

# 阿里云 负载均衡

历史文档

文档版本：20190815

# 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
	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。	 <b>禁止：</b> 重置操作将丢失用户配置数据。
	该类警示信息可能导致系统重大变更甚至故障，或者导致人身伤害等结果。	 <b>警告：</b> 重启操作将导致业务中断，恢复业务所需时间约10分钟。
	用于补充说明、最佳实践、窍门等，不是用户必须了解的内容。	 <b>说明：</b> 您也可以通过按Ctrl + A选中全部文件。
>	多级菜单递进。	设置 > 网络 > 设置网络类型
<b>粗体</b>	表示按键、菜单、页面名称等UI元素。	单击 <b>确定</b> 。
<code>courier</code> 字体	命令。	执行 <code>cd /d C:/windows</code> 命令，进入Windows系统文件夹。
<code>##</code>	表示参数、变量。	<code>bae log list --instanceid Instance_ID</code>
<code>[ ]</code> 或者 <code>[a b]</code>	表示可选项，至多选择一个。	<code>ipconfig [-all -t]</code>
<code>{ }</code> 或者 <code>{a b}</code>	表示必选项，至多选择一个。	<code>swich {stand   slave}</code>

# 目录

---

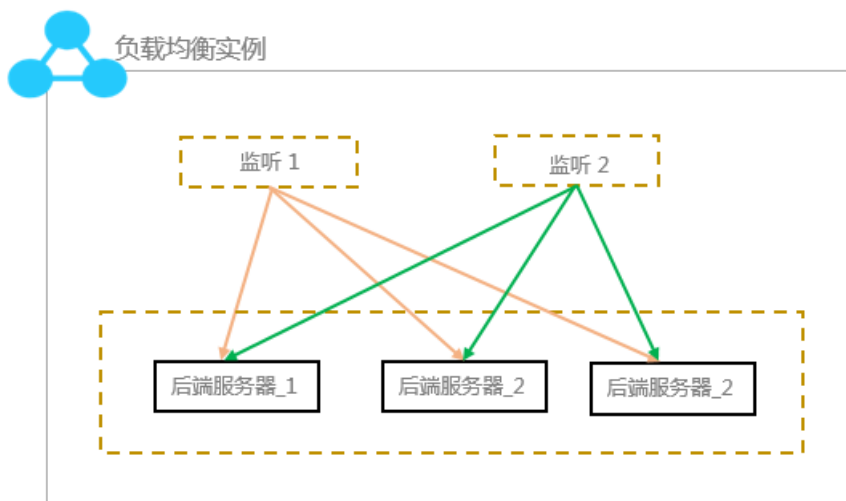
法律声明.....	I
通用约定.....	I
1 用户指南（旧版控制台） .....	1
1.1 负载均衡实例.....	1
1.1.1 什么是负载均衡实例.....	1
1.1.2 性能保障型实例.....	3
1.1.3 网络流量路径说明.....	14
1.1.4 创建实例.....	16
1.1.5 创建IPv6实例.....	17
1.1.6 管理实例.....	20
1.1.7 管理标签.....	20
1.1.8 绑定EIP.....	22
1.1.9 按量付费实例变配.....	22
1.1.10 包年包月实例变配.....	24
1.2 监听.....	25
1.2.1 监听介绍.....	25
1.2.2 四层监听.....	27
1.2.3 七层监听.....	27
1.2.4 健康检查.....	27
1.2.5 共享实例带宽.....	27
1.3 后端服务器.....	29
1.3.1 后端服务器概述.....	29
1.3.2 添加默认服务器.....	30
1.3.3 创建虚拟服务器组.....	31
1.3.4 创建主备服务器组.....	32
1.4 DDoS基础防护.....	33
1.5 各地域带宽峰值限制.....	34
1.6 多可用区.....	35

# 1 用户指南 (旧版控制台)

## 1.1 负载均衡实例

### 1.1.1 什么是负载均衡实例

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



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。

阿里云负载均衡服务



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具有负载均衡的私网访问权限的客户端。

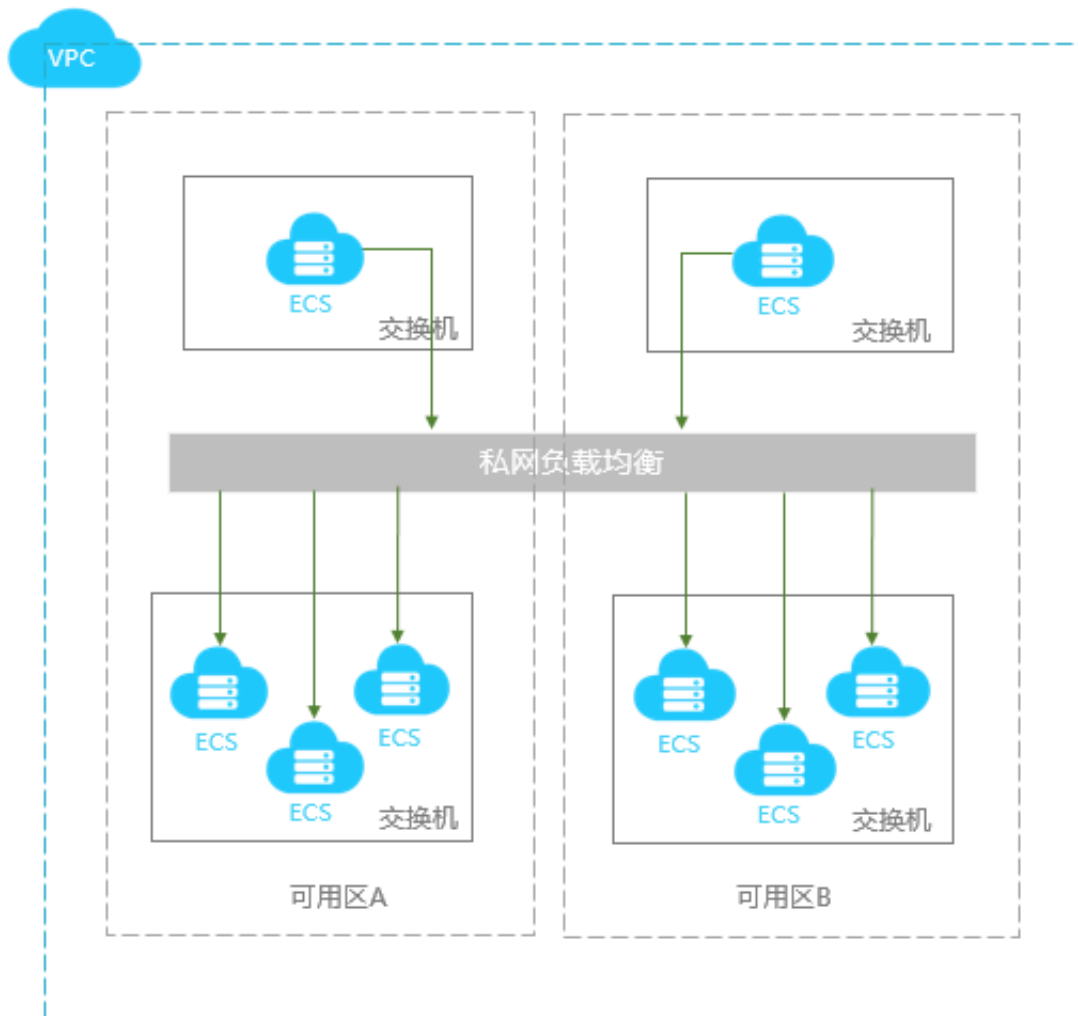
私网负载均衡实例可以进一步对网络类型进行选择：

- 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。



### 1.1.2 性能保障型实例

阿里云负载均衡计划将于2018年4月1日开始针对性能保障型实例收取规格费，同时继续保留性能共享型实例的售卖。

#### 1. 什么是负载均衡性能保障型实例？

负载均衡性能保障型实例提供了可保障的性能指标。与之相对的是负载均衡性能共享型实例，资源是所有实例共享的，所以不保障实例的性能指标。

在推出负载均衡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问号图标，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开放了如下六种实例规格（各地域因资源情况不同，开放的规格可能略有差异，请以控制台购买页为准）。

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)	5,000	3,000	1,000
规格 2	标准型I (slb.s2.small)	50,000	5,000	5,000
规格 3	标准型II (slb.s2.medium)	100,000	10,000	10,000
规格 4	高阶型I (slb.s3.small)	200,000	20,000	20,000
规格 5	高阶型II (slb.s3.medium)	500,000	50,000	30,000
规格 6	超强型I (slb.s3.large)	1,000,000	100,000	50,000

## 2. 性能保障型实例如何收费？

负载均衡性能保障型实例需要收取规格费用，收费模型如下：



性能保障型费用 = 实例费 + 流量/带宽费 + 规格费



说明:

负载均衡私网实例也可以选择性能共享型实例或性能保障型实例，性能保障型私网实例，也需要收取规格费用，收费方式与公网性能保障型实例一致，但不收取流量费/带宽费和实例费。

负载均衡分为两种计费模式，预付费和按量付费。在不同的计费模式下，性能保障型实例的规格费收取规则不同：

- 预付费模式

性能保障型实例规格费按照预付费模式收取，即在实例的付费周期内，实例规格费按照固定的价格收取。假设您选择的是高阶型I (slb.s3.small)规格，并且选择购买时长为3个月，则规格费用 = slb.s3.small规格费月价 x 3月。

· 按量付费模式

性能保障型实例规格费按使用量收取，即不论您选择的何种规格，实例规格规格费均会按照您实际使用的规格收取。

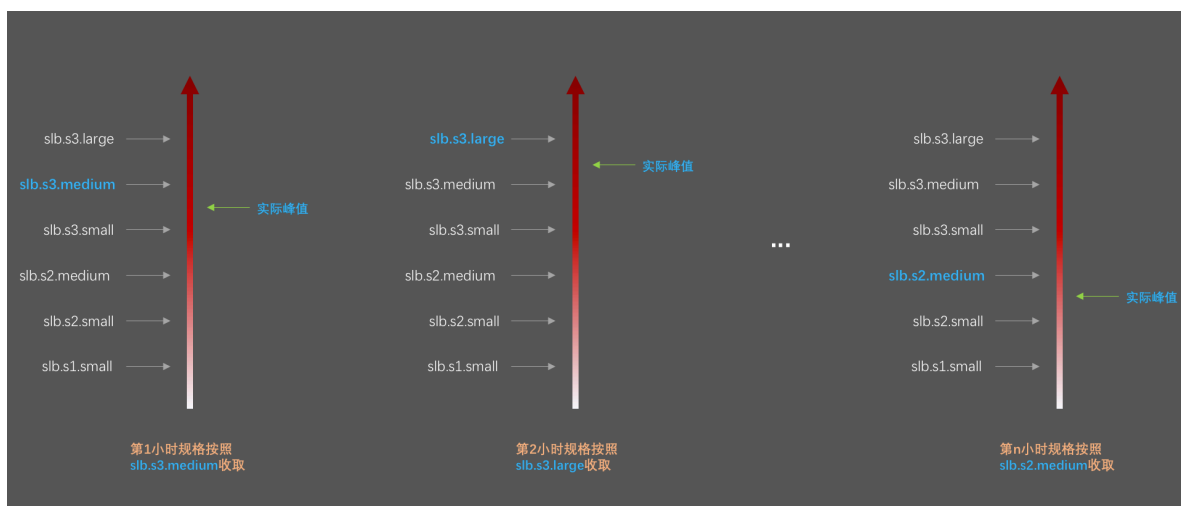
例如，您选择了超强型I (slb.s3.large)规格（最大连接数1,000,000；CPS 500,000；QPS 50,000）。您的实例在某个小时内各项指标产生的实际峰值如下：

最大连接数	每秒新建连接数 (CPS)	每秒查询数 (QPS)
90000	4000	11000

- 从最大连接数维度看，90,000超过slb.s2.small规格中最大连接数50,000的上限，但未达到slb.s2.medium规格中最大连接数的100,000上限，因此从最大连接数维度计算，该小时规格为slb.s2.medium。
- 从每秒新建连接数 (CPS) 维度看，4,000超过slb.s1.small规格中CPS 3,000的上限，但未到达slb.s2.small规格中CPS 5,000的上限，因此从CPS维度计算，该小时规格为slb.s2.small。
- 从每秒查询数 (QPS) 维度看，11,000超过slb.s2.medium规格中QPS 10,000的上限，但未达到slb.s3.small中QPS 20,000的上限，因此从QPS维度计算，该小时规格为slb.s3.small。

综合以上三个维度，QPS指标的规格 (slb.s3.small) 最大，因此将QPS维度的规格作为该小时实例的综合规格，该小时内该实例将按照slb.s3.small规格进行计费。

以后每小时规格费均按照上述方式计算，如下图所示：



因此，按量付费的性能保障型实例具有自动弹性伸缩（或计费）的能力。您在购买时所选的规格，是性能的上限，比如您选择高阶型II (slb.s3.medium)，那么意味着，您的实例最大可以达到的规格上限就是高阶型II (slb.s3.medium)。

## 3. 性能保障型实例规格费的定价

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费和流量/带宽费正常收取。更多详细信息，参考[计费说明](#)。

地域	规格	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(元/月)	按量付费(元/时)
华东1 (杭州)	规格 1	简约型I (slb.s1.small)	5,000	3,000	1,000	免费	免费
华北3 (张家口)	规格 2	标准型I (slb.s2.small)	50,000	5,000	5,000	190.00	0.32
华北5 (呼和浩特)	规格 3	标准型II (slb.s2.medium)	100,000	10,000	10,000	380.00	0.63
华北1 (青岛)	规格 4	高阶型I (slb.s3.small)	200,000	20,000	20,000	760.00	1.27
华北2 (北京)	规格 5	高阶型II (slb.s3.medium)	500,000	50,000	30,000	1,143.00	1.91
华东2 (上海)	规格 6	超强型I (slb.s3.large)	1,000,000	100,000	50,000	1,908.00	3.18
华南1 (深圳)							

地域	规格	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(元/月)	按量付费(元/时)
亚太东南 1 (新加坡)	规格 1	简约型I (slb.s1.small)	5,000	3,000	1,000	免费	免费
亚太东南 3 (吉隆坡)	规格 2	标准型I (slb.s2.small)	50,000	5,000	5,000	228.00	0.38
亚太东南 5 (雅加达)	规格 3	标准型II (slb.s2.medium)	100,000	10,000	10,000	456.00	0.76
亚太南部1 (孟买)	规格 4	高阶型I (slb.s3.small)	200,000	20,000	20,000	912.00	1.52
美国西部1 (硅谷)	规格 5	高阶型II (slb.s3.medium)	500,000	50,000	30,000	1,372.00	2.29
美国东部 1 (弗吉尼亚)	规格 6	超强型I (slb.s3.large)	1,000,000	100,000	50,000	2,290.00	3.82
香港							
亚太东北1 (东京)							
亚太东南1 (悉尼)							
中东东部1 (迪拜)							
欧洲中部 1 (法兰克福)							
英国 (伦敦)							

#### 4. 如何选择性能保障型实例？

- 如果您购买的是按量付费实例，如上文所描述，规格费是按量（弹性）计费的，因此建议您直接选择您可以买到的最大规格，对于大多数用户而言，即高阶型I(slb.s3.large)，这样可以保证较好的业务灵活性（弹性），且不会让您额外多付出成本。但如果您认为您的业务量不太可能到达超强型I(slb.s3.large)，也可以设置一个合理的弹性上限，比如高阶型II(slb.s3.medium)。
- 如果您购买的是预付费实例，情况会略微复杂一点。因为规格费按照固定费率恒定收取，而您不希望购买一个超出您实际业务量很多的规格，并因此付出不必要的成本，因此您需要评估您的实

际业务量，并合理的考虑一些冗余，然后选择一个较合适的规格，对于业务量评估来说，主要参考下面几个原则：

- 如果是四层监听，关注的重点是长连接的并发连接数，那么最大（并发）连接数应当作为一个关键指标来参考。根据不同的业务场景，您需要预估一个负载均衡实例需要承载的最大并发连接数，并选择相应的规格。
- 如果是七层监听，关注的重点是QPS的性能，QPS决定了一个七层应用系统的吞吐量。同样，您也需要根据经验对QPS进行预估。在初步选定一个规格后，在业务压测和实测过程中对规格进行微调。
- 结合与性能保障型实例一起推出的其它关键监控指标，查看实际业务流量的走势、峰值情况，对性能规格进行更加精确的选取。更多详细信息，参考[监控数据](#)。

### 并发连接数监控示例



### 新建连接数监控示例



### QPS监控示例



### 5. 是否可以调整性能保障型实例的规格?

您可在控制台对性能保障型实例进行变配，如下图所示。

负载均衡ID/名称	可用区	服务地址(全部)	状态	网络类型(全部)	端口/健康检查	后端服务器	实例规格	带宽计费方式(全部)	付费方式(全部)	操作
226w90b0...	华北 2 可用区 B(主) 华北 2 可用区 A(备)	47.138.101.140 公网	运行中	经典网络	未配置(配置)	未配置(配置)	性能保障型 slb.s1.small	按固定带宽	包年包月 2017-10-05 00:00:00	管理   更多
226w90b0...	华北 2 可用区 B(主) 华北 2 可用区 A(备)	47.138.101.140 公网	运行中	经典网络	未配置(配置)	未配置(配置)	性能共享型	按使用流量	按量付费 2017-09-01	启动 停止 编辑标签 变更配置 续费
226w90b0...	华北 2 可用区 B(主) 华北 2 可用区 A(备)	47.138.101.140 公网	运行中	经典网络	未配置(配置)	未配置(配置)	性能共享型	按使用流量	按量付费 2017-09-01	

**当前配置**

实例名称: 226w90b0...

计费项: 配置费(公网IP+性能规格)+带宽费    实例规格: 简约型(slb.s1.small)    主可用区: cn-beijing-b    计费周期: 1个月

备可用区: cn-beijing-a    带宽值: 483Mbps    流量: 按固定带宽计费    实例: 公网

地域: 华北 2(北京)    slb服务: 是    云盾: 是    可用区类型: 多可用区

到期时间: 2017-10-05 00:00:00

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需要开通白名单才可以降配。详情参考[包年包月实例变配](#)。

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。



#### 说明:

- 将性能共享型实例变更为性能保障型实例后，无法再将其变更回性能共享型。

- 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。
- 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，其他变配操作均不会影响业务。因此建议在业务低谷期进行此类变配。
- 所有的变配操作都不影响负载均衡实例的IP地址。

## 警告!!!

### 请注意：

如将性能共享型实例变更为性能保障型实例，SLB将有小概率出现短暂的业务中断（1秒），建议在业务低谷期进行变配，或者使用GSLB将业务调度至其他的SLB实例后，再仅对计费方式和带宽进行变更，业务不受上述影响。  
性能共享型实例变配为性能保障型实例后，无法再变回性能共享型实例！

我已知晓上述风险，

## 6. 性能保障型实例何时收费？

阿里云负载均衡计划将于2018年4月1日开始针对性能保障型实例收取规格费，同时继续保留性能共享型实例的售卖。

性能保障型实例的规格费收取将按地域分批次生效：

- 第一批：

生效时间：4月1日至4月10日陆续生效

生效地域：亚太东南1（新加坡）、亚太东南3（吉隆坡）、亚太东南5（雅加达）、亚太南部1（孟买）、美国西部1（硅谷）、美国东部1（弗吉尼亚）

- 第二批：

生效时间：4月11日至4月20日陆续生效

生效地域：华东1（杭州）、华北3（张家口）、华北5（呼和浩特）、香港



- 第三批:

生效时间: 4月21日至4月30日陆续生效

生效地域: 华北1 (青岛)、华北2 (北京)、华东2 (上海)、华南1 (深圳)

#### 7. 收取规格费以后, 性能共享型实例需也会额外收取费用吗?

不会。

原有的性能共享型实例 (如果您不将其变配性能保障型) 将继续保持为性能共享型实例, 不收取规格费。您也可以通过变配, 将性能共享型实例升级成性能保障型实例。变更成性能保障型后, 当性能保障型实例开始正式收费时, 该实例将收取规格费。

#### 8. 为何有时性能保障型实例看起来达不到规格中的性能指标上限?

短木板原理。

性能保障型实例并不保障三个指标 (包含带宽指标) 同时达到指定规格的指标上限。即规格中哪个指标先达到峰值, 就以哪个指标开始限速。

同样, 如果购买了按带宽付费的实例, 当实例带宽达到峰值上限时, 也可能出现因为带宽限速而导致某些指标达不到规格上限的情况。

比如某用户选择高阶型I (slb.s3.small) 实例, 当实例的QPS已经达到20000, 但并发连接数确远未达到20万, 那么该实例最大连接数可能永远都不会达到规格上限, 因为新建的连接请求会因为QPS达到上限而被丢弃。

#### 9. 还可以购买性能共享型实例吗?

可以。

当前继续开放性能共享型实例的售卖, 后续性能共享型实例有可能会下线, 届时会通过官网公告、邮件等方式通知。

#### 10. 私网负载均衡实例也会收取规格费吗?

如果您选择的是性能共享型私网实例, 则不会收取规格费; 如果您选择的是性能保障型私网实例, 则需要收取规格费。规格费收取方式与公网实例规格费计费规则一致。私网实例免收实例费和流量费。

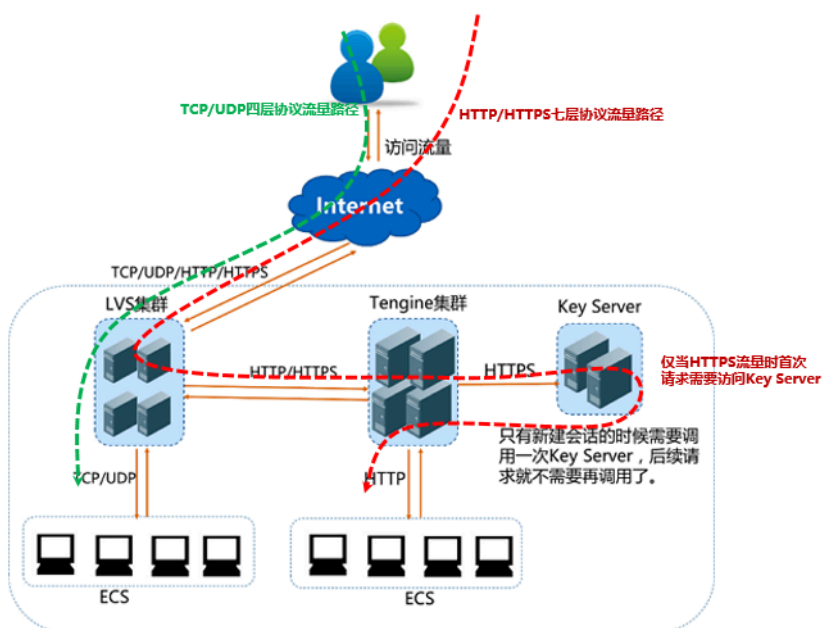
### 1.1.3 网络流量路径说明

负载均衡作为流量转发服务，将来自客户端的请求通过负载均衡集群转发至后端服务器，后端服务器再将响应通过内网返回给负载均衡。

#### 入网流量路径

对于入网流量，负载均衡会根据用户在控制台或API上配置的转发策略，对来自前端的访问请求进行转发和处理，数据流转如图 1-1: 入网流量路径所示。

图 1-1: 入网流量路径




1. TCP/UDP协议和HTTP/HTTPS协议的流量都需要经过LVS集群进行转发。
2. LVS集群内的每一台节点服务器均匀地分配海量访问请求，并且每一台节点服务器之间都有会话同步策略，以保证高可用。
  - 如果相应的负载均衡实例服务端口使用的是四层协议（TCP或UDP），那么LVS集群内每个节点都会根据负载均衡实例负载均衡策略，将其承载的服务请求按策略直接分发到后端ECS服务器。
  - 如果相应的负载均衡实例服务端口使用的是七层HTTP协议，那么LVS集群内每个节点会先将其承载的服务请求均分到Tengine集群，Tengine集群内的每个节点再根据负载均衡策略，将服务请求按策略最终分发到后端ECS服务器。
  - 如果相应的负载均衡实例服务端口使用的是七层HTTPS协议，与上述HTTP处理过程类似，差别是在按策略将服务请求最终分发到后端ECS服务器前，先调用Key Server进行证书验证及数据包加解密等前置操作。

### 出网流量路径

负载均衡SLB和后端ECS之间是通过内网进行通信的。

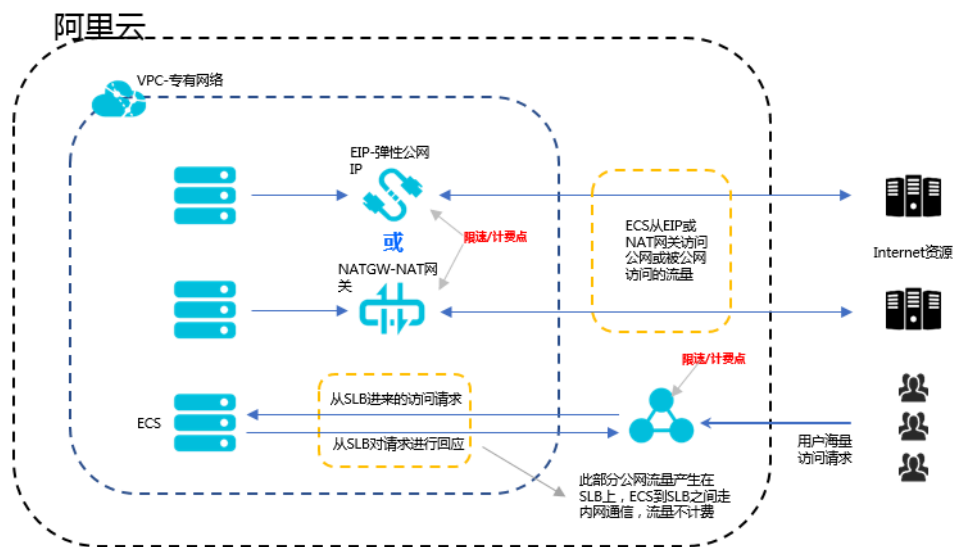
- 如果ECS仅仅处理来自负载均衡的请求，可以不购买公网带宽（ECS公网IP/弹性公网IP/NAT网关等）。

 **说明:**  
 早期创建的一些ECS上直接分配了公网IP（ifconfig中可见接口上分配的公网ip地址），此类ECS如果仅通过SLB对外提供服务，即便在公网接口（网卡）上看到有流量统计，也不会产生ECS的公网费用。

- 如果需要直接通过后端ECS对外提供服务，或后端ECS有访问外网的需求，那么需要相应的配置或购买ECS公网IP/弹性公网IP/NAT网关等服务。

ECS的公网流量访问路径如图 1-2: 出网流量路径所示。

图 1-2: 出网流量路径



总体原则：流量从哪里进来，就从哪里出去。

1. 通过负载均衡进入的流量在负载均衡SLB上限速/计费，仅收取出方向流量费用，入方向流量不收取（在未来可能会改变），SLB到ECS之间是阿里云内网通信，不收取流量费用。
2. 来自弹性公网IP/NAT网关的流量，分别在弹性公网IP/NAT网关上进行限速/计费，如果在购买ECS时选择了公网带宽，限速/计费点在ECS上。
3. 负载均衡SLB仅提供被动访问公网的能力，即后端ECS只能在收到通过负载均衡SLB转发来的公网的请求时，才能访问公网回应该请求，如后端ECS希望主动发起公网访问，则需要配置/购买ECS公网带宽、弹性公网IP或NAT网关来实现。

4. ECS公网带宽（购买ECS时配置）、弹性公网IP、NAT网关均可以实现ECS的双向公网访问（访问或被访问），但没有流量分发和负载均衡的能力。

## 1.1.4 创建实例

在配置负载均衡服务前，您必须先创建一个负载均衡实例。

### 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[准备工作](#)。

### 操作步骤

1. 登录[负载均衡管理控制台](#)。
2. 在左侧导航栏，单击实例管理，然后单击左上角的创建负载均衡。
3. 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。

4. 根据如下信息，配置负载均衡实例。

配置	说明
地域	<p>选择负载均衡实例的所属地域。</p> <div style="border: 1px solid #ccc; padding: 5px; margin-top: 10px;">  说明：            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。         </div>
可用区类型	<p>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。</p> <ul style="list-style-type: none"> <li>· 单可用区：负载均衡实例只部署在一个可用区上。</li> <li>· 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。</li> </ul>
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。
实例规格	<p>选择一个性能规格。</p> <p>不同的性能规格所提供的性能指标也不同，详情查看<a href="#">4. 如何选择性能保障型实例？</a>。</p>

配置	说明
实例类型	<p>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<a href="#">什么是负载均衡实例</a>。</p> <ul style="list-style-type: none"> <li>· 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。</li> <li>· 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。</li> </ul>
IP版本	选择IPv4。
网络类型	<p>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。</p> <ul style="list-style-type: none"> <li>· 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。</li> <li>· 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。</li> </ul>
计费方式	选择一种计费方式。
购买数量	选择购买数量。

5. 单击立即购买，完成支付。

### 1.1.5 创建IPv6实例

负载均衡SLB支持创建IPv6实例。创建后，系统会为实例分配一个公网IPv6地址，可以被互联网上的IPv6客户端访问。

#### 前提条件



说明:

- 目前，仅有华东1地域的E、F两个可用区支持创建IPv6实例且实例类型必须为性能保障型实例。
- 互联网IPv6网络大环境还处于建设初期，当前可能存在部分线路访问不通，如有请工单反馈，同时IPv6公测期间，不提供SLA保障。

#### 背景信息

##### IPv6介绍

IPv6是Internet Protocol Version 6的缩写，其中Internet Protocol译为互联网协议。IPv6是IETF（互联网工程任务组，Internet Engineering Task Force）设计的用于替代现行版本IP协议（IPv4）的下一代IP协议，通过将IPv4中32位的地址长度扩展为128位，使得地址空间扩大了79,228,162,514,264,337,593,543,950,336倍。使用IPv6，可以让全世界的每一粒沙子都能分配到一个IP地址。

## SLB IPv6 特点

负载均衡IPv6支持有以下特点：

- 平滑迁移IPv6，业务无感知

IPv6 SLB后端可以直接挂载使用IPv4地址的ECS，无需对原有系统做改造，就可以平滑地将业务迁移到IPv6。

通过新增IPv6入口，对原有IPv4业务无任何影响，仅需要在业务总量增加的情况下，适量对后端ECS进行横向扩容即可。

- IPv6访问控制让业务部署更加安全可靠

阿里云负载均衡SLB支持IPv6访问控制，您可以根据业务需要灵活地配置访问控制策略。

- 访问控制黑名单可有效阻断恶意地址对负载均衡业务的访问。
- 访问控制白名单仅允许白名单中授权的地址访问负载均衡业务。

## 操作步骤

1. 登录[负载均衡管理控制台](#)。
2. 选择实例 > 实例管理。
3. 在实例管理页面，单击左上角的创建负载均衡。
4. 配置负载均衡实例，IP版本选择IPv6。

其他配置和普通实例配置相同，参考[SLB实例配置说明](#)。



说明：

目前，仅有华东1地域的E、F两个可用区支持创建IPv6实例且实例类型必须为性能保障型实例。

**基本配置**

地域

华北 1(青岛)	华北 2(北京)	华北 3(张家口)	华北 5(呼和浩特)	华东 1(杭州)	华东 2(上海)
华南 1(深圳)	香港	亚太东北 1(东京)	亚太东南 1(新加坡)	亚太东南 2(悉尼)	亚太东南 3(吉隆坡)
亚太东南 5(雅加达)	亚太南部 1(孟买)	美东 1(弗吉尼亚)	美西 1(硅谷)	中东东部 1(迪拜)	欧洲中部 1(法兰克福)

不同地域之间的产品内网不互通；订购后不支持更换地域，请谨慎选择教我选择>> 查看我的产品地域>> 各区域黑洞触发阈值>>

可用区类型 **多可用区**

单可用区指实例只在一个可用区存在；多可用区指实例在两个可用区存在,当主可用区不可用时会在备可用区恢复服务。详情参考>>

主可用区 **华东 1 可用区 F**

主可用区是当前承载流量的可用区，备可用区默认不承载流量，主可用区不可用时才承载流量 教我选择>>

备可用区 **华东 1 可用区 E**

实例名称

长度限制为1-80个字符，允许包含中文、字母、数字、'-','/','.'、'\_'这些字符

---

**网络与实例类型**

实例规格 **标准型I (slb.s2.small)**

该规格最大可以支持连接数: 50000，新建连接数 (CPS): 5000，每秒查询数 (QPS): 5000  
性能保障型实例2018年4月起正式收取规格费  
【按量付费模式下可选择最大规格，规格费将根据每小时使用的实际规格进行收取，闲时免规格费】  
点击查看具体收费详情>>

实例 **公网** [实例类型详解>>](#) [?](#)

负载均衡实例仅提供公网IP，可以通过Internet访问的负载均衡服务

IP版本 **IPv6**

### 5. 返回实例列表页面，查看已创建的IPv6实例。

### 预期结果

创建后，系统会为该实例分配一个IPv6地址。

关于负载均衡收费的说明：  
阿里云负载均衡计划将于2018年4月1日开始针对性能保障型实例收取规格费，具体收费方式以及如何使用性能保障型实例，[查看详情](#)

**实例管理** 刷新 创建负载均衡

负载均衡名称  请输入实例名称进行模糊查询.多个按,分隔 搜索 标签 管理自动续费 ≡ 🔍 ⚙️

负载均衡ID/名称	可用区	服务地址(全部)	状态	网络类型(全部)	端口/健康检查/后端服务器	实例规格	带宽计费方式(全部)	操作
<input type="checkbox"/> lb-bp15fsc... (未设置)	华东 1 可用区 F(主) 华东 1 可用区 E(备)	2400:3200:1500::37(公网)	运行中	经典网络	无监听(配置)	未配置(配置)	性能保障型 slb.s2.small	按使用流量 <span>管理</span>   <span>更多</span>

## 1.1.6 管理实例

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，或选择 管理 > 停止。

- 启动负载均衡实例

勾选已经停止运行的负载均衡实例，单击页面下方的启动，或选择 管理 > 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，或选择 管理 > 释放设置。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[标签](#)。

- 变更计费

单击 更多 > 变更配置，修改实例的类型和带宽。

- 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。

- 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](#)。
- 单击服务器，查看或添加后端服务器。详情参考[服务器](#)。
- 单击监控，查看监控信息，设置报警机制。详情参考[监控](#)。

## 1.1.7 管理标签

负载均衡提供标签管理功能，方便您通过标签对负载均衡实例进行分类。

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

完成以下操作，添加标签：

1. 登录[负载均衡管理控制台](#)。
2. 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。
3. 单击编辑标签。



4. 在编辑标签窗口，单击新建，然后输入新建标签的标签键和值，单击确定。

### 搜索实例

完成以下操作，搜索指定标签关联的实例：

1. 登录[负载均衡管理控制台](#)。
2. 在实例管理页面，选择地域，查看该地域的所有实例。
3. 单击请选择标签，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。



4. 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。

### 删除标签

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。

完成以下操作，删除标签：

1. 登录[负载均衡管理控制台](#)。
2. 在实例管理页面，选择地域，查看该地域的所有实例。
3. 选择要删除标签的实例，单击更多 > 编辑标签。
4. 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。



说明:

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。

### 1.1.8 绑定EIP

您可以为专有网络类型的SLB实例绑定一个EIP。绑定后，SLB实例便可以转发来自公网的请求。

#### 操作步骤

1. 登录[负载均衡管理控制台](#)。
2. 在左侧导航栏，单击实例管理。
3. 选择负载均衡实例的地域，找到目标实例。



说明:

确保负载均衡实例的网络类型为专有网络。

4. 单击更多 > 绑定EIP。
5. 选择一个EIP，然后单击确认。

### 1.1.9 按量付费实例变配

您可以更改后付费实例的带宽、实例规格和计费方式（按流量计费或按带宽计费）。

#### 背景信息

在变更前，请注意：

- 性能共享型实例变更为性能保障型实例会有10-30秒业务中断，其他变配操作均不会影响业务，也不会变更负载均衡实例IP。建议在业务低谷期进行此类变配，或先通过[全球负载均衡（GSLB）](#)实现实例间的负载均衡后，再进行变配。

- 将性能共享型实例变更为性能保障型实例后，无法再将其变更回性能共享型。

您可以选择使用简约型I (slb.s1.small)规格，该规格免收规格费。

- 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。
- 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。

## 操作步骤

1. 登录**负载均衡管理控制台**。
2. 选择目标实例的所属地域。
3. 找到目标实例，然后单击**更多 > 变更配置**。

关于负载均衡收费的说明：  
阿里云负载均衡国内5大区域全面开启性能保障型实例，目前用户可以在北京、杭州、深圳、上海、青岛5大区域通过控制台购买性能保障型实例（API暂未开放）同时继续保留共享型实例售卖，现计划延长公测期，继续保持规格免费，后续收费计划将通过官网公告、短信、邮件方式进行通知，关于性能保障型实例如何使用请 [查看详情](#)

实例管理 华北1 华北2 华北3 华北5 华东1 华东2 华南1 香港 亚太东北1(东京) 亚太东南1(新加坡) 亚太东南2(悉尼) 亚太东南3(吉隆坡)  
美国东部1(弗吉尼亚) 美国西部1(硅谷) 中东东部1(迪拜) 欧洲中部1(法兰克福)

刷新 创建负载均衡

负载均衡名称 请输入实例名称进行模糊查询,多个按分隔 搜索 标签

负载均衡ID/名称	可用区	服务地址(全部)	状态	网络类型(全部)	端口/健康检查	后端服务器	标签	实例规格	带宽计费方式(全部)	付费方式(全部)	操作
lb-bp1wz7h8w... (未设置)	华东1可用区F(主) 华东1可用区E(备)	120.15(公网)	运行中	经典网络	未配置(配置)	未配置(配置)	未绑定	性能保障型 slb.s2.small	按固定带宽	2017-11-30 19:41	管理 更多
lb-bp1kjekef... acs-slb-c4e1afcd9e...	华东1可用区F(主) 华东1可用区E(备)	118.29(公网)	运行中	经典网络	TCP: 80 正常	c4e1afcd9efe547... c4e1afcd9efe547...	未绑定	性能共享型	按使用流量	2017-15:3	启动 停止 释放 编辑标签 变更配置
	华东1可用区					ca5ce0cca9da346...					

4. 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:

- 将性能共享型实例变更为性能保障型实例后，无法再将其变更回性能共享型。

- 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。

配置变更

实例规格 **标准型 (slb.s2.small)**

连接数: 50000, 新建连接数 (CPS): 5000, 每秒查询数 (QPS): 5000  
性能保障型实例公测期间, 免收性能规格费, **正式收费后, 该规格费用为: ¥0.32/小时** 详情点击>>

实例类型 **公网** [实例类型详解>>](#) (?)

计费方式 **按使用流量计费** **按固定带宽计费**

开通后即开始按固定带宽计费, 和实例状态及使用流量无关  
进行变配操作时, 若选择的计费方式与当前计费方式不同, 则代表变更计费方式, 变更计费方式将在第二天0点生效  
进行变配操作时, 若选择的计费方式与当前计费方式相同, 则代表变更带宽规格, 变更带宽规格即时生效  
阿里云最高提供5Gbps的恶意流量攻击防护, [了解更多>>](#) [提升防护能力>>](#)

带宽值 **1250Mbps** 2500Mbps 5000Mbps **9** Mbps

开通后即开始按固定带宽计费, 和实例状态及使用流量无关

服务监听设置 **每个服务监听都需要设置带宽峰值限制, 并且只能为大于0的整数, 总和不能大于带宽值。**

## 1.1.10 包年包月实例变配

### 背景信息

在变配前，请注意：

- 将性能共享型实例变更为性能保障型实例后，无法再将其变更回性能共享型。
- 将性能共享型实例变更为性能保障型实例时，可能会出现10-30秒的业务中断（其他变配操作均不会影响业务）。因此建议在业务低谷期进行此类变配，或通过DNS实现实例间的负载均衡后，再进行变配。
- 所有的变配操作都不影响负载均衡实例的IP地址。
- 如果您需要对包年包月实例进行降配操作，请登录[配额管理](#)页面，自主申请预付费实例实时降配特权。

### 操作步骤

1. 登录[负载均衡管理控制台](#)。
2. 选择目标实例的所属地域。
3. 找到目标实例，然后单击更多 > 降配或升配。

4. 在配置变更区域，选择新的带宽值或规格，完成支付。

在变更配置时，您可以为实例下的监听配置带宽，若不配置则所有监听共享实例的带宽。详情参见[共享实例带宽](#)。



说明：

将性能共享型实例变更为性能保障型实例后，无法再将其变更回性能共享型。

## 1.2 监听

### 1.2.1 监听介绍

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/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	<ul style="list-style-type: none"> <li>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</li> <li>基于源地址的会话保持</li> <li>在网络层可直接看到来源地址</li> <li>数据传输快</li> </ul>	<ul style="list-style-type: none"> <li>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</li> <li>无特殊要求的Web应用</li> </ul> <p>详情参见<a href="#">添加TCP监听</a>。</p>
UDP	<ul style="list-style-type: none"> <li>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</li> <li>可靠性相对低；数据传输快</li> </ul>	<p>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。</p> <p>详情参见<a href="#">添加UDP监听</a>。</p>
HTTP	<ul style="list-style-type: none"> <li>应用层协议，主要解决如何包装数据</li> <li>基于Cookie的会话保持</li> <li>使用X-Forward-For获取源地址</li> </ul>	<p>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。</p> <p>详情参见<a href="#">添加HTTP监听</a>。</p>

协议	说明	使用场景
HTTPS	<ul style="list-style-type: none"><li>· 加密传输数据，可以阻止未经授权的访问</li><li>· 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</li></ul>	<p>需要加密传输的应用。</p> <p>详情参见<a href="#">添加HTTPS监听</a>。</p>



说明:

负载均衡已在全部地域支持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FAQ](#)。

### 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](#)。

### 添加监听

1. 基本配置 2. 健康检查配置 3. 配置成功

健康检查方式:  TCP  HTTP

检查端口:   
默认使用后端服务器的端口进行健康检查

收起高级配置

响应超时时间: \*  秒  
每次健康检查响应的最大超时时间; 输入范围1-300秒, 默认为5秒

健康检查间隔: \*  秒  
进行健康检查的时间间隔; 输入范围1-50秒, 默认为2秒

不健康阈值: \*  2 3 4 5 6 7 8 9 10  
表示云服务器从成功到失败的连续健康检查失败次数。

健康阈值: \*  2 3 4 5 6 7 8 9 10  
表示云服务器从失败到成功的连续健康检查成功次数。

上一步 确认 取消

## 1.2.2 四层监听

## 1.2.3 七层监听

## 1.2.4 健康检查

## 1.2.5 共享实例带宽

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一直没有出流量，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一直没有出流量，而监听B的出流量很大，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 的出流量，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也有出流量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出流量同等大小，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。

## 1.3 后端服务器

### 1.3.1 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。

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。



说明：

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。

#### 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中的ECS。

参见[添加默认服务器](#)创建一个默认服务器组。

#### 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

添加后端服务器时，注意：

- 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。



### 说明：

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：

1. 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。
  2. 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。）
- ECS进行热迁移时，可能导致SLB长连接断开，重连即可恢复，请做好应用的重连工作。

## 1.3.2 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组。

### 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。

### 操作步骤

1. 登录[负载均衡管理控制台](#)。

2. 在实例管理页面，选择目标实例的所属地域。
3. 单击目标实例的ID。
4. 选择默认服务器。
5. 单击添加。
6. 在未添加的服务器页面，单击添加。

系统显示我的服务器页面。

7. 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。
8. 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。

### 1.3.3 创建虚拟服务器组

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。

#### 前提条件

- 您已经[创建负载均衡实例](#)。
- 您创建了 ECS 实例，用来接收转发的请求。

#### 背景信息

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。

如果您在您同时为一个七层监听添加了后端服务器、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。

在创建虚拟服务器组时，请注意：

- 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加端口组成。

#### 操作步骤

1. 登录 [负载均衡管理控制台](#)。
2. 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。
3. 单击负载均衡实例的 ID 链接，进入负载均衡的详情页面。
4. 在负载均衡实例菜单栏，选择虚拟服务器组。
5. 在虚拟服务器组页面，单击 [创建虚拟服务器组](#)。
6. 在 [创建虚拟服务器组](#) 对话框，完成如下操作：
  - a) 在 [虚拟服务器组名称](#) 文本框中，输入虚拟服务器组名称。
  - b) 单击 [添加](#)，在我的服务器列表选择要添加的服务器。
7. 在已添加服务器，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。

### 1.3.4 创建主备服务器组

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，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。

#### 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。

#### 背景信息

将主备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主备服务器组中的后端服务器，不会再将流量转发给其他后端服务器。



注意：

主备服务器组只支持四层监听（TCP和UDP协议）。

#### 操作步骤

1. 登录[负载均衡管理控制台](#)。
2. 在实例管理页面，选择目标实例的所属地域。
3. 单击目标实例的ID。

4. 选择主备服务器页签。
5. 在主备服务器组页面，单击创建主备服务器组。
6. 在创建主备服务器组对话框，完成如下操作：
  - a) 在主备服务器组名称文本框中，输入主备服务器组名称。
  - b) 单击添加，在我的服务器列表中，选择目标ECS实例。
  - c) 在已添加服务器中，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。

## 1.4 DDoS基础防护

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，云盾将会针对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。

### 查看防护阈值

使用子账号登录阿里云控制台后，若无法在控制台上查看防护阈值，需要先对子账号授权。详情参见[授予云盾基础防护只读权限](#)。

完成以下操作查看防护阈值：

1. 登录[负载均衡管理控制台](#)。
2. 选择地域，查看该地域的所有实例。

3. 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。

#### 授权云盾基础防护只读权限

完成以下操作授予子账号只读访问云盾DDoS基础防护(Anti-DDoS Basic)的权限：



说明：

使用主账号进行授权。

1. 使用主账号登录访问控制RAM管理控制台。
2. 在左侧导航栏，单击用户管理，找到目标子账号，然后单击管理。
3. 单击用户授权策略，然后单击编辑授权策略。
4. 在弹出的对话框，在可授权策略列表中搜索AliyunYundunDDoSReadOnlyAccess，将其加入到已授权策略列表。单击确定。

## 1.5 各地域带宽峰值限制

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。



说明：

所有地域的私网带宽峰值都为5 Gbps。

地域	带宽峰值
华北 1（青岛）	5 Gbps
华东 1（杭州）	5 Gbps
华北 2（北京）	5 Gbps
华东 2（上海）	5 Gbps
华南 1（深圳）	5 Gbps
华北 3（张家口）	5 Gbps

地域	带宽峰值
华北 5 (呼和浩特)	5 Gbps
香港	2 Gbps
美国东部 1 (弗吉尼亚)	1 Gbps
美国西部 1 (硅谷)	2 Gbps
亚太东北 1 (东京)	1 Gbps
亚太东南 1 (新加坡)	5 Gbps
亚太东南 2 (悉尼)	1 Gbps
亚太东南 3 (吉隆坡)	5 Gbps
中东东部 1 (迪拜)	500 Mbps
欧洲中部 1 (法兰克福)	1 Gbps
亚太南部 1 (孟买)	5 Gbps

## 1.6 多可用区

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。

### 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。

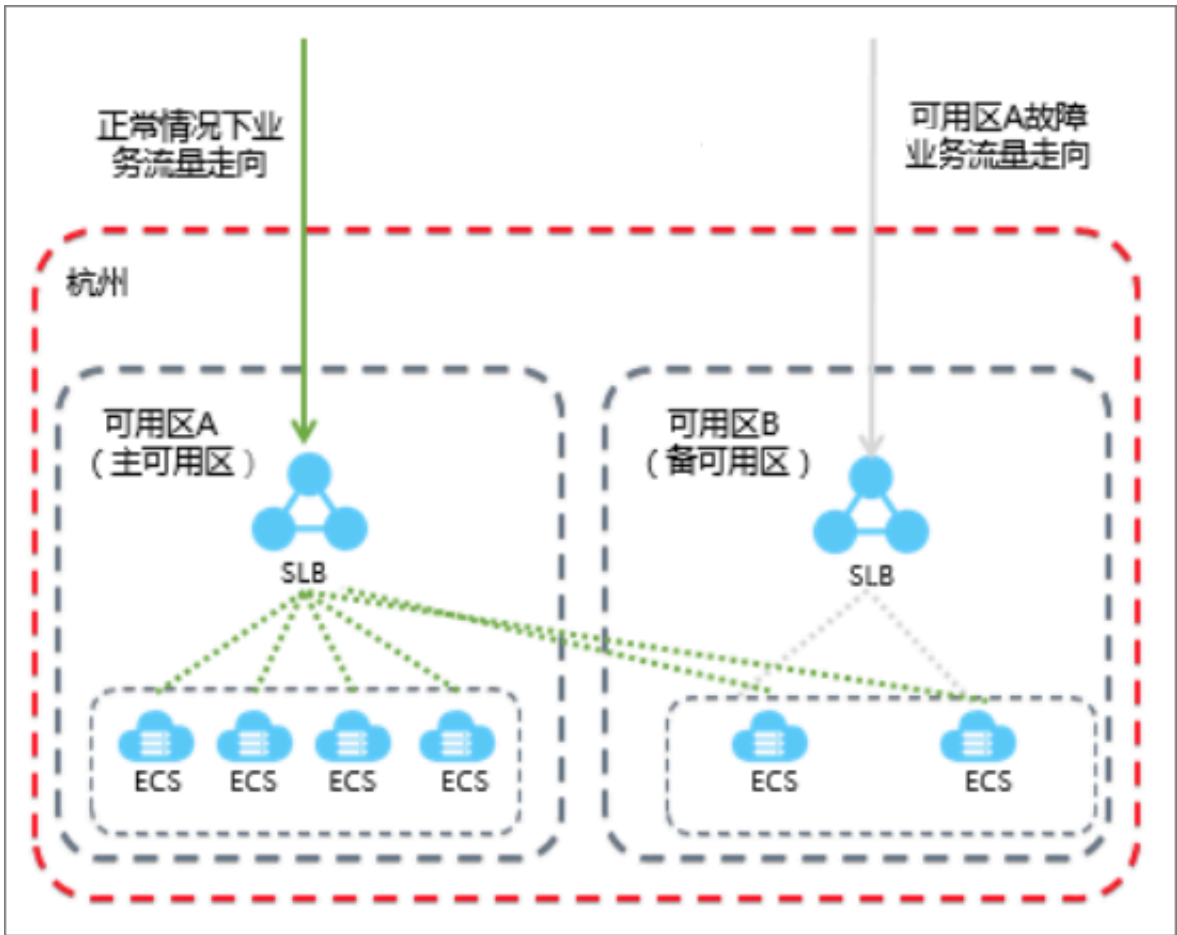
为了向广大用户提供更加稳定可靠的负载均衡服务，阿里云负载均衡已在各地域（Region）部署了多可用区以实现同地域下的跨机房容灾。当主可用区的机房故障或不可用时，负载均衡仍然有能力在非常短的时间内（约30秒）切换到另外一个备可用区的机房并恢复服务的能力；当主可用区恢复时，负载均衡同样会自动切换到主可用区的机房提供服务。

关于负载均衡主备可用区，请注意：

- SLB支持跨可用区挂载后端ECS，即只要ECS和SLB实例在同一个地域即可。SLB可以同时将流量分发至不同可用区的ECS上。
- 正常情况下，备可用区的SLB实例处于待机状态。您不可以手动切换SLB实例的主备工作状态，只有当阿里云检测到整个可用区不可用时如机房整体断电、机房出口光缆中断等，负载均衡才会切换到备可用区。而并非某个实例出现故障，就切换到备可用区。
- SLB和ECS是不同的集群。可用区A的SLB不可用时，ECS并不一定不可用，因此如果仅因为SLB集群故障导致的SLB主备倒换，备可用区的SLB依然可以将流量分发至不同可用区的ECS。

但当整个可用区的所有集群断电或光缆中断时，那么可用区的所有服务（包括但不限于SLB、ECS等）就都无法正常工作了。

更多信息，参见[负载均衡高可用最佳实践](#)。



### 主备可用区列表

下表列举了各地域的主备可用区，您可以通过DescribeZones接口查询可用的主备可用区。

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



地域	可用区类型	可用区	
		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	可用区A
华北5	单可用区	可用区A	可用区A
欧洲中部1	单可用区	可用区A	可用区A
中东东部1	单可用区	可用区A	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	可用区A
亚太东南3	单可用区	可用区A	可用区A
亚太东北1	单可用区	可用区A	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	可用区A

地域	可用区类型	可用区	
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