除访问日志	accesslog	delete access log
创建证书	certificate	create certificate
更新证书	certificate	update certificate
删除证书	certificate	delete certificate
创建健康检查	healthmonitor	create healthmonitor
更新健康检查	healthmonitor	update healthmonitor
删除健康检查	healthmonitor	delete healthmonitor
创建转发策略	l7policy	create forwarding policy
更新转发策略	l7policy	update forwarding policy
删除转发策略	l7policy	delete forwarding policy
创建转发规则	l7rule	create forwarding rule
更新转发规则	l7rule	update forwarding rule

删除转发规则	l7rule	delete forwarding rule
创建监听器	listener	create listener
更新监听器	listener	update listener
删除监听器	listener	delete listener
创建负载均衡器	loadbalancer	create loadbalancer
更新负载均衡器	loadbalancer	update loadbalancer
删除负载均衡器	loadbalancer	delete loadbalancer
添加后端云主机	member	add backend ecs
更新后端云主机	member	update backend ecs
移除后端云主机	member	remove backend ecs
创建后端云主机组	pool	create backend member group
更新后端云主机组	pool	update backend member group
删除后端云主机组	pool	delete backend member group

5.2 查看审计日志

操作场景

在您开通了云审计服务后，系统开始记录云服务资源的操作。云审计服务管理控制台保存最近 7 天的操作记录。

本节介绍如何在云审计服务管理控制台查看最近 7 天的操作记录。

操作步骤

- 登录管理控制台。
- 单击“服务列表”，选择“管理与部署 > 云审计服务”，进入云审计服务信息页面。
- 单击左侧导航树的“事件列表”，进入事件列表信息页面。
- 事件列表支持通过筛选来查询对应的操作事件。当前事件列表支持四个维度的组合查询，详细信息如下：

  - 事件来源、资源类型和筛选类型。

    在下拉框中选择查询条件。

    其中筛选类型选择事件名称时，还需选择某个具体的事件名称。

    选择资源 ID 时，还需选择或者手动输入某个具体的资源 ID。

    选择资源名称时，还需选择或手动输入某个具体的资源名称。

  - 操作用户：在下拉框中选择某一具体的操作用户，此操作用户指用户级别，而非租户级别。

  - 事件级别：可选项为“所有事件级别”、“normal”、“warning”、“incident”，只可选择其中一项。

  - 时间范围：可选择查询最近七天内任意时间段的操作事件。

- 在需要查看的记录左侧，单击展开该记录的详细信息，展开记录如下图所示。

- 在需要查看的记录右侧，单击“查看事件”，弹出一个窗口，如下图所示，显示了该操作事件结构的详细信息。

6 常见问题

6.1 弹性负载均衡（增强型）是什么？

弹性负载均衡（Elastic Load Balancing ，简称 ELB）增强型是将访问流量根据转发策略分发到后端多台弹性云主机的流量分发控制服务。弹性负载均衡（增强型）可以通过流量分发扩展应用系统对外的服务能力，实现更高水平的应用程序容错性能。

用户通过基于浏览器、统一化视图的云计算管理图形化界面，可以创建 ELB（增强型），为服务配置需要监听的端口，配置云主机。消除单点故障，提高整个系统的可用性。

6.2 弹性负载均衡服务是否收费？

负载均衡服务本身不收费，但绑定弹性 IP 时会收取弹性 IP 及带宽的费用。

6.3 弹性负载均衡支持哪些转发方式？

当前 ELB（增强型）支持加权轮询、加权最少连接和源 IP 三种模式的转发规则。

- 加权轮询算法：按顺序依次将请求分发给不同的云主机。它用相应的权重表示云主机的处理性能，按照权重的高低以及轮询方式将请求分配给各云主机，相同权重的云主机处理相同数目的连接数。
- 加权最少连接：通过当前活跃的连接数来估计云主机负载情况的一种动态调度算法。
- 源 IP 算法：将请求的源 IP 地址作为散列键（HashKey），从静态分配的散列表找出对应的云主机。

经典型负载均衡支持如下三种算法

- 轮询算法：将请求轮流发送给后端云主机，常用于短连接服务，例如 HTTP 等服务。
- 最少连接：优先将请求发给拥有最少连接数的后端云主机，常用于长连接服务，例如数据库连接等服务。

- 源 IP：将请求的源 IP 地址作为散列键（HashKey），从静态分配的散列表找出对应的云主机。这可以使得同一个客户端 IP 的请求始终被派发至某特定的云主机。该方式适合负载均衡无 cookie 功能的 TCP 协议。

6.4 弹性负载均衡是否支持 IPv6?

弹性负载均衡（经典型、增强型-性能共享型）不支持 IPv6；弹性负载均衡（增强型-性能保障型）支持 IPv6，可在创建实例时将实例创建在具有双栈子网的 VPC 中，IPv6 地址分配在双栈子网中，即可获得 IPv6 地址。

如果您已经创建性能保障型负载均衡器，但当时未分配 IPv6 地址，可将其所在子网先开启 IPv6，再到负载均衡器页面为其绑定新的 IPv6 地址。

6.5 弹性负载均衡性能保障型的性能指标是多少？

flavor 分类	规格类型	规格信息		
		单 flavor 并发最大连接数	单 flavor 每秒新建连接数 (CPS)	单 flavor 每秒查询数 (QPS)
L4_flavor	小型 I	50000	1000	-
	小型 II	100000	2000	-
	中型 I	200000	4000	-
	中型 II	400000	8000	-
	大型 I	1000000	20000	-
	大型 II	2000000	40000	-
L7_flavor	小型 I	200000	2000	4000
	小型 II	400000	4000	8000
	中型 I	800000	8000	16000
	中型	2000000	20000	40000
	大型 I	4000000	40000	80000
	大型	8000000	80000	160000

6.6 弹性负载均衡是否可以添加不同操作系统的云主机？

可以。ELB 本身不会限制后端的云主机使用哪种操作系统，只要您的 2 台云主机中的应用服务部署是相同且保证数据的一致性即可。但是，我们建议您选择 2 台相同操作系统的云主机进行配置，以便您日后的管理维护。

6.7 单个用户支持保有多少个弹性负载均衡？

单个用户默认可创建 5 个增强型负载均衡器，5 个经典型负载均衡器。如果需要创建更多弹性负载均衡器，请申请更高配额，申请弹性负载均衡器个数不超过 255 个。

6.8 什么是配额？

为防止资源滥用，平台限定了各服务资源的配额，对用户的资源数量和容量做了限制。如果当前资源配额限制无法满足使用需要，您可以申请扩大配额。

6.9 弹性负载均衡如何支持多证书？

每个监听器只支持一个证书/证书链，如果要支持多证书/证书链，需创建多个监听器。

6.10 如何配置私网或公网负载均衡？

创建一个增强型负载均衡，系统分配一个私有 IP，默认是私网负载均衡，如果为这个私有 IP 绑定一个公网的 IP，则可作为公网负载均衡。增强型负载均衡同时可支持私网、公网访问。

6.11 监听器是什么？

承担弹性负载均衡具体的协议和端口配置，云主机协议和端口配置，监听策略配置。

6.12 监听器中分配算法和会话保持算法是什么关系？

会话保持功能，目的是将同一个用户的会话分发到相同的后端节点，增强型负载均衡支持情况如下表所示。

增强型会话保持支持情况

分配策略	会话保持	L4（TCP）	L7（HTTP/HTTPS）
加权轮询算法	源 IP 地址	支持	不支持
	HTTP cookie	不涉及	支持
	应用程序 cookie	不涉及	支持
加权最少连接	源 IP 地址	不支持	不支持
	HTTP cookie	不涉及	不支持
	应用程序 cookie	不涉及	不支持
源 IP 地址	源 IP 地址	支持	支持
	HTTP cookie	不涉及	不支持
	应用程序 cookie	不涉及	不支持

一般建议：算法可以使用轮询算法，四层会话保持使用源 IP 地址，7 层使用 HTTP cookie 方式。

6.13 什么是负载均衡协议（端口）？

增强型负载均衡系统支持 4 层（TCP、UDP）和 7 层（HTTP、HTTPS）协议的负载均衡，经典型负载均衡系统支持 4 层（TCP、UDP）和 7 层（HTTP、HTTPS）协议的负载均衡，可通过具体提供的服务能力选择对应的协议以及该协议对外呈现的端口。

监听器协议	用途
TCP	TCP 的应用部署
UDP	UDP 的应用部署
HTTP	Web 应用
HTTPS	基于 HTTPS 的 Web 应用

6.14 什么是云主机协议（端口）？

后端云主机自身提供的网络服务的协议以及协议的端口，如使用 windows 操作系统上安装的 IIS（webservice），该服务默认的协议为 HTTP，端口为 80。

6.15 弹性负载均衡分配的弹性 IP 是否为独占？

在您购买使用 ELB 服务的整个生命周期内：

经典型负载均衡：分配的弹性公网 IP 都是由您所购买的服务独占。

增强型负载均衡：分配的弹性公网 IP 支持解绑，解绑后的增强型负载均衡变成私网型负载均衡，解绑后的公网 IP 可被其他资源绑定。

6.16 删除弹性负载均衡有什么影响？

如果您的 ELB 服务地址（IP）已经正常解析到域名且对外提供服务，除非必要请不要删除您创建的 ELB 服务，删除了 ELB 服务以后相应的服务配置和服务地址（IP）将会被释放掉，数据一旦删除，不可恢复。如果您重新创建 ELB 服务，可以重新由系统重新给您分配一个新的服务地址（IP），也可以指定原 IP 地址申请。

6.17 健康检查异常如何排查？

ELB（的健康检查是通过 ELB（增强型）系统向后端云主机发起心跳检查的方式来实现的，而 ELB（增强型）系统和云主机之间是通过内网进行通信的，为了确保健康检查工作的正常进行，您需要确保能够通过内网访问您的云主机，请按照以下方法排查。

- 单击对应的负载均衡名称，进入负载均衡基本信息页面。切换到后端云主机组页签，单击对应的后端云主机组名称，在其基本信息页面，单击【健康检查】右侧的配置按钮。执行如下操作：
  - 检查“健康检查协议”和“健康检查端口”，确保后端云主机已配置相应协议并开启端口；
  - 检查“检查路径”：如果是使用 HTTP 协议进行健康检查，还应检查后端主机的健康检查路径是否正确。
- 检查云主机中防火墙等软件是否有对来自健康检查源 IP 的屏蔽；
- 检查后端云主机所在安全组规则是否配置放行 100.125.0.0/16，并配置 ELB（增强型）用于健康检查的协议和端口。健康检查的协议和端口在弹出的健康检查配置项提示框中获取。

- 若采用默认的健康检查方式：需要放行后端云主机业务端口；
- 若配置了不同于云主机业务端口的健康检查端口：需要放行云主机业务端口与健康检查端口。

- 如果以上配置检查均正常但问题依然存在，请通过 400-810-9889 或在线工单进行报障；

6.18 为什么很多访问 ELB 实例后端云主机的 IP 是 100.125 开头？

这是由于 ELB 系统进行健康检查引起的。

ELB 系统除了会通过系统云主机的内网 IP 将来自外部的访问请求转到后端云主机上之外，还会对云主机进行健康检查，并对后端服务进行可用性监控，这些访问的来源都是由 ELB 系统发起的，具体包含的 IP 地址段是：100.125.0.0/16。

为了确保您对外服务的可用性，请确保在云主机所在的安全组上对上述地址的访问配置放行规则。

6.19 如何获得来访者的真实 IP？

针对 7 层（HTTP 协议）服务，ELB（增强型）通过 Http Header:X-Forwarded-For 获取来访者真实 IP，该功能已经默认开启，无需配置，也不能修改。

针对 4 层（TCP 协议）服务，建议采用 TOA 方案。

6.20 ELB 支持什么类型的会话保持？

增强型负载均衡器支持源 IP、HTTP cookie、APP_COOKIE 三种会话保持类型，经典型负载均衡器支持源 IP、HTTP cookie 两种会话类型。

6.21 使用 UDP 协议有什么注意事项？

- 负载均衡健康检查是通过 UDP 报文和 Ping 报文探测来获取后端云主机的状态信息。针对此种情况，用户需要确保后端云主机开启 ICMP 协议，确认方法如下：

用户登录后端云主机，以 root 权限执行以下命令：

cat /proc/sys/net/ipv4/icmp_echo_ignore_all

若返回值为 1，表示 ICMP 协议关闭；若为 0，则表示开启。

- 当前 UDP 协议服务健康检查可能存在服务真实状态与健康检查不一致的问题：

  如果后端 ECS 云主机是 Linux 云主机，在大并发场景下，由于 Linux 的防 ICMP 攻击保护机制，会限制云主机发送 ICMP 的速度。此时，即便服务已经出现异常，但由于无法向前端返回“port XX unreachable”报错信息，会导致负载均衡由于没收到 ICMP 应答进而判定健康检查成功，最终导致服务真实状态与健康检查不一致。

- 当负载均衡的类型为经典型私网负载均衡时，不允许创建 UDP 协议的监听器。

6.22 什么是 UDP 健康检查？

UDP 是面向非连接的一种协议，在发送数据前不会通过进行三次握手建立连接，UDP 健康检查的实现过程如下：

1. 健康检查的节点根据健康检查配置，向后端云主机发送 ICMP request 消息。

- 如果健康检查节点收到了后端云主机返回的 ICMP reply 消息，则认为服务正常，继续进行健康检查。

- 如果健康检查节点没有收到后端云主机返回的 ICMP reply 消息，则认为服务异常，判定健康检查失败。

2. 健康检查的节点收到 ICMP reply 消息后，会给后端云主机发送 UDP 探测报文。

- 如果在【超时时间】之内，健康检查的节点服务器收到了后端云主机返回的 port unreachable 的 ICMP 消息，则认为服务异常，判定健康检查失败。

- 如果在【超时时间】之内，健康检查的节点服务器没有收到后端云主机返回的 ICMP 错误信息，则认为服务正常，判定健康检查成功。

当您配置 UDP 健康检查时，推荐使用配置页面默认的各项数值。

如果后端云主机上端口状态和页面健康检查显示不一致，请您按照一下方法排查。

1. 检查健康检查超时时间是否过小。

可能的原因：后端云主机回复的 reply 或 port unreachable 类型的 ICMP 消息未能在超时时间内到达健康检查的节点，导致健康检查结果不准确。

建议采取的措施：将超时时间调整为更大的值。

由于 UDP 健康检查的原理不同于其他健康检查，建议健康检查超时时间不要过小，否则后端云主机能会反复上线或下线。

2. 后端云主机是否限制了 ICMP 消息产生的速率。

Linux 系统下，请用以下命令检查 ICMP 消息速率的限制。

```
sctl -q net.ipv4.icmp_ratelimit
```

默认值为：1000

```
sysctl -q net.ipv4.icmp_ratemask
```

默认值为：6168

请确认第一条命令返回值为默认值或 0，并用以下命令放开 port unreachable 消息产生的速率限制。

```
sysctl -w net.ipv4.icmp_ratemask=6160
```

更详细的信息请参考 Linux Programmer's Manual 相关页面：

```
man 7 icmp
```

或者访问地址：http://man7.org/linux/man-pages/man7/icmp.7.html

【注】放开 port unreachable 类型 ICMP 消息的速率限制，会让暴露在公网上的云主机在端口扫描时，不受限制次数地产生 port unreachable 消息。

6.23 如何检查弹性负载均衡会话保持不生效问题？

- 查看后端云主机组上是否开启了会话保持。

- 查看后端云主机的健康检查状态是否正常，如果异常，流量会切换到其他后端云主机，导致会话保持失效。
- 如果选择的是源 ip 算法，需要注意请求到达弹性负载均衡之前 ip 是否发生变化。
- 如果是 HTTP 或 HTTPS 监听器，配置了会话保持，需要注意发送的请求是否带有 cookie，如果带有 cookie，则观察该 cookie 值是否发生了变化（因为 7 层会话保持基于 cookie）。

6.24 如何检查弹性负载均衡业务访问延时大？

- 将弹性 IP 绑定到后端云主机，不经过弹性负载均衡直接访问后端服务，查看访问延时。用来判断是弹性负载均衡的问题，还是前端网络问题或者后端服务问题。
- 查看业务流量是否超过了 EIP 的带宽限制。
- 如果直接访问后端存在业务访问延时大，需要排查后端服务是否压力过大，是否配置了安全策略等。
- 查看异常主机数的监控来判断后端云主机的健康检查状态是否有跳变。在后端服务状况不稳定时，因为弹性负载均衡的重试机制，如果连接一台后端超时，请求会重新发往下一台后端，请求成功，这样业务就表现为访问成功，但是延时很大。
- 如果问题依然存在，请联系天翼云技术支持。

6.25 如何检查弹性负载均衡服务不通或异常中断？

- 检查后端云主机的健康检查状态是否正常。
- 客户后端服务安全策略中是否放通了 100.125 网段，比如 DNS 服务中的 acl 规则。
- 弹性负载均衡与客户端的 TCP 连接，空闲超时时间是 300s，超过 300s，弹性负载均衡会向客户端和服务端发送 RST 断开连接。

6.26 如何检查弹性负载均衡前后端流量不一致？

检查客户端请求是否有失败的请求，特别是返回码是 4xx 的请求。因为这些请求可能因为是异常请求被弹性负载均衡拒绝，没有转发至后端云主机。

6.27 如何检查请求不均衡？

- 检查是否开启了会话保持。如果配置了会话保持，而客户端的个数又比较少时，很容易导致不均衡。
- 检查后端云主机的健康检查状态是否正常，特别要关注下是否有健康检查状态一会正常一会异常的情况。健康检查异常或者状态切换都会导致流量不均衡。
- 检查负载均衡算法是否是源 ip 算法。此时同一个 ip 发过来的请求都会分发到同一个后端，导致流量不均衡。
- 后端服务是否开启了 TCP keepalive 保持长连接。如果开启，则有可能因为长连接上的请求数不同导致流量不均衡。
- 将云主机添加到 ELB 后端时是否设置了权重，权重不同，分发的流量也不同。

6.28 如何检查弹性负载均衡健康检查异常？

- 后端云主机的安全组是否放通了 100.125.0.0/16 网段。
- 后端服务是否正常。
  - 如果是 TCP 健康检查，检查对应端口是否在监听。
  - 如果是 HTTP 健康检查，在本地访问健康检查配置中设置的 url，查看返回码是否正常。
- 查看后端服务的负载。如果负载很高，可能会导致健康检查的连接或请求超时。
- 检查云主机内部是否开启了防火墙或其他安全类防护软件，这些软件可能会屏蔽 100.125.0.0/16 网段的 ip。
- 是否手动修改了云主机内部的路由。查看主网卡（比如 eth0）上是否配置有 192.168.0.0/16 网段的默认路由。如果路由更改，可能导致健康检查报文无法到达后端云主机。
- 如果是 UDP 监听器，需要确保后端云主机未关闭 ICMP 协议，详情参考 6.16 使用 UDP 协议有什么注意事项？。

6.29 如何检测压测性能上不去？

- 检查后端云主机的负载状态，如果 CPU 达到 100%，可能是后端应用达到性能瓶颈。
- 查看流量是否超过绑定到弹性负载均衡的 EIP 的带宽，带宽超限后，会有大量丢包和请求失败，影响压测性能。
- 如果是短连接测试，可能是客户端端口不足导致建立连接失败，可以通过客户端处于 time_wait 状态的连接数量来判断。可通过增加客户端 ip 来解决。
- 后端云主机的监听队列 backlog 满了，导致后端云主机不回复 syn_ack 报文，使得客户端连接超时。可以通过调整 net.core.somaxconn 参数来调大 backlog 的上限值。

7 附录

7.1 TOA 插件配置

操作场景：

ELB 可以针对客户访问的业务为访问者提供个性化的管理策略，制定策略之前需要获取来访者的真实 IP。TOA 内核模块主要用来获取 ELB 转化过的访问者真实 IP 地址（仅支持 IPv4），该插件安装在 ELB 后端云主机。

当客户需要在操作系统中编译 TOA 内核模块时，可参考本文档进行配置。

Linux 内核版本为 2.6.32 和 Linux 内核版本为 3.0 以上的操作系统，在配置 TOA 内核模块的操作步骤上有所区别，具体操作请参照相应的操作步骤进行配置。

说明：

TOA 模块在以下操作系统中验证可以正常工作，但不支持 UDP 协议的弹性负载均衡器。

- CentOS 6.8（Kernel version 2.6.32）
- Suse 11 sp3（Kernel version 3.0.76）
- CentOS 7/7.2（Kernel version 3.10.0）
- Ubuntu 16.04.3（Kernel version 4.4.0）
- OpenSUSE 42.2（Kernel version 4.4.36）
- CoreOS 10.10.5（Kernel version 4.9.16）

前提条件

- 编译内核模块开发环境需与当前内核版本开发环境一致。
- 确保虚拟机可以访问开放源。
- 如果是非 root 用户，需拥有 sudo 权限。

操作步骤

以下操作步骤是针对 Linux 内核版本为 3.0 以上的操作系统。

1. 准备编译环境。

说明：

安装内核模块开发包的过程中，如果源里面找不到对应内核版本的安装包，需要自行去网上下载需要的安装包。

以下是不同 Linux 发行版本的操作说明，请根据环境选择对应的方案。

CentOS 环境下的操作步骤。

i. 执行如下命令，安装 gcc 编译器。

sudo yum install gcc

ii. 执行如下命令，安装 make 工具。

sudo yum install make

iii. 执行如下命令，安装内核模块开发包，开发包头文件与库的版本需要与内核版本一致。

sudo yum install kernel-devel-`uname -r`

说明：

如果自带源里没有对应的内核开发包，可以到如下地址中去下载对应的 rpm 包。

地址：https://mirror.netcologne.de/oracle-linux-repos/ol7_latest/getPackage/

以 3.10.0-693.11.1.el7.x86_64 为例，下载后执行以下命令安装：

rpm -ivh kernel-devel-3.10.0-693.11.1.el7.x86_64.rpm。

Ubuntu、Debian 环境下的操作步骤。

i. 执行如下命令，安装 gcc 编译器。

sudo apt-get install gcc

ii. 执行如下命令，安装 make 工具。

sudo apt-get install make

iii. 执行如下命令，安装内核模块开发包，开发包头文件与库的版本需要与内核版本一致。

sudo apt-get install linux-headers-`uname -r`

SUSE 环境下的操作步骤。

i. 执行如下命令，安装 gcc 编译器。

sudo zypper install gcc

ii. 执行如下命令，安装 make 工具。

```
sudo zypper install make
```

iii. 执行如下命令，安装内核模块开发包，开发包头文件与库的版本需要与内核版本一致。

```
sudo zypper install kernel-default-devel
```

CoreOS 环境下的操作步骤。

CoreOS 环境下在容器内进行内核模块的编译时，需要先启动一个用于内核模块开发的容器，然后再进行编译。

详细过程参见 CoreOS 官方文档，获取方式如下链接所示。

https://coreos.com/os/docs/latest/kernel-modules.html

2. 编译内核模块

a. 使用 git 工具，执行如下命令，下载 TOA 内核模块源码。

git clone https://github.com/huaweicloud/elb-toa.git

说明：

如果未安装 git 工具，请进入以下链接下载 TOA 模块源代码。

https://github.com/huaweicloud/elb-toa

b. 执行如下命令，进入源码目录，编译模块。

```
cd src
make
```

编译过程未提示 warning 或者 error，说明编译成功，检查当前目录下是否已经生成 toa.ko 文件。

3. 加载内核模块

a. 执行如下命令，加载内核模块。

```
sudo insmod toa.ko
```

b. 执行如下命令，验证模块加载情况，查看内核输出信息。

```
dmesg | grep TOA
```

若提示信息包含“TOA: toa loaded”，说明内核模块加载成功。

说明：

CoreOS 在容器中编译完内核模块后，需要将内核模块复制到宿主系统，然后在宿主系统中加载内核模块。由于编译内核模块的容器和宿主系统共享/lib/modules

目录，可以在容器中将内核模块复制到该目录下，以供宿主系统使用。

4. 自动加载内核模块

为了使 TOA 内核模块在系统启动时生效，可以将加载 TOA 内核模块的命令加到客户的启动脚本中。

自动加载内核模块的方法有以下两种方法：

- 客户可以根据自身需求，在自定义的启动脚本中添加加载 TOA 内核模块的命令。

- 参考以下操作步骤配置启动脚本。

  i. 在“/etc/sysconfig/modules/”目录下新建 toa.modules 文件。该文件包含了 TOA 内核模块的加载脚本。

  toa.modules 文件内容，请参考如下示例：

  ```
  #!/bin/sh
  /sbin/modinfo -F filename /root/toa/toa.ko > /dev/null 2>&1
  if [ $? -eq 0 ]; then
  /sbin/insmod /root/toa/toa.ko
  fi
  ```

  其中“/root/toa/toa.ko”为 TOA 内核模块文件的路径，客户需要将其替换为自己编译的 TOA 内核模块路径。

  ii. 执行以下命令，为 toa.modules 启动脚本添加可执行权限。

  ```
  sudo chmod +x /etc/sysconfig/modules/toa.modules
  ```

说明：

客户升级内核后，会导致现有 TOA 内核模块不匹配，因此需要重新编译 TOA 内核模块。

5. 安装多节点

如果要在相同的客户操作系统中加载此内核模块，可以将 toa.ko 文件拷贝到需要加载此模块的虚拟机中，然后参照 3 步骤加载内核模块。

内核模块加载成功以后，应用程序可以正常获取访问者的真实源 IP 地址。

说明：

节点的操作系统发行版与内核版本必须相同。

- **操作场景：**

以下操作步骤针对 Linux 内核版本为 2.6.32 的操作系统

说明：

TOA 插件支持 2.6.32-xx 内核版本的操作系统（CentOS 6.8 镜像）。参考如下步骤，进行配置。

1. 从以下网站中获取含有 TOA 模块的内核源代码包（Linux-2.6.32-220.23.1.el6.x86_64.rs.src.tar.gz）。

   http://kb.linuxvirtualserver.org/images/3/34/Linux-2.6.32-220.23.1.el6.x86_64.rs.src.tar.gz

2. 解压 TOA 模块的内核源码包。
3. 修改编译相关参数。

   a. 进入“linux-2.6.32-220.23.1.el6.x86_64.rs”文件夹。

   b. 编辑“net/toa/toa.h”文件。

   将#define TCPOPT_TOA200 配置项修改为#define TCPOPT_TOA254

   c. 在 shell 页面，执行以下命令。

   ```
   sed -i 's/CONFIG_IPV6=m/CONFIG_IPV6=y/g' .config
   echo -e '\n# toa\nCONFIG_TOA=m' >> .config
   ```

   配置之后 IPV6 模块将会被编译进内核中，TOA 会被编译成单独内核模块，可以单独启动和停止。

d. 编辑 Makefile。

可在“EXTRAVERSION =”等号后加上自定义的一些说明，将会在“uname -r”中显示，例如-toa。

4. 执行以下命令，编译软件包。

make -j n

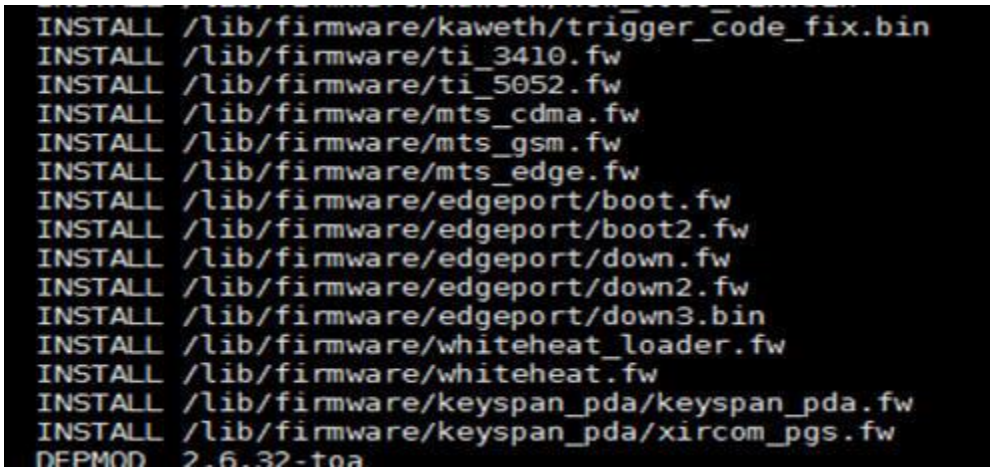
说明：

n 可以依据系统 CPU 核数配置相应的参数，例如：4 核 CPU，可配置为 4，从而加快编译速度。

5. 执行以下命令，安装内核模块。

make modules_install

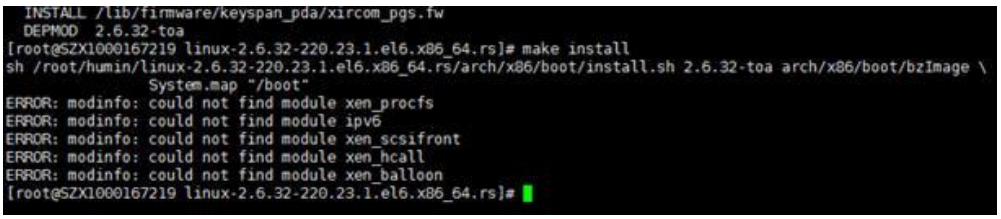
命令执行结果如下图所示。



6. 执行如下命令，安装内核。

make install

命令执行结果如下图所示。



7. 打开“/boot/grub/grub.conf”文件，配置开机默认启动，如下图所示。

a. 将开机默认启动内核由第一个内核修改为第零个内核，即“default=1”修改为“default=0”。

b. 在新增的含有 toa 模块的 vmlinuz-2.6.32-toa 内核行末尾添加“nohz=off”参数。如果不关闭 nohz，大压力下 CPU0 可能会消耗过高，导致压力不均匀

```
default=1
timeout=5
splashimage=(hd0,1)/boot/grub/splash.xpm.gz
hiddenmenu
title Red Hat Enterprise Linux Server (2.6.32-toa)
        root (hd0,1)
        kernel /boot/vmlinuz-2.6.32-toa ro root=UUID=
et nohz=off
        initrd /boot/initramfs-2.6.32-toa.img
```

c. 修改完成后保存退出，重启操作系统。

重启系统时，系统将加载 vmlinuz-2.6.32-toa 内核。

8. 待系统重启完成之后，执行以下命令加载 TOA 模块。

modprobe toa

建议将 modprobe toa 命令加入开机启动脚本，以及系统定时监控脚本中，如下图所示。

```
[root@SZX1000167219 ~]# modprobe toa
[root@SZX1000167219 ~]# lsmod |grep toa
toa                     4203  0
[root@SZX1000167219 ~]#
```

TOA 模块加载完成后，查询内核信息如下图所示。

```
[root@SZX1000167219 ~]# uname -a
Linux SZX1000167219 2.6.32-toa #1 SMP Sat Oct 15 11:50:05 CST 2016 x86_64 x86_64 x86_64 GNU/Linux
```