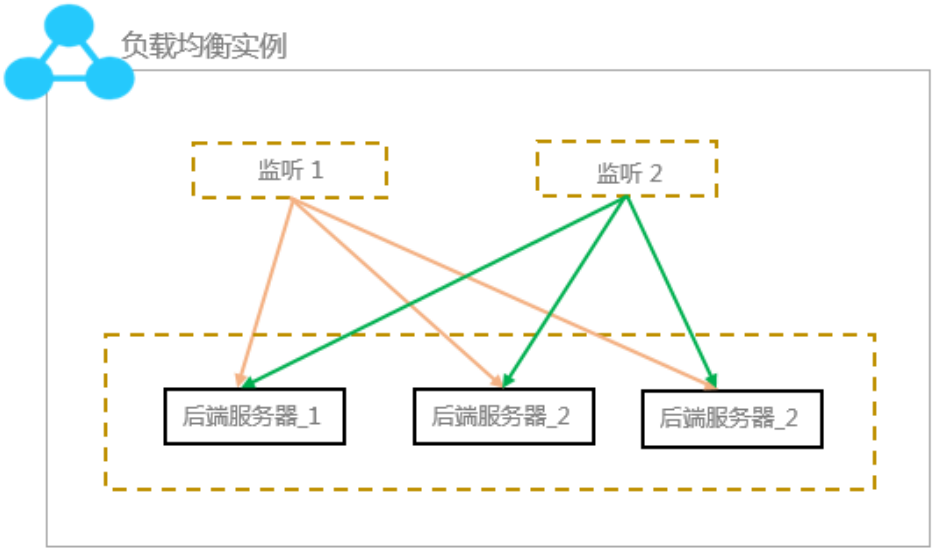
# 负载均衡

## 用户指南

# 用户指南

## 负载均衡实例

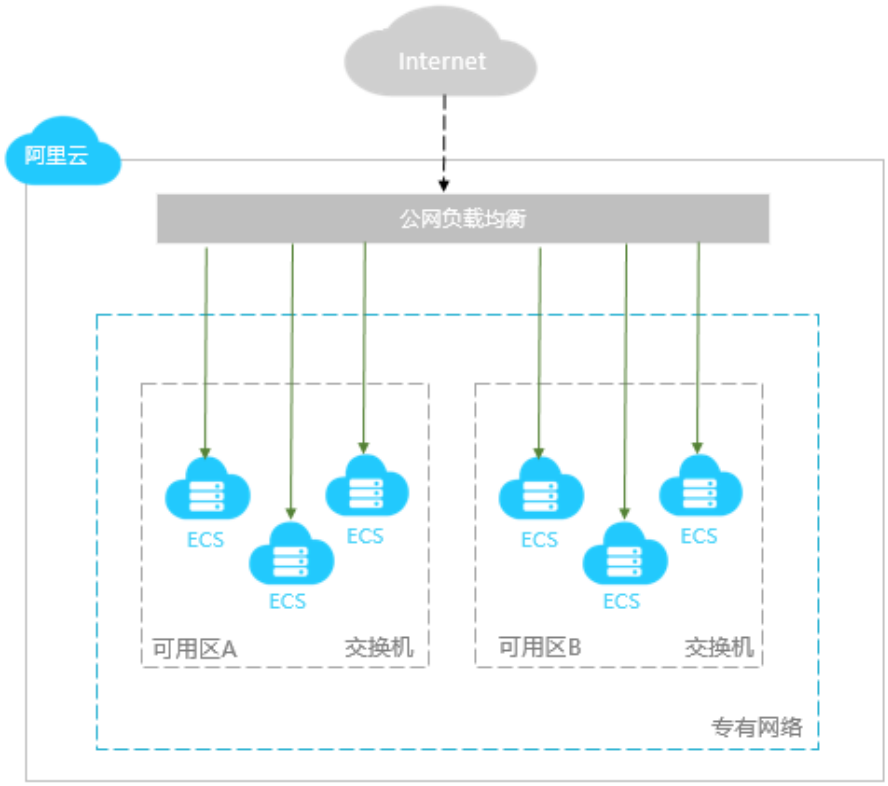
负载均衡实例是一个运行的负载均衡服务实体。要使用负载均衡服务，您必须创建一个负载均衡实例，然后在实例中添加监听和后端服务器。详情参考创建负载均衡实例。




阿里云提供公网和私网两种类型的负载均衡服务。您可以根据业务场景来选择配置对外公开或对内私有的负载均衡，系统会根据您的选择分配公网或私网服务地址。

# 公网负载均衡实例

公网类型的负载均衡实例可以通过Internet将客户端请求按照您制定的监听规则分发到添加的后端服务器ECS上。在您创建公网负载均衡实例后，系统会为其分配一个公网服务地址，您可以将您的域名和该公网服务地址进行绑定，对外提供服务。




# 私网负载均衡实例

私网类型的负载均衡实例只能在阿里云内部使用，可以转发的请求只能来自对负载均衡的私网具有访问权限的客户端。

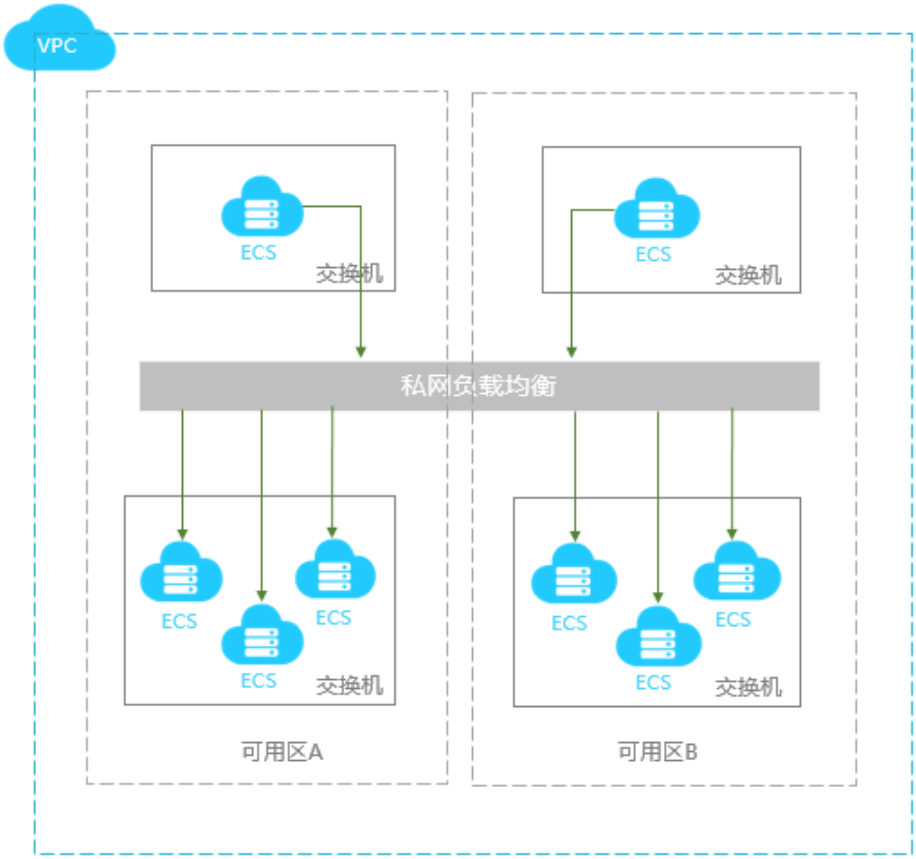
私网负载均衡实例可以进一步对网络类型进行选择：

**经典网络**

如果您选择的私网负载均衡实例的网络类型是经典网络，那么您的私网负载均衡实例的服务地址由阿里云统一分配和管理。该私网负载均衡服务只能被相同地域的经典网络ECS实例访问。

**专有网络**

如果您选择的私网负载均衡实例的网络类型是专有网络，那么您的私网负载均衡实例的服务地址会从您指定的专有网络的交换机网段内分配。该私网负载均衡服务只能被相同VPC内的ECS实例访问。




# 什么是性能保障型实例

性能保障型实例提供了可保障的性能指标（性能SLA）。与之相对的是性能共享型实例，即不保障实例的性能指标，资源是所有实例共享的。

在阿里云负载均衡推出性能保障型实例之前，您所有购买的实例均为性能共享型实例。在控制台上，您可以查看已购实例的类型。

把鼠标移至性能保障型实例的绿色图标，可查看具体的性能指标，如下图所示。


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/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。

阿里云负载均衡性能保障型实例提供了如下六种实例规格。

| 规格 | | 最大连接数 | 每秒新建连接数(CPS) | 每秒查询数(QPS) |
|---|---|---|---|---|
| 规格 1 | 简约型I (slb.s1.small) | 5000 | 3000 | 1000 |
| 规格 2 | 标准型I (slb.s2.small) | 50000 | 5000 | 5000 |
| 规格 3 | 标准型II (slb.s2.medium) | 100000 | 10000 | 10000 |
| 规格 4 | 高阶型I (slb.s3.small) | 200000 | 20000 | 20000 |
| 规格 5 | 高阶型II (slb.s3.medium) | 500000 | 50000 | 30000 |
| 规格 6 | 超强型I (slb.s3.large) | 1000000 | 100000 | 50000 |

# 性能保障型实例的变配操作限制

您可在控制台对性能保障型实例进行变配，如下图所示。


按量付费的性能保障型实例的规格可以升配也可以降配，包年包月的性能保障型实例目前暂时只允许升配，不允许降配。

因此，建议您先使用按量付费的实例进行业务测试，确认好规格后再购买所需规格的包年包月实例。

另外，变更性能保障型实例规格时，如果同时变更计费方式(按流量计费或按带宽计费)，则规格变更需要到次日零点才能生效。如果仅仅是对实例规格进行变更，变更立即生效。建议您在变更规格时，尽量不要变更计费方式。

> **注意**：由于历史存量原因，部分实例可能存在于较老的集群。此部分实例在变配到性能保障型实例时，因为需要将实例迁移，因此可能出现10-30秒的业务中断，因此建议在业务低谷期进行此类变配，或通过GSLB来做实例间的负载均衡后，再进行变配。


# 性能保障型实例的定价

性能保障型实例根据不同规格收取规格费，如下表所示。阿里云负载均衡提供了一种免费的规格，可满足以前绝大部分共享型实例用户的需求。

> **注意**：下表中所列的只是规格费用。除规格费以外，负载均衡实例的实例配置费用和流量费保持不变。更多详细信息，参考计费说明。

| 规格 | 最大连接数 | 每秒新建连接数 (CPS) | 每秒查询数 (QPS) | 包年包月 | | 按量付费 |
|---|---|---|---|---|---|---|
| | | | | 月价(元/月) | 年价(元/年) | 小时价(元/时) |

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 规格 1 | 简约型I (slb.s1.small) | 5000 | 3000 | 1000 | 免费 | 免费 | 免费 |
| 规格 2 | 标准型I (slb.s2.small) | 50000 | 5000 | 5000 | 190.00 | 1,938.00 | 0.32 |
| 规格 3 | 标准型II (slb.s2.medium) | 100000 | 10000 | 10000 | 380.00 | 3,876.00 | 0.63 |
| 规格 4 | 高阶型I (slb.s3.small) | 200000 | 20000 | 20000 | 760.00 | 7,752.00 | 1.27 |
| 规格 5 | 高阶型II (slb.s3.medium) | 500000 | 50000 | 30000 | 1,143.00 | 11,658.60 | 1.91 |
| 规格 6 | 超强型I (slb.s3.large) | 1000000 | 100000 | 50000 | 1,908.00 | 19,461.60 | 3.18 |

# 前提条件

在您创建负载均衡实例前，确保您已经做好了相关规划，详情参考**规划和准备**。

# 操作步骤

登录**负载均衡管理控制台**。

在**实例管理**页面，单击右上角的**创建负载均衡**。

在购买页面选择一种付费方式。本教程选择**按量付费**。

参考**计费说明**了解负载均衡的计费模式。

根据如下信息，配置负载均衡实例。

| **配置** | | **说明** |
|---|---|---|
| 基本配置 | 地域 | 选择负载均衡实例的所属地域。<br>**注意**：确保负载均衡实例的地域和后端添加的云服务器ECS的地域相同。 |

| | | |
|---|---|---|
| | 可用区类型 | 显示所选地域的可用区类型。云产品的可用区指的是一套独立的基础设施，常用数据中心IDC表示。不同的可用区之间具有基础设施（网络、电力、空调等）的独立性，就是说一个可用区的基础设施故障不影响另外一个可用区。可用区是属于某个地域的，一个地域下可能有一个或者多个可用区。负载均衡已经在大部分地域部署了多可用区。<br>- 单可用区：负载均衡实例只部署在一个可用区上。<br>- 多可用区：负载均衡实例会部署在两个可用区上。默认启用主可用区的实例。当主可用区出现故障时，将会自动切换到备可用区继续提供负载均衡服务，可以大大提升本地可用性。 |
| | 主可用区 | 选择负载均衡实例的主可用区，主可用区是当前承载流量的可用区。 |
| | 备可用区 | 选择负载均衡实例的备可用区。备可用区默认不承载流量，主可用区不可用时才承载流量。 |
| 网络与实例类型 | 实例规格 | 选择一个性能规格。<br>不同的性能规格所提供的性能指标也不同，详情查看**如何使用性能保障型实例？**。 |
| | 实例类型 | 根据业务场景选择配置对外公开或对内私有的负载均衡服务，系统会根据您的选择分配公网或私网服务地址。更多详细信息，参考**实例与网络类型**。<br>- 公网：公网负载均衡实例仅提供公网IP，可以通过Internet访问负载均衡。<br>- 私网：私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。 |
| | 网络类型 | 如果您选择的实例类型是私网，您还需要选择该负载均衡实例的网络类型。<br>- 经典网络：经典网络的负 |

| | | |
|---|---|---|
| | | 载均衡实例的服务地址由阿里云统一分配和管理。<br>- 专有网络：专有网络的负载均衡实例的服务地址会从您指定的专有网络的交换机网段内分配。 |
| | 计费方式 | 选择一种计费方式。 |
| 购买量 | 购买数量 | 选择购买数量。 |

单击**立即购买**。

在**确认订单**页面，核对配置信息，单击**去开通**完成创建。

在 **实例管理**页面，选择负载均衡实例的所属地域，您可以查看该地域的所有负载均衡实例。此外，您还可以：

**修改负载均衡实例名称**

将光标移至负载均衡ID区域，单击出现的铅笔图标，输入实例名称。

**暂停负载均衡实例**

勾选负载均衡实例，单击页面下方的**停止**，或单击**更多>停止**。

**启动暂停的负载均衡实例**

勾选已经停止运行的负载均衡实例，单击页面下方的**启动**，或单击 **更多>启动**。

**释放负载均衡实例**

勾选负载均衡实例， 单击页面下方的**释放设置**，或单击 **更多>释放**。在**释放设置**对话框，选择立即释放或在某个特定时刻释放实例。

**设置标签**

您可以通过标签实例进行分类和统一管理。详情参考**管理标签**。

**变更计费**

预付费模式下，您可以变更购买实例的带宽规格，但只支持升级带宽，不支持降低带宽：单

击**更多>变更带宽规格**。

后付费模式下，您可以在按使用流量和按公网带宽两种计费方式间切换，单击**更多>变更计费方式**。

更多详细信息，参考变配流程。

**查看负载均衡实例详情**

单击负载均衡实例的ID链接或**管理**，查看负载均衡实例详情。

在详情页面，您可以单击**消费明细**，查看负载均衡服务的费用明细。


单击**监听**，查看或添加负载均衡监听。详情参考监听介绍。

单击**服务器**，查看或添加后端服务器。详情参考后端服务器概述。

单击**监控**， 查看监控信息，设置报警机制。详情参考设置报警规则。

您可以更改后付费实例的带宽、实例规格和计费方式。

完成以下操作，更改后付费实例的配置：

登录负载均衡管理控制台。

选择目标实例的所属地域。

找到目标实例，然后单击**更多>变更配置**。


在**配置变更**区域，选择新的带宽值、实例规格或计费方式后，完成支付。

您可以更改按固定带宽计费实例的带宽。

在变更带宽时，您还可以为实例中的每个监听指定一个带宽峰值，监听带宽峰值总和不能大于实例的带宽值。如果不开启带宽峰值限定，那么该实例下的所有监听共享指定的带宽。

后付费实例支持按流量计费和按带宽计费。您可以变更后付费实例的计费方式，计费方式的变更会在次日零点生效。

您可以更改性能保障型实例的规格，变更实时生效。

**说明**：如果您更改实例规格时，也变更了计费方式（按带宽计费和按流量计费的变更），那么规格的变更会同计费方式的变更一起在次日零点生效。

您可以升高或降低包年包月实例的带宽，也可以升高或降低其性能规格。

**说明**：您需要提交工单开通白名单后，方可使用降配功能。

完成以下操作，更改包年包月实例配置：

登录负载均衡管理控制台。

选择目标实例的所属地域。

找到目标实例，然后单击**更多**>**降配**或**升配**。


在**配置变更**区域，选择新的带宽值或规格，完成支付。


# 标签概述

负载均衡提供标签管理功能，方便您通过对负载均衡实例添加标签进行负载均衡服务分类。

每个标签都由一对键值对组成，负载均衡标签的使用限制如下：

目前不支持未绑定实例的空标签存在，标签必须绑定在某个负载均衡实例上。

一个实例最多可以绑定10个标签。

一个实例上的每个标签的标签键必须唯一，相同标签键的标签会被覆盖。

每个地域中的的标签信息不互通，例如在华东1地域创建的标签在华东2地域不可见。

# 添加标签

登录负载均衡管理控制台。

在**实例管理**页面，选择地域，然后勾选需要添加同一标签的实例。

单击**编辑标签**。

在**编辑标签**窗口，单击**新建标签**，然后输入新建标签的标签键和值，单击**确定**。

# 标签搜索实例

登录负载均衡管理控制台。

在**实例管理**页面，选择地域，查看该地域的所有实例。

单击**标签**，然后选择要搜索的实例绑定的标签键和标签值。

符合您选择条件的实例会显示在实例列表。

单击标签旁边已选的标签键的删除图标，清除标签过滤条件。

## 移除实例标签

负载均衡不支持批量删除多个实例的标签，您只能单独对某一个实例进行标签移除。

登录负载均衡管理控制台。

在**实例管理**页面，选择地域，查看该地域的所有实例。

选择要删除标签的实例，单击**更多** > **编辑标签**。

在**编辑标签**窗口，单击要移除的标签的删除图标，然后单击**确定**。

注意：当一个标签从一个实例上移除后，如果该标签没有和其他实例绑定，系统会将该标签删除。


# 后端服务器

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。

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。

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。默认后端服务器是在实例维度上维护的，即负载均衡实例下的所有监听都只能够将流量转发到相同服务器的相同端口上。

您也可以通过服务器组的方式添加ECS。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的不同监听就可以将请求转发给不同的服务器组内不同端口的ECS。

> **注意**：如果您在配置监听时，选择使用服务器组，那么该监听会将请求转发到关联的服务器组中的ECS，而不会再将请求转发给后端服务器池中的ECS。

# 主备服务器组

当您有传统的主备需求时，即后端服务器中有一台主机和一台备机。当主机工作正常时，流量将直接走主机；当主机宕机时，流量将走到备机。此时，可以使用主备服务器组，避免服务中断。

由于备机不会做健康检查，所以只要主机健康检查失败，系统会直接将流量切到备机。当主机健康检查成功恢复服务后，流量会自动切到主机。

主备服务器组是在监听维度上维护的，并且只支持四层监听，详情参考创建主备服务器组。

# 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。详情参考创建虚拟服务器组。

# 注意事项

负载均衡不支持跨地域部署，确保ECS实例的所属地域和负载均衡实例的所属地域相同。

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。负载均衡监听的前端端口对应后端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，您可以根据后端ECS实例的对外服务能力和情况来区别设定。

**注意**：如果您同时开启了会话保持功能，那么有可能会造成对后端应用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，建议您可以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，可以参考以下方法检查处理：

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。

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。（如设置会话保持，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。）

# 前提条件

您已创建负载均衡实例。

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。如果您以前未使用过ECS，参考快速入门（Linux）和 快速入门（Windows）创建ECS实例。

# 操作步骤

登录负载均衡管理控制台。

在**实例管理**页面，选择目标实例的所属地域。

单击目标实例的ID链接，进入负载均衡实例的详情页面。

在左侧导航栏，单击**服务器** > **后端服务器**。

在**负载均衡服务器池**页面，单击**未添加的服务器**页签。

单击目标ECS实例对应的**添加**，或者勾选多个目标ECS实例，然后单击页面下方的**批量添加**。

**注意**： ECS实例的网络类型要和该负载均衡实例的类型匹配。详情查看负载均衡实例类型。

- 经典网络的公网负载均衡实例，可添加经典网络类型的ECS或者同属于同一VPC的ECS；

- 专有网络的私网负载均衡实例，仅能添加和负载均衡实例相同VPC内的ECS；
- 经典网络的私网负载均衡实例，仅能添加经典网络类型的ECS。

在**添加后端服务器**对话框，指定添加的ECS实例的权重，然后单击**确定**。

权重越高的ECS实例将被分配到更多的访问请求。您可以根据后端ECS实例的对外服务能力和情况来区别设定。

**注意**：权重设置为0，该服务器不会再接受新请求。

添加后的实例会显示在**已添加的服务器**页签下，您可以移除或者修改添加的ECS实例的权重。

当您有传统的主备需求时，即后端服务器中有一台主机和一台备机，可选择使用主备服务器组。当主机正常工作时，流量将直接走主机；当主机不可用时，流量将走到备机，避免服务中断。

主备服务器组和虚拟服务器组都是在监听维度上维护的，即实例下的不同监听可将流量转发到不同的服务器组。但是一个虚拟服务器组可以添加多个ECS实例，而一个主备服务器组只允许添加两个ECS实例，其中一个作为主机，另外一个作为备机。

**注意**：主备服务器组只支持四层监听（TCP和UDP协议）。

# 前提条件

您已创建负载均衡实例。

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。如果您以前未使用过ECS，参考快速入门（Linux）和 快速入门（Windows）创建ECS实例。

# 操作步骤

登录负载均衡管理控制台。

在**实例管理**页面，选择目标实例的所属地域。

单击目标实例的ID链接，进入负载均衡实例的详情页面。

在左侧导航栏，单击**服务器** > **主备服务器组**。

在**主备服务器组**页面，单击**创建主备服务器组**。

在**创建主备服务器组**对话框，完成如下操作：

在**分组名称**文本框中，输入主备服务器组名称。

选择目标ECS实例的网络类型。

**注意**： ECS实例的网络类型要和该负载均衡实例的类型匹配。详情查看负载均衡实例类型。

i. 经典网络的公网负载均衡实例，可添加经典网络类型的ECS或者同属于同一VPC的ECS；
ii. 专有网络的私网负载均衡实例，仅能添加和负载均衡实例相同VPC内的ECS；
iii. 经典网络的私网负载均衡实例，仅能添加经典网络类型的ECS。

在**可选服务器列表**中，单击目标ECS实例。

在**已选服务器列表**中，输入ECS实例的端口和权重，并选择作为主机使用的ECS实例，单击**确定**。

# 概述

虚拟服务器组（VServer group）是一组 ECS 实例。虚拟服务器组允许您在监听维度上个性化定义服务器组，即实例下的不同监听可使用不同的后端服务器组，可满足域名和 URL 转发个性化需求。

如果您没有创建虚拟服务器组，负载均衡实例会将请求按照您设置的权重和监听规则转发给所有添加的后端服务器（ECS 实例）；如果您创建了虚拟服务器组，负载均衡实例会将请求按照您设置的监听规则转发给关联的虚拟服务器组，默认实例维度添加的独立的后端服务器不再接收请求。

**注意**：如果您在一个负载均衡实例的 HTTP/HTTPS 监听下，添加了后端服务器、虚拟服务器组和转发规则，请求转发的顺序如下：

判断请求其是否能够匹配上某条转发规则，如果匹配，则将流量转发到该规则的虚拟服务器组。

若不匹配并且在该监听上设置了虚拟服务器组，那么将流量转发到监听关联的虚拟服务器组。

若您没有在该监听上设置虚拟服务器组，即将流量转发到实例级别添加的各后端服务器。

**使用限制**

虚拟服务器组只能添加监听所在地域的后端服务器。

一个后端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。

虚拟服务器组由服务器+端口组成。

# 前提条件

您已经创建了负载均衡实例。

您创建了 ECS 实例，用来接收转发的请求。如果您以前未使用过阿里云 ECS，参考 快速入门（Linux）或 快速入门（Windows）创建 ECS 实例。

# 操作步骤

登录 负载均衡管理控制台。

在 **实例管理** 页面，选择地域，查看该地域下的所有负载均衡实例。

单击负载均衡实例的 ID 链接，进入负载均衡的详情页面。

在负载均衡实例菜单栏，单击 **服务器** > **虚拟服务器组**。

在 **虚拟服务器组** 页面，单击 **创建虚拟服务器组**。

在 **创建虚拟服务器组** 对话框，完成如下操作：

在 **分组名称** 文本框中，输入虚拟服务器组名称。

选择您要添加的 ECS 实例的网络类型。

**注意**： ECS 实例的网络类型要和负载均衡实例的类型匹配，如果不匹配，添加按钮则会灰掉。详情查看 负载均衡实例类型 。

i. 经典网络的公网负载均衡实例，可添加经典网络 ECS 实例或者同一 VPC 实例 ID 的 ECS 实例;
ii. 专有网络的私网负载均衡实例，仅能添加同一 VPC 实例 ID 的 ECS 实例;
iii. 经典网络的私网负载均衡实例，仅能添加经典网络的 ECS 实例。

在 **可选服务器列表** ，单击要添加的实例。

选中添加的实例会显示在右侧的 **已选服务器列表** 中。

在 **已选服务器列表**，输入每个 ECS 实例的端口和权重，单击 **确定**。

创建的虚拟服务器组会显示在 **虚拟服务器组** 页面，您可以删除或者增加虚拟服务器的 ECS 实例数量 （单击 **编辑**）。您也可以将该虚拟服务器组和实例下的监听或者转发策略关联。


# 监听

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。

负载均衡监听包括监听配置和健康检查配置。

## 监听配置

负载均衡提供四层（TCP/UDP协议）和七层（HTTP/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 | - 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<br>- 基于源地址的会话保持<br>- 在网络层可直接看到 | - 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<br>- 无特殊要求的Web应 |

| | | |
|---|---|---|
| | 来源地址<br>- 数据传输快 | 用 |
| UDP | - 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<br>- 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 | - 应用层协议，主要解决如何包装数据<br>- 基于Cookie的会话保持<br>- 使用X-Forward-For获取源地址 | 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 | - 加密传输数据，可以阻止未经授权的访问<br>- 统一的证书管理服务，用户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 |

**说明**：负载均衡已在欧洲中部（法兰克福）与亚太东南3（吉隆坡）地域支持HTTP/2和WSS/WS协议，详情参见HTTP/2协议支持常见问题和WS/WSS协议支持常见问题。

# 健康检查配置

负载均衡对后端服务器提供健康检查，提高服务的可用性。详情参见健康检查原理和健康检查配置。


# 四层监听

## 四层监听概述

阿里云提供四层（TCP协议和UDP协议）的负载均衡服务。四层监听将请求直接转发到后端ECS实例，而且不修改请求标头。

### TCP协议

TCP是面向连接的协议，在正式收发数据前，必须和对方建立可靠的连接。TCP协议适用于注重可靠性，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、远程登录和无特殊要求的Web应用。

### UDP协议

UDP是面向非连接的协议，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，可靠性相对低但数据传输快。UDP协议多用于关注实时性而相对不注重可靠性的场景，如视频聊天、金融实时行情推送等。

UDP协议监听有如下限制：

- 每个监听最大连接数限制：100,000。
- 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。

# 四层监听配置

| 监听配置 | 说明 |
| --- | --- |
| 前端协议 [端口] | 用来接收请求并向后端服务器进行请求转发的前端协议和端口。<br>配置四层监听，协议选择TCP或UDP，端口为1-65535。<br>**注意**： 在同一个负载均衡实例内，前端端口不可重复。 |
| 后端协议 [端口] | 后端服务器（ECS实例）开放用来接收请求的后端端口。<br>后端的协议类型和前端相同，端口为1-65535。<br>**注意**：在同一个负载均衡实例内，后端端口可重复。 |
| 带宽峰值 | 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。<br>当不限制监听带宽时，各监听共享实例的总带宽。更多详细信息，参考**共享实例带宽**。 |
| 调度算法 | 负载均衡支持轮询、加权轮询（WRR）、加权最小连接数（WLC）三种调度算法。<br>- 轮询：按照访问顺序依次将外部请求依序分发到后端服务器。<br>- 加权轮询：权重值越高的后端服务器，被轮询到的次数（概率）也越高。<br>- 加权最小连接数：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。 |
| 使用服务器组 | 选择是否使用服务器组。使用服务器组，可以在监听维度上个性化定义服务器组，即实例下的不同监听可使用不同的服务器组。<br>**注意**：使用服务器组后，该监听会将流量转发到选择的服务器组，实例维度的后端服务器不再生效。如果不开启服务器组，监听会将流量转发到后端服务器池内添加的服务器上。详情参考添加后端服务器。 |
| 服务器组类型 | 如果选择使用服务器组，选择您要使用的服务器组类型：<br>- 虚拟服务器组：一个虚拟服务器组（VServer group）由多个后端服务器组成，且后端服务器的端口可以不同。您可以为不同的监听配置不同的虚拟服务器组，这样就可以将请求转发至不同的后端服务器。详情参考创建虚拟服务器组。<br>- 主备服务器组：一个主备服务器组由两台后端服务器组成，即一台主服务器，一台备服务器。当您有传统的主备需求时，可以使用主备服务器组。当主机工作正常时，将流量转发至主服务器；当主机宕机时，会将流量转发至备服务器，避免服务中断。详情参考创建主备服务器组。 |
| 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。 |
| **高级配置** | |
| 获取真实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。<br>**注意**：经典网络的负载均衡实例的UDP监听暂不支持查看源地址。 |
| 会话保持 | 是否开启会话保持。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。<br>针对TCP监听，负载均衡是基于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。<br>**注意**：UDP监听不支持会话保持。 |
| 连接超时时间 | 指定TCP连接的超时时间。可选值为10-900秒。 |

| | |
|---|---|
| | **注意**：此配置只适用于TCP监听。 |

# 七层监听

## 七层监听概述

阿里云提供七层（HTTP协议和HTTPS协议）的负载均衡服务，负载均衡七层监听原理上是反向代理的一种实现。客户端HTTP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协议访问后端服务器，而不是直接转发报文到后端服务器。

**HTTP协议**

HTTP是应用层协议，主要解决如何包装数据。适用于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协议**

HTTPS是以安全为目标的HTTP通道，即HTTP下加入SSL层来保证数据安全。负载均衡支持HTTPS单向和双向认证。提供证书管理功能，无需在后端服务器上进行证书配置，详情查看配置HTTPS监听。

## 七层监听配置

| 监听配置 | 说明 |
|---|---|
| 前端协议 [端口] | 用来接收请求并向后端服务器进行请求转发的前端协议和端口。<br>配置七层监听，协议选择HTTP或HTTPS，端口为1-65535。<br>**注意**：在同一个负载均衡实例内前端端口不可重复。 |
| 后端协议 [端口] | 后端服务器（ECS实例）开放用来接收请求的后端端口。<br>后端的协议类型为HTTP，端口为1-65535。<br>**注意**：在同一个负载均衡实例内后端端口可重复。 |
| 带宽峰值 | 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。 |

| | |
|---|---|
| | 当不限制监听带宽时，各监听共享实例的总带宽。更多详细信息，参考共享实例带宽。 |
| 调度算法 | 负载均衡支持轮询、加权轮询（WRR）、加权最小连接数（WLC）三种调度算法。<br>- 轮询：按照访问顺序依次将外部请求依序分发到后端服务器。<br>- 加权轮询：权重值越高的后端服务器，被轮询到的次数（概率）也越高。<br>- 加权最小连接数：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。 |
| 使用服务器组 | 开启配置后，可以在监听维度上个性化定义服务器组，即实例下的不同监听可使用不同的后端服务器组。<br>一个虚拟服务器组（VServer group）由多个后端服务器组成，且后端服务器的端口可以不同。您可以为不同的监听配置不同的虚拟服务器组，这样就可以将请求转发至不同的后端服务器。详情参考创建虚拟服务器组。<br>**注意**：使用服务器组后，该监听会将流量转发到选择的服务器组，实例维度的后端服务器不再生效。如果不开启服务器组，监听会将流量转发到后端服务器池内添加的服务器上。详情参考添加后端服务器。 |
| 双向认证 | 开启该配置后支持在服务端和客户端进行HTTPS双向认证，您需要上传服务器证书和CA证书。<br>不开启，单向认证只需上传服务器证书。<br>**注意**：该选项只适用于HTTPS监听。 |
| 服务器证书 | 用于用户浏览器检查服务器发送的证书是否是由自己信赖的中心签发的。<br>服务器证书可以到阿里云云盾证书服务购买，也可以到其它服务商购买。服务器证书需要上传到负载均衡的证书管理系统。详情参考上传证书。<br>**注意**：该选项只适用于HTTPS监听。 |
| CA证书 | 服务器用CA证书验证收到的客户端证书。如果没有通过验证，拒绝连接。开启双向认证功能后，CA证书和服务器证书都需要上传到负载均衡的证书管理系统。详情参考生成证书。<br>**注意**：该选项只适用于HTTPS监听。 |
| 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。 |
| **高级配置** | |

| | |
|---|---|
| 获取真实IP | 针对七层服务，负载均衡通过HTTP Header: X-Forwarded-For获取来访者真实IP。详情参考**获取来访者真实IP**。 |
| 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。<br>针对七层（HTTP协议和HTTPS协议）监听，负载均衡使用Cookie进行会话保持。负载均衡提供了两种Cookie处理方式：<br>- 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/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。<br>- 重写Cookie：可以根据需要指定HTTPS/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。详情参考**会话保持规则配置**。<br>查看**会话保持常见问题**了解更多信息。 |
| Gzip数据压缩 | 开启该配置对特定文件类型进行压缩。<br>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、application/xml。 |
| 附加HTTP头字段 | 选择您要添加的自定义HTTP header字段：<br>- X-Forwarded-For: 添加该字段获取客户端的IP地址。<br>- X-Forwarded-Proto: 添加该字段获取客户端与监听连接时所用的协议 (HTTP或HTTPS)。<br>- SLB-IP: 添加该字段获取负载均衡实例的公网IP。<br>- SLB-ID: 添加该字段获取负载均衡实例的ID。 |

为了满足数据传输的安全需求，负载均衡提供了HTTPS监听，支持单向和双向认证。

在使用HTTPS监听时，注意：

在使用HTTPS监听前，您需要将需要的证书上传到负载均衡系统。详情查看上传证书。

| 证书 | 说明 | 单向认证是否需要 | 双向认证是否需 |
| --- | --- | --- | --- |
| 服务器证书 | 用来证明服务器的身份。<br>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<br>服务器证书需要上传到负载均衡的证书管理系统。 | 是<br>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。<br>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<br>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<br>服务器证书需要上传到负载均衡的证书管理系统。 |

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PEM证书文件中含BEGIN DH PARAMETERS字段的字串上传。更多详细信息，参考证书要求。

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SNI功能。

HTTPS监听的会话ticket保持时间设置为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。

本指南提供配置HTTPS监听（单向认证）的完整教程。完成以下三个任务完成配置：

1. 上传服务器证书
2. 配置负载均衡实例
3. 测试负载均衡服务

# 上传服务器证书

在配置HTTPS监听（单向认证）前，您需要购买服务器证书，并将服务器证书上传到负载均衡的证书管理系统。上传后，无需在后端ECS上进行其它证书配置。

登录负载均衡管理控制台。

在左侧导航栏，单击**证书管理**，然后单击**创建证书**。

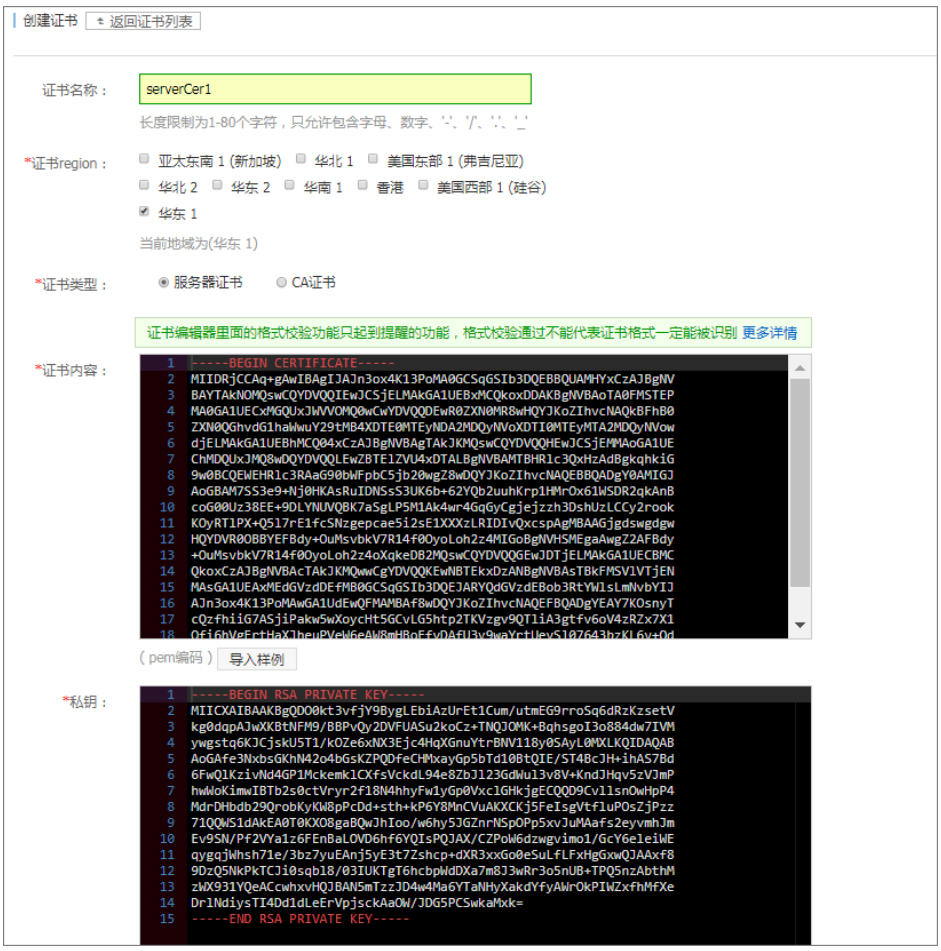
按照以下信息，配置证书：

证书Region: 选择**华东 1**。

**注意**：证书的地域和负载均衡实例的地域要相同。

证书类型：选择**服务器证书**。

证书内容和私钥：复制服务器证书的内容和私钥。单击**导入样例**查看合法的证书格式。上传的证书必须是PEM格式，详情查看**证书格式要求**。




单击**确定**完成上传。

# 配置负载均衡实例

登录负载均衡管理控制台。

在**实例管理**页面，单击**创建负载均衡**。

配置负载均衡实例，单击**立即购买**完成支付。

注意：网络类型选择：**公网**，地域选择**华东1**。详细配置信息参考创建负载均衡实例。

创建成功后，返回**实例管理**页面，单击**华东1**地域，然后单击已创建的负载均衡实例ID链接。

在**详情**左侧导航栏，单击**监听**，然后单击**添加监听**。

在**添加监听**窗口，完成如下配置。

前端协议 [端口]：HTTPS 443。

后端协议 [端口]：HTTP 80。

调度算法：轮询。

服务器证书：选择已上传的服务器证书。

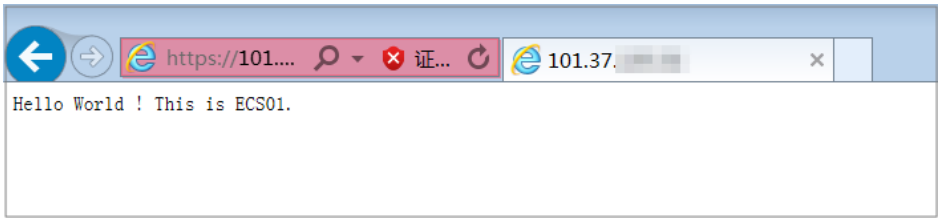

在左侧导航栏，单击**服务器** > **后端服务器**，然后单击**添加后端服务器**，添加ECS服务器。

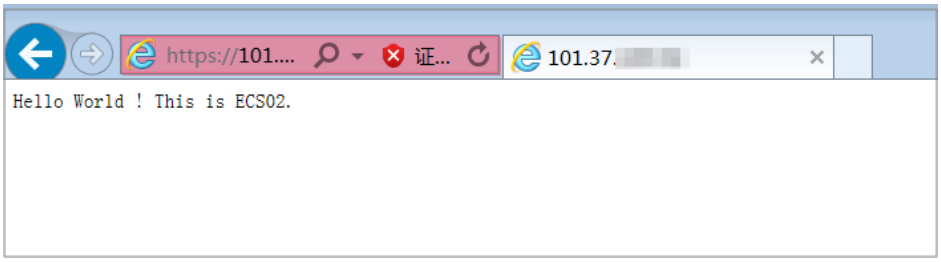
# 测试负载均衡服务

负载均衡实例配置完成后，在**实例管理**页面，查看健康检查状态。当状态为**正常**时，表示后端服务器可以正常接收处理负载均衡监听转发的请求。

在浏览器中，输入负载均衡的公网服务地址。刷新浏览器，您可以观察到请求在两台ECS服务器之间转换。

因为使用了自建的服务器证书，所以下图示例中会有不信任提示。




本指南将引导您配置HTTPS双向认证的负载均衡服务。本指南中使用自签名的CA证书为客户端证书签名。

完成以下操作配置HTTPS监听（双向认证）：

1. 准备服务器证书
2. 使用OpenSSL生成CA证书
3. 生成客户端证书
4. 上传服务器证书和CA证书
5. 安装客户端证书
6. 配置负载均衡实例
7. 测试负载均衡服务

# 步骤一: 准备服务器证书

服务器证书用于用户浏览器检查服务器发送的证书是否是由自己信赖的中心签发的，服务器证书可以到阿里云云盾证书服务购买，也可以到其他服务商处购买。

# 步骤二: 使用OpenSSL生成CA证书

运行以下命令在/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(证书备份目录)。
- **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。

在**conf**目录下新建一个包含如下信息的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
```

运行以下命令生成私钥key文件。

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

运行结果如下图所示。

```
root@iZbp1hfvivcqx1jbwap31iZ:~/ca/conf# cd /root/ca
root@iZbp1hfvivcqx1jbwap31iZ:~/ca# sudo openssl genrsa -out private/ca.key
Generating RSA private key, 2048 bit long modulus
.....+++
..+++
e is 65537 (0x10001)
```

运行以下命令并按命令后的示例提供需要输入的信息，然后回车，生成证书请求csr文件。

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

注意：Common Name请输入您的负载均衡服务的域名。

```
root@iZbp1hfvivcqx1jbwap31iZ:~/ca# sudo openssl req -new -key private/ca.key -ou
t 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@alibaba.com

Please enter the following 'extra' attributes
to be sent with your certificate request
A challenge password []:
An optional company name []:
root@iZbp1hfvivcqx1jbwap31iZ:~/ca#
```

运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

运行以下命令为CA的key设置起始序列号，可以是任意四个字符。

$ sudo echo FACE > serial

运行以下命令创建CA键库。

$ sudo touch index.txt

运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 -crldays 7 -config "/root/ca/conf/openssl.conf"

输出为：

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

# 步骤三: 生成客户端证书

运行以下命令在**ca**目录内创建一个存放客户端key的目录**users**。

$ sudo mkdir users

运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

> 注意：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。

运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，提供对应的信息。

> 注意：**A challenge password**是客户端证书口令(请注意将它和client.key的口令区分开，本教程设置密码为test)，可以与服务器端证书或者根证书口令一致。

```
root@iZbp1hfvivcqx1jbwap31iZ:~/ca# sudo openssl req -new -key /root/ca/users/cli
ent.key -out /root/ca/users/client.csr
Enter pass phrase for /root/ca/users/client.key:
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@alibaba.com

Please enter the following 'extra' attributes
to be sent with your certificate request
A challenge password []:test
An optional company name []:Alibaba
root@iZbp1hfvivcqx1jbwap31iZ:~/ca#
```

运行以下命令使用步骤二中的CA Key为刚才的客户端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@iZbp1hfvivcqx1jbwap31iZ:~/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@alibaba.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@iZbp1hfvivcqx1jbwap31iZ:~/ca#
```

运行以下命令将证书转换为大多数浏览器都能识别的PKCS12文件。

$ sudo openssl pkcs12 -export -clcerts -in /root/ca/users/client.crt -inkey /root/ca/users/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。

```
root@iZbp1hfvivcqx1jbwap31iZ:~/ca# sudo openssl pkcs12 -export -clcerts -in /roo
t/ca/users/client.crt -inkey /root/ca/users/client.key -out /root/ca/users/clien
t.p12
Enter pass phrase for /root/ca/users/client.key:
Enter Export Password:
Verifying - Enter Export Password:
root@iZbp1hfvivcqx1jbwap31iZ:~/ca#
```

运行以下命令查看生成的客户端证书。

```
cd users
ls
```

```
root@iZbp1hfvivcqx1jbwap31iZ:~/ca# cd users
root@iZbp1hfvivcqx1jbwap31iZ:~/ca/users# ls
client.crt  client.csr  client.key  client.p12
root@iZbp1hfvivcqx1jbwap31iZ:~/ca/users#
```

# 步骤四: 上传服务器证书和CA证书

登录负载均衡管理控制台。

在**实例管理**页面，单击**创建负载均衡**。

配置负载均衡实例，单击**立即购买**完成支付。

注意：网络类型选择：**公网**，地域选择**华东1**。详细配置信息参考创建负载均衡实例。

创建成功后，在**实例管理**页面，将鼠标移至实例名称区域，单击出现的铅笔图标，修改负载均衡实例名称。

在**负载均衡**左侧导航栏，单击**证书管理**，然后单击**创建证书**，上传服务器证书。

在**创建证书**页面，完成如下配置后，单击**确定**。

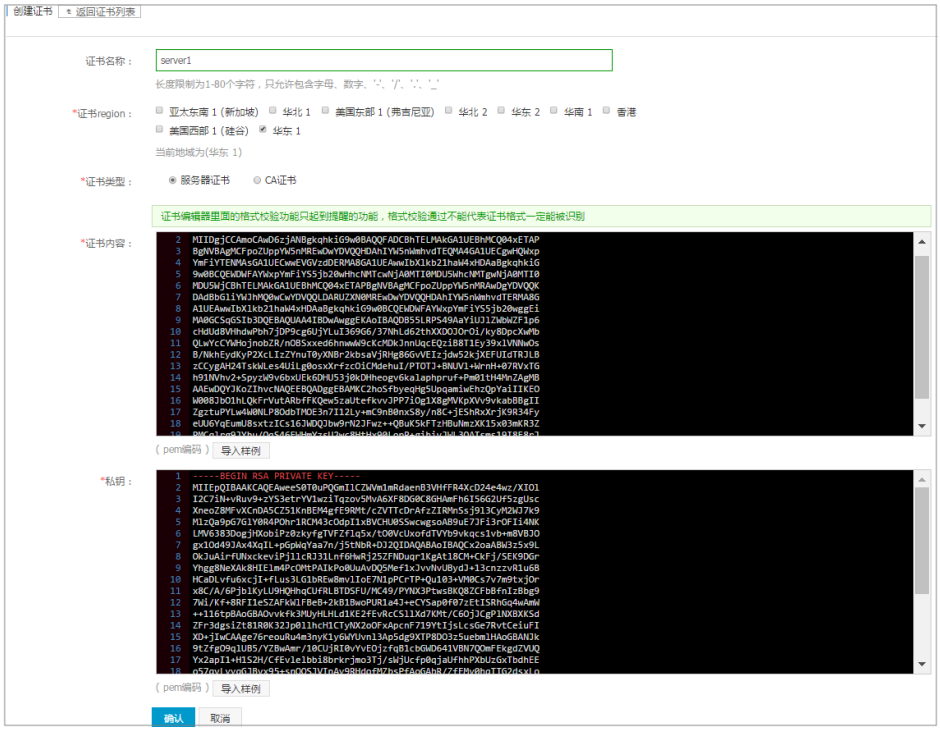
证书地域：本教程中**选择华东1**。

注意：证书的地域和负载均衡实例的地域要相同。

证书类型：选择**服务器证书**。

证书内容和私钥：复制您的服务器证书内容和私钥。

在复制内容前，您可以单击**导入样式**，查看正确的证书和私钥格式。更多详细信息查看证书要求。




在负载均衡左侧导航栏，单击**证书管理**，然后单击**创建证书**，上传CA证书。

在**创建证书**页面，完成如下配置后，单击**确定**。

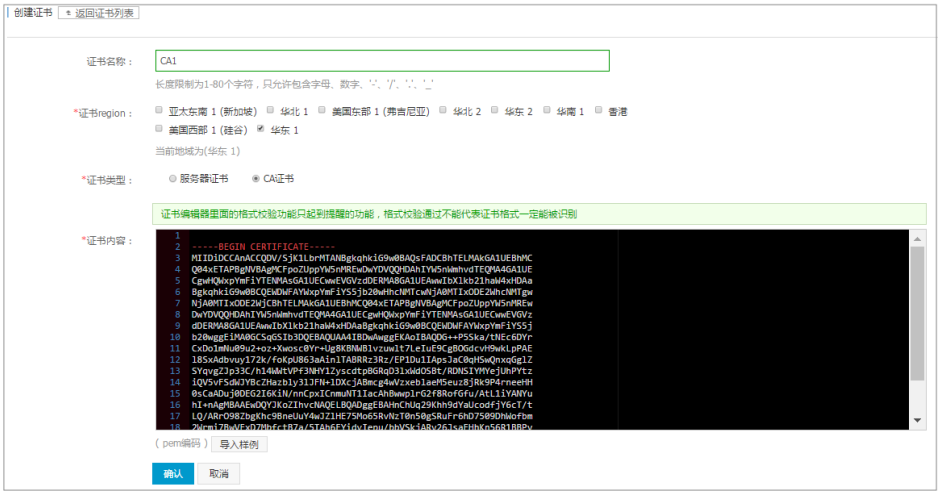
证书地域：本教程中**选择华东1**。

注意：证书的地域和负载均衡实例的地域要相同。

证书类型：选择**CA证书**。

证书内容：复制您的CA证书内容。

在复制内容前，您可以单击**导入样式**，查看正确的CA证书格式。更多详细信息查看**证书要求**。




# 步骤五: 安装客户端证书

将生成的客户端证书安装到客户端。本教程以Windows客户端，IE浏览器为例。

打开Git Bash命令行窗口，运行以下命令导出步骤三中生成的客户端证书。

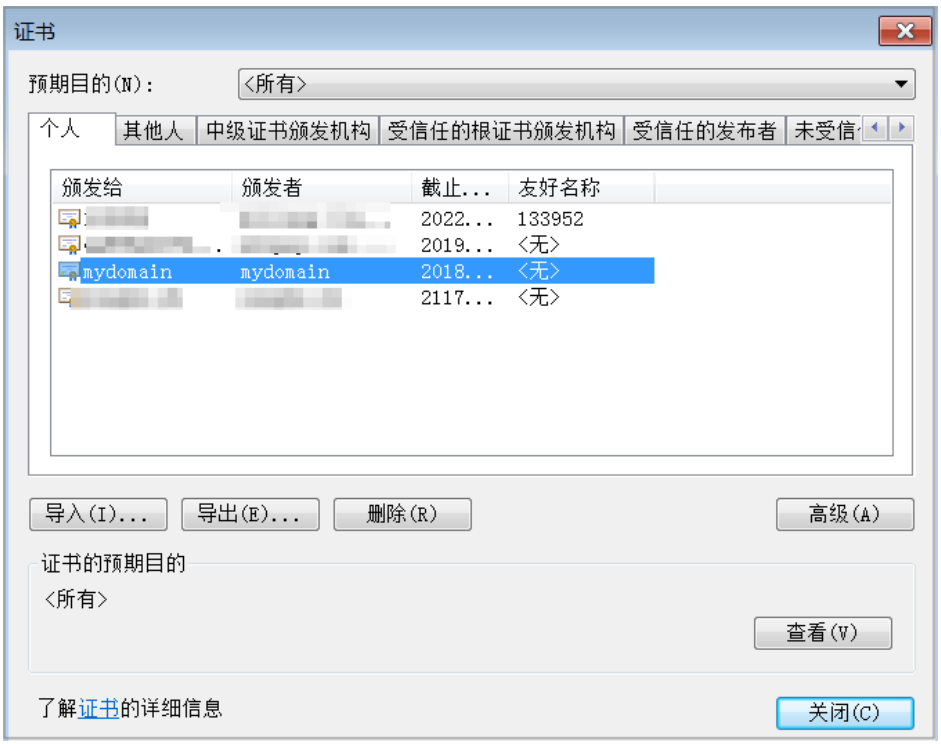
scp root@IPaddress:/root/ca/users/client.p12 ./

注意：IPaddress是生成客户端证书的服务器的IP地址。

在IE浏览器中导入下载的客户端证书。

打开IE浏览器，单击**设置** > **Internet选项**。

单击**内容**页签，然后单击**证书**，导入下载的客户端证书。在导入证书时需要输入在步骤三时生成PKCS12文件的密码。




# 步骤六: 配置HTTPS双向认证监听

登录负载均衡管理控制台。

在**实例管理**页面，单击**华东1**地域，然后单击已创建的负载均衡实例ID链接。

在**详情**左侧导航栏，单击**监听**，然后单击**添加监听**。

在**添加监听**窗口，完成如下配置。

前端协议 [端口]：HTTPS 443。

后端协议 [端口]：HTTP 80。

带宽峰值: 输入带宽峰值。

调度算法：轮询。

双向认证：开启。

服务器证书：选择已上传的服务器证书。

CA证书： 选择已上传的CA证书。

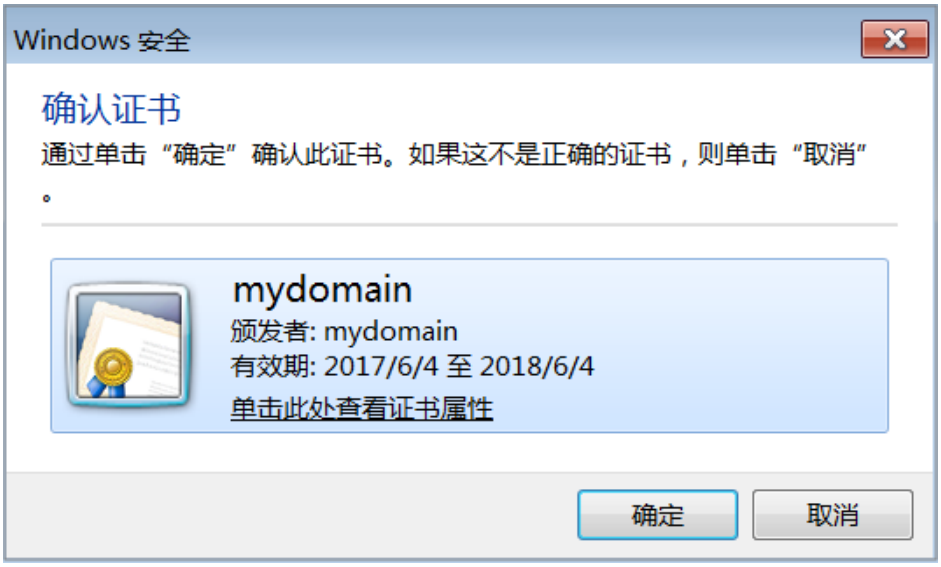
单击**下一步**，然后单击**确认**完成配置。


在**详情**左侧导航栏，单击**服务器** > **后端服务器**，然后单击**添加后端服务器**，添加ECS服务器。

# 步骤七: 测试HTTPS双向认证

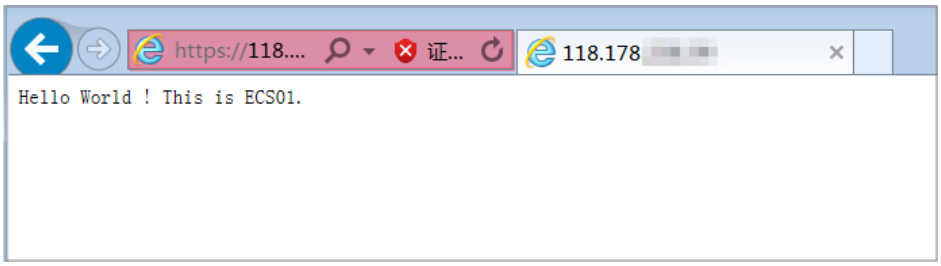
在**实例管理**页面，查看健康检查状态。当状态为**正常**时，表示后端服务器可以正常接收处理负载均衡监听转发的请求。

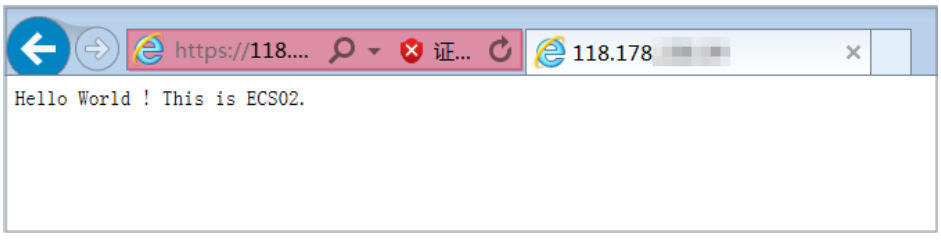
在浏览器中，输入负载均衡的公网服务地址，当提示是否信任客户端证书时，选择信任。




刷新浏览器，您可以观察到请求在两台ECS服务器之间转换。

因为使用了自建的服务器证书，所以下图示例中会有不信任提示。




# 域名或URL转发规则

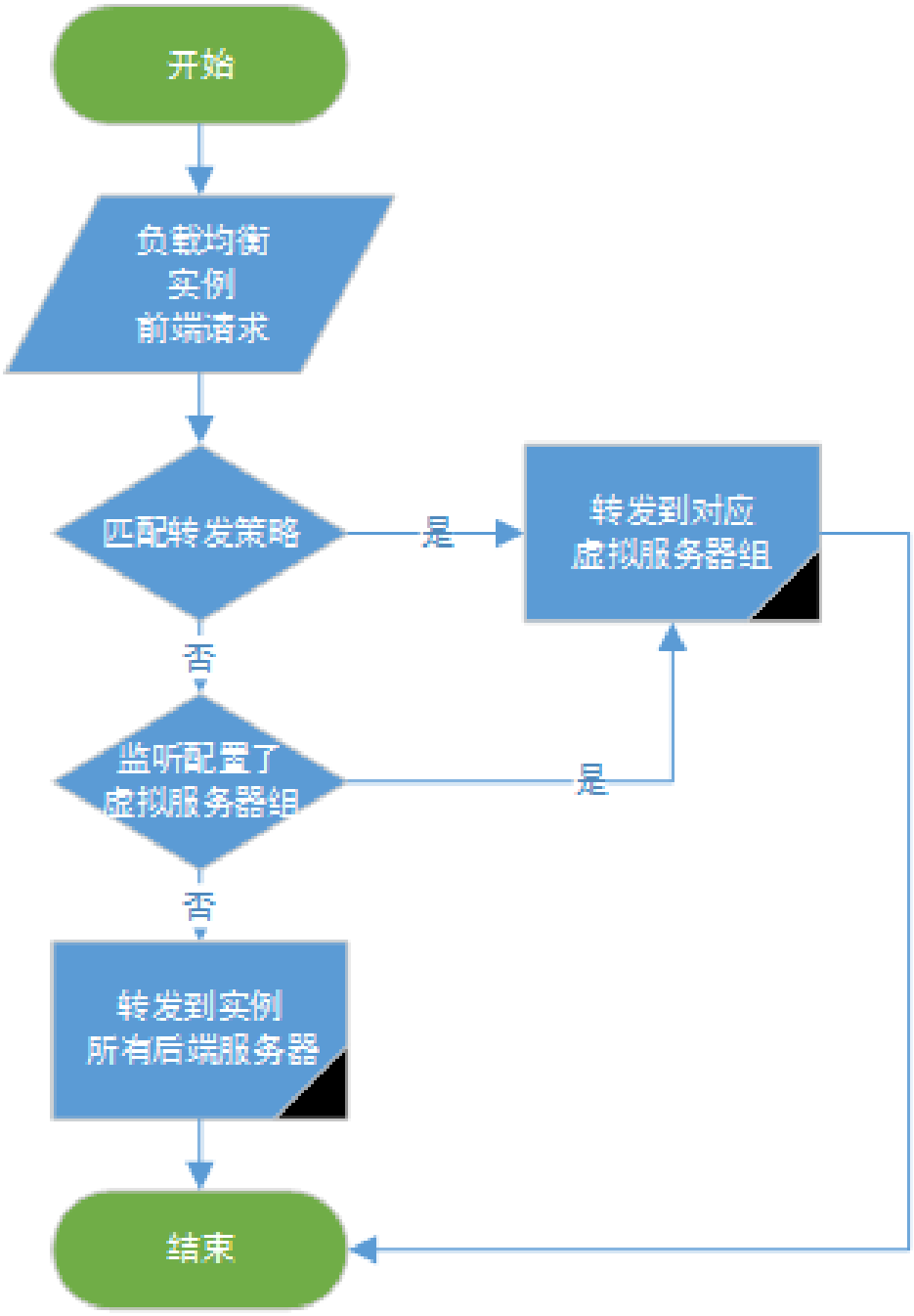
七层负载均衡服务支持配置域名或者URL转发策略，将来自不同域名或者URL的请求转发给不同的ECS处理。您可以在一个监听下添加多条转发策略，每条转发策略关联不同的虚拟服务器组（一个虚拟服务器组由一组ECS实例组成）。比如您可以将所有读请求转发到一组后端服务器上而将写请求转发到另一组后端服务器上，这样可以更灵活地适配业务需求，合理分配资源。

如下图所示，在配置了转发策略后，负载均衡系统将按照以下规则转发前端请求：

如果能匹配到相应监听关联的转发策略，则按转发策略，将请求转发到对应的虚拟服务器组。

如果未匹配，而对应监听启用并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，则转发到负载均衡实例后端服务器池中的ECS。




您不需要在转发规则上单独配置健康检查，下表对比描述了三个维度的健康检查机制。

| 维度 | 健康检查配置 | 健康检查目标服务器 |
| --- | --- | --- |
| 后端服务器 | 使用配置监听时的健康检查配置 | 所有后端ECS |
| 虚拟服务器组 | 使用配置监听时的健康检查配置 | 相应虚拟服务器组包含的服务器 |
| 转发策略 | 使用配置监听时的健康检查配置 | 相应虚拟服务器组包含的服务器 |

**注意**：由于虚拟服务器组中可以对ECS配置不同的端口，因此在配置健康检查时不要设置检查端口，否则会导致采用了不一致端口承载服务的服务器健康检查失败。

# 域名或URL转发规则说明

负载均衡支持分别添加域名或URL转发规则，也支持添加域名+URL组合的转发规则。

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/）。域名只能使用字母、数字、连字符（-）、点（.）。


支持精确匹配和通配符匹配两种模式：

精确域名：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 > 小范围通配符匹配 > 大范围通配符匹配，如下表所示。

| 模式 | 请求测试URL | 域名规则与匹配情况 | | |
|---|---|---|---|---|
| | | www.aliyun.com | *.aliyun.com | *.market.aliyun.com |
| 精确匹配 | www.aliyun.com | √ | | |
| 泛域名匹配 | market.aliyun.com | | √ | |
| 泛域名匹配 | info.market. | | | √ |

| | | | | |
|---|---|---|---|---|
| | aliyun.com | | | |

URL转发规则配置

单独配置URL转发规则时，域名配置项留空。参考以下原则添加URL：

URL只能包含字母、数字和以下特殊字符：

-./%?#&。

URL必须以斜杆（/）开头。

**注意**：如果您在URL中只输入了一个斜杆（/），则URL转发规则失效。

URL转发支持字符串匹配，按照顺序匹配原则。比如 /admin、/bbs、/test。


域名+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，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。参见如何实现相同域名不同路径的流量转发。

比如有两个域名分别是www.aaa.com和www.bbb.com，要求访问www.aaa.com/index.html时，将请求转发给ServerGroup1处理，其它来自xxx.html的请求转发给ServerGroup2处理。您需要配置如下转发策略，否则匹配到www.aaa.com的域名但没有相关策略匹配会返回404的响应码。


# 配置域名URL转发策略

## 前提条件

您已经创建了七层（HTTP/HTTPS）监听，详情参考配置监听。

您已经创建了接收请求的虚拟服务器组。详情参考创建虚拟服务器组。

## 操作步骤

登录负载均衡管理控制台。

在**实例管理**页面，选择地域，查看该地域的所有负载均衡实例。

单击负载均衡实例的ID链接。

在**详情**页面的左侧导航栏，单击**监听**。

单击目标七层监听的**操作**列内的**更多 > 添加转发策略**。

在**转发策略**页面，单击**添加转发策略**。

在**添加转发策略**对话框，配置转发规则，然后单击**确认**。

单击**添加转发策略 +**添加另一条转发策略，单击**确认**完成配置。

# 健康检查

## 配置健康检查

您可以通过控制台或API配置监听的健康检查。关于健康检查的原理参考负载均衡健康检查原理。其它健康检查问题，参考健康检查常见问题。

注意：负载均衡监听为TCP协议时，健康检查方式可选TCP或HTTP。

在负载均衡实例的详情页面，单击**监听** > **添加监听**，在添加监听的第二步可进行健康检查配置。

# 健康检查参数说明

在配置健康检查时，建议您使用默认值。

| 健康检查配置 | 说明 |
| --- | --- |
| 域名和检查路径<br>(仅限HTTP方式) | HTTP健康检查默认由负载均衡系统通过后端ECS内网IP地址向该服务器应用配置的缺省首页发起http head请求。<br>- 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，需要指定域名和具体的检查路径。<br>- 如果您对http head请求限定了host字段的参数，您只需要指定检查路径，即用于健康检查页面文件的URI。 |
| 正常状态码<br>(仅限HTTP方式) | 选择健康检查正常的HTTP状态码。<br>默认值为http_2xx和http_3xx。 |
| 检查端口 | 健康检查服务访问后端时的探测端口。<br>默认值为配置监听时指定的后端端口。<br>**注意**：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。 |
| 响应超时时间 | 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。<br>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。 |
| 健康检查间隔 | 进行健康检查的时间间隔。<br>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。<br>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。 |
| 不健康阈值 | 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。<br>可选值2-10，默认为3次。 |
| 健康阈值 | 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。 |

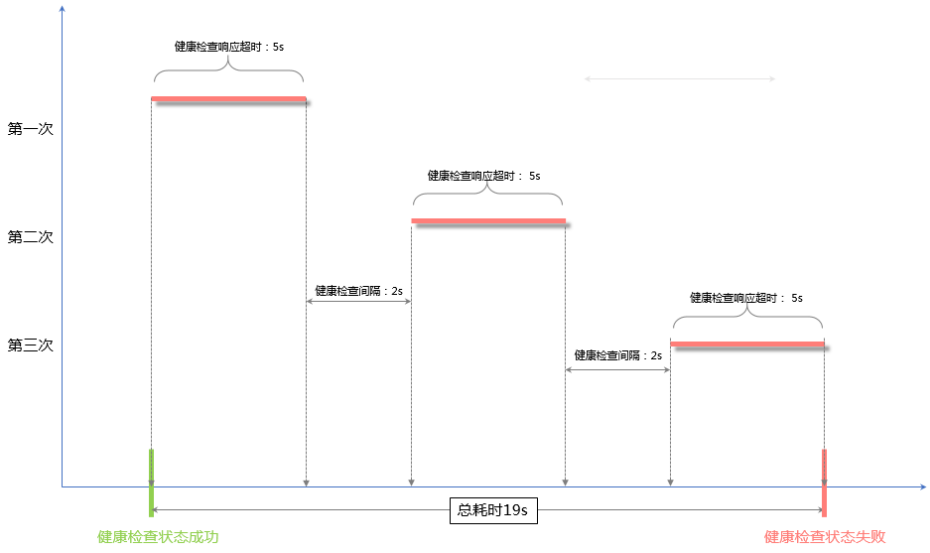
| | |
|---|---|
| | 可选值 2-10，默认为3次。 |
| 健康检查请求和健康检查返回结果 | 为UDP监听配置健康检查时，您可以在**健康检查请求**中输入请求的内容（比如youraccountID），在**健康检查返回结果**中输入预期的返回结果（比如slb123）。<br><br>同时在后端服务器的应用逻辑中加入相应的健康检查应答逻辑，如收到youraccountID的请求时，回应slb123。<br><br>此时，当负载均衡收到后端服务器发来的正确响应时，则认为健康检查成功，否则认为健康检查失败。此方式能最大程度确保健康检查的可靠性。 |

# 健康检查响应超时和健康检查间隔示例

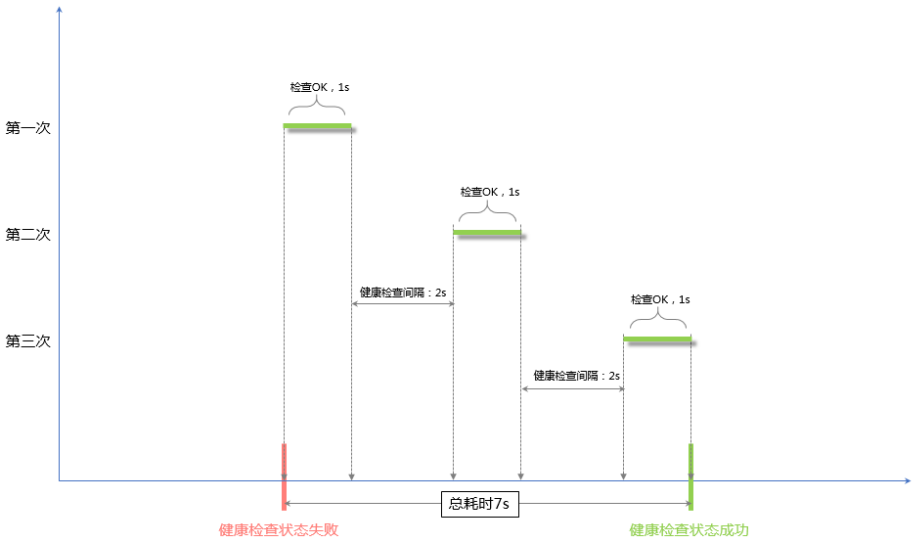
以如下健康检查配置为例：

- 响应超时时间：10秒
- 健康检查间隔：5秒
- 健康检查间隔：3次
- 不健康阈值：3次

从健康到不健康的检查过程如下图所示：




从不健康到健康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）：




**说明**：健康检查间隔是从上一次健康检查请求发起后开始计算，而不是上一次健康请求响应后开始计算。

# 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。综上原因，如果您的应用服务器需要校验请求的host字段校，那么则需要配置相关的域名，确保健康检查正常工作。

**特别提醒**：访问控制白名单现已经全网开放，无需提工单开通。

白名单是一种访问控制方式，可以为负载均衡监听设置仅允许哪些IP访问，适用于应用只允许特定IP访问的场景。

注意：

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。

设置白名单的过程中可能会引起访问负载均衡监听短时中断。

# 操作步骤

登录负载均衡管理控制台。

选择地域，查看该地域下的负载均衡实例。

单击需要设置访问控制的负载均衡实例的ID链接，打开详情页。

在负载均衡实例菜单栏，单击**监听**，打开监听配置页面。

在**监听**页面，单击**更多 > 设置访问控制**。


在**访问控制设置**对话框，进行如下配置：

单击**是否开启访问控制**开关，打开开关。

在**白名单设置**区域内输入允许访问该监听的IP地址。

多个IP地址以逗号隔开且不可重复，最多允许输入300个IP地址。支持输入单个IP地址或者IP网段。

单击**确认**，完成配置。

# 后续操作

如果您想关闭白名单访问控制，单击**更多 >设置访问控制**，然后关闭**是否开启访问控制**开关。

如果你想修改白名单中的IP地址，单击 **更多 > 设置访问控制**， 在**白名单设置**区域内修改IP地址。

负载均衡支持按带宽计费的负载均衡实例下的所有监听共享实例的总带宽。在创建监听时，您可以设置带宽峰值也可以选择不设置。

配置：您可以对监听的带宽进行限制，但所有监听带宽峰值的总和不能超过实例的带宽峰值。

不限制：不限制带宽的情况下，实例下的监听共享实例带宽。


# 如何共享带宽？

假如您购买了一个带宽峰值为 10MB 的负载均衡实例，并在该实例下创建了三个监听（监听A、监听B和监听C）。监听A的带宽峰值设置为 4MB，另外两个监听没有设置带宽峰值。三个监听的带宽使用可能出现如下几种情况：

如果监听A和监听C一直没有流量进入，那么监听B最多也只能跑满剩余的 6MB 带宽（10MB - 4MB)。

如果监听C一直没有流量进入，而监听B的入流量很大，超过了剩余的 6MB 带宽。此时，监听B已经产生丢包，而监听A只有 4MB 的入流量，没有超过设置的带宽峰值，所以不会产生丢包。

如果监听A一直是满速在跑（监听峰值 4MB），而后监听B和监听C也有流量进入并且两个监听的流量很大，那么监听B和监听C就会共享（竞争）剩余的 6MB 带宽。此时，监听A的流量不会受监听B和监听C的影响，始终能达到预留的 4MB 峰值；如果监听B和监听C入流量同等大小，两个监听占用的带宽去会趋近于均分。

因此，对监听带宽的限制值是资源预留，这是为了保证核心的业务始终有足够的带宽。非核心的业务可以不设置监听带宽值，它们竞争实例剩余的带宽资源。

# 证书管理

负载均衡只支持PEM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；请将这些内容一并上传。

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。

下图为PEM格式的证书示例。


# 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。详情参见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-----
```

# 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。详情参见RFC1421。

**注意**：如果您的私钥是加密的，比如私钥的开头和结尾是-----BEGIN PRIVATE KEY-----, -----END PRIVATE KEY-----或-----BEGIN ENCRYPTED PRIVATE KEY-----, -----END ENCRYPTED 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 rsa -in old_server_key.pem -out new_server_key.pem

下图为RSA私钥示例。


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

运行以下命令在/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。

在**conf**目录下新建一个包含如下信息的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
```

运行以下命令生成私钥key文件。

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

运行结果如下图所示。

```
root@iZbp1hfvivcqx1jbwap31iZ:~/ca/conf# cd /root/ca
root@iZbp1hfvivcqx1jbwap31iZ:~/ca# sudo openssl genrsa -out private/ca.key
Generating RSA private key, 2048 bit long modulus
.....+++
..+++
e is 65537 (0x10001)
```

运行以下命令并按提示输入所需信息，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

注意：Common Name需要输入负载均衡的域名。

```
root@iZbp1hfvivcqx1jbwap31iZ:~/ca# sudo openssl req -new -key private/ca.key -ou
t 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@alibaba.com

Please enter the following 'extra' attributes
to be sent with your certificate request
A challenge password []:
An optional company name []:
root@iZbp1hfvivcqx1jbwap31iZ:~/ca#
```

运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

运行以下命令为CA的key设置起始序列号，可以是任意四个字符。

$ sudo echo FACE > serial

运行以下命令创建CA键库。

$ sudo touch index.txt

运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 -crldays 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

运行以下命令在**ca**目录内创建一个存放客户端key的目录**users**。

$ sudo mkdir users

运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

> 注意：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。

运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

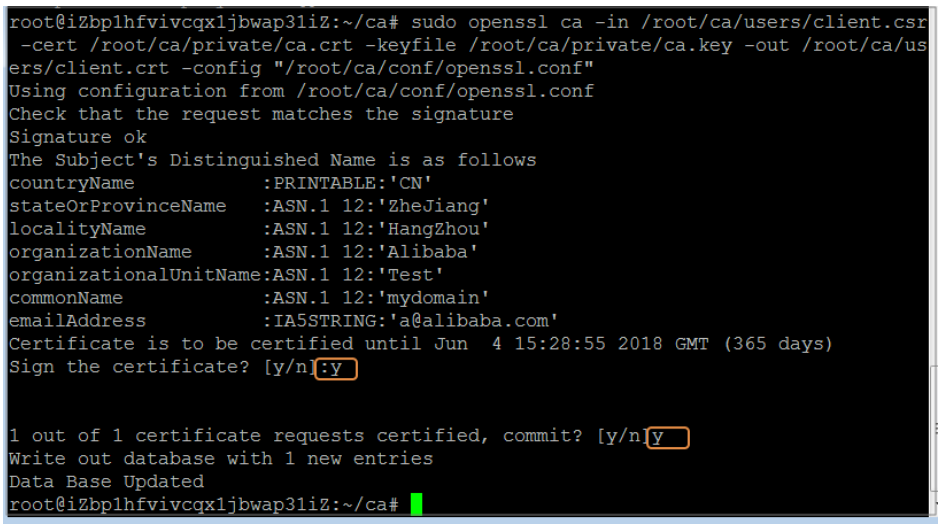
输入该命令后，根据提示输入上一步输入的pass phrase，然后根据提示输入对应的信息。

> 注意：**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。

运行以下命令使用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**。




运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/users/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。

运行以下命令查看生成的客户端证书。

```
cd users
ls
```

负载均衡只支持PEM格式的证书，其它格式的证书需要转换成PEM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。

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

为了支持HTTPS数据传输加密认证，负载均衡提供证书管理功能。您可以将证书存储到负载均衡证书管理系统，而不需要将证书直接部署到后端ECS服务器上，上传到证书管理系统的私钥都会加密存储。一个证书可以应用于一个或多个监听。

- 负载均衡证书是分地域管理的，即一个证书如果要在多个地域使用，那么创建、上传证书时就需要选择多个地域。
- 每个账号最多可以创建100个证书。

# 前提条件

已经购买了服务器证书。

已经生成了CA证书和客户端证书。详情参考生成证书。

# 操作步骤

登录负载均衡控制台。

在**负载均衡**左侧导航栏，单击**证书管理**。

在**证书管理**页面，单击**创建证书**。

在**创建证书**页面，完成如下配置后，单击**确定**。

证书名称：输入证书名称。

证书地域: 选择要使用证书的地域。如果您的证书需要在多个地域使用，请把对应地域都选上。

证书类型：

服务器证书：HTTPS单向认证，只需要上传服务器证书和私钥。

CA证书：如果您要配置HTTPS双向认证，除了上传服务器证书外，您还需选择CA证书进行上传。

证书内容：复制您的服务器或者CA证书内容。您可以单击**导入样式**，查看正确的证书样式。更多详情查看**证书要求**。

私钥：复制您的服务器证书的私钥。可以单击**导入样式**，查看正确的证书样式。更多详情查看**证书要求**。

私钥只有**服务器**类型的证书才需要。

# 相关操作

上传证书后，您可以在负载均衡管理控制台的**证书管理**页面中查看证书、编辑证书名称、删除证书。

单击**修改名称**，修改证书的名称。

单击**删除**，删除该证书。

注意：证书如果在HTTPS监听中被引用了，则无法删除该证书。


## 应用场景

- 证书过期，需要创建新的证书。
- 负载均衡添加证书报错，可能是私钥内容错误，需要替换为新的满足需求的证书。

## 操作步骤

新建并上传一个新的证书。

详情参见生成证书和上传证书。

在HTTPS监听中配置新的证书。

详情参见配置HTTPS。

打开**证书管理**页面，找到目标证书，然后单击**删除**。

在弹出的对话框中，单击**确认**。

# 日志管理

负载均衡提供日志管理功能，您可以查看某个实例的操作日志和健康检查日志。

操作日志

负载均衡提供近一个月的操作日志信息。目前只支持控制台查看，不支持通过Open API获取这些日志。

健康检查日志

目前负载均衡只存储三天内的健康检查日志信息。如果您想存储更多的健康检查日志，需要设置日志

存储。详情查看管理健康检查日志。

# 操作步骤

登录负载均衡控制台。

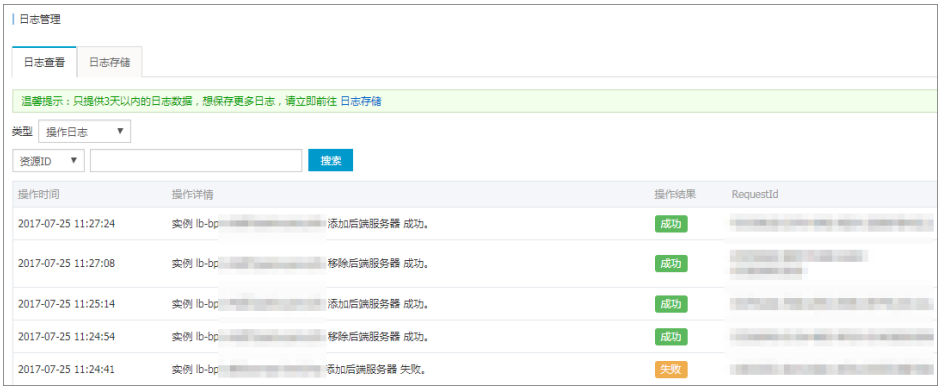
在左侧导航栏，单击**日志管理**。

在**日志管理**页面，单击**日志查看**页签。

类型选择**操作日志**。

查看所有负载均衡实例的操作日志，比如启动、添加、删除监听， 添加后端服务器等操作日志。

您可以通过操作结果（成功或失败）过滤要查看的实例操作日志，也可以通过具体的资源ID快速地找到指定资源的相关操作日志。资源ID是这次操作的主体对象的ID，例如负载均衡实例ID、证书ID、虚拟服务器组ID、转发策略ID等。




您可以在**日志管理**页面，查看三天内的健康检查日志。如需要更久的健康检查日志，您需要将健康检查日志存储到OSS中，并可以下载完整的健康检查日志。

- 存储健康检查日志
- 查看健康检查日志
- 下载健康检查日志

# 存储健康检查日志

您可以通过负载均衡提供的日志管理功能，查看负载均衡实例后端ECS的健康检查日志。当前，负载均衡只存储三天内的健康检查日志信息，您可以通过开通OSS服务，将所有的健康检查日志存储到创建的bucket中。

您可以随时开启和关闭日志存储功能。开启日志存储后，负载均衡会在所选bucket中创建一个名称为**AliyunSLBHealthCheckLogs**的文件夹用来存储健康检查日志文件。负载均衡的健康检查日志每小时生成一次，系统会自动创建一个以日期为名称的子文件夹用来存储当天的健康检查日志文件，如20170707。

当天每小时生成的日志文件以生成的截止时间命名。比如在00:00-01:00生成的健康检查日志，日志文件名为01.txt；在01:00-02:00生成的健康检查日志，日志文件名为02.txt。

> 注意：只有检查到后端ECS出现异常时，才会生成健康检查日志。健康检查日志每小时生成一次，若该小时内后端ECS未检测到异常，则无健康检查日志。

配置健康检查日志存储，您需要执行以下操作：

创建Bucket

授权日志访问

设置日志存储

## 创建Bucket

在设置日志存储前，您需要开通OSS服务并创建用来存储健康检查日志的bucket。

打开对象存储OSS产品页面，单击**立即开通**。

开通OSS服务后，登录OSS管理控制台。

单击**新建Bucket**。


在**新建Bucket**对话框，配置Bucket信息，单击**确定**。

> **注意**：确保bucket的地域和负载均衡实例的地域相同。

更多Bucket配置的详细信息，查看**创建存储空间**。

## 授权日志访问

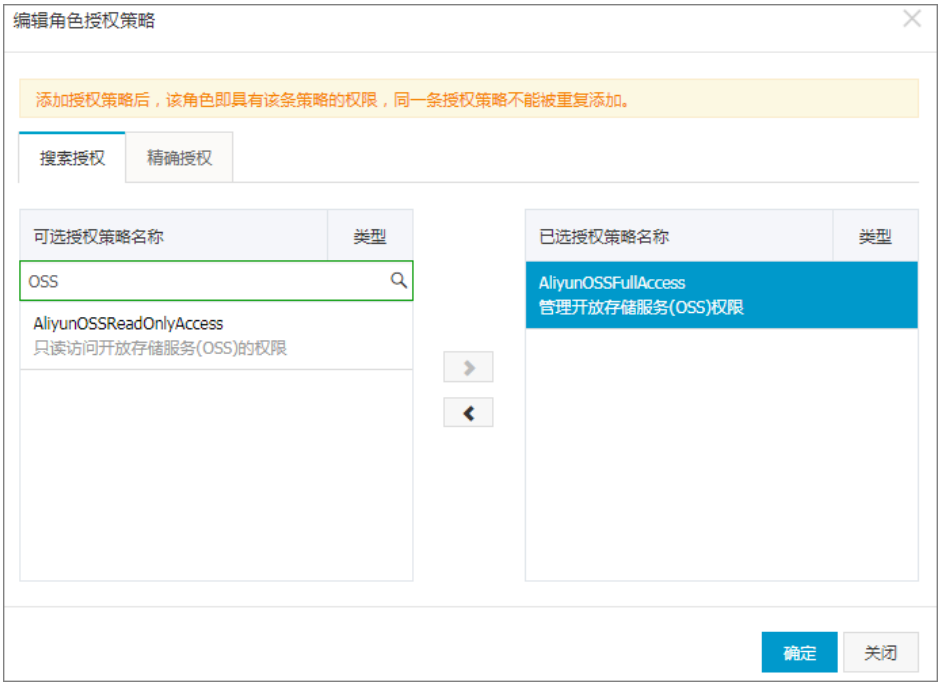
创建好Bucket后，您还需要对负载均衡的日志角色（SLBLogDefaultRole）授权，允许该角色访问OSS的相关资源。

登录**访问控制管理控制台**。

在左侧导航栏，单击**角色管理**。

找到名称为**SLBLogDefaultRole**的角色，然后单击**授权**。


在**编辑角色授权策略**对话框，选择**AliyunOSSFullAccess**，然后单击**确定**完成授权。




授权完成后，单击**SLBLogDefaultRole**，然后单击**角色授权策略**，查看授权策略。

## 设置日志存储

登录负载均衡控制台。

在左侧导航栏，单击**日志管理**。

在**日志管理**页面，单击**日志存储**页签。

单击目标地域的**设置日志存储**链接。


在**设置日志存储**对话框，选择用来存储健康检查日志的Bucket，然后单击**确认**。

单击状态栏下的开关，开启日志存储。


# 查看健康检查日志

您可以在负载均衡管理控制台，查看三天内的健康检查日志。

登录负载均衡控制台。

在左侧导航栏，单击**日志管理**。

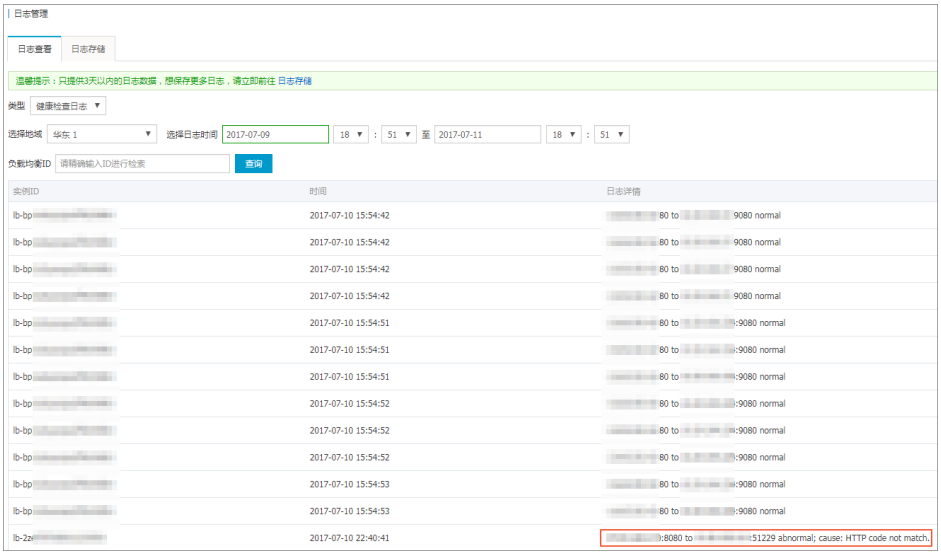
在**日志管理**页面，单击**日志查看**页签。

类型选择**健康检查日志**，然后选择要查看的日志时间。

> **注意**：只有检查到后端ECS出现异常时，才会生成健康检查日志。健康检查日志每小时生成一次，若该小时内后端ECS未检测到异常，则无健康检查日志。

当健康检查日志的信息为SLB_instance_IP:port to Added_ECS_instance_IP:port abnormal; cause:XXX时，代表后端ECS实例健康检查异常，您可以根据提示的异常原因进行排查。

当健康检查日志的信息为SLB_instance_IP:port to Added_ECS_instance_IP:port normal时，代表后端ECS实例恢复正常。



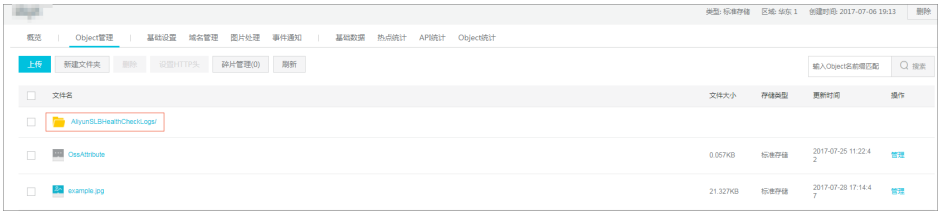

# 下载健康检查日志

您可以在OSS管理控制台中，下载存储的完整的健康检查日志。

登录OSS管理控制台。

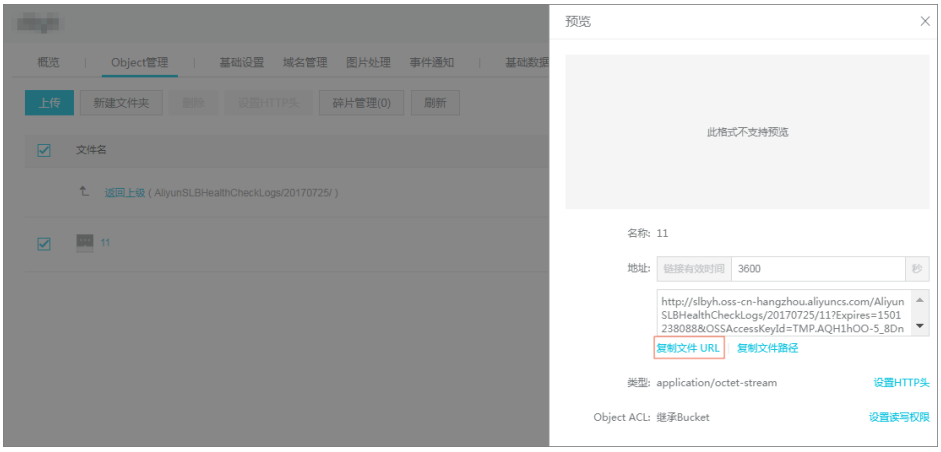
在**概览**页面，单击目标Bucket，然后单击**Object管理**。

在**Object管理**页面，单击健康检查日志文件夹**AliyunSLBHealthCheckLogs/**。




单击您要下载的健康检查日志的文件夹。


单击目标文件的**管理**，然后单击**复制文件 URL**。




在浏览器中输入复制的URL，下载日志文件。

您可以使用API、SDK等其它方式下载日志文件，详情参考**简单下载**。

# 监控

如果负载均衡开启了健康检查功能，并且后端ECS实例的健康检查状态正常，您就可以在控制台查看负载均衡服各项监控指标的实时数据和历史数据。

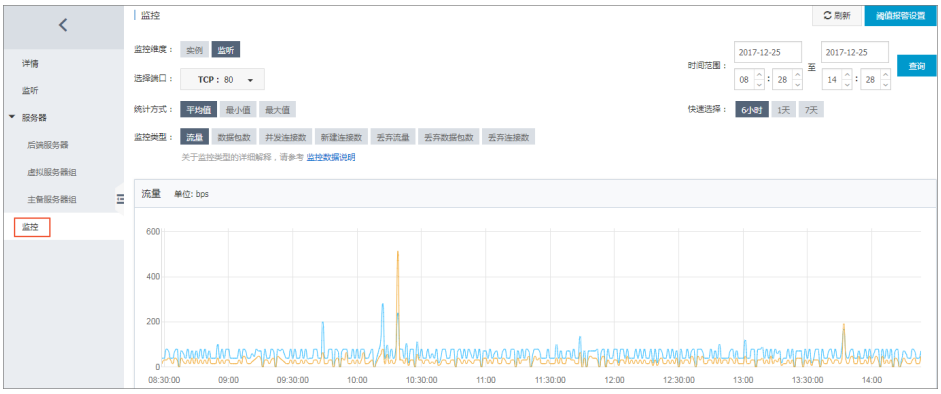
## 查看监控数据

登录负载均衡管理控制台。

单击目标实例的ID链接。

在左侧导航栏，单击**监控**。

选择要查看的监控指标，查看监控数据。




# 监控指标

| 监控指标 | 说明 |
| --- | --- |
| 流量 | - 流入流量：从外部访问负载均衡所消耗的流量。<br>- 流出流量：负载均衡访问外部所消耗的流量。 |
| 数据包 | - 流入数据包数：负载均衡每秒接到的请求数据包数量。<br>- 流出数据包数：负载均衡每秒发出的数据包数量。 |
| 并发连接数 | - 活跃连接数：所有ESTABLISHED状态的TCP连接。因为如果您采用的是长连接的情况，一个连接会同时传输多个文件请求。<br>- 非活跃连接数：表示指除ESTABLISHED状态的其它所有状态的TCP连接数。Windows和Linux服务器都可以使用netstat -an命令查看。<br>- 并发连接数：所有建立的TCP连接数量。 |
| 新建连接数 | 在统计周期内，新建立的从客户端连接到负载均衡的连接请求的平均数。 |

| | |
|---|---|
| 丢弃流量 | - 丢弃入流量：每秒丢失的入流量。<br>- 丢弃出流量：每秒丢失的出流量。 |
| 丢弃数据包 | - 丢弃流入数据包：每秒丢弃的流入数据包的数量。<br>- 丢弃流出数据包：每秒丢弃的流出数据包的数量。 |
| 丢弃连接数 | 每秒丢弃的连接数。 |
| 7层协议QPS | 每秒可以处理的HTTP/HTTPS请求。<br>**注意**：只有7层（HTTP/HTTPS）监听才有该监控指标。 |
| 7层协议RT | 负载均衡的平均响应时间。<br>**注意**：只有7层（HTTP/HTTPS）监听才有该监控指标。 |
| 7层协议返回码(2XX)/(3xx)/(4xx)(5xx)(Others) | 监听返回的HTTP响应代码的数量。<br>**注意**：只有7层（HTTP/HTTPS）监听才有该监控指标。 |
| 7层协议UpstreamCode4XX/5XX | 后端服务器返回的HTTP响应代码的数量。<br>**注意**：只有7层（HTTP/HTTPS）监听才有该监控指标。 |
| 7层协议UpstreamRT | 后端服务器的平均响应时间。<br>**注意**：只有7层（HTTP/HTTPS）监听才有该监控指标。 |

开通云监控服务后，您可以在云监控控制台配置监控报警规则。云监控支持“短信”、“邮件”、“旺旺” 三种报警方式，暂不支持电话报警。关于云监控的更多详细信息，参考云监控文档。

注意：负载均衡的监听或实例被删除，其在云监控设置的报警规则也会相应删除。

# 操作步骤

登录负载均衡控制台。

选择地域，然后单击目标实例的ID链接。

确保该实例已经配置了监听，开启了健康检查。

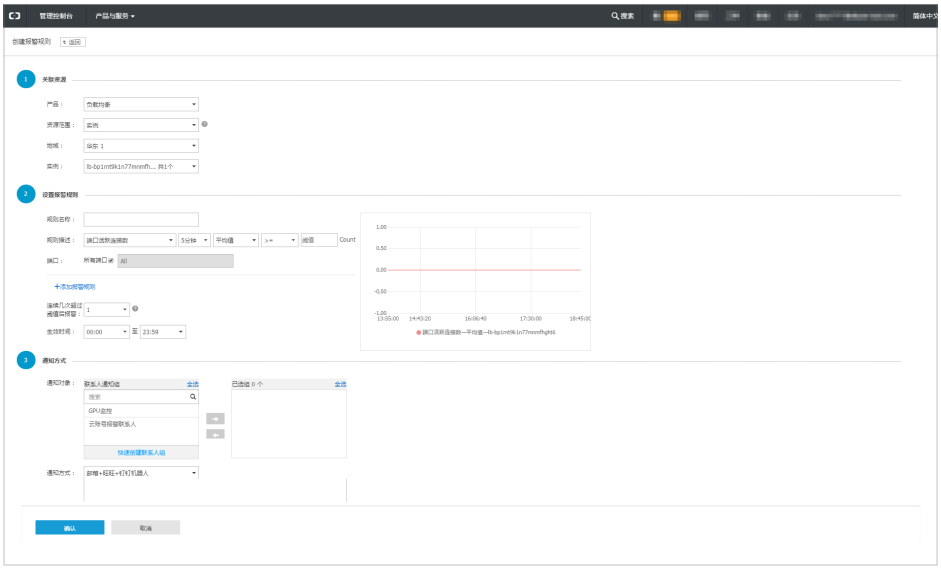
在**详情**页面左侧导航栏，单击**监控**，进入**监控**页面。

单击**阈值报警设置**，进入云服务监控页面。

单击**创建报警规则**。

配置报警规则。

详情参考负载均衡监控配置。



# 什么是多可用区

云产品的可用区指的是一套独立的基础设施，不同的可用区之间基础设施（网络，电力和空调等）相互独立，即一个可用区出现基础设施故障不影响另外一个可用区。

为了向广大用户提供更加稳定可靠的负载均衡服务，阿里云负载均衡已在各地域（Region）部署了多可用区以实现同地域下的跨机房容灾。当主可用区的机房故障或不可用时，负载均衡仍然有能力在非常短的时间内（约30秒）切换到另外一个备可用区的机房并恢复服务的能力；当主可用区恢复时，负载均衡同样会自动切换到主可用区的机房提供服务。

在创建负载均衡实例时，您可以选择将负载均衡创建在支持多可用区的地域，提高服务的可用性。

**注意：**

每个地域只有一种属性，不是多可用区就是单可用区。

实例创建后，不能修改主备可用区。

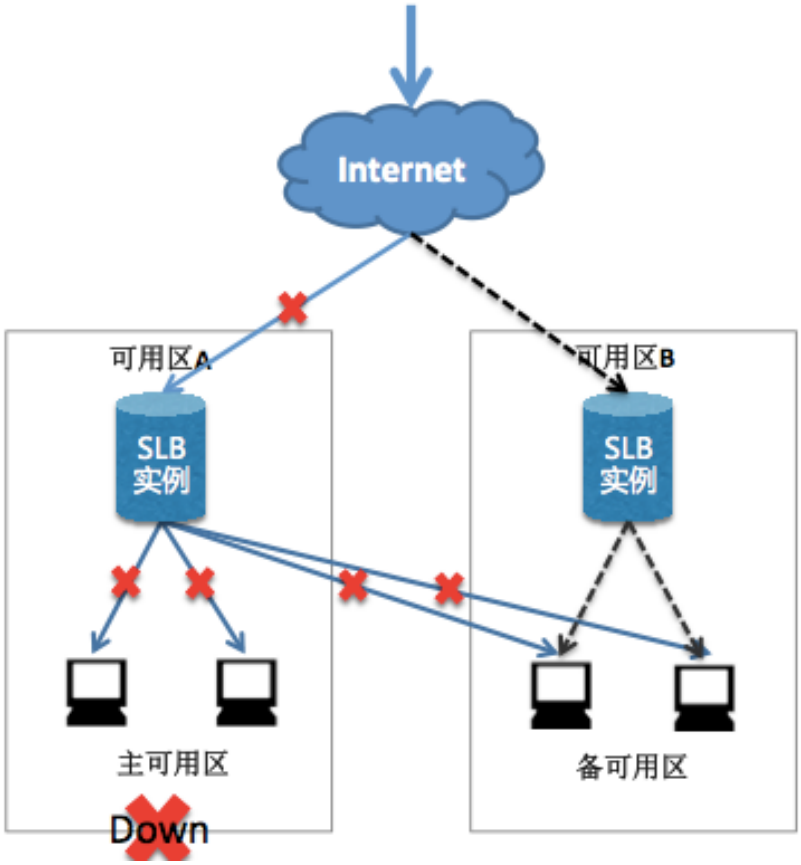
| **地域** | **可用区类型** | **可用区** | |
|---|---|---|---|
| 华东1 | 多可用区 | **主可用区** | **备可用区** |
| | | 可用区B | 可用区D |
| | | 可用区D | 可用区E |
| | | 可用区E | 可用区F |
| | | 可用区F | 可用区E |
| 华东2 | 多可用区 | **主可用区** | **备可用区** |
| | | 可用区A | 可用区B |
| | | 可用区B | 可用区A或可用区D |
| | | 可用区C | 可用区B |
| | | 可用区D | 可用区B |
| 华南1 | 多可用区 | **主可用区** | **备可用区** |
| | | 可用区A | 可用区B |
| | | 可用区B | 可用区A |
| | | 可用区C | 可用区B |
| 华北1 | 多可用区 | **主可用区** | **备可用区** |
| | | 可用区B | 可用区C |
| | | 可用区C | 可用区B |
| 华北2 | 多可用区 | **主可用区** | **备可用区** |
| | | 可用区A | 可用区B或可用区D |
| | | 可用区B | 可用区A或可用区C |
| | | 可用区C | 可用区B |
| | | 可用区D | 可用区A |
| 华北3 | 单可用区 | 可用区A | |
| 华北5 | 单可用区 | 可用区A | |
| 欧洲中部1 | 单可用区 | 可用区A | |
| 中东东部1 | 单可用区 | 可用区A | |

| | | | |
|---|---|---|---|
| 新加坡 | 多可用区 | **主可用区** | **备可用区** |
| | | 可用区A | 可用区B |
| | | 可用区B | 可用区A |
| 亚太东南2 | 单可用区 | 可用区A | |
| 亚太东南3 | 单可用区 | 可用区A | |
| 亚太东北1 | 单可用区 | 可用区A | |
| 香港 | 多可用区 | **主可用区** | **备可用区** |
| | | 可用区B | 可用区C |
| | | 可用区C | 可用区B |
| 美东1 | 单可用区 | 可用区A | |
| 美西1 | 多可用区 | **主可用区** | **备可用区** |
| | | 可用区A | 可用区B |
| | | 可用区B | 可用区A |

# 多可用区与高可用

除了选择具有多可用区的地域实现同城容灾外，也建议用户可以结合自身的应用需要，综合考虑后端服务器部署来实现更可靠的同地域高可用的方案。

比如，您可以将后端服务器ECS实例分别部署在主备可用区内，如下图所示。




当可用区A未出现故障时，用户访问流量如上图蓝色实线所示；当可用区A发生故障时，用户访问流量的分发将变成如上图黑色虚线所示，这样就可以避免因为单个可用区的故障而导致对外服务的不可用。

负载均衡各地域的包年包月可售卖最大带宽和按流量实例带宽峰值如下表所示。

> **注意**：所有地域的私网带宽峰值都为5 GB。

| 地域 | 带宽峰值 |
| --- | --- |
| 华北 1（青岛） | 5 GB |
| 华东 1（杭州） | 5 GB |
| 华北 2（北京） | 5 GB |
| 华东 2（上海） | 5 GB |
| 华南 1（深圳） | 5 GB |
| 华北 3（张家口） | 5 GB |
| 华北 5（呼和浩特） | 5 GB |
| 香港 | 2 GB |
| 美国东部 1（弗吉尼亚） | 1 GB |
| 美国西部 1（硅谷） | 2 GB |

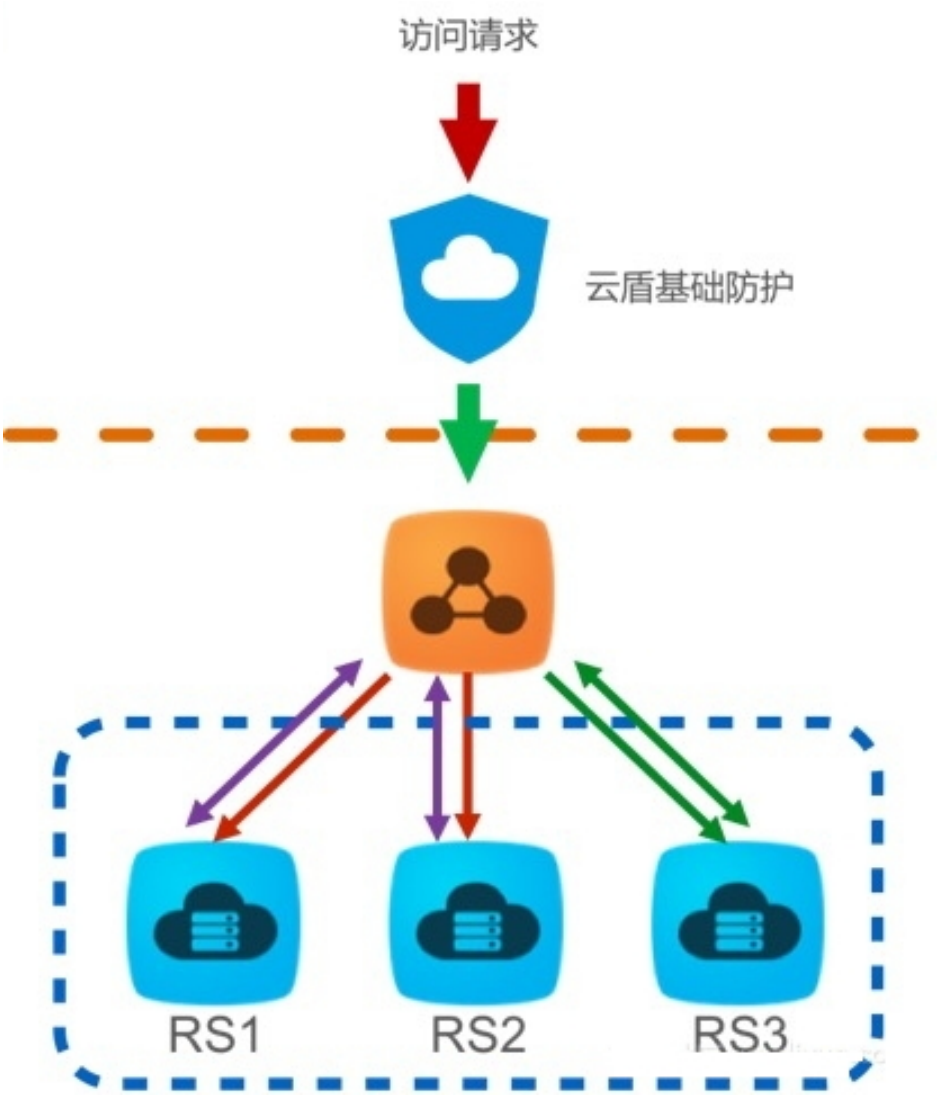
| | |
|---|---|
| 亚太东北 1（东京） | 1 GB |
| 亚太东南 1（新加坡） | 5 GB |
| 亚太东南 2（悉尼） | 1 GB |
| 亚太东南 3（吉隆坡） | 5 GB |
| 中东东部 1（迪拜） | 500 M |
| 欧洲中部 1（法兰克福） | 1 GB |
| 亚太南部 1 (孟买) | 5 GB |

负载均衡控制台可以查看公网负载均衡实例的云盾阈值。

**说明：** 该功能目前已在青岛、北京、杭州、上海、深圳、香港、新加坡、美东和美西地域上线。

# DDoS基础防护介绍

阿里云免费为负载均衡服务提供最高5G的DDoS基础防护。如下图所示，所有来自Internet的流量都要先经过云盾再到达负载均衡，云盾会针对常见的攻击进行清洗过滤。云盾DDoS基础防护可以防御SYN Flood、UDP Flood、ACK Flood、ICMP Flood 和DNS Flood等DDoS攻击。




云盾DDoS基础防护根据公网负载均衡实例的带宽设定清洗阈值和黑洞阈值。当入方向流量达到阈值上限时，触发清洗和黑洞：

清洗：当来自Internet的攻击流量较大或符合某些特定攻击流量模型特征时，云盾将会针攻击流量启动清洗操作，清洗包括攻击报文过滤、流量限速、包限速等。

黑洞：当来自Internet的攻击流量非常大时，为保护整个集群的安全，流量将会被黑洞处理，即所有入流量全部被丢弃。

更多信息，查看DDoS基础防护文档。

# 查看防护阈值

登录负载均衡管理控制台。

选择地域，查看该地域的所有实例。

将鼠标移至目标实例的云盾图标，查看BPS清洗阈值、PPS清洗阈值和黑洞阈值。您可以单击DDoS控制台链接查看更多信息。

BPS清洗阈值：入方向流量超过了BPS清洗阈值时，触发清洗。

PPS清洗阈值：入方向数据包数超过了PPS清洗阈值时，触发清洗。

黑洞阈值：入方向流量超过黑洞阈值时将触发黑洞。