

ALIBABA CLOUD

阿里云

专有云企业版

负载均衡
用户指南

产品版本：v3.16.2

文档版本：20220915

 阿里云

法律声明

阿里云提醒您在阅读或使用本文档之前仔细阅读、充分理解本法律声明各条款的内容。如果您阅读或使用本文档，您的阅读或使用行为将被视为对本声明全部内容的认可。

1. 您应当通过阿里云网站或阿里云提供的其他授权通道下载、获取本文档，且仅能用于自身的合法合规的业务活动。本文档的内容视为阿里云的保密信息，您应当严格遵守保密义务；未经阿里云事先书面同意，您不得向任何第三方披露本手册内容或提供给任何第三方使用。
2. 未经阿里云事先书面许可，任何单位、公司或个人不得擅自摘抄、翻译、复制本文档内容的部分或全部，不得以任何方式或途径进行传播和宣传。
3. 由于产品版本升级、调整或其他原因，本文档内容有可能变更。阿里云保留在没有任何通知或者提示下对本文档的内容进行修改的权利，并在阿里云授权通道中不时发布更新后的用户文档。您应当实时关注用户文档的版本变更并通过阿里云授权渠道下载、获取最新版的用户文档。
4. 本文档仅作为用户使用阿里云产品及服务的参考性指引，阿里云以产品及服务的“现状”、“有缺陷”和“当前功能”的状态提供本文档。阿里云在现有技术的基础上尽最大努力提供相应的介绍及操作指引，但阿里云在此明确声明对本文档内容的准确性、完整性、适用性、可靠性等不作任何明示或暗示的保证。任何单位、公司或个人因为下载、使用或信赖本文档而发生任何差错或经济损失的，阿里云不承担任何法律责任。在任何情况下，阿里云均不对任何间接性、后果性、惩戒性、偶然性、特殊性或刑罚性的损害，包括用户使用或信赖本文档而遭受的利润损失，承担责任（即使阿里云已被告知该等损失的可能性）。
5. 阿里云网站上所有内容，包括但不限于著作、产品、图片、档案、资讯、资料、网站架构、网站画面的安排、网页设计，均由阿里云和/或其关联公司依法拥有其知识产权，包括但不限于商标权、专利权、著作权、商业秘密等。非经阿里云和/或其关联公司书面同意，任何人不得擅自使用、修改、复制、公开传播、改变、散布、发行或公开发表阿里云网站、产品程序或内容。此外，未经阿里云事先书面同意，任何人不得为了任何营销、广告、促销或其他目的使用、公布或复制阿里云的名称（包括但不限于单独为或以组合形式包含“阿里云”、“Aliyun”、“万网”等阿里云和/或其关联公司品牌，上述品牌的附属标志及图案或任何类似公司名称、商号、商标、产品或服务名称、域名、图案标示、标志、标识或通过特定描述使第三方能够识别阿里云和/或其关联公司）。
6. 如若发现本文档存在任何错误，请与阿里云取得直接联系。

通用约定

格式	说明	样例
 危险	该类警示信息将导致系统重大变更甚至故障，或者导致人身伤害等结果。	 危险 重置操作将丢失用户配置数据。
 警告	该类警示信息可能会导致系统重大变更甚至故障，或者导致人身伤害等结果。	 警告 重启操作将导致业务中断，恢复业务时间约十分钟。
 注意	用于警示信息、补充说明等，是用户必须了解的内容。	 注意 权重设置为0，该服务器不会再接受新请求。
 说明	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 说明 您也可以通过按Ctrl+A选中全部文件。
>	多级菜单递进。	单击设置> 网络> 设置网络类型。
粗体	表示按键、菜单、页面名称等UI元素。	在结果确认页面，单击确定。
Courier字体	命令或代码。	执行 <code>cd /d C:/window</code> 命令，进入Windows系统文件夹。
斜体	表示参数、变量。	<code>bae log list --instanceid</code> <code>Instance_ID</code>
[] 或者 [a b]	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
{ } 或者 {a b}	表示必选项，至多选择一个。	<code>switch {active stand}</code>

目录

1.什么是负载均衡	06
2.登录SLB控制台	08
3.快速入门	09
3.1. 入门概述	09
3.2. 准备工作	09
3.3. 创建负载均衡实例	11
3.4. 配置负载均衡实例	12
3.5. 释放负载均衡实例	14
4.负载均衡实例	15
4.1. 实例概述	15
4.2. 创建负载均衡实例	18
4.3. 启动和暂停实例	20
4.4. 释放实例	20
5.监听	21
5.1. 监听概述	21
5.2. 添加TCP监听	21
5.3. 添加UDP监听	23
5.4. 添加HTTP监听	25
5.5. 添加HTTPS监听	28
5.6. 配置转发策略	31
5.7. 开启访问控制	32
5.8. 关闭访问控制	33
6.后端服务器	34
6.1. 后端服务器概述	34
6.2. 默认服务器组	35
6.2.1. 添加默认服务器	35

6.2.2. 添加IDC作为默认服务器	35
6.2.3. 编辑后端服务器的权重	36
6.2.4. 移除后端服务器	36
6.3. 虚拟服务器组	36
6.3.1. 添加ECS实例作为虚拟服务器	37
6.3.2. 添加IDC作为虚拟服务器	37
6.3.3. 编辑虚拟服务器组	38
6.3.4. 删除虚拟服务器组	39
6.4. 主备服务器	39
6.4.1. 添加ECS实例作为主备服务器	39
6.4.2. 添加IDC服务器作为主备服务器	39
6.4.3. 删除主备服务器组	40
6.5. 通过弹性网卡添加后端服务器	40
7.健康检查	42
7.1. 健康检查概述	42
7.2. 配置健康检查	49
7.3. 关闭健康检查	50
8.证书管理	52
8.1. 证书概述	52
8.2. 证书要求	52
8.3. 上传证书	53
8.4. 生成CA证书	54
8.5. 转换证书格式	58
8.6. 替换证书	58

1.什么是负载均衡

负载均衡（Server Load Balancer）是将访问流量根据转发策略分发到后端多台云服务器（ECS实例）的流量分发控制服务。负载均衡扩展了应用的服务能力，增强了应用的可用性。

概述

负载均衡通过设置虚拟服务地址，将添加的同一地域的多台ECS实例虚拟成一个高性能和高可用的后端服务池，并根据转发规则，将来自客户端的请求分发给后端服务器池中的ECS实例。

负载均衡默认检查云服务器池中的ECS实例的健康状态，自动隔离异常状态的ECS实例，消除了单台ECS实例的单点故障，提高了应用的整体服务能力。

组成部分

负载均衡由以下三个部分组成：

- 负载均衡实例（Server Load Balancer instances）

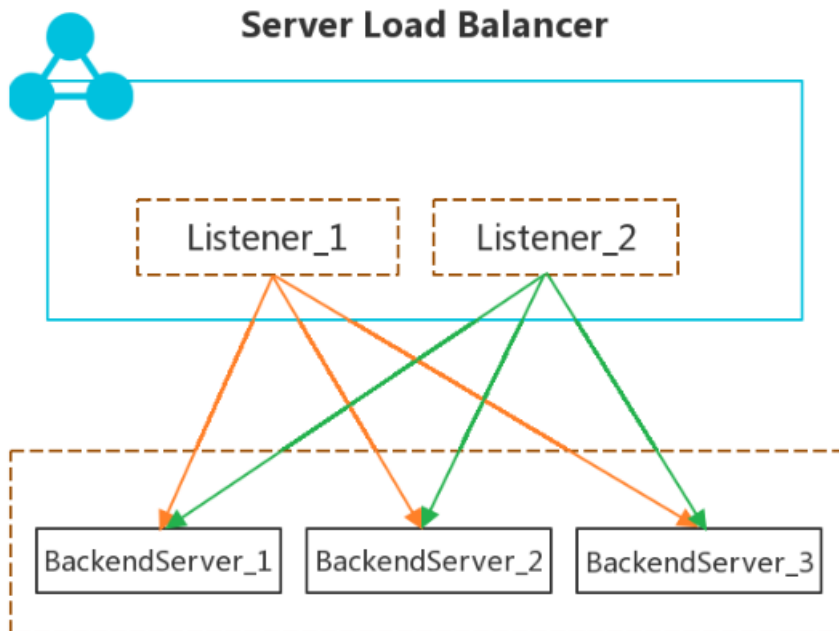
一个负载均衡实例是一个运行的负载均衡服务，用来接收流量并将其分配给后端服务器。要使用负载均衡服务，您必须创建一个负载均衡实例，并至少添加一个监听和两台ECS实例。

- 监听（Listeners）

监听用来检查客户端请求并将请求转发给后端服务器。监听也会对后端服务器进行健康检查。

- 后端服务器（Backend Servers）

一组接收前端请求的ECS实例。您可以单独添加ECS实例到后端服务器池，也可以通过虚拟服务器组或主备服务器组来批量添加和管理。



产品优势

- 高可用

采用全冗余设计，无单点，支持同城容灾。搭配DNS可实现跨地域容灾，可用性高达99.95%。

根据应用负载进行弹性扩容，在流量波动情况下不中断对外服务。

- 可扩展

您可以根据业务的需要，随时增加或减少后端服务器的数量，扩展应用的服务能力。

- 低成本

与传统硬件负载均衡系统高投入相比，成本可下降60%。

- 安全

结合云盾，可提供5Gbps的防DDoS攻击能力。

- 高并发

集群支持亿级并发连接，单实例提供千万级并发能力。

2. 登录SLB控制台

本节以Chrome浏览器为例，介绍负载均衡用户如何登录到Apsara Uni-manager运营控制台。


前提条件

- 登录Apsara Uni-manager运营控制台前，确认您已从部署人员处获取Apsara Uni-manager运营控制台的服务域名地址。
- 推荐使用Chrome浏览器。

操作步骤

1. 在浏览器地址栏中，输入Apsara Uni-manager运营控制台的访问地址，按回车键。
2. 输入正确的用户名及密码。

请向运营管理员获取登录控制台的用户名和密码。

 **说明** 首次登录Apsara Uni-manager运营控制台时，需要修改登录用户名的密码，请按照提示完成密码修改。为提高安全性，密码长度必须为 8~20 位，且至少包含以下两种类型：


- 英文大写或小写字母（A~Z、a~z）
- 阿拉伯数字（0~9）
- 特殊符号（感叹号（！）、at（@）、井号（#）、美元符号（\$）、百分号（%）等）

3. 单击登录。
4. 在页面顶部的菜单栏中，单击产品 > 网络 > 负载均衡 SLB。

3.快速入门

3.1. 入门概述

本教程指引您快速创建一个公网负载均衡实例，将来自客户端的请求转发到两台后端ECS上。

 **说明** 在开始搭建负载均衡服务前，您需要确定负载均衡实例的地域、类型、付费模式等配置，更多信息，请参见[准备工作](#)。

本教程包含以下操作：

1. **创建负载均衡实例**

创建负载均衡实例。负载均衡实例是一个运行的负载均衡服务实体。

2. **配置负载均衡实例**

配置负载均衡实例，添加监听规则和后端服务器。

3. **释放负载均衡实例**

如果您不需要负载均衡服务了，为避免不必要的计费，可以将其删除。

3.2. 准备工作

在使用负载均衡前，您需要根据您的业务确定负载均衡的监听类型和网络类型等。

规划实例地域

在选择地域时，请注意：

- 为了减少延迟并提高下载速度，建议选择离您最近的地域。
- 为了提供更加稳定可靠的服务，阿里云负载均衡已在大部分地域提供主备可用区，实现同地域下的跨机房容灾。建议您选择提供主备可用区的地域。
- 由于负载均衡不支持跨地域部署，因此应选择与后端ECS实例相同的地域。

选择实例类型（公网或内网）

负载均衡提供面向公网和内网的负载均衡服务：

- 如果您需要使用负载均衡分发来自公网的请求，选择创建公网负载均衡实例。
公网负载均衡实例提供一个公网IP，用来接收来自Internet的请求。
- 如果您需要使用负载均衡分发来自内网的请求，选择创建内网负载均衡实例。
内网负载均衡实例仅提供阿里云私网IP，只能通过阿里云内部网络访问该负载均衡服务，无法从Internet访问。

选择实例规格

负载均衡性能共享型实例，资源是所有实例共享的，不保障实例的性能指标。

负载均衡还推出了性能保障型实例，您可以独享已购实例的资源，更好地保障服务的可用性。负载均衡提供6种实例规格供您选择：

规格	最大连接数	每秒新建连接数（CPS）	每秒查询数（QPS）
简约型I（slb.s1.small）	5,000	3,000	1,000
标准型I（slb.s2.small）	50,000	5,000	5,000
标准型II（slb.s2.medium）	100,000	10,000	10,000
高阶型I（slb.s3.small）	200,000	20,000	20,000
高阶型II（slb.s3.medium）	500,000	50,000	30,000
超强型I（slb.s3.large）	1,000,000	100,000	50,000

选择协议类型

阿里云提供基于四层协议（TCP和UDP）和七层协议（HTTP和HTTPS）的负载均衡：

- 四层监听将请求直接转发给后端服务器，不会修改报头。客户端请求到达负载均衡监听后，负载均衡服务器会使用监听中配置的后端端口与后端服务器建立TCP连接。
- 七层监听原理上是反向代理的一种实现。客户端请求到达负载均衡监听后，负载均衡服务器会通过与后端服务器建立TCP连接，即再次通过新TCP连接HTTP协议访问后端，而不是直接转发报文到后端ECS。

由于七层监听比四层监听在底层实现上多了一个Tengine处理环节，因此，七层监听性能没有四层好。此外，客户端端口不足、后端服务器连接过多等场景也可能导致七层服务性能不高，如果您对性能有很高的要求，建议您使用四层监听。

准备后端服务器

在使用负载均衡服务前，您需要创建ECS实例并部署相关应用，然后将ECS实例添加到负载均衡实例中来处理转发的客户端请求。

创建ECS时，请注意：

- ECS实例的地域和可用区

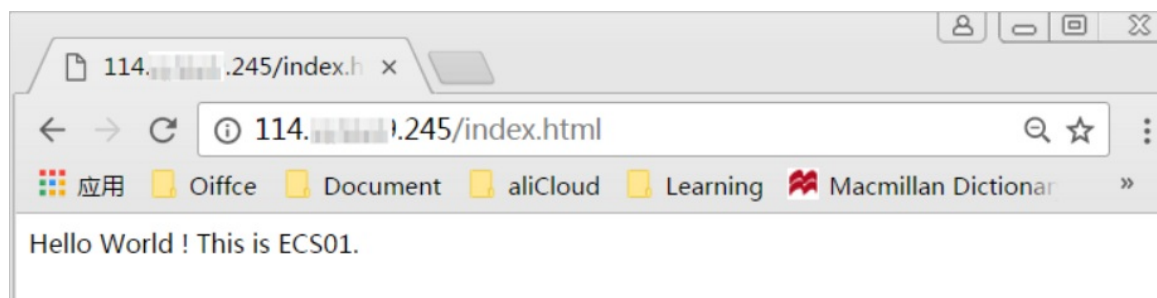
确保ECS实例的地域和负载均衡实例的地域相同。

本文在华东1（杭州）地域创建了两个ECS实例，为了便于辨识，将实例分别命名为ECS01和ECS02。

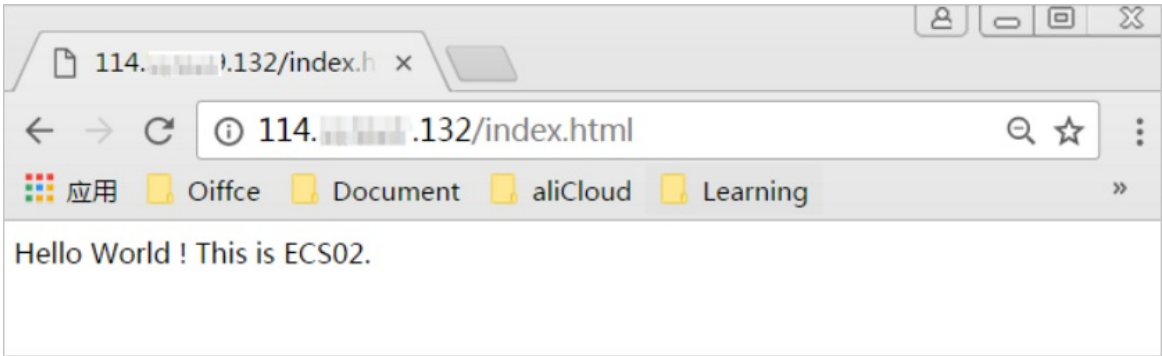
- 应用配置

分别在ECS01和ECS02两个实例上使用Apache搭建了两个静态网页，如下图所示。

- 在浏览器中输入ECS01实例绑定的弹性公网IP地址：



- 在浏览器中输入ECS02实例绑定的弹性公网IP地址：



在ECS上部署好应用后，不需要再进行特别的配置。但如果您要配置一个四层监听（TCP协议或UDP协议），并且ECS使用的是Linux系统，确保ECS实例上`/etc/sysctl.conf`目录下`net.ipv4.conf`文件中的以下三个参数的值为零：

```
net.ipv4.conf.default.rp_filter = 0
net.ipv4.conf.all.rp_filter = 0
net.ipv4.conf.eth0.rp_filter = 0
```

3.3. 创建负载均衡实例

负载均衡实例是运行的负载均衡服务实体。一个负载均衡实例可以添加多个监听和后端服务器。

前提条件

- 您需要创建好ECS实例，并搭建应用服务。
- 确保ECS实例的部门和负载均衡实例的组织相同，并且ECS实例的安全组允许端口80/443上的HTTP/HTTPS访问。

操作步骤

- 登录SLB控制台。
- 在左侧导航栏，选择实例 > 实例管理。
- 在实例管理页面，单击创建负载均衡。
- 配置负载均衡实例，然后单击提交。

配置	说明
区域	
组织	从下拉列表中选择负载均衡实例的所属组织。 <div>? 说明 确保负载均衡实例的组织 and 后端服务器ECS实例的组织相同。</div>
资源集	输入负载均衡实例所在的资源集。
地域	输入负载均衡实例所在的地域。

配置	说明
可用区	从下拉列表中选择负载均衡实例所在的可用区。
集群	从下拉列表中选择负载均衡实例所在的集群。
基本设置	
创建数量	选择创建负载均衡实例的数量。
负载均衡名称	输入实例名称。 如果创建实例数量大于1，实例名称将有系统自动设置。
规格类型	性能保障性 ：独享已购实例的资源。不同规格的性能保障型实例，性能指标不同。
规格	性能保障型 实例规格。更多信息，请参见 实例概述 。
网络与示例类型	
实例类型	选择实例类型，分为 内网 和 公网 ，本示例选择内网。
网络类型	选择网络类型，分为 专有网络 和 经典网络 ，本示例选择专有网络。
IP版本	选择IP版本。
专有网络vpc	选择一个专有网络。
交换机vswitch	选择一台交换机。
服务IP	输入服务IP，请保证服务IP的有效性，否则无法成功创建负载均衡实例。 如果不设置，则使用系统自动分配的IP。
计费方式	计费方式按实例类型分类： <ul style="list-style-type: none">内网：计费方式为默认计费方式。公网：计费方式为默认和按带宽计费方式。
带宽	带宽峰值，单位Mbps。 取值范围：1~5120。 <div><div>?</div><div>说明</div><div>该参数仅在实例类型为公网，且计费方式为按带宽计费方式时有效。</div></div>

后续步骤

配置负载均衡实例


3.4. 配置负载均衡实例

创建负载均衡实例后，您需要对负载均衡实例进行配置才能进行流量转发，您需要添加至少一个监听和一组后端服务器。本文指引您配置一个TCP监听并添加部署了静态网页的两个ECS实例（ECS01和ECS02）作为后端服务器。

操作步骤

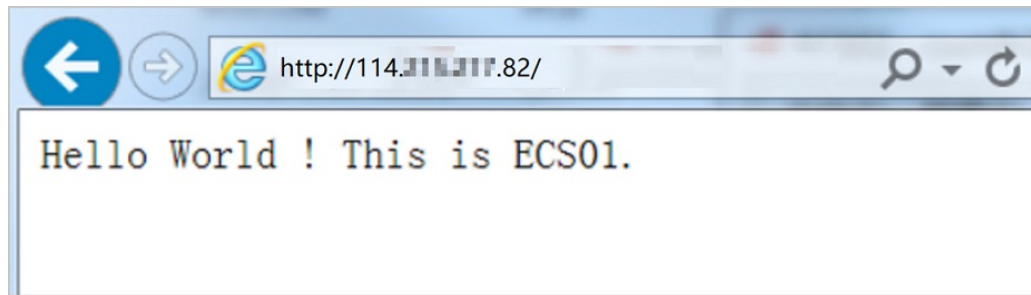
1. [登录SLB控制台](#)。
 2. 在实例管理页面，找到目标实例，在操作列单击**监听配置向导**。
 3. 在**协议&监听**配置向导，根据以下信息，配置监听规则，其它配置保持默认选项。然后单击**下一步**。
 - **选择负载均衡协议**：本文选择**TCP**协议。
 - **监听端口**：用来接收请求并向后端服务器进行请求转发的负载均衡系统的前端协议和端口。本文端口设置为80。

负载均衡对外提供服务的端口，通常HTTP协议使用80端口，HTTPS协议使用443端口。单击**高级配置**的**修改**，完成以下配置：
 - **开启监听带宽限速**：设定不同的带宽峰值来限定后端ECS实例的不同应用所能对外提供的服务能力。
 - **调度算法**：负载均衡支持如下三种调度算法，本文选择**轮询**。
 - **加权轮询（WRR）**：将访问请求依序分发后端服务器，后端服务器的权重越高，被分发的几率也越大。
 - **轮询（RR）**：按照访问顺序将访问请求分发给后端服务器。
 - **一致性哈希（CH）**：仅性能保障型实例支持一致性哈希（CH）调度算法。
 - **源IP**：基于源IP地址的一致性hash，相同的源地址会调度到相同的后端服务器。
 - **四元组**：基于四元组的一致性hash（源IP、目的IP、源端口和目的端口），相同的流会调度到相同的后端服务器。
 - **加权最小连接数（WLC）**：除了根据每台后端服务器设定的权重值来进行轮询，同时还考虑后端服务器的实际负载（即连接数）。当权重值相同时，当前连接数越小的后端服务器被轮询到的次数（概率）也越高。
 4. 在**后端服务器**配置向导，选择**默认服务器组**，单击**继续添加**，添加后端服务器。
 - i. 在**我的服务器**面板，选择之前创建的ECS01和ECS02实例，单击**下一步**。
 - ii. 配置**权重**，权重越大转发的请求越多，默认为100，保持默认值即可。
 - iii. 单击**添加**。
 - iv. 在**默认服务器**页签下，配置**后端协议端口**，ECS实例上开放的用来接收请求的后端端口，在同一个负载均衡实例内可重复。本文端口设置为80。
 5. 单击**下一步**，配置**健康检查**，本文使用默认值。

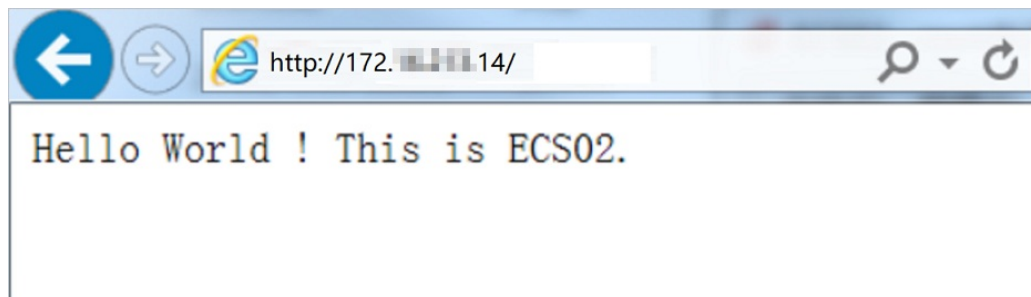
开启健康检查功能后，当后端某个ECS健康检查出现问题时，负载均衡服务会将请求转发到其它健康检查正常的ECS上，而当该ECS恢复正常运行时，负载均衡会自动恢复它的请求转发。
 6. 单击**下一步**，进入**配置审核**配置向导，单击**提交**。
 7. 单击**知道了**，返回**实例管理**页面，单击。
- 当后端ECS的健康检查状态为**运行中**时，表示后端ECS可以正常处理负载均衡转发的请求了。

8. 在浏览器中输入负载均衡实例的服务地址，测试负载均衡服务。

ECS01



ECS02



3.5. 释放负载均衡实例

您可以根据需要删除负载均衡实例，避免不必要的计费。删除负载均衡实例不会删除后端ECS，也不会影响后端ECS的运行。

操作步骤

1. [登录SLB控制台](#)。
2. 在**实例管理**页面，选择目标实例，在操作列选择... > **释放设置**，或者勾选目标实例，单击列表底部的**释放设置**。
3. 在**释放设置**页面，单击**下一步**。

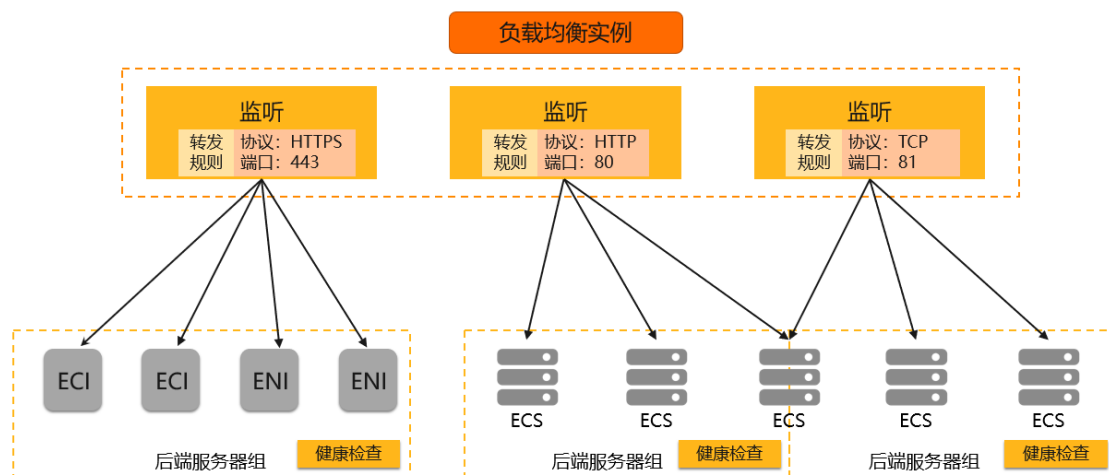
 **说明** 系统执行释放时间是每整点和每半点，但是系统会按照您设置的释放时间停止计费。

4. 单击**确定**，完成负载均衡实例的释放。

4. 负载均衡实例

4.1. 实例概述

负载均衡SLB（Server Load Balancing）实例接收来自客户端的请求，并将请求分发给后端服务器。使用负载均衡服务，您需要创建一个负载均衡实例，在实例中添加监听和后端服务器。



实例类型

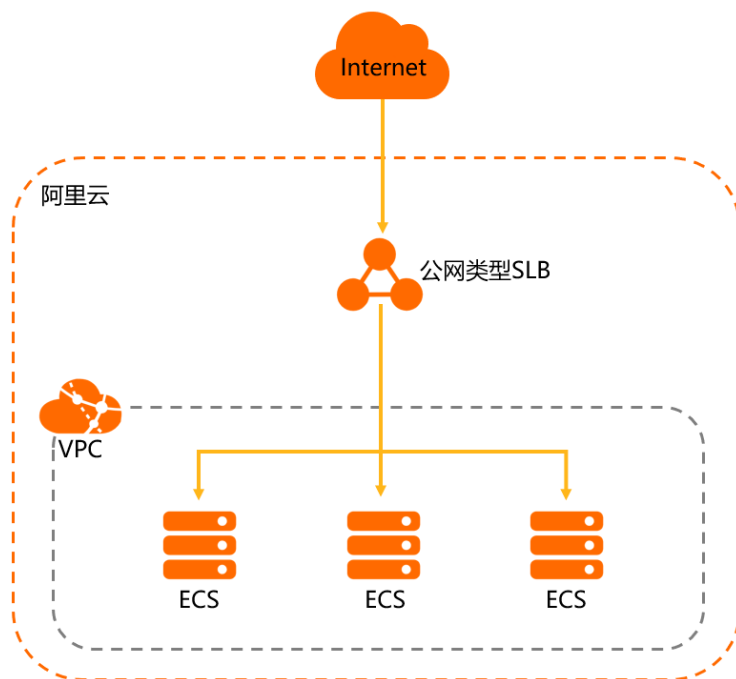
阿里云提供公网和私网两种类型的SLB实例。

公网类型的SLB

在创建公网类型的SLB实例时，系统会为其分配一个公网IP。您可以将您的域名和该公网IP进行绑定，SLB实例通过互联网接受客户端的访问请求，按照监听规则将客户端的访问请求分发给后端服务器。

公网类型SLB提供的公网能力具有以下特点：

- SLB会被分配一个公网IP地址，该IP地址与SLB强绑定，不能解绑。
- 计费类型为默认计费方式和按带宽计费方式。

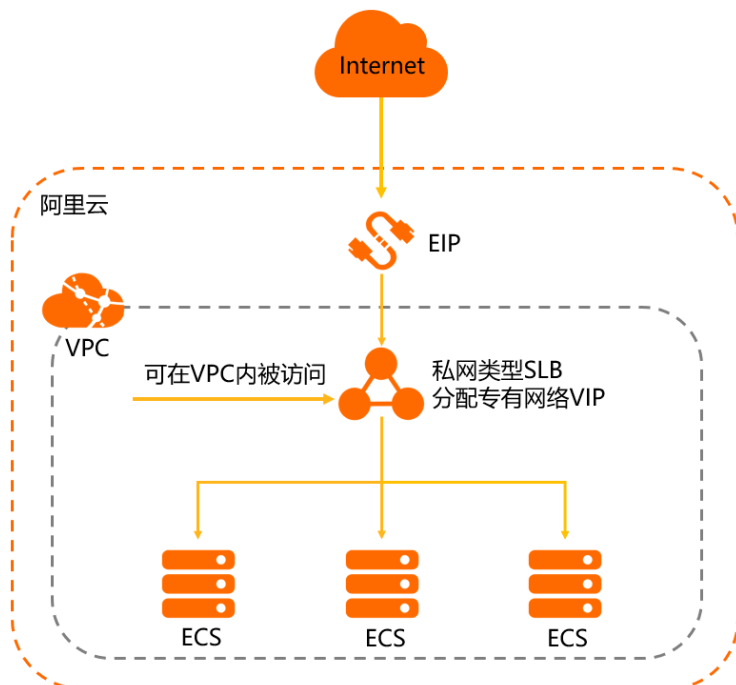


私网类型的SLB

私网类型的SLB通过私网IP对外提供服务，将来自阿里云内部网络的访问请求按照监听规则分发至后端服务器。

私网类型的SLB实例可配合EIP提供公网负载均衡能力。私网类型SLB提供的公网能力具有以下特点：

- SLB可以与EIP绑定来提供公网访问能力，EIP和SLB可以灵活解绑。
- EIP加入共享带宽，除包年包月和按量计费外，还可以使用95带宽峰值计费。



不同计费类型的私网类型的SLB支持的网络能力不同：


- 包年包月的实例支持专有网络和经典网络：

- 专有网络

如果私网SLB实例的网络类型是专有网络，那么私网SLB实例的私网IP地址会从您指定的专有网络的交换机网段内分配。该SLB实例只能被同VPC内的ECS实例访问。

- 经典网络

如果私网SLB实例的网络类型是经典网络，那么私网SLB实例的私网IP地址由阿里云统一分配和管理。该SLB实例只能被经典网络ECS实例访问。

 **注意** 经典网络类型的私网SLB实例已不支持新购。

- 按量付费的私网类型的SLB实例仅支持专有网络。

实例规格

性能保障型实例

性能保障型实例的三个关键指标如下：

- 最大连接数-Max Connection

最大连接数定义了一个负载均衡实例能够承载的最大连接数量。当实例上的连接超过规格定义的最大连接数时，新建连接请求将被丢弃。

- 每秒新建连接数-Connection Per Second（CPS）

每秒新建连接数定义了新建连接的速率。当新建连接的速率超过规格定义的每秒新建连接数时，新建连接请求将被丢弃。

- 每秒查询数-Query Per Second（QPS）


每秒请求数是七层监听特有的概念，指的是每秒可以完成的HTTP或HTTPS的查询（请求）的数量。当请求速率超过规格所定义的每秒查询数时，新建连接请求将被丢弃。

性能保障型实例规格

规格	最大连接数	每秒新建连接数（CPS）	每秒查询数（QPS）
简约型I（slb.s1.small）	5,000	3,000	1,000
标准型I（slb.s2.small）	50,000	5,000	5,000
标准型II（slb.s2.medium）	100,000	10,000	10,000
高阶型I（slb.s3.small）	200,000	20,000	20,000

性能共享型实例

资源在所有性能共享型实例中共享，实例的性能指标无法保障。

 **注意** 性能共享型实例已不支持新购。

性能保障型实例与性能共享型实例区别

特性	性能保障型实例	性能共享型实例
资源分配	资源独享	资源共享
可用性SLA	99.95%	不提供
IPv6	√	-
支持SNI多证书	√	-
支持黑白名单	√	-
支持挂载ENI	√	-
添加ECS弹性网卡辅助IP	√	-
HTTP重定向HTTPS	√	-
一致性HASH	√	-
TLS安全策略	√	-
HTTP2	√	-
WS（WebSocket）或WSS（Web Socket Secure）	√	-

② 说明 上表中，“√”表示支持，“-”表示不支持。

4.2. 创建负载均衡实例

负载均衡实例是运行的负载均衡服务实体。一个负载均衡实例可以添加多个监听和后端服务器。

前提条件


- 您需要创建好云服务器ECS（Elastic Compute Service）实例，并搭建应用服务。
- 确保ECS实例的组织和负载均衡实例的组织相同，并且ECS安全组允许后端服务被访问。

操作步骤

1. [登录SLB控制台](#)。
2. 在左侧导航栏，选择实例 > 实例管理。
3. 在实例管理页面，单击创建负载均衡，完成以下配置，单击提交。

配置	说明
区域	

配置	说明
组织	<p>从下拉列表中选择负载均衡实例的所属组织。</p> <p> 说明 确保负载均衡实例的组织和后端服务器ECS实例的组织相同。</p>
资源集	输入负载均衡实例所在的资源集。
地域	输入负载均衡实例所在的地域。
可用区	从下拉列表中选择负载均衡实例所在的可用区。
集群	从下拉列表中选择负载均衡实例所在的集群。
基本设置	
创建数量	选择创建负载均衡实例的数量。
负载均衡名称	<p>输入实例名称。</p> <p>长度限制为2~128个字符，以字母或中文开头，支持数字、全角字符、短划线（-）、半角冒号（:）、半角句号（.）和下划线（_）字符，支持换行和空格，开头不能为 <code>http://</code> 或 <code>https://</code>。</p>
规格类型	选择负载均衡实例规格类型为性能保障型。
规格	选择一个性能规格。
网络与示例类型	
实例类型	选择实例类型，分为内网和公网。
网络类型	选择网络类型，分为经典网络和专有网络。
专有网络vpc	是独有的云上虚拟网络。当网络类型为专有网络时，此参数必选。
交换机vswitch	是组成专有网络的基础网络设备。当网络类型为专有网络时，此参数必选。
IP版本	选择IP版本。
服务IP	<p>输入服务IP，请保证服务IP的有效性，否则无法成功创建负载均衡实例。如果不设置，则使用系统自动分配的IP。</p> <p> 说明 您指定负载均衡实例的私网IP地址时，该地址必须包含在交换机的目标网段下。</p>

配置	说明
计费方式	计费方式按实例类型分类： <ul style="list-style-type: none">内网：计费方式为默认计费方式。公网：计费方式为默认计费方式和按带宽计费方式。
带宽	带宽峰值，单位Mbps。 取值范围：1~5120。 <div> 说明 该参数仅在实例类型为公网，且计费方式为按带宽计费方式时有效。</div>

4.3. 启动和暂停实例

您可以随时启动或暂停负载均衡实例。实例暂停后不再接收和转发客户端流量。

操作步骤

1. [登录SLB控制台](#)。
2. 在左侧导航栏，选择**实例 > 实例管理**。
3. 在**实例管理**页面，找到目标实例，在操作列选择... > **启动**或**停止**。
4. 如果您想批量启动或停止多个实例，选择实例后，在页面下方单击**启动**或**停止**。

4.4. 释放实例

您可以根据需求设置立即释放实例。

操作步骤

1. [登录SLB控制台](#)。
2. 在左侧导航栏，选择**实例 > 实例管理**。
3. 找到目标实例，在操作列选择... > **释放设置**。
4. 在**释放设置**面板，单击**下一步**。
5. 单击**确定**，输入验证信息，确认释放实例。

5. 监听

5.1. 监听概述

创建负载均衡实例后，您需要为实例配置监听。负载均衡实例监听负责检查连接请求，然后根据调度算法定义的转发策略将请求流量分发至后端服务器。

负载均衡提供四层（TCP或UDP协议）和七层（HTTP或HTTPS协议）监听，您可根据应用场景选择监听协议：

协议	说明	使用场景
TCP	<ul style="list-style-type: none">面向连接的协议，在正式收发数据前，必须和对方建立可靠的连接。基于源地址的会话保持。在网络层可直接看到来源地址。数据传输快。	<ul style="list-style-type: none">适用于注重可靠性，对数据准确性要求高，速度可以相对较慢的场景，如文件传输、发送或接收邮件、远程登录。无特殊要求的Web应用。
UDP	<ul style="list-style-type: none">面向非连接的协议，在数据发送前不与对方进行三次握手，直接进行数据包发送，不提供差错恢复和数据重传。可靠性相对低；数据传输快。	关注实时性而相对不注重可靠性的场景，如视频聊天、金融实时行情推送。
HTTP	<ul style="list-style-type: none">应用层协议，主要解决如何包装数据。基于Cookie的会话保持。使用X-Forward-For获取客户真实IP地址。	需要对数据内容进行识别的应用，如Web应用、小的手机游戏等。
HTTPS	<ul style="list-style-type: none">加密传输数据，可以阻止未经授权的访问。统一的证书管理服务，您可以将证书上传到负载均衡，解密操作直接在负载均衡上完成。	需要加密传输的应用。

5.2. 添加TCP监听

TCP协议适用于注重可靠性、对数据准确性要求高和速度可以相对较慢的场景，如文件传输、发送或接收邮件和远程登录等。您可以添加一个TCP监听转发来自TCP协议的请求。

步骤一：打开监听配置向导

完成以下操作，打开监听配置向导。

1. [登录SLB控制台](#)。
2. 在左侧导航栏，选择**实例 > 实例管理**。
3. 选择以下一种方法，打开监听配置向导。
 - 在**实例管理**页面，找到目标实例，单击**操作**下的**监听配置向导**。

- 在实例管理页面，单击目标实例ID。在监听页面，单击添加监听。

步骤二：配置协议监听

完成以下操作，配置协议监听。

- 完成以下操作，配置协议监听，然后单击下一步。

监听配置	说明
选择负载均衡协议	选择监听的协议类型。 本文，选择TCP。
监听端口	用来接收请求并向后端服务器进行请求转发的监听端口。 端口范围为1~65535。
高级配置	
调度算法	<p>负载均衡支持以下调度算法。</p> <ul style="list-style-type: none"> 加权轮询（WRR）：权重值越高的后端服务器，被轮询到的次数（概率）也越高。 轮询（RR）：按照访问顺序依次将外部请求分发到后端服务器。 一致性哈希（CH）： <ul style="list-style-type: none"> 源IP：基于源IP地址的一致性hash，相同的源地址会调度到相同的后端服务器。 四元组：基于四元组的一致性hash（源IP、目的IP、源端口和目的端口），相同的流会调度到相同的后端服务器。 <div>  说明 仅性能保障型实例支持一致性哈希（CH）调度算法。 </div> <ul style="list-style-type: none"> 加权最小连接数（WLC）：除了根据每台后端服务器设定的权重值来进行轮询，同时还考虑后端服务器的实际负载（即连接数）。当权重值相同时，当前连接数越小的后端服务器被轮询到的次数（概率）也越高。
开启会话保持	<p>是否开启会话保持。</p> <p>开启会话保持后，负载均衡监听会把来自同一客户端的访问请求分发到同一台后端服务器上。</p> <p>TCP协议是基于IP地址的会话保持，即来自同一IP地址的访问请求转发到同一台后端服务器上。</p>
开启监听带宽限速	<p>选择是否配置监听带宽。</p> <p>对于按带宽计费的负载均衡实例，您可以针对不同监听设定不同的带宽峰值来限定监听的流量。实例下所有监听的带宽峰值总和不能超过该实例的带宽。</p> <p>默认不开启，各监听共享实例的总带宽。</p>

监听配置	说明
连接超时时间	指定TCP连接的超时时间，范围10~900秒。
获取客户端真实IP	针对四层监听，后端服务器可直接获得来访者的真实IP，无需采用其它手段获取。
创建完毕自动启动监听	是否在监听配置完成后启动负载均衡监听，默认开启。

步骤三：添加后端服务器

添加处理前端请求的后端服务器。您可以使用实例配置的默认服务器组，也可以为监听配置一个虚拟服务器组或主备服务器组。

1. 在**后端服务器配置向导**，选择**默认服务器组**，单击**继续添加**。
2. 在**我的服务器面板**，选择要添加的ECS实例，然后单击**下一步**。
3. 在**配置端口和权重配置向导**，配置添加的后端服务器的权重，权重越高的ECS实例将被分配到更多的访问请求。

 **说明** 权重设置为0，该服务器不会再接受新请求。

4. 单击**添加**，配置后端服务器（ECS实例）开放用来接收请求的端口，端口范围为1~65535。
同一个负载均衡实例内，后端服务器端口可以相同。
5. 单击**下一步**。

步骤四：配置健康检查

负载均衡通过健康检查来判断后端服务器（ECS实例）的业务可用性。健康检查机制提高了前端业务整体可用性，避免了后端ECS异常对总体服务的影响。单击**修改**更改健康检查配置，详情请参见[配置健康检查](#)。

步骤五：提交配置

完成以下操作，确认监听配置。

1. 在**配置审核配置向导**，检查监听配置，您可以单击**修改**更改配置。
2. 确认无误后，单击**提交**。
3. 在**负载均衡业务配置向导对话框**，配置成功后，单击**知道了**。

配置成功后，您可以在监听页面查看已创建的监听。

5.3. 添加UDP监听

UDP协议多用于关注实时性而对可靠性要求相对较低的场景，如视频聊天和金融实时行情推送等。您可以添加一个UDP监听转发来自UDP协议的请求。

背景信息

在添加UDP监听前，注意如下限制：

- UDP监听的250、4789和4790三个端口为系统保留端口，暂时不对外开放。
- 暂不支持分片包。
- 经典网络负载均衡实例的UDP监听暂不支持查看源地址。

- 在以下两种情况下，UDP协议监听配置需要五分钟才能生效：
 - 移除后端服务器。
 - 健康检查检测到异常后，将后端服务器的权重设置为0。

步骤一：打开监听配置向导

完成以下操作，打开监听配置向导。

1. [登录SLB控制台](#)。
2. 在左侧导航栏，选择**实例 > 实例管理**。
3. 选择以下一种方法，打开监听配置向导。
 - 在**实例管理**页面，找到目标实例，单击**操作**下的**监听配置向导**。
 - 在**实例管理**页面，单击目标实例ID。在**监听**页面，单击**添加监听**。

步骤二：配置协议监听

完成以下操作，配置协议监听：

1. 在**协议&监听**配置向导，根据以下信息配置监听，然后单击**下一步**。

监听配置	说明
选择负载均衡协议	选择监听的协议类型。 本操作选择UDP。
监听端口	用来接收请求并向后端服务器进行请求转发的监听端口。 端口范围为1-65535。
高级配置	
调度算法	负载均衡支持以下三种调度算法。 <ul style="list-style-type: none">○ 加权轮询（WRR）：权重值越高的后端服务器，被轮询到的次数（概率）也越高。○ 轮询（RR）：按照访问顺序依次将外部请求分发到后端服务器。○ 一致性哈希（CH）：<ul style="list-style-type: none">■ 源IP：基于源IP地址的一致性hash，相同的源地址会调度到相同的后端服务器。■ 四元组：基于四元组的一致性hash（源IP+目的IP+源端口+目的端口），相同的流会调度到相同的后端服务器。
开启会话保持	是否开启会话保持。 UDP会话保持基于源IP的一致性HASH算法实现。

监听配置	说明
开启监听宽度限速	<p>选择是否配置监听带宽。</p> <p>对于按带宽计费的负载均衡实例，您可以针对不同监听设定不同的带宽峰值来限定监听的流量。实例下所有监听的带宽峰值总和不能超过该实例的带宽。</p> <p>默认不开启，各监听共享实例的总带宽。</p>
获取客户端真实IP	<p>UDP协议监听的后端服务器可直接获取客户端的真实IP。</p> <div><p> 说明 经典网络实例的UDP协议暂不支持查看源地址。</p></div>
创建完毕自动启动监听	是否在监听配置完成后启动负载均衡监听，默认开启。

步骤三：添加后端服务器

添加处理前端请求的后端服务器。您可以使用实例配置的默认服务器组，也可以为监听配置一个虚拟服务器组或主备服务器组。

本操作中，以默认后端服务器组为例：

- 1. 在**后端服务器配置向导**，选择**默认服务器组**，单击**继续添加**。
- 2. 在**我的服务器面板**，选择要添加的ECS实例，然后单击**下一步**。
- 3. 配置添加的后端服务器的权重。

权重越高的ECS实例将被分配到更多的访问请求。

 **说明** 权重设置为0，该服务器不会再接受新请求。

- 4. 单击**添加**，配置后端服务器（ECS实例）开放用来接收请求的端口，端口范围为1~65535。
同一个负载均衡实例内，后端服务器端口可以相同。
- 5. 单击**下一步**。

步骤三：配置健康检查

负载均衡通过健康检查来判断后端服务器（ECS实例）的业务可用性。健康检查机制提高了前端业务整体可用性，避免了后端ECS异常对总体服务的影响。

在**健康检查配置向导**，单击**修改**更改健康检查配置。具体操作，请参见[配置健康检查](#)。

步骤四：提交配置

- 1. 在**配置审核配置向导**，检查监听配置，您可以单击**修改**更改配置。
- 2. 确认无误后，单击**提交**。
- 3. 等待配置成功后，单击**知道了**。

配置成功后，您可以在监听页面查看已创建的监听。

5.4. 添加HTTP监听

HTTP协议适用于需要对数据内容进行识别的应用，如Web应用和手机小游戏等。您可以添加一个HTTP监听转发来自HTTP协议的请求。

步骤一：配置监听

- 1. 登录SLB控制台。
- 2. 选择以下一种方法，打开监听配置向导。
 - 在实例管理页面，单击目标实例操作列的监听配置向导。
 - 在实例管理页面，单击目标实例ID，然后在监听页签单击添加监听。
- 3. 配置协议监听。

监听配置	说明
选择负载均衡协议	选择监听的协议类型。 本操作，选择HTTP。
监听端口	用来接收请求并向后端服务器进行请求转发的监听端口。端口范围为1~65535。
高级配置	单击修改展开高级配置。
调度算法	选择调度算法。 <ul style="list-style-type: none">◦ 加权轮询(WRR)：权重值越高的后端服务器，被轮询到的次数（概率）也越高。◦ 轮询(RR)：按照访问顺序依次将外部请求分发到后端服务器。
监听转发	选择是否将HTTP监听的流量转发到HTTPS监听。 <div> 说明 如果开启监听转发，需要选择目的监听。</div>
开启会话保持	选择是否开启会话保持。 开启会话保持功能后，负载均衡会把来自同一客户端的访问请求分发到同一台后端服务器上进行处理。 HTTP协议会话保持基于Cookie。负载均衡提供了两种Cookie处理方式： <ul style="list-style-type: none">◦ 植入Cookie：您只需要指定Cookie的过期时间。 客户端第一次访问时，负载均衡会在返回请求中植入Cookie（即在HTTP或HTTPS响应报文中插入ServerId），下次客户端携带此Cookie访问，负载均衡服务会将请求定向转发给之前记录到的后端服务器上。◦ 重写Cookie：可以根据需要指定HTTPS或HTTP响应中插入的Cookie。您需要在后端服务器上维护该Cookie的过期时间和生存时间。 负载均衡服务发现用户自定义了Cookie，将会对原来的Cookie进行重写，下次客户端携带新的Cookie访问，负载均衡服务会将请求定向转发给之前记录到的后端服务器。

监听配置	说明
开启监听带宽限速	选择是否配置监听带宽，取值范围为0~5120 Mbps。 对于按带宽计费的负载均衡实例，您可以针对不同监听设定不同的带宽峰值来限定监听的流量。实例下所有监听的带宽峰值总和不能超过该实例的带宽。默认不开启，各监听共享实例的总带宽。
连接空闲超时时间	指定连接空闲超时时间，取值范围为1~60秒。 在超时时间内一直没有访问请求，负载均衡会暂时中断当前连接，直到下一次请求来临时重新建立新的连接。
连接请求超时时间	指定请求超时时间，取值范围为1~180秒。 在超时时间内后端服务器一直没有响应，负载均衡将放弃等待，给客户端返回HTTP 504错误码。
Gzip数据压缩	开启该配置对特定文件类型进行压缩。 目前Gzip支持压缩的类型包括：text/xml、text/plain、text/css、application/javascript、application/x-javascript、application/rss+xml、application/atom+xml和application/xml。
附加HTTP头字段	选择您要添加的自定义HTTP header字段： <ul style="list-style-type: none">添加 X-Forwarded-For 字段获取客户端的真实IP地址。添加 SLB-ID 字段获取实例ID。添加 SLB-IP 字段获取实例的公网IP。添加 X-Forwarded-Proto 字段获取实例的监听协议。
获取客户端真实IP	获取来访者的真实IP地址，默认开启。
创建完毕自动启动监听	是否在监听配置完成后启动负载均衡监听，默认开启。

4. 单击下一步。

步骤二：添加后端服务器

添加处理前端请求的后端服务器。您可以使用实例配置的默认服务器组，也可以为监听配置一个虚拟服务器组或主备服务组。

本操作中，以默认后端服务器组为例。

- 1. 在后端服务器配置向导，选择默认服务器组，然后单击继续添加。
- 2. 在我的服务器面板，选择要添加的ECS实例，然后单击下一步。
- 3. 在配置端口和权重页签下，配置添加的后端服务器的权重，权重越高的ECS实例将被分配到更多的访问请求。

❓ 说明 权重设置为0，该服务器不会再接受新请求。

- 4. 单击添加，配置后端服务器（ECS实例）开放用来接收请求的端口，端口范围为1~65535。

同一个负载均衡实例内，后端服务器端口可以相同。

5. 单击下一步。

步骤三：配置健康检查

负载均衡通过健康检查来判断后端服务器（ECS实例）的业务可用性。健康检查机制提高了前端业务整体可用性，避免了后端ECS异常对总体服务的影响。

在健康检查配置向导，单击修改更改健康检查配置。具体操作，请参见[配置健康检查](#)。

步骤四：提交配置

- 1. 在配置审核配置向导，检查监听配置，您可以单击修改更改配置。
- 2. 确认无误后，单击提交。
- 3. 等待配置成功后，单击知道了。

配置成功后，您可以在监听页面查看已创建的监听。

5.5. 添加HTTPS监听

HTTPS协议适用于需要加密传输的应用。您可以添加一个HTTPS监听转发来自HTTPS协议的请求。

步骤一：配置监听

- 1. [登录SLB控制台](#)。
- 2. 选择以下一种方法，打开监听配置向导。
 - 在实例管理页面，单击目标实例操作列的监听配置向导。
 - 在实例管理页面，单击目标实例ID，然后选择监听页签，单击添加监听。
- 3. 完成以下配置，然后单击下一步。

监听配置	说明
选择负载均衡协议	选择监听的协议类型。 本操作，选择HTTPS。
监听端口	用来接收请求并向后端服务器进行请求转发的监听端口。端口范围为1~65535。
高级配置	单击修改展开高级配置。
调度算法	选择调度算法。 <ul style="list-style-type: none">◦ 加权轮询(WRR)：权重值越高的后端服务器，被轮询到的次数（概率）也越高。◦ 轮询(RR)：按照访问顺序依次将外部请求分发到后端服务器。◦ 加权最小连接数（WLC）：除了根据每台后端服务器设定的权重值来进行轮询，同时还考虑后端服务器的实际负载（即连接数）。当权重值相同时，当前连接数越小的后端服务器被轮询到的次数（概率）也越高。

监听配置	说明
开启会话保持	<p>选择是否开启会话保持。</p> <p>开启会话保持功能后，负载均衡会把来自同一客户端的访问请求分发到同一台后端服务器上进行处理。</p> <p>HTTP协议会话保持基于Cookie。负载均衡提供了两种Cookie处理方式：</p> <ul style="list-style-type: none"> ◦ 植入Cookie：您只需要指定Cookie的过期时间。 <p>客户端第一次访问时，负载均衡会在返回请求中植入Cookie（即在HTTP或HTTPS响应报文中插入ServerId），下次客户端携带此Cookie访问，负载均衡服务会将请求定向转发给之前记录到的后端服务器上。</p> <ul style="list-style-type: none"> ◦ 重写Cookie：可以根据需要指定HTTPS或HTTP响应中插入的Cookie。您需要在后端服务器上维护该Cookie的过期时间和生存时间。 <p>负载均衡服务发现用户自定义了Cookie，将会对原来的Cookie进行重写，下次客户端携带新的Cookie访问，负载均衡服务会将请求定向转发给之前记录到的后端服务器。</p>
启用HTTP 2.0	选择是否开启SLB前端协议版本为HTTP 2.0。
开启监听带宽限速	<p>选择是否配置监听带宽。</p> <p>对于按带宽计费的负载均衡实例，您可以针对不同监听设定不同的带宽峰值来限定监听的流量。实例下所有监听的带宽峰值总和不能超过该实例的带宽。默认不开启，各监听共享实例的总带宽。</p>
连接空闲超时时间	<p>指定连接空闲超时时间，取值范围为1~60秒。</p> <p>在超时时间内一直没有访问请求，负载均衡会暂时中断当前连接，直到下一次请求来临时重新建立新的连接。</p>
连接请求超时时间	<p>指定请求超时时间，取值范围为1~180秒。</p> <p>在超时时间内后端服务器一直没有响应，负载均衡将放弃等待，给客户端返回HTTP 504错误码。</p>
Gzip数据压缩	<p>开启该配置对特定文件类型进行压缩。</p> <p>目前Gzip支持压缩的类型包括：text/xml、text/plain、text/css、application/javascript、application/x-javascript、application/rss+xml、application/atom+xml和application/xml。</p>
附加HTTP头字段	<p>选择您要添加的自定义HTTP header字段：</p> <ul style="list-style-type: none"> ◦ 添加 X-Forwarded-For 字段获取客户端的真实IP地址。 ◦ 添加 SLB-ID 字段获取实例ID。 ◦ 添加 SLB-IP 字段获取实例的公网IP。 ◦ 添加 X-Forwarded-Proto 字段获取实例的监听协议。
获取客户端真实IP	获取来访者的真实IP地址，默认开启。

监听配置	说明
创建完毕自动启动监听	是否在监听配置完成后启动负载均衡监听，默认开启。

步骤二：配置SSL证书

添加HTTPS监听，您需要上传服务器证书或CA证书，如下表所示。

证书	说明	单向认证是否需要	双向认证是否需要
服务器证书	用来证明服务器的身份。 用户浏览器用来检查服务器发送的证书是否是由自己信赖的中心签发的。	是 服务器证书需要上传到负载均衡的证书管理系统。	是 服务器证书需要上传到负载均衡的证书管理系统。
客户端证书	用来证明客户端的身份。 用于证明客户端用户的身份，使得客户端用户在与服务器端通信时可以证明其真实身份。您可以用自签名的CA证书为客户端证书签名。	否	是 需要客户端进行安装。
CA 证书	服务器用CA证书验证客户端证书的签名。 如果没有通过验证，拒绝连接。	否	是 CA证书需要上传到负载均衡的证书管理系统。

在上传证书前，请注意：

- 目前阿里云负载均衡支持的公钥算法：RSA 1024、RSA 2048、RSA 4096、ECDSA P-256、ECDSA P-384 和 ECDSA P-521。
- 上传的证书格式必须是NGINX。
- 证书上传到负载均衡后，负载均衡即可管理证书，不需要在后端ECS上绑定证书。
- 因为证书的上传、加载和验证都需要一些时间，所以使用HTTPS协议的实例生效也需要一些时间。一般一分钟后就会生效，最长不会超过三分钟。
- HTTPS监听使用的ECDHE算法簇支持前向保密技术，不支持将DHE算法簇所需要的安全增强参数文件上传，即客户端CA公钥证书文件中含 `BEGIN DH PARAMETERS` 字段的字串上传。
- 目前负载均衡HTTPS监听不支持SNI（Server Name Indication），您可以改用TCP监听在后端ECS上实现SN功能。
- HTTPS监听的会话ticket保持时间默认为300秒。
- HTTPS监听实际产生的流量会比账单流量更多一些，因为会使用一些流量用于协议握手。
- 在新建连接数很高的情况下，会占用较大的流量。
 1. 在SSL证书配置页面，选择已上传的服务器证书，或单击新建服务器证书上传一个服务器证书。
 2. 如果您要开启HTTPS双向认证或者设置TLS安全策略，单击高级配置后面的修改。
 3. 打开双向认证，并选择一个已上传的CA证书，或新建一个CA证书。

步骤三：添加后端服务器

添加处理前端请求的后端服务器。您可以使用实例配置的默认服务器组，也可以为监听配置一个虚拟服务器组或主备服务组。

本操作中，以默认后端服务器组为例。

1. 在**后端服务器配置**页面，选择**默认服务器组**，然后单击**继续添加**。
2. 在**我的服务器**面板，选择要添加的后端服务器，然后单击**下一步**。
3. 在**权重**列下，配置添加的后端服务器的权重。

说明

- 权重越大ECS实例将被分配到更多的访问请求，默认为100。可通过单击**重置**修改权重为默认值。
- 权重设置为0，该服务器不会再接受新请求。

4. 单击**添加**，配置后端服务器用来接收请求的端口，端口范围为1~65535。然后单击**下一步**。

同一个负载均衡实例内，后端服务器端口可以相同。

步骤四：配置健康检查

CLB

通过健康检查来判断后端服务器（ECS实例）的业务可用性。健康检查机制提高了前端业务整体可用性，避免了后端ECS异常对总体服务的影响。

步骤五：提交配置

1. 在**配置审核**配置向导，检查监听配置，您可以单击**修改**更改配置。
2. 确认无误后，单击**提交**。
3. 等待配置成功后，单击**知道了**。

配置成功后，您可以在监听页面查看已创建的监听。

5.6. 配置转发策略

七层负载均衡支持配置域名或URL转发策略，将来自不同域名或URL的请求转发给不同的ECS处理。

背景信息

您可以在一个监听下添加多条转发策略，每条转发策略关联不同的服务器组（一个服务器组由一组ECS实例组成）。例如您可以将所有的请求转发到一组后端服务器上而将写请求转发到另一组后端服务器上，这样可以更灵活地适配业务需求，合理分配资源。

负载均衡的请求转发判断规则如下：

- 如果能匹配到相应监听关联的域名或URL转发规则，则按转发规则，将请求转发到对应的服务器组。
- 如果未匹配到域名或URL转发规则，而且对应监听配置了虚拟服务器组，则将请求转发到对应的虚拟服务器组。
- 如果均未匹配，则按照监听配置将请求转发到负载均衡实例后端默认服务器池中的ECS实例。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID，在**监听**页签，找到目标监听。

只有HTTP和HTTPS监听支持配置域名或URL转发规则。

- 3. 在目标监听的**操作列**单击**配置转发策略**。
- 4. 根据以下信息配置转发规则。
 - 域名转发规则配置
 - 单独配置域名转发规则时，URL配置项留空，不用输入正斜线（/）。域名只能使用字母、数字、短划线（-）、半角句号（.）。
 - 支持精确匹配和通配符匹配两种模式。例如，精确域名为：www.aliyun.com，通配符域名（泛域名）：*.aliyun.com, *.market.aliyun.com。当前端请求同时匹配多条域名规则时，规则的匹配优先级为精确匹配高于通配符匹配，如下表所示。

域名匹配规则

模式	请求URL	域名规则（√代表对应请求与该域名规则匹配，x代表不匹配）		
		www.aliyun.com	*.aliyun.com	*.market.aliyun.com
精确匹配	www.aliyun.com	√	x	x
泛域名匹配	market.aliyun.com	x	√	x
	info.market.aliyun.com	x	x	√

- URL转发规则配置
 - 单独配置URL转发规则时，域名配置项留空。
 - URL只能包含字母、数字和以下特殊字符：`-./%?#&`
 - URL必须以正斜线（/）开始。
- 说明 如果您在URL中只输入了一个正斜线（/），则URL转发规则失效。
- URL转发支持字符串匹配，遵循顺序匹配原则。例如 `/admin`、`/bbs_` 和 `/ino_test`。
- 域名和URL转发规则配置

当需要根据相同域名下不同的URL路径进行流量转发时，您可以对域名和URL转发规则进行组合。建议您配置一个默认转发策略（URL留空），以免未匹配到的其它URL访问出错。

例如一个网站的域名为 `www.example.com`，要求将来自 `www.example.com/index.html` 的请求转发给服务组1处理，其他请求都转发给服务器组2处理。要满足该需求，您需要配置两条转发规则。否则匹配到 `www.example.com` 的域名但没有相关策略匹配会返回404的响应码。

- 5. 单击**保存**。

5.7. 开启访问控制


负载均衡提供监听级别的访问控制。您可以针对不同的监听设置访问白名单。

操作步骤

- 1. [登录SLB控制台](#)。
- 2. 单击需要设置访问控制的实例ID。

3. 单击**监听**页签，在目标监听**操作**列选择... > **设置访问控制**。

4. 完成以下配置，然后单击**确定**。

配置	说明
启用访问控制	开启访问控制。
白名单设置	<p>转发来自所选访问控制策略组中设置的IP地址或地址段的请求。</p> <div><p> 说明 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。如果开启了白名单访问，但访问策略组中没有添加任何IP，则负载均衡监听会转发全部请求。</p></div>

5.8. 关闭访问控制

如果不需要设置访问限制，您可以关闭访问控制。

操作步骤

1. **登录SLB控制台**。
2. 单击需要设置访问控制的实例ID。
3. 单击**监听**页签，找到目标监听，在**操作**列选择... > **设置访问控制**。
4. 在**访问控制**设置页面，关闭访问控制，然后单击**确定**。


6. 后端服务器

6.1. 后端服务器概述

在使用负载均衡服务前，您需要添加ECS实例作为负载均衡实例的后端服务器，用来接收负载均衡监听转发的请求。

后端服务器简介

负载均衡服务通过设置虚拟服务地址，将添加的同一地域的多台ECS实例虚拟成一个高性能、高可用的应用服务池。您也可以通过虚拟服务器组管理后端服务器。不同的监听可以关联不同的服务器组，这样一个负载均衡实例可以将请求根据不同监听转发给不同的服务器组内不同端口的后端服务器。

 **说明** 如果您在配置监听时，选择使用虚拟服务器组，那么该监听会将请求转发到关联的服务器组中的ECS，而不会将请求转发给默认服务器组中的ECS实例。

限制说明

您可以在任意时刻增加或减少负载均衡实例的后端ECS数量，还可以在不同ECS实例之间进行流量分发切换。但是为了保证您对外服务的稳定性，确保在执行上述操作时，开启了负载均衡的健康检查功能并同时保证负载均衡实例中至少有一台正常运行的ECS。

添加后端服务器时，有以下限制条件：

- 负载均衡本身不会限制后端ECS实例使用哪种操作系统，只要您的两台ECS实例中的应用服务部署是相同的且保证数据的一致性即可。建议您选择相同操作系统的ECS实例作为后端服务器，以便日后管理和维护。
- 一个负载均衡实例最多支持添加50个监听，每个监听对应后端ECS实例上的一个应用。负载均衡的监听端口对应后端ECS实例上的应用服务端口。
- 您可以指定后端服务器池内各ECS实例的转发权重。权重越高的ECS实例将被分配到更多的访问请求。
- 如果您同时开启了会话保持功能，那么有可能会造成后端服务器的访问并不是完全相同的。如果出现了访问不均衡的情况，建议您暂时关闭会话保持功能，观察一下是否依然存在这种情况。

当负载均衡服务分发请求不均匀时，请参考以下方法检查处理：

- i. 统计一个时间段内，后端ECS实例的Web服务访问日志记录数据量。
 - ii. 按照负载均衡的配置，对比多台ECS实例日志的数量是否有相差。（开启会话保持后，需要剥离相同IP的访问日志。如果负载均衡配置了权重，要根据权重比例计算日志中访问比例是否正常。）
- ECS进行热迁移时，可能导致SLB长连接断开。重新连接后即可恢复，请做好应用的重连工作。

默认服务器组

用来接收前端请求的ECS实例。如果监听没有设置虚拟服务器组或主备服务器组，默认将请求转发至默认服务器组中的ECS。

在使用负载均衡服务前，必须至少添加一台默认服务器接收负载均衡转发的客户端请求。具体操作，请参见[添加默认服务器](#)。

虚拟服务器组

当您需要将不同的请求转发到不同的后端服务器上时，或需要通过域名和URL进行请求转发时，可以选择使用虚拟服务器组。具体操作，请参见[添加ECS实例作为虚拟服务器](#)。

主备服务器组

一个主备服务器组只包括两台ECS实例，一台作为主服务器，一台作为备服务器。由于备服务器不会做健康检查，所以只要主服务器健康检查失败，系统会直接将流量切到备机。当主服务器健康检查成功恢复服务后，流量会自动切到主服务器。具体操作，请参见[添加ECS实例作为主备服务器](#)。

 **说明** 只有TCP和UDP监听支持添加主备服务器组。

6.2. 默认服务器组

6.2.1. 添加默认服务器

在使用负载均衡服务前，必须至少添加一台默认服务器接收负载均衡转发的客户端请求。

前提条件

在向默认服务器组中添加ECS实例前，请确保：

- 您已创建负载均衡实例。具体操作，请参见[创建负载均衡实例](#)。
- 您已创建了ECS实例并部署了相关应用，用来接收转发的请求。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID。
3. 单击**默认服务器组**页签。
4. 单击**添加**。
5. 在**我的服务器**页面的**选择服务器**页签下，选择目标ECS实例。
6. 单击**下一步**。
7. 在**配置端口和权重**页签下，指定添加的ECS实例的后端服务权重。

权重：权重越高的ECS实例将被分配到更多的访问请求。

注意

- 权重范围0~100。如果一台服务器的权重设置为0，该服务器不再接收转发申请。
- 如果开启会话保持，可能会造成后端服务器的请求分配不均匀。

8. 单击**添加**。
9. 选择目标ECS实例，单击**确定**。

6.2.2. 添加IDC作为默认服务器

在使用负载均衡服务前，必须至少添加一台默认服务器接收负载均衡转发的客户端请求。

前提条件

已具有部署了相关应用的IDC服务器，用来接收转发的请求。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID。
3. 单击默认服务器组页签。
4. 单击添加IDC服务器。
5. 在我的服务器面板，单击添加。
6. 配置与IDC连通的VPC、IDC服务器名称和IDC服务器IP，然后单击下一步。
7. 在配置端口和权重配置向导，指定添加的IDC后端服务权重。

权重：权重越高的IDC服务器将被分配到更多的访问请求。



注意 如果权重设置为0，该服务器不会再接受新请求。

8. 单击添加。
9. 选择目标IDC服务器，单击确定。

6.2.3. 编辑后端服务器的权重

后端服务器添加完成后，需要修改后端服务器流量分发权重。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID。
3. 单击默认服务器组页签。
4. 找到目标后端服务器，在权重列单击对应数字。
5. 修改权重，然后单击确定。

权重越高的ECS实例或IDC服务器将被分配到更多的访问请求。



注意 权重范围为0~100，如果权重设置为0，该服务器不会再接受新请求。

6.2.4. 移除后端服务器

如果后端服务器不需要用来转发流量了，可以移除对应的后端服务器。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID。
3. 单击默认服务器组页签。
4. 在操作列单击移除，移除后端服务器。
5. 在确认对话框，单击确定。

6.3. 虚拟服务器组

6.3.1. 添加ECS实例作为虚拟服务器

虚拟服务器组（VServer group）是一组ECS实例服务器。将虚拟服务器组和一个监听关联后，监听只会将流量转发给关联的虚拟服务器组的后端服务器，不会再将流量转发给其他后端服务器。

前提条件

- 您已创建负载均衡实例。具体操作，请参见[创建负载均衡实例](#)。
- 您已创建了ECS实例并部署了相关应用，用来接收转发的请求。

背景信息

在创建虚拟服务器组时，请注意：

- 一个ECS实例可以属于多个虚拟服务器组。
- 一个虚拟服务器组可绑定在一个实例的多个监听上。
- 虚拟服务器组由ECS实例和应用端口组成。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID。
3. 单击**虚拟服务器组**页签。
4. 在**虚拟服务器组**页签，单击**创建虚拟服务器组**。
5. 在**创建虚拟服务器组**页面，配置相关参数。
 - i. 在**虚拟服务器组名称**文本框中，输入虚拟服务器组名称。
 - ii. 单击**添加**，在**我的服务器**面板选择要添加的服务器。
 - iii. 单击**下一步**。
 - iv. 输入每个ECS实例的端口和权重，单击**添加**。

端口和权重配置说明如下：

- **端口**：ECS实例开放用来接收请求的后端端口。

在同一个负载均衡实例内，后端端口可重复且同一个后端服务器实例可以通过单击**添加端口**，设置多个端口号。

- **权重**：权重越高的ECS实例将被分配到更多的访问请求。

 **注意** 权重设置为0，该服务器不会再接受新请求。

- v. 单击**添加**。
6. 勾选目标虚拟服务器，单击**创建**。

6.3.2. 添加IDC作为虚拟服务器

虚拟服务器组（VServer group）是一组ECS实例或IDC服务器。将虚拟服务器组和一个监听关联后，监听只会将流量转发给关联的虚拟服务器组的后端服务器，不会再将流量转发给其他后端服务器。

前提条件

在创建虚拟服务器组前，确保您已具有部署了相关应用的IDC服务器，用来接收转发的请求。

背景信息

在创建虚拟服务器组时，请注意：

- 一个IDC服务器可以属于多个虚拟服务器组。
- 一个虚拟服务器组可绑定在一个实例的多个监听上。
- 虚拟服务器组由IDC服务器和应用端口组成。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID。
3. 单击虚拟服务器组页签。
4. 在虚拟服务器组页面，单击创建虚拟服务器组。
5. 在创建虚拟服务器组页面，配置相关参数。
 - i. 在虚拟服务器组名称文本框中，输入虚拟服务器组名称。
 - ii. 单击添加IDC服务器。
 - iii. 在 我的服务器，单击添加。
 - iv. 配置参数与IDC连通的VPC、IDC服务器名称和IDC服务器IP。

其中，IDC服务器IP地址必须能与云上VPC互通。
 - v. 单击下一步。
 - vi. 输入每个IDC服务器的端口和权重，单击添加。

端口和权重配置说明如下：

- **端口**：IDC服务器开放用来接收请求的后端端口，一个IDC服务器可以添加多个端口。

在同一个负载均衡实例内，后端端口可重复。
- **权重**：权重越高的IDC服务器将被分配到更多的访问请求。

 **注意** 权重设置为0，该服务器不会再接受新请求。

6. 勾选目标服务器，单击创建。

6.3.3. 编辑虚拟服务器组

虚拟服务器组创建完成后，您可以修改虚拟服务器组中的ECS实例或IDC服务器配置。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID。
3. 单击虚拟服务器组页签。
4. 单击目标虚拟服务器组对应的编辑选项。
5. 修改ECS实例或IDC服务器的端口和权重，然后单击保存。

6.3.4. 删除虚拟服务器组

虚拟服务器组不用于流量转发时，您可以删除对应的虚拟服务器组。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID。
3. 单击**虚拟服务器组**页签。
4. 单击目标虚拟服务器组对应的**删除**选项。
5. 在弹出的对话框中，单击**确定**。

6.4. 主备服务器

6.4.1. 添加ECS实例作为主备服务器

在传统的主备场景下，即后端服务器中有一台主机和一台备机，流量默认被直接转发至主机；当主机宕机时，流量将被转发至备机。

前提条件

在创建主备服务器组前，确保：

- 您已创建负载均衡实例。具体操作，请参见[创建负载均衡实例](#)。
- 您已创建了ECS实例并部署了相关应用，用来接收转发的请求。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID。
3. 单击**主备服务器组**页签。
4. 在主备服务器组页面，单击**创建主备服务器组**。
5. 在**创建主备服务器组**页面，配置创建主备服务器组参数。
 - i. 在**主备服务器组名称**文本框中，输入主备服务器组名称。
 - ii. 单击**添加**，在**我的服务器**的**选择服务器**配置向导选择要添加的服务器。

主备服务器组只能添加两台ECS实例。
 - iii. 单击**下一步**。
 - iv. 配置ECS实例开放用来接收请求的后端端口，然后单击**添加**。

一个ECS实例可以设置多个端口。
 - v. 选择将一台服务器作为主服务器。
 - vi. 单击**创建**。

6.4.2. 添加IDC服务器作为主备服务器

在传统的主备场景下，即后端服务器中有一台主机和一台备机，流量默认被直接转发至主机；当主机宕机时，流量将被转发至备机。

前提条件

您已具有部署了相关应用的IDC服务器，用来接收转发的请求。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID。
3. 单击主备服务器组页签。
4. 在主备服务器组页面，单击创建主备服务器组。
5. 在创建主备服务器组页面，配置创建主备服务器组参数。
 - i. 在主备服务器组名称文本框中，输入主备服务器组名称。
 - ii. 单击添加IDC服务器，在我的服务器面板单击添加。
主备服务器组最多只能添加两台ECS实例。
 - iii. 配置参数与IDC连通的VPC、IDC服务器名称和IDC服务器IP。
其中，IDC服务器IP地址必须能与云上VPC互通。
 - iv. 单击下一步。
 - v. 配置ECS实例开放用来接收请求的后端端口，然后单击添加。
一个IDC服务器可以设置多个端口。
 - vi. 在创建主备服务器组页面，在机器类型列选择一台服务器作为主服务器。
 - vii. 单击创建。

6.4.3. 删除主备服务器组

主备服务器组不用于流量转发时，您可以删除对应的主备服务器组。

操作步骤

1. [登录SLB控制台](#)。
2. 单击目标实例的ID。
3. 单击主备服务器组页签。
4. 在主备服务器组页签下，单击目标主备服务器组对应操作列的删除。
5. 在弹出的确认对话框中，单击确定。

6.5. 通过弹性网卡添加后端服务器

弹性网卡ENI（Elastic Network Interface）是一种可以附加到专有网络VPC类型ECS实例上的虚拟网卡，通过ENI，您可以实现高可用集群搭建、低成本故障转移和精细化的网络管理。SLB后端服务器支持添加ENI上的主IP及其辅助IP。

背景信息

负载均衡实例添加后端服务器组时，如果ECS实例绑定多个ENI，支持挂载ENI上的主网卡及其辅助网卡。

② 说明

- 仅性能保障型实例后端服务器支持添加ENI上的主网卡及其辅助网卡。
- 目前控制台只支持挂载ENI上的辅助网卡，如果要挂载ENI上的主网卡，需要调用API进行挂载。



操作步骤

1. 登录SLB控制台。
2. 在顶部菜单栏，选择SLB实例的所属地域。
3. 在实例管理页面，单击需要添加后端服务器组的实例ID。
4. 单击虚拟服务器组、默认服务器组或主备服务器组页签。

② 说明 默认服务器组、虚拟服务器组 and 主备服务器组均支持挂载ENI上主网卡及其辅助网卡，本文以虚拟服务器组为例，介绍如何添加后端服务器。

5. 在虚拟服务器组页签，单击创建虚拟服务器组。
6. 在创建虚拟服务器组页面，单击添加。
7. 在选择服务器配置向导的服务器类型下拉列表中选择云服务器名称，然后打开高级模式开关。
8. 在服务器列表中，选中目标云服务器，单击下一步。
9. 在配置端口和权重配置向导，设置服务器的端口和权重，然后单击添加。

如果添加的是监听的后端服务器组，返回实例管理页面，可以看到挂载了ENI及其辅助网卡的后端服务器组如下：

- ：表示ECS实例。
- ：表示ENI及其辅助网卡。

7. 健康检查

7.1. 健康检查概述

负载均衡通过健康检查来判断后端服务器（ECS实例）的业务可用性。健康检查机制提高了前端业务整体可用性，避免了后端ECS异常对总体服务的影响。

开启健康检查功能后，当后端某台ECS健康检查出现异常时，负载均衡会自动将新的请求分发到其它健康检查正常的ECS上；而当该ECS恢复正常运行时，负载均衡会将其自动恢复到负载均衡服务中。

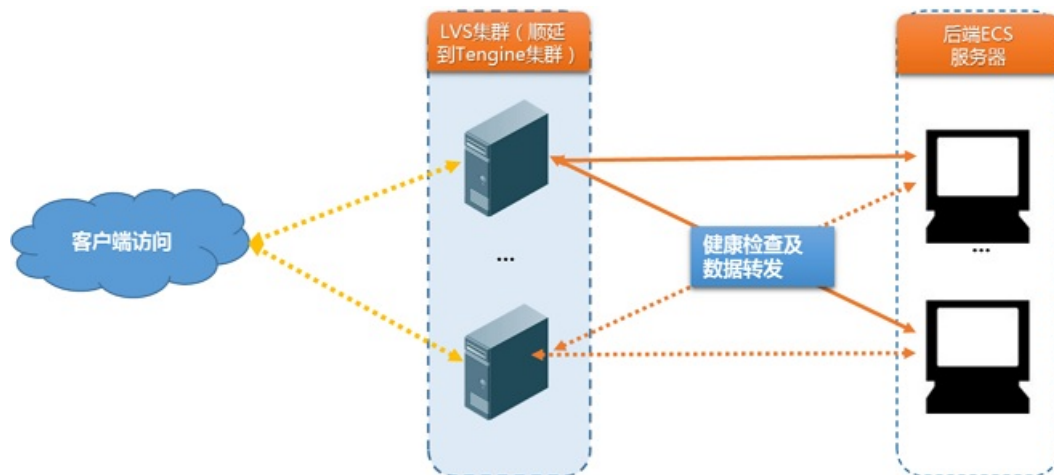
如果您的业务对负载敏感性高，高频率的健康检查探测可能会对正常业务访问造成影响。您可以结合业务情况，通过降低健康检查频率、增大健康检查间隔、七层检查修改为四层检查等方式，来降低对业务的影响。但为了保障业务的持续可用，不建议关闭健康检查。

健康检查过程

负载均衡采用集群部署。LVS集群或Tengine集群内的相关节点服务器同时承载了数据转发和健康检查职责。

LVS集群内不同服务器分别独立、并行地根据负载均衡策略进行数据转发和健康检查操作。如果某一台LVS节点服务器对后端某一台ECS健康检查失败，则该LVS节点服务器将不会再将新的客户端请求分发给相应的异常ECS。LVS集群内所有服务器同步进行该操作。

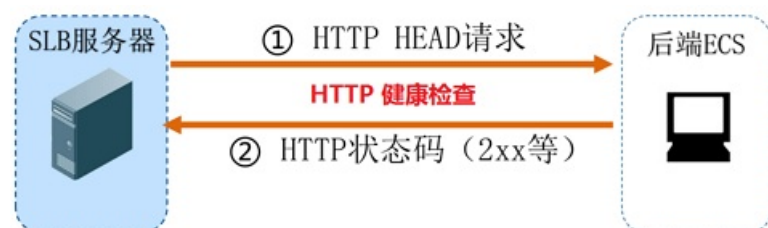
如下图所示，负载均衡健康检查使用的地址段是100.64.0.0/10，后端服务器务必不能屏蔽该地址段。您无需在ECS安全组中额外针对该地址段配置放行策略，但如有配置iptables等安全策略，请务必放行（100.64.0.0/10 是阿里云保留地址，其他用户无法分配到该网段内，不会存在安全风险）。



HTTP/HTTPS监听健康检查机制

针对七层（HTTP或HTTPS协议）监听，健康检查通过HTTP HEAD探测来获取状态信息，如下图所示。

对于HTTPS监听，证书在负载均衡系统中进行管理。负载均衡与后端ECS之间的数据交互（包括健康检查数据和业务交互数据），不再通过HTTPS进行传输，以提高系统性能。

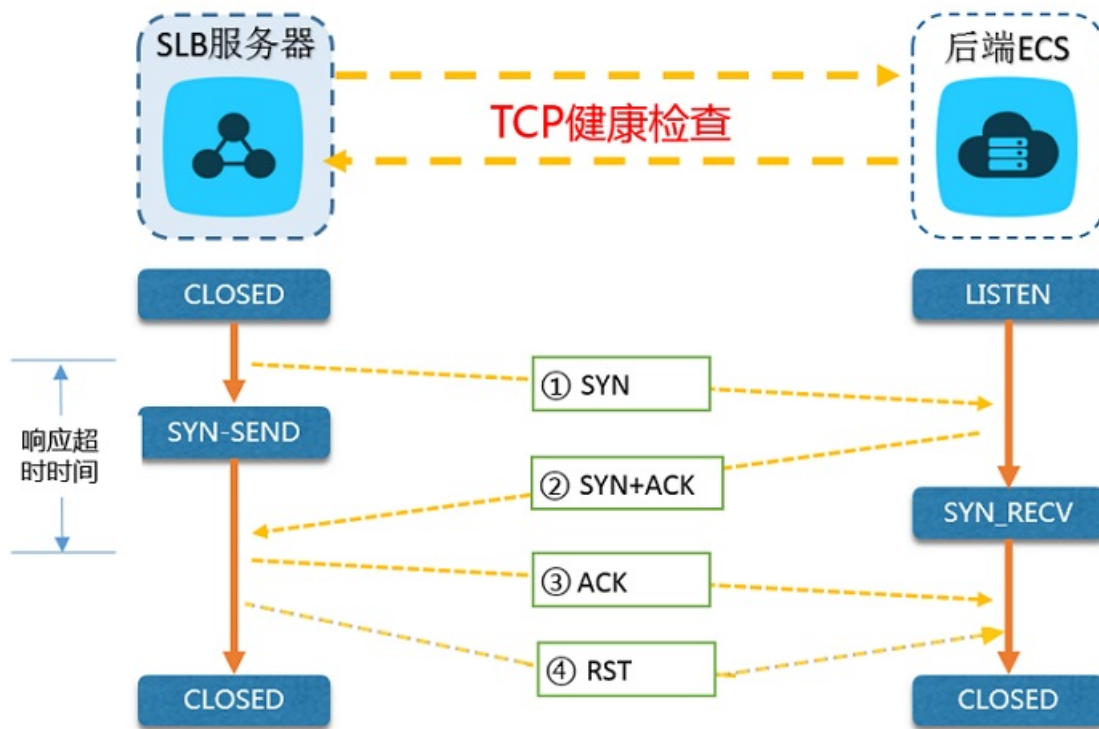


七层监听的检查机制如下：

1. Tengine节点服务器根据监听的健康检查配置，向后端ECS的内网IP+【健康检查端口】+【检查路径】发送HTTP HEAD请求（包含设置的【域名】）。
2. 后端ECS收到请求后，根据相应服务的运行情况，返回HTTP状态码。
3. 如果在【响应超时时间】之内，Tengine节点服务器没有收到后端ECS返回的信息，则认为服务无响应，判定健康检查失败。
4. 如果在【响应超时时间】之内，Tengine节点服务器成功接收到后端ECS返回的信息，则将该返回信息与配置的状态码进行比对。如果匹配则判定健康检查成功，反之则判定健康检查失败。

TCP监听健康检查机制

针对四层TCP监听，为了提高健康检查效率，健康检查通过定制的TCP探测来获取状态信息，如下图所示。



TCP监听的检查机制如下：

1. LVS节点服务器根据监听的健康检查配置，向后端ECS的内网IP+【健康检查端口】发送TCP SYN数据包。
2. 后端ECS收到请求后，如果相应端口正在正常监听，则会返回SYN+ACK数据包。
3. 如果在【响应超时时间】之内，LVS节点服务器没有收到后端ECS返回的数据包，则认为服务无响应，判定健康检查失败；并向后端ECS发送RST数据包中断TCP连接。
4. 如果在【响应超时时间】之内，LVS节点服务器成功收到后端ECS返回的数据包，则认为服务正常运行，判定健康检查成功，而后向后端ECS发送RST数据包中断TCP连接。

② 说明 正常的TCP三次握手，LVS节点服务器在收到后端ECS返回的SYN+ACK数据包后，会进一步发送ACK数据包，随后立即发送RST数据包中断TCP连接。

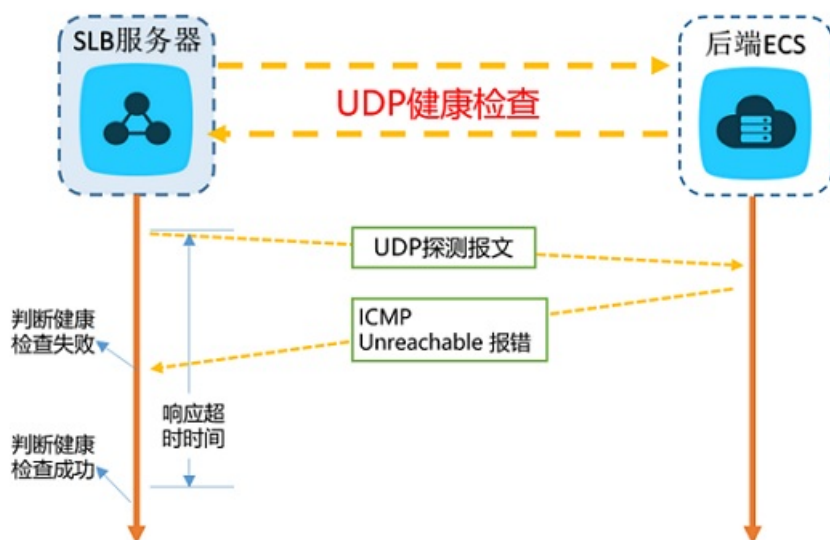
该实现机制可能会导致后端ECS认为相关TCP连接出现异常（非正常退出），并在业务软件如Java连接池等日志中抛出相应的错误信息，如 `Connection reset by peer`。

解决方案：

- TCP监听采用HTTP方式进行健康检查。
- 在后端ECS配置了获取客户端真实IP后，忽略来自前述负载均衡服务地址段相关访问导致的连接错误。

UDP监听健康检查

针对四层UDP监听，健康检查通过UDP报文探测来获取状态信息，如下图所示。



UDP监听的检查机制如下：

1. LVS节点服务器根据监听的健康检查配置，向后端ECS的内网IP+【健康检查端口】发送UDP报文。
2. 如果后端ECS相应端口未正常监听，则系统会返回类似 `port XX unreachable` 的ICMP报错信息，反之不做任何处理。
3. 如果在【响应超时时间】之内，LVS节点服务器收到了后端ECS返回的上述错误信息，则认为服务异常，判定健康检查失败。
4. 如果在【响应超时时间】之内，LVS节点服务器没有收到后端ECS返回的任何信息，则认为服务正常，判定健康检查成功。

② 说明 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是Linux服务器，在大并发场景下，由于Linux的防ICMP攻击保护机制，会限制服务器发送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。

解决方案：

负载均衡通过发送您指定的字符串到后端服务器，必须得到指定应答后才认为检查成功。但该实现机制需要客户端程序配合应答。

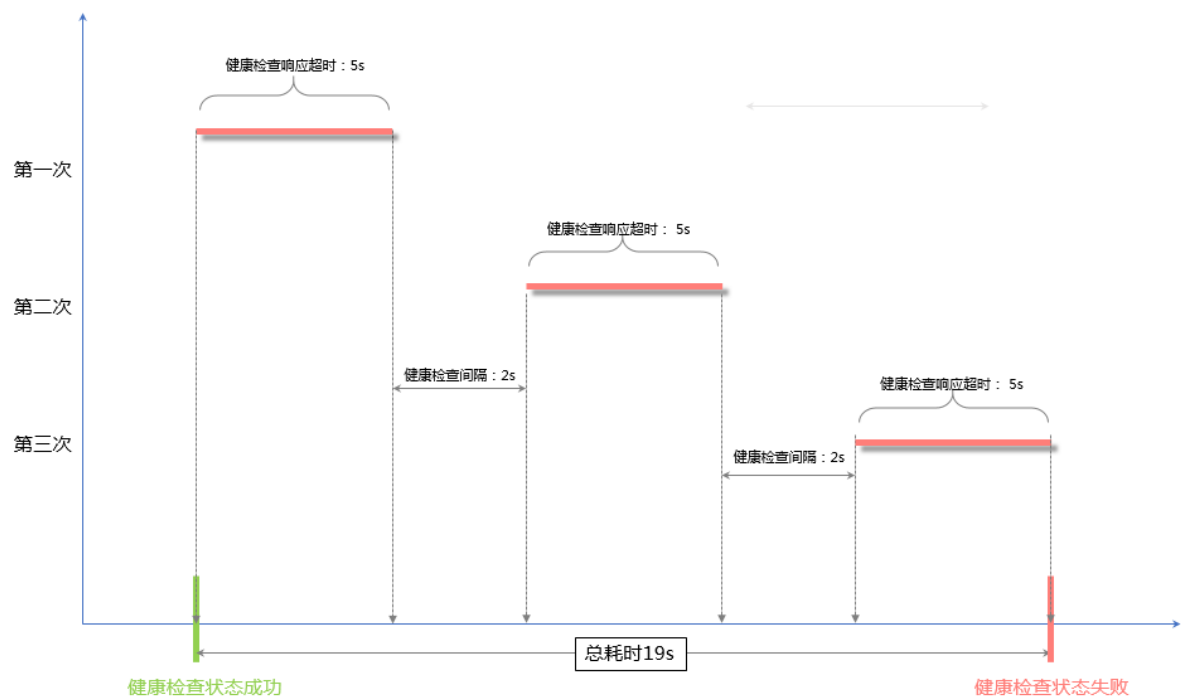
健康检查时间窗

健康检查机制的引入，有效提高了业务服务的可用性。但是，为了避免频繁的健康检查失败引起的切换对系统可用性的冲击，健康检查只有在健康检查时间窗内连续多次检查成功或失败后，才会进行状态切换。健康检查时间窗由以下三个因素决定：

- 健康检查间隔（每隔多久进行一次健康检查）
- 响应超时时间（等待服务器返回健康检查的时间）
- 检查阈值（健康检查连续成功或失败的次数）

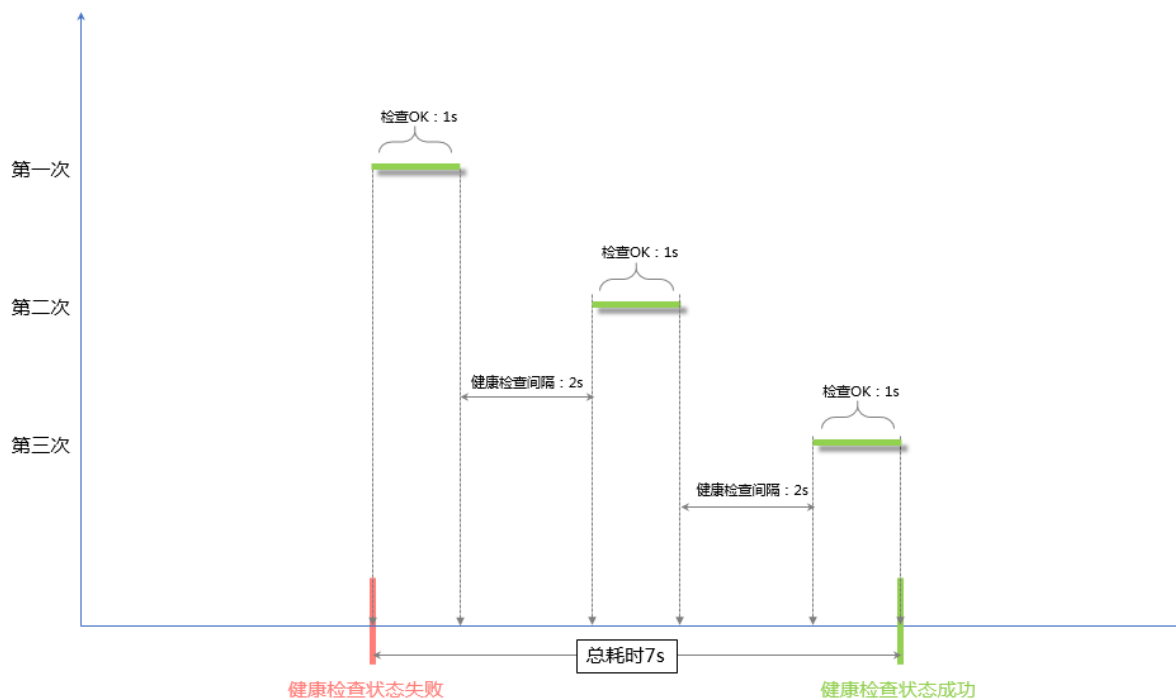
健康检查时间窗口的计算方法如下：

- 健康检查失败时间窗口=响应超时时间×不健康阈值+检查间隔×（不健康阈值-1）



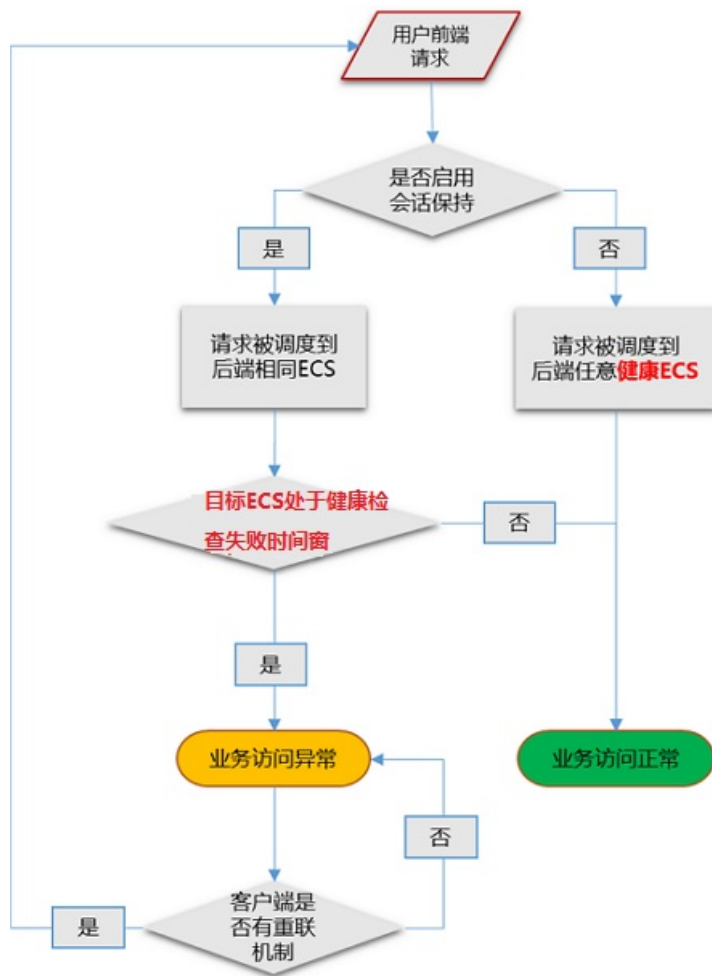
- 健康检查成功时间窗口=（健康检查成功响应时间×健康阈值）+检查间隔×（健康阈值-1）

② 说明 健康检查成功响应时间是一次健康检查请求从发出到响应的的时间。当采用TCP方式健康检查时，由于仅探测端口是否存活，因此该时间非常短，几乎可以忽略不计。当采用HTTP方式健康检查时，该时间取决于应用服务器的性能和负载，但通常都在秒级以内。



健康检查状态对请求转发的影响如下：

- 如果目标ECS的健康检查失败，新的请求不会再分发到相应ECS上，所以对前端访问没有影响。
- 如果目标ECS的健康检查成功，新的请求会分发到该ECS上，前端访问正常。
- 如果目标ECS存在异常，正处于健康检查失败时间窗，而健康检查还未达到检查失败判定次数（默认为三次），则相应请求还是会被分发到该ECS，进而导致前端访问请求失败。



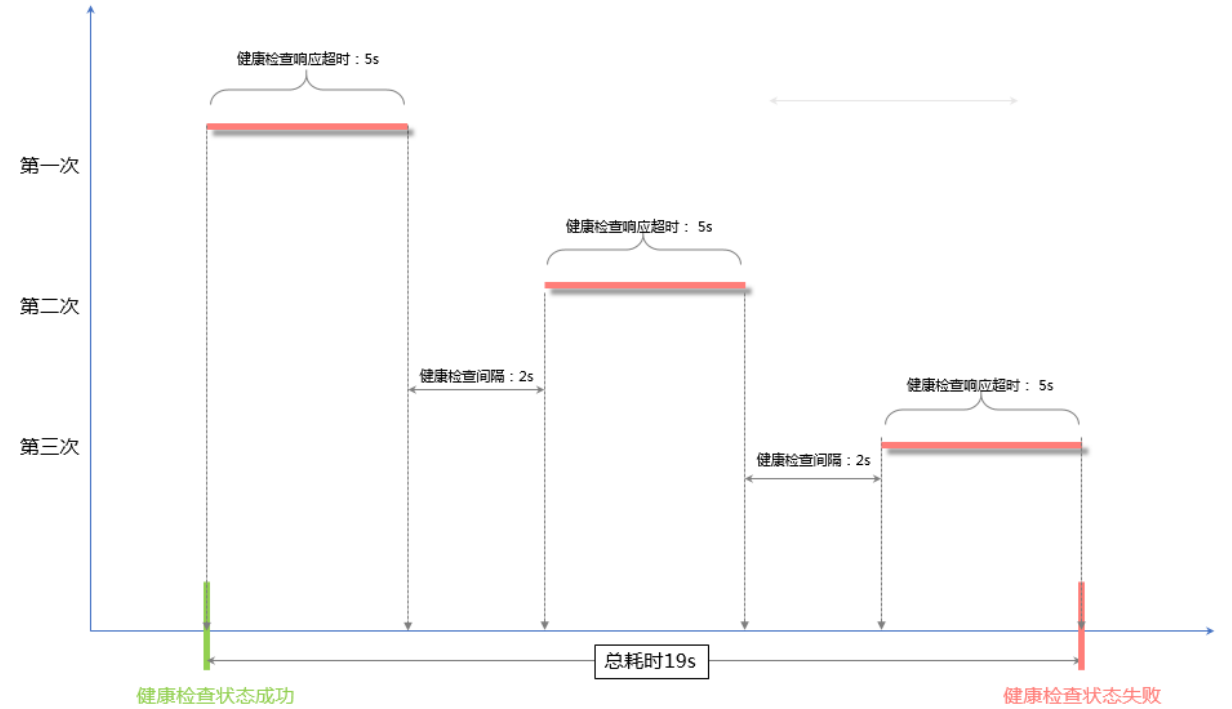
健康检查响应超时和健康检查间隔示例

以如下健康检查配置为例：

- 响应超时时间：5秒
- 健康检查间隔：2秒
- 健康阈值：3次
- 不健康阈值：3次

健康检查失败时间窗口=响应超时时间×不健康阈值+检查间隔×（不健康阈值-1）， $5 \times 3 + 2 \times (3 - 1) = 19s$ ，即以19s为一个时间窗，健康检查响应时间超过19s，健康检查状态为不健康。

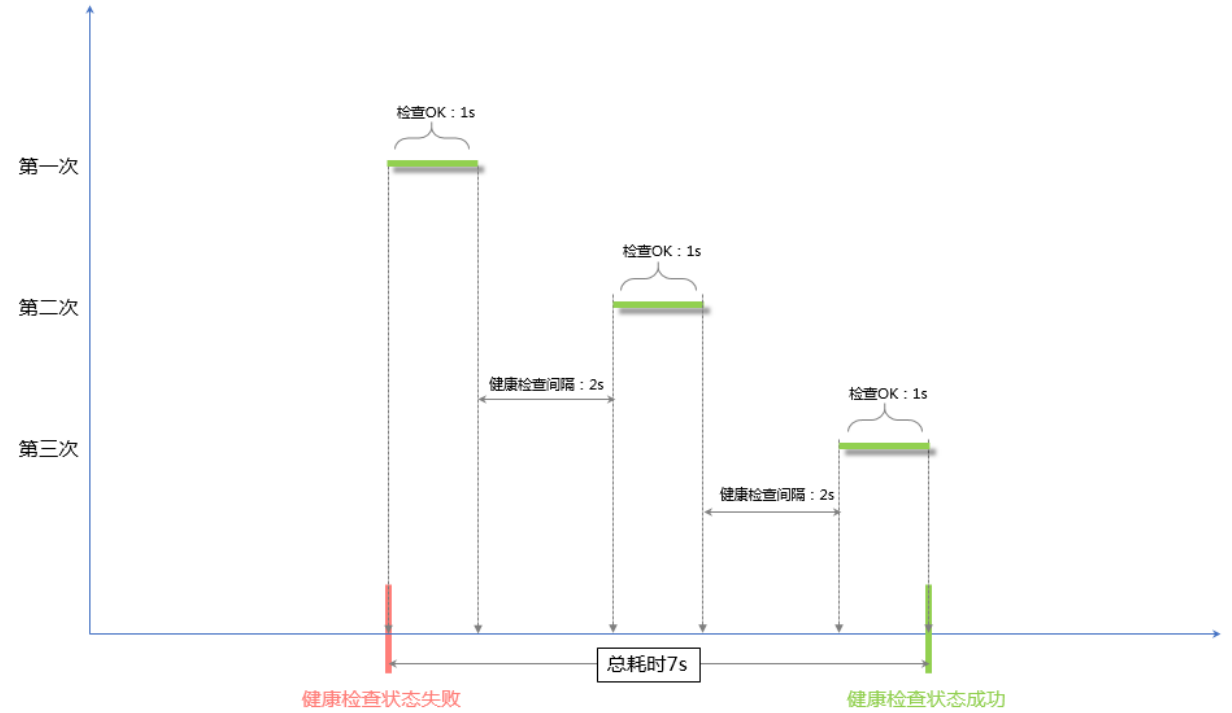
从健康状态到不健康状态的检查过程如下图所示：



健康检查成功时间窗口=（健康检查成功响应时间×健康阈值）+检查间隔×（健康阈值-1），
（1×3）+2×（3-1）=7s，即以7s为一个时间窗，健康检查成功响应时间低于7s，健康检查状态为健康。

❓ 说明 健康检查成功响应时间是一次健康检查请求从发出到响应的时间。当采用TCP方式健康检查时，由于仅探测端口是否存活，因此该时间非常短，几乎可以忽略不计。当采用HTTP方式健康检查时，该时间取决于应用服务器的性能和负载，但通常都在秒级以内。

从不健康状态到健康的状态检查过程如下图所示（假设服务器响应健康检查请求需要耗时1s）：



HTTP健康检查中域名的设置

当使用HTTP方式进行健康检查时，可以设置健康检查的域名，但并非强制选项。因为有些应用服务器会对请求中的host字段做校验，即要求请求头中必须存在host字段。如果在健康检查中配置了域名，则SLB会将域名配置到host字段中去，反之，如果没有配置域名，SLB则不会在请求中附带host字段，因此健康检查请求就会被服务器拒绝，可能导致健康检查失败。综上所述，如果您的应用服务器需要校验请求的host字段，那么则需要配置相关的域名，确保健康检查正常工作。

7.2. 配置健康检查

您可以在添加监听时配置健康检查，通常，使用默认的健康检查配置即可。

操作步骤

- 1. 登录SLB控制台。
- 2. 单击负载均衡实例的ID。
- 3. 单击监听页签。
- 4. 单击添加监听，或找到目标监听在操作列单击修改监听配置。
- 5. 单击下一步至健康检查页签，配置健康检查，单击下一步，完成健康检查的配置。

在配置健康检查时，建议您使用默认值。

健康检查配置说明


健康检查配置	说明
健康检查协议	选择健康检查协议类型，监听为TCP协议时，健康检查方式可选TCP或HTTP模式。 <ul style="list-style-type: none">○ TCP模式的健康检查是基于网络层探测，通过发送SYN握手报文来检测服务器端口是否存活。○ HTTP模式的健康检查是通过发送head请求，通过发送HEAD或GET请求模拟浏览器的访问行为来检查服务器应用是否健康。
健康检查方法 (仅HTTP和HTTPS健康检查协议支持)	七层监听（HTTP/HTTPS）健康检查支持GET方法。 使用GET方法时，如果Response长度超过8K，会被截断，但不会影响健康检查结果的判定。
健康检查路径和健康检查域名（可选） (仅HTTP健康检查协议支持)	如果您用来进行健康检查的页面并不是应用服务器的缺省首页，需要指定具体的检查路径。 因为有些应用服务器会对请求中的host字段做校验，即要求请求头中必须存在host字段。如果在健康检查中配置了域名，则SLB会将域名配置到host字段中去，反之，如果没有配置域名，SLB则不会在请求中附带host字段，因此健康检查请求就会被服务器拒绝，可能导致健康检查失败。综上所述，如果您的应用服务器需要校验请求的host字段，则需要配置相关域名，确保健康检查正常工作。
正常状态码 (仅HTTP健康检查协议支持)	选择健康检查正常的HTTP状态码。 默认值为http_2xx和http_3xx。

健康检查配置	说明
健康检查端口	<p>健康检查服务访问后端时的探测端口。</p> <p>默认值为配置监听时指定的后端端口。</p> <div><p> 说明 如果该监听配置了虚拟服务器组或主备服务器组，且组内的ECS实例的端口都不相同，此时不需要配置检查端口。负载均衡系统会使用各自ECS的后端端口进行健康检查。</p></div>
健康检查响应超时时间	<p>接收来自运行状况检查的响应需要等待的时间。如果后端ECS在指定的时间内没有正确响应，则判定为健康检查失败。</p> <p>范围是1~300秒，UDP监听的默认值为10秒，HTTP/HTTPS/TCP监听的默认值为5秒。</p>
健康检查间隔时间	<p>进行健康检查的时间间隔。</p> <p>LVS集群内所有节点，都会独立、并行地遵循该属性对后端ECS进行健康检查。由于各LVS节点的检查时间并不同步，所以，如果从后端某一ECS上进行单独统计，会发现来自负载均衡的健康检查请求在时间上并不会遵循上述时间间隔。</p> <p>范围是1~50秒，UDP监听的默认值为5秒，HTTP/HTTPS/TCP监听的默认值为2秒。</p>
健康检查健康阈值	<p>同一LVS节点服务器针对同一ECS服务器，从失败到成功的连续健康检查成功次数。</p> <p>可选值2~10，默认为3次。</p>
健康检查不健康阈值	<p>同一LVS节点服务器针对同一ECS服务器，从成功到失败的连续健康检查失败次数。</p> <p>可选值2~10，默认为3次。</p>
健康检查请求	<p>自定义健康检查请求，只允许包含字母、数字字符，最大长度限制为500字符。</p> <div><p> 说明 该参数仅在选择负载均衡协议选择UDP时有效。</p></div>
健康检查返回结果	<p>自定义健康检查返回结果，只允许包含字母、数字字符，最大长度限制为500字符。</p> <div><p> 说明 该参数仅在选择负载均衡协议选择UDP时有效。</p></div>

7.3. 关闭健康检查

您可以关闭健康检查功能，但关闭健康检查后，当后端某个ECS健康检查出现异常时，负载均衡还是会把请求转发到该异常的ECS上，造成部分业务不可访问。所以建议一般情况下不要关闭健康检查。

背景信息

 **说明** 只有HTTP和HTTPS监听支持关闭健康检查。UDP和TCP监听无法关闭健康检查。

操作步骤

1. [登录SLB控制台](#)。
2. 在**实例管理**页面，单击负载均衡实例的ID。
3. 在**监听**页签下，找到目标监听，在**操作**列单击**修改监听配置**。
4. 在**配置监听**对话框，连续单击**下一步至健康检查**。
5. 关闭健康检查开关，单击**下一步**，单击**提交**，然后单击**知道了**。

8. 证书管理

8.1. 证书概述

配置HTTPS监听，您可以将所需的第三方签发的服务器证书和CA证书上传到负载均衡。上传后，无需在后端服务器再配置证书。

负载均衡支持第三方签发的证书，上传第三方签发证书，您需要持有证书的公钥/私钥文件。

支持HTTPS服务器证书及客户端CA证书。

在创建证书前，注意每个账号最多可以创建100个证书。

8.2. 证书要求

负载均衡只支持NGINX格式的证书。在上传证书前，确保您的证书、证书链和私钥符合格式要求。

Root CA机构颁发的证书

如果是通过Root CA机构颁发的证书，您拿到的证书是唯一的一份，不需要额外的证书，配置的站点即可被浏览器等访问设备认为可信。

证书格式必须符合如下要求：

- 以 `-----BEGIN CERTIFICATE-----`，`-----END CERTIFICATE-----` 开头和结尾。
- 每行64个字符，最后一行长度可以不足64个字符。
- 证书内容不能包含空格。

中级机构颁发的证书

如果是通过中级CA机构颁发的证书，您拿到的证书文件包含多份证书，需要将服务器证书与中级证书合并在一起上传。

证书链格式必须符合如下要求：

- 服务器证书放第一位，中级证书放第二位，中间不能有空行。
- 证书内容不能包含空格。
- 证书之间不能有空行，并且每行64字节。更多信息，请参见[RFC1421](#)。
- 符合证书的格式要求。一般情况下，中级机构在颁发证书时会有对应说明，证书要符合证书机构的格式要求。

中级机构颁发的证书链示例。

```
-----BEGIN CERTIFICATE-----
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
-----END CERTIFICATE-----
-----BEGIN CERTIFICATE-----
-----END CERTIFICATE-----
```

证书公钥

目前阿里云负载均衡支持如下公钥算法：

- RSA 1024
- RSA 2048
- RSA 4096
- ECDSA P-256
- ECDSA P-384
- ECDSA P-521

RSA私钥格式要求

在上传服务器证书时，您也需要上传证书的私钥。


RSA私钥格式必须符合如下要求：

- 以 `-----BEGIN RSA PRIVATE KEY-----`，`-----END RSA PRIVATE KEY-----` 开头和结尾，请将这些内容一并上传。
- 字串之间不能有空行，每行64字符，最后一行长度可以不足64字符。更多信息，请参见[RFC1421](#)。

如果您的私钥是加密的，例如私钥中包含 `Proc-Type: 4,ENCRYPTED`，需要先运行以下命令进行转换：

```
openssl rsa -in old_server_key.pem -out new_server_key.pem
```

EC私钥格式要求

 说明 目前仅英国（伦敦）地域支持。

在上传服务器证书时，您也需要上传证书的私钥。

EC私钥格式必须符合如下要求：

- 以 `-----BEGIN EC PARAMETERS-----`，`-----END EC PARAMETERS-----` 开头和结尾，请将这些内容一并上传。
- 字串之间不能有空行，每行64字符，最后一行长度可以不足64字符。更多信息，请参见[RFC1421](#)。

如果您的私钥是加密的，例如私钥的开头和结尾是 `-----BEGIN EC PRIVATE KEY-----`，`-----END EC PRIVATE KEY-----` 或者私钥中包含 `Proc-Type: 4,ENCRYPTED`，需要先运行以下命令进行转换：

```
openssl ec -in old_server_key.pem -out new_server_key.pem
```

下图为EC私钥示例。

```
-----BEGIN EC PARAMETERS-----
Bgqg*****Bw==
-----END EC PARAMETERS-----
-----BEGIN EC PRIVATE KEY-----
MHcCAQEIEICo9b+vQUhqFUWgWjE0YY4h0b3bE/udcubxVwcVY99MuoAoGCCqGSM49
*****4xz0SHsuQc/7XBmgmrMpAmE80c0DR
5HcMHFxRpTGLv22T62e5KqN1W3uN9Hp1gg==
-----END EC PRIVATE KEY-----
```

8.3. 上传证书

配置HTTPS监听，您必须将所需的服务器证书和CA证书上传到负载均衡。上传后，无需在后端服务器再配置证书。

前提条件

- 您已经购买了服务器证书。
- 您已经生成了CA证书和客户端证书。

背景信息

每个账号最多可以创建100个证书。

操作步骤

1. [登录SLB控制台](#)。
2. 在左侧导航栏，单击[证书管理](#)。
3. 单击[创建证书](#)。
4. 在创建证书面板，根据以下信息上传证书内容，然后单击[创建](#)。

配置	说明
证书名称	输入证书名称。 长度为1~80个字符，只能包含大小写英文字母、数字、短划线（-）、正斜线（/）、半角句号（.）、下划线（_）和星号（*）。
组织	选择证书所属的组织。
资源集	选择证书所属的资源集。
证书标准	选择要上传的证书标准：国际证书或国密证书。
证书类型	选择证书的类型为服务器证书。
公钥证书	复制服务器证书内容。 单击 查看样例 查看正确的证书样式。更多信息，请参见 证书要求 。
私钥	复制服务器证书的私钥内容。 单击 查看样例 查看正确的证书样式。更多信息，请参见 证书要求 。 <div> 注意 只有上传服务器证书时，才需要上传私钥。</div>
证书部署地域	选择证书的部署地域。

8.4. 生成CA证书

在配置HTTPS监听时，您可以使用自签名的CA证书，并且使用该CA证书为客户端证书签名。

使用Open SSL生成CA证书

1. 上传证书到Linux服务器，执行以下命令，在 `/root` 目录下新建一个 `ca` 文件夹，并在 `ca` 文件夹下创建四个子文件夹。

```
sudo mkdir ca
cd ca
sudo mkdir newcerts private conf server
```

- `newcerts` 目录将用于存放CA签署过的数字证书。
 - `private` 目录用于存放CA的私钥。
 - `conf` 目录用于存放一些简化参数用的配置文件。
 - `server` 目录存放服务器证书文件。
2. 在 `conf` 目录下新建一个包含以下信息的 `openssl.conf` 文件。

```
[ ca ]
default_ca = foo
[ foo ]
dir = /root/ca
database = /root/ca/index.txt
new_certs_dir = /root/ca/newcerts
certificate = /root/ca/private/ca.crt
serial = /root/ca/serial
private_key = /root/ca/private/ca.key
RANDFILE = /root/ca/private/.rand
default_days = 365
default_crl_days = 30
default_md = md5
unique_subject = no
policy = policy_any
[ policy_any ]
countryName = match
stateOrProvinceName = match
organizationName = match
organizationalUnitName = match
localityName = optional
commonName = supplied
emailAddress = optional
```

3. 执行以下命令，生成私钥key文件。

```
cd /root/ca
sudo openssl genrsa -out private/ca.key
```

执行结果如下图所示。

```
root@iZb1...:~/ca/conf# cd /root/ca
root@iZb1...:~/ca# sudo openssl genrsa -out private/ca.key
Generating RSA private key, 2048 bit long modulus
.....+++
..+++
e is 65537 (0x10001)
```

4. 执行以下命令，按照提示输入所需信息，然后按下回车键生成证书请求 `.csr` 文件。

```
sudo openssl req -new -key private/ca.key -out private/ca.csr
```

? 说明

Common Name需要输入负载均衡的域名。

```
root@iZb1...iZ:~/ca# sudo openssl req -new -key private/ca.key -out private/ca.csr
You are about to be asked to enter information that will be incorporated
into your certificate request.
What you are about to enter is what is called a Distinguished Name or a DN.
There are quite a few fields but you can leave some blank
For some fields there will be a default value,
If you enter '.', the field will be left blank.
-----
Country Name (2 letter code) [AU]:CN
State or Province Name (full name) [Some-State]:ZheJiang
Locality Name (eg, city) []:HangZhou
Organization Name (eg, company) [Internet Widgits Pty Ltd]:Alibaba
Organizational Unit Name (eg, section) []:Test
Common Name (e.g. server FQDN or YOUR name) []:mydomain
Email Address []:a@lib.com

Please enter the following 'extra' attributes
to be sent with your certificate request
A challenge password []:
An optional company name []:
root@iZb1...iZ:~/ca#
```

5. 执行以下命令，生成凭证`crt`文件。

```
sudo openssl x509 -req -days 365 -in private/ca.csr -signkey private/ca.key -out private/ca.crt
```

6. 执行以下命令，为CA的key设置起始序列号，可以是任意四个字符。

```
sudo echo FACE > serial
```

7. 执行以下命令，创建CA键库。

```
sudo touch index.txt
```

8. 执行以下命令，为移除客户端证书创建一个证书撤销列表。

```
sudo openssl ca -gencrl -out /root/ca/private/ca.crl -crl days 7 -config "/root/ca/conf/openssl.conf"
```

输出为：

```
Using configuration from /root/ca/conf/openssl.conf
```

为客户端证书签名

1. 上传证书到Linux服务器，执行以下命令，在`ca`目录内创建一个存放客户端key的目录`users`。

```
sudo mkdir users
```

2. 执行以下命令，为客户端创建一个key。


```
sudo openssl genrsa -des3 -out /root/ca/users/client.key 1024
```

? 说明

创建key时要求输入pass phrase，这个是当前key的口令，以防止本密钥泄漏后被人盗用。两次输入同一个密码。

3. 执行以下命令，为客户端key创建一个证书签名请求.csr文件。

```
sudo openssl req -new -key /root/ca/users/client.key -out /root/ca/users/client.csr
```

输入该命令后，根据提示输入上一步输入的pass phrase，然后根据提示输入对应的信息。

? 说明

A challenge password是客户端证书口令。注意将它和client.key的口令进行区分。

4. 执行以下命令，使用CA证书的key为客户端key签名。

```
sudo openssl ca -in /root/ca/users/client.csr -cert /root/ca/private/ca.crt -keyfile /root/ca/private/ca.key -out /root/ca/users/client.crt -config "/root/ca/conf/openssl.conf"
```

当出现确认是否签名的提示时，两次都输入y。

```
root@iz1z1z1z1z:~/ca# sudo openssl ca -in /root/ca/users/client.csr
-cert /root/ca/private/ca.crt -keyfile /root/ca/private/ca.key -out /root/ca/us
ers/client.crt -config "/root/ca/conf/openssl.conf"
Using configuration from /root/ca/conf/openssl.conf
Check that the request matches the signature
Signature ok
The Subject's Distinguished Name is as follows
countryName             :PRINTABLE:'CN'
stateOrProvinceName     :ASN.1 12:'ZheJiang'
localityName            :ASN.1 12:'HangZhou'
organizationName        :ASN.1 12:'Alibaba'
organizationalUnitName   :ASN.1 12:'Test'
commonName               :ASN.1 12:'mydomain'
emailAddress             :IA5STRING:'a@163.com'
Certificate is to be certified until Jun  4 15:28:55 2018 GMT (365 days)
Sign the certificate? [y/n]:y

1 out of 1 certificate requests certified, commit? [y/n]:y
Write out database with 1 new entries
Data Base Updated
root@iz1z1z1z1z:~/ca#
```

5. 执行以下命令，将证书转换为PKCS12文件。

```
sudo openssl pkcs12 -export -clcerts -in /root/ca/users/client.crt -inkey /root/ca/user
s/client.key -out /root/ca/users/client.p12
```

按照提示输入客户端client.key的pass phrase。再输入用于导出证书的密码。这个是客户端证书的保护密码，在安装客户端证书时需要输入这个密码。

6. 执行以下命令，查看生成的客户端证书。

```
cd users
ls
```

8.5. 转换证书格式

负载均衡只支持NGINX格式的证书，其它格式的证书需要转换成以.pem或者.crt为后缀的格式后，才能上传到负载均衡。建议使用Open SSL进行转换。

DER转换为PEM

DER格式通常使用在Java平台中，证书文件后缀一般为.der、.cer或者.crt。

- 运行以下命令进行证书转化：

```
openssl x509 -inform der -in certificate.cer -out certificate.pem
```

- 运行以下命令进行私钥转化：

```
openssl rsa -inform DER -outform PEM -in privatekey.der -out privatekey.pem
```

P7B转换为PEM

P7B格式通常使用在Windows Server和Tomcat中。

运行以下命令进行证书转化：

```
openssl pkcs7 -print_certs -in incertificate.p7b -out outcertificate.cer
```

PFX转换为PEM

PFX格式通常使用在Windows Server中。

- 运行以下命令提取证书：

```
openssl pkcs12 -in certname.pfx -nokeys -out cert.pem
```

- 运行以下命令提取私钥：

```
openssl pkcs12 -in certname.pfx -nocerts -out key.pem -nodes
```

8.6. 替换证书

为避免证书过期对您的服务产生影响，请在证书过期前替换证书。

操作步骤

1. 新建并上传一个新的证书。
具体操作，请参见[上传证书](#)。
2. 在HTTPS监听中配置新的证书。
具体操作，请参见[添加HTTPS监听](#)。
3. 打开证书管理页面，找到过期目标证书，然后单击删除。
4. 在弹出的对话框中，单击确定。