阿里云
负载均衡

常见问题

文档版本：20190717

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。	 **禁止:** 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。	 **警告:** 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。	 **说明:** 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
粗体	表示按键、菜单、页面名称等UI元素。	单击 确定。
`courier` 字体	命令。	执行 `cd /d C:/windows` 命令，进入Windows系统文件夹。
*## *	表示参数、变量。	`bae log list --instanceid` *`Instance_ID`*
[]或者[a|b]	表示可选项，至多选择一个。	`ipconfig` *`[-all|-t]`*
{}或者{a|b}	表示必选项，至多选择一个。	`swich` *`{stand | slave}`*

目录

1 为什么无法访问负载均衡

本文主要介绍通过客户端无法访问负载均衡的可能原因和处理方法。

说明:

本次示例中负载均衡前端端口是80，ECS后端端口为80，ECS内网IP是10.11.192.1。对于无法访问负载均衡问题进行排查时，需要根据实际情况配置端口和内网IP信息。

序号	可能原因	处理方法
1	后端服务器无法访问SLB，对于四层负载均衡服务，目前不支持负载均衡后端ECS实例直接为客户端提供服务的同时，又作为负载均衡的后端服务器。	-
2	健康检查异常。	健康检查方法请参见四层监听（TCP/UDP）健康检查异常排查和七层监听（HTTP/HTTPS）健康检查异常排查。
3	不支持通过SLB搭建FTP、tftp、h323和sip等	· 如果是Linux系统，您可以尝试配置22端口的转发，使用sftp连接传输数据。 · 支持通过EIP可见模式将EIP绑定到FTP服务器上对外提供FTP服务，配置详情请参见使用EIP部署FTP服务器。

4	服务器内网防火墙设置没有放行80端口。	可以选择执行如下命令，暂时关闭防火墙进行测试。
		· Windows 服务器上运行：
		`firewall.cpl`
		· Linux 服务器上运行：
		`/etc/init.d/iptables stop`

序号	可能原因	处理方法
5	后端端口异常。	· 对于四层负载均衡，使用telnet测试有响应即为正常。 示例：使用`telnet 10.11.192.1 80`来测试。 · 对于七层负载均衡，HTTP状态码需要是200等代表正常的状态码，检验方法如下： - Windows：直接在ECS上访问ECS的内网IP测试是否正常。 示例：`http://10.11.192.1` - Linux：使用`curl -I`命令查看状态是否为`HTTP/1.1 200 OK`。 示例：`curl -I 10.11.192.1`

6	rp_filter特性和负载均衡底层LVS的策略路由产生冲突，导致访问出现异常。	1. 登录四层负载均衡后端添加的Linux系统的ECS实例。
		2. 编辑*/etc/sysctl.conf*文件，将系统配置文件中的以下三个参数值设置为0。
		`net.ipv4.conf.default.rp_filter = 0`
		`net.ipv4.conf.all.rp_filter = 0`
		`net.ipv4.conf.eth0.rp_filter = 0`
		3. 执行`sysctl -p`命令，使配置生效。

7	监听功能异常	在服务器上执行以下命令，如果能看到10.11.192.1:80的监听信息，或者0.0.0.0:80的监听信息，说明这部分端口的监听正常。
		· Windows 服务器上运行：
		`netstat -ano | findstr :80`
		· Linux 服务器上运行：
		`netstat -anp | grep :80`

8	创建负载均衡实例后，没有添加监听。	请配置监听，详情请参见配置监听。
9	负载均衡通过域名访问不通，可能为用户域名解析错误导致。	-

序号	可能原因	处理方法
10	客户端本地网络或运营商中间链路异常。	从不同地域及不同网络环境，对负载均衡相应服务端口做访问测试。 如果只有本地网络访问时出现异常，则判定是网络异常导致的问题，此时可以继续通过持续进行ping测试或MTR路由跟踪等手段做进一步排查分析。
11	客户端IP被云盾拦截。	1. 在客户端网络环境下访问`http://ip.taobao.com`，获取客户端网络环境对应的公网IP。 2. 将获取的IP配置为白名单，该操作将会对来自相应IP到负载均衡的所有访问全部放行。 说明: 该操作可能会带来安全风险，确保白名单中的IP不会对负载均衡进行恶意攻击。
12	用户使用完高防IP之后切换回普通模式，未关闭访问控制白名单功能。	关闭ACL白名单。
如果还未能解决问题，请在提交工单时提供如下信息，以便我们更高效地协助您解决问题。 · 负载均衡实例ID或负载均衡服务IP地址。 · 访问`ip.taobao.com`时获取的客户端对应的公网IP。 · 公网客户端对负载均衡IP长时间ping及MTR路由跟踪测试截图。		

2 为什么请求不均衡

可能原因

负载均衡请求不均衡可能有以下几种原因：

- ECS实例请求连接数较少。
- 不同ECS实例的性能不同导致请求不均衡。

  说明：

  ECS实例内存使用情况不能准确的判断请求转发是否均衡。
- 开启了会话保持功能 。

  配置了会话保持，当访问负载均衡实例的客户端又很少时，容易导致不均衡，尤其在使用少量客户端对负载均衡进行测试的时候。比如TCP监听，开启了会话保持（四层是基于来源地址做会话保持），使用一台客户端对负载均衡实例进行压测，就会导致不均衡。
- ECS健康检查异常 。

  后端服务器ECS的健康建状态异常会导致不均衡，尤其在压测的时候容易忽略后端服务器ECS的健康检查状态，如果有后端服务器ECS健康检查失败或者健康检查状态经常跳跃（好到坏，又从坏到好，反复变化）必然会导致不均衡。
- TCP Keepalive保持长连接 。

  后端服务器ECS有些开启了TCP Keepalive保持长连接，而有些又没有开启，则连接会在保持长连接的后端服务器上堆积，造成不均衡。

排查和解决方法

- 查看后端各台ECS的权重是否相同。
- 在相关时间段内是否有健康检查失败或波动现象，查找波动的原因；或者健康检查没有配置正确的响应码2xx，3xx导致了健康检查显示正常，但后端服务有异常。
- 是否同时使用了加权最小连接数（WLC）调度方式和会话保持，如果是，尝试改为加权轮询（WRR）算法和会话保持。

3 如何获取客户端真实IP

阿里云负载均衡服务支持获取客户端真实IP地址的功能。

负载均衡服务获取真实IP说明

负载均衡提供获取客户端真实IP地址的功能，该功能默认是开启的。

- 四层负载均衡（TCP协议）服务可以直接在后端ECS上获取客户端的真实IP地址，无需进行额外的配置。
- 七层负载均衡（HTTP/HTTPS协议）服务需要对应用服务器进行配置，然后使用`X-Forwarded-For`的方式获取客户端的真实IP地址。

  真实的客户端IP会被负载均衡放在HTTP头部的X-Forwarded-For字段，格式如下：

  ```
  X-Forwarded-For: 用户真实IP，代理服务器1-IP， 代理服务器2-IP，...
  ```

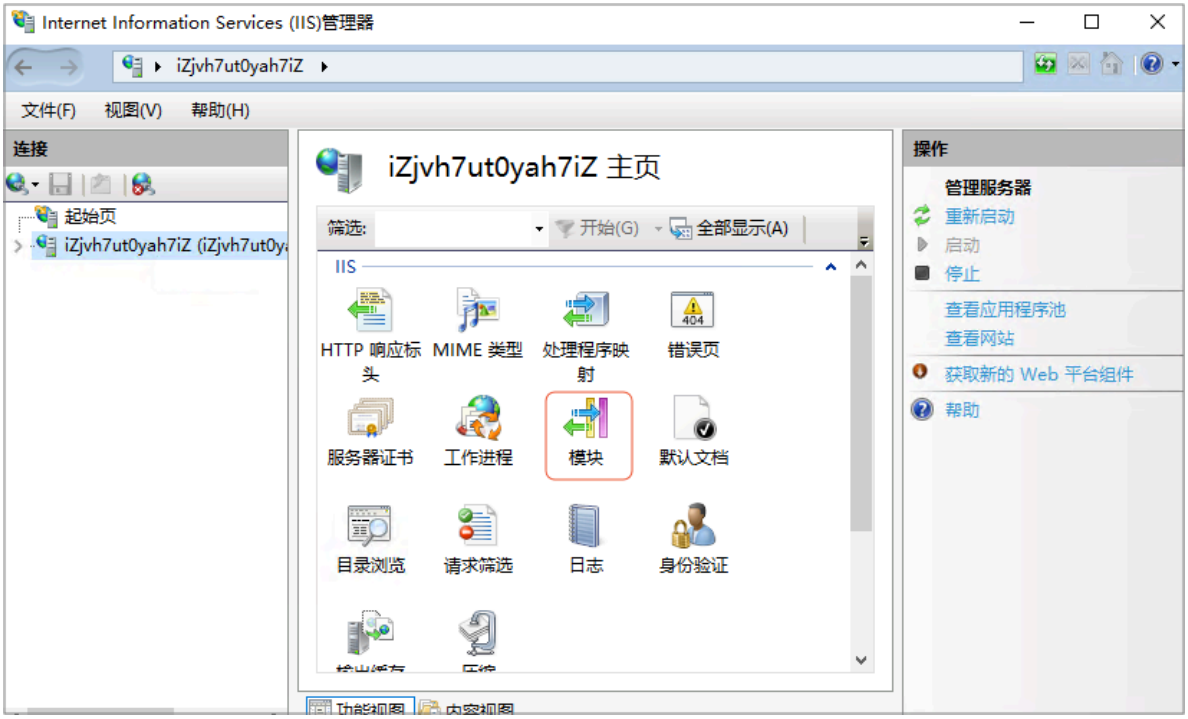
  当使用此方式获取客户端真实IP时，获取的第一个地址就是客户端真实IP。

  > **说明:**
  > 负载均衡的HTTPS监听是在负载均衡服务上的加密控制，后端仍旧使用HTTP协议，因此，在Web应用服务器上配置HTTPS和HTTP监听没有区别。

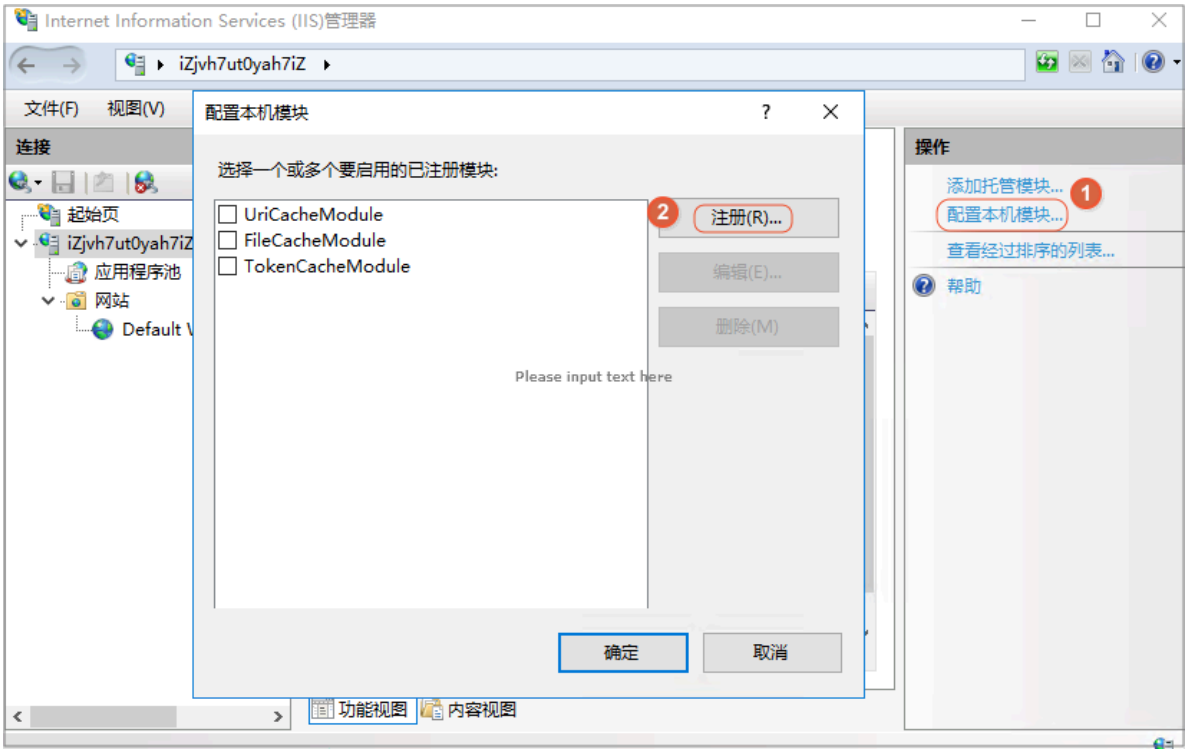
配置IIS7/IIS8服务器

1. 下载并解压 *F5XForwardedFor*文件。
2. 根据自己的服务器操作系统版本将*x86\Release*或*x64\Release*目录下的*F5XFFHttpModule.dll*和*F5XFFHttpModule.ini*拷贝到某个目录，比如*C:\F5XForwardedFor*。确保IIS进程对该目录有读取权限。

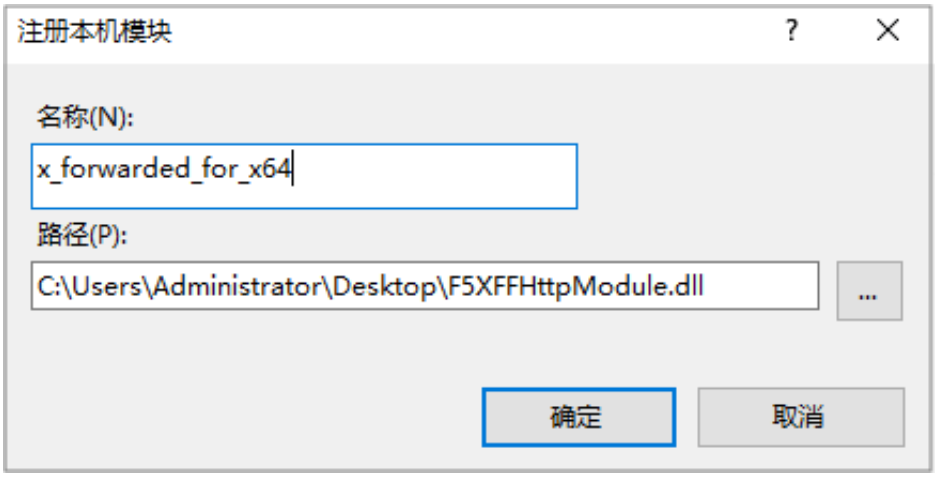
3. 打开IIS管理器，双击模块功能。



4. 单击配置本机模块，然后在弹出的对话框中，单击注册。

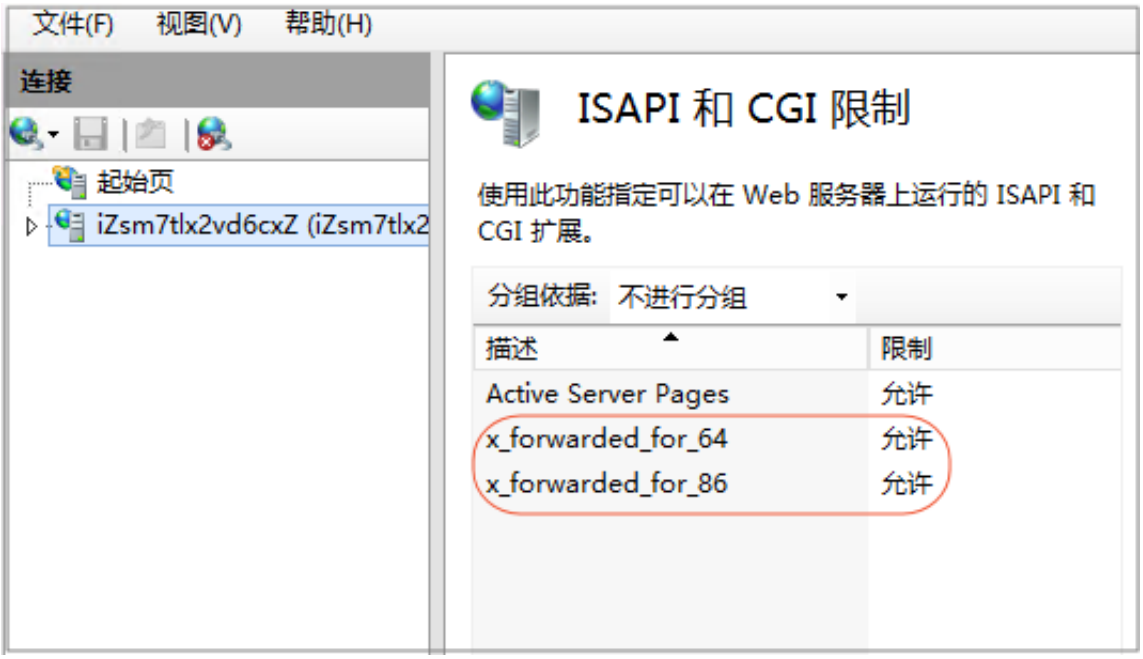


5. 添加下载的.dll文件。



6. 为添加的两个文件授权允许运行ISAPI和CGI扩展。

> **说明:**
> 确保您已经安装了ISAPI和CGI应用程序。



7. 重启IIS服务器，等待配置生效。

配置Apache服务器

1. 运行以下命令安装Apache的一个第三方模块mod_rpaf。

```
wget https://github.com/gnif/mod_rpaf/archive/v0.6.0.tar.gz
tar zxvf mod_rpaf-0.6.tar.gz
cd mod_rpaf-0.6
```

```
 /alidata/server/httpd/bin/apxs -i -c -n mod_rpaf-2.0.so mod_rpaf-2.
0.c
```

2. 修改Apache的配置文件*/alidata/server/httpd/conf/httpd.conf*，在最末尾添加以下配置信息。

```
 LoadModule rpaf_module modules/mod_rpaf-2.0.so
 RPAFenable On
 RPAFsethostname On
 RPAFproxy_ips  <IP_address>
 RPAFheader X-Forwarded-For
```

说明:

如果您要获取代理服务器的地址，可以将代理服务器的网段添加到`RPAFproxy_ips <IP_address>`，如负载均衡的IP地址段100.64.0.0/10（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）和高防IP地址段。多个IP地址段用逗号分隔。

3. 添加完成后重启Apache。

```
/alidata/server/httpd/bin/apachectl restart
```

配置Nginx服务器

1. 运行以下命令安装http_realip_module。

```
 wget http://nginx.org/download/nginx-1.0.12.tar.gz
 tar zxvf nginx-1.0.12.tar.gz
 cd nginx-1.0.12
 ./configure --user=www --group=www --prefix=/alidata/server/
nginx --with-http_stub_status_module --without-http-cache --with-
http_ssl_module --with-http_realip_module
 make
 make install
 kill -USR2 `cat /alidata/server/nginx/logs/nginx.pid`
 kill -QUIT `cat /alidata/server/nginx/logs/ nginx.pid.oldbin`
```

2. 打开*nginx.conf*文件。

```
vi /alidata/server/nginx/conf/nginx.conf
```

3. 在以下配置信息后添加新的配置字段和信息。

```
 fastcgi connect_timeout 300;
 fastcgi send_timeout 300;
 fastcgi read_timeout 300;
 fastcgi buffer_size 64k;
 fastcgi buffers 4 64k;
 fastcgi busy_buffers_size 128k;
```

```
fastcgi temp_file_write_size 128k;
```

需要添加的配置字段和信息为：

```
set_real_ip_from IP_address
real_ip_header X-Forwarded-For;
```

说明：

如果您要获取代理服务器的地址，可以将代理服务器的网段添加到`set_real_ip_from <IP_address>`，如负载均衡的IP地址段100.64.0.0/10（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）和高防IP地址段。多个IP地址段用逗号分隔。

4. 重启Nginx。

```
/alidata/server/nginx/sbin/nginx -s reload
```

4 如何处理健康检查导致的大量日志

负载均衡的日志管理功能会自动保存三天内的健康检查日志，如果健康检查日志过多，对您的运维工作造成不便，您可以选择以下方案来减少或禁止某些场景下健康日志的产生。

说明:

减少健康检查日志的数量可能会导致您无法及时发现负载均衡实例运行时所出现的问题，请您谨慎权衡每种方案所带来的风险，根据您的实际情况进行选择。

- 获取访问日志
- 调整健康检查频率
- 关闭7层负载均衡下的健康检查
- 将7层负载均衡切换为4层负载均衡
- 关闭健康检查页面的应用日志

获取访问日志

HTTP协议健康检查默认使用HEAD请求方法，因此过滤掉HEAD的请求，就可以获得实际的访问日志。

调整健康检查频率

通过延长健康检查的间隔时间来减少健康检查的次数，降低健康检查产生的日志数量。

方案风险说明:

延长健康检查的间隔时间后，后端ECS实例出现故障时，负载均衡发现故障ECS实例的时间也会变长。

操作步骤:

1. 登录负载均衡管理控制台。
2. 在实例管理页面中找到相应的负载均衡实例，单击实例ID。
3. 在监听页签下，找到对应监听，单击监听操作列的配置。
4. 在配置监听对话框中，单击下一步，再单击下一步，进入健康检查配置。

5. 调整健康检查间隔时间，范围为1～50秒，间隔越大，健康检查的频率就越低，后端服务器产生的日志也会相应减少。请根据您的实际情况进行修改。

6. 单击下一步至确定，完成修改。

关闭7层负载均衡下的健康检查

使用7层HTTP/HTTPS负载均衡模式时，健康检查由HTTP Head请求实现，后端服务器的应用日志会记录相应的健康检查请求信息，可能导致大量的日志信息。

风险说明

HTTP/HTTPS模式下关闭健康检查后，负载均衡不再检查后端服务器，一旦某台后端服务器发生故障，则无法实现访问流量自动切换至其它正常的后端服务器。

操作步骤

1. 登录负载均衡管理控制台。

2. 在实例管理页面中找到对应的负载均衡实例，单击实例ID。
3. 在监听页签下，单击操作列的进单击配置。
4. 在配置监听对话框中单击下一步，再单击下一步，进入健康检查配置。
5. 关闭开启健康检查。

6. 单击下一步至确定，完成修改。

将7层负载均衡切换4层负载均衡

4层TCP模式下的的健康检查仅仅使用TCP的三次握手实现，不会生成应用日志。如果您的业务可以切换为4层TCP模式，采用该方法可以减少应用日志的产生。

风险说明

将HTTP/HTTPS模式的负载均衡修改为TCP模式后，负载均衡将只检查监听端口状态，不检查HTTP状态，会导致负载均衡无法实时获知HTTP应用是否出现问题。

操作步骤

1. 登录负载均衡管理控制台。
2. 在实例管理页面中找到对应的负载均衡实例，单击实例ID。
3. 在监听页签下，找到对应监听，单击配置。
4. 在配置监听对话框中单击下一步，再单击下一步，进入健康检查配置。
5. 将健康检查协议修改为TCP。

6. 单击下一步至确定，完成修改。

关闭健康检查页面的应用日志

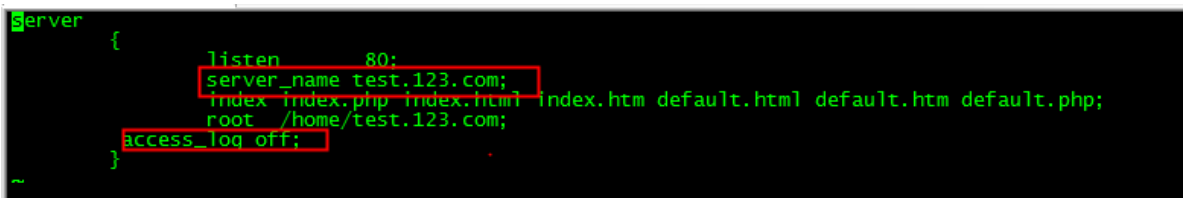
在业务站点之外，独立配置健康检查站点，并关闭健康检查页面的应用日志，可以减少健康检查日志的数量。例如业务站点为abc.123.com，则使用test.123.com作为健康检查站点，并关闭test.123.com站点的日志记录。

风险说明

如果健康检查的站点正常，但是业务站点出现异常时，健康检查则无法检测到业务站点的异常。

操作步骤

1. 在后端服务器上新建一个健康检查站点和健康检查页面，并关闭日志记录。本操作以nginx为例进行说明。



2. 登录负载均衡管理控制台。
3. 在实例管理页面中找到对应的负载均衡实例，单击实例ID。
4. 在击监听页签下，找到对应监听，单击操作列的配置。
5. 在配置监听对话框中单击下一步，再单击下一步，进入健康检查配置。

6. 在健康检查域名（可选）中输入健康检查站点的域名，在健康检查路径中输入健康检查页面的的相对路径。

7. 单击下一步至确定，完成修改。

5 如何排查ECS实例异常

在负载均衡服务中开启健康检查功能后，当后端某个ECS健康检查出现问题时，会将请求转发到其他健康检查正常的ECS上。当该ECS恢复正常运行时，负载均衡会将其自动恢复到对外或对内的服务中。

针对七层负载均衡服务，当监听获取到健康检查异常的信息时，可以从以下方面对ECS实例进行健康排查：

- 确保您能够直接通过ECS访问到您的应用服务。
- 确保后端服务器开启了相应的端口，该端口必须与您在负载均衡监听配置中配置的后端端口保持一致。
- 检查后端ECS内部是否开启了防火墙或其他的安全类防护软件，这类软件很容易将负载均衡服务的本地 IP地址屏蔽掉，导致负载均衡服务无法跟后端服务器进行通讯。
- 检查负载均衡健康检查参数设置是否正确，建议参照缺省提供的健康检查参数进行设置。
- 建议使用静态页面来进行健康检查，如果您用于健康检查的页面在后端ECS应用服务器上并不是缺省首页，需要您在健康检查配置中指定该页面的URL。健康检查指定的检测文件，建议是html形式的简单页面，只用于检查返回结果，不建议用php等动态脚本语言。
- 检查后端ECS资源是否有较高负载，降低了ECS对外提供服务的响应速度。

另外，由于七层负载均衡服务跟后端ECS之间通过内网通讯，因此需要ECS监听内网或者全网端口。您可使用以下方法进行检查：

1. 检查监听功能是否正常。

   假设负载均衡前端端口是80，ECS后端端口也是80，ECS内网IP是10.11.192.1。在服务器上运行以下命令，如果能看到10.11.192.1:80的监听信息，或者0.0.0.0:80的监听信息，说明这部分端口的监听正常。

   - Windows 服务器上运行：`netstat -ano | findstr :80`
   - Linux 服务器上运行：`netstat -anp | grep :80`

2. 检查服务器内网防火墙是否放行80端口，可以暂时关闭防火墙进行测试。输入以下命令关闭防火墙。

   - Windows：`firewall.cpl`
   - Linux：`/etc/init.d/iptables stop`

3. 检查后端端口是否正常。

   - 对于四层负载均衡，使用telnet测试有响应即为正常。本例中使用`telnet 10.11.192.1 80`来测试。
   - 对于七层负载均衡，HTTP状态码需要是200等代表正常的状态码，检验方法如下：
     - Windows：直接在ECS上访问ECS的内网IP测试是否正常，本例中为：`http://10.11.192.1`。
     - Linux：使用`curl -I`命令查看状态是否为HTTP/1.1 200 OK，本例是：`curl -I 10.11.192.1`。

6 如何排查四层监听（TCP/UDP）健康检查异常

健康检查用于探测您的后端服务器是否处于正常工作状态，当健康检查出现异常时，通常说明您的后端服务器出现了异常，但也有可能是您的健康检查配置不正确导致，本文主要介绍对四层监听（TCP/UDP）健康检查异常进行排查的详细步骤。

操作步骤

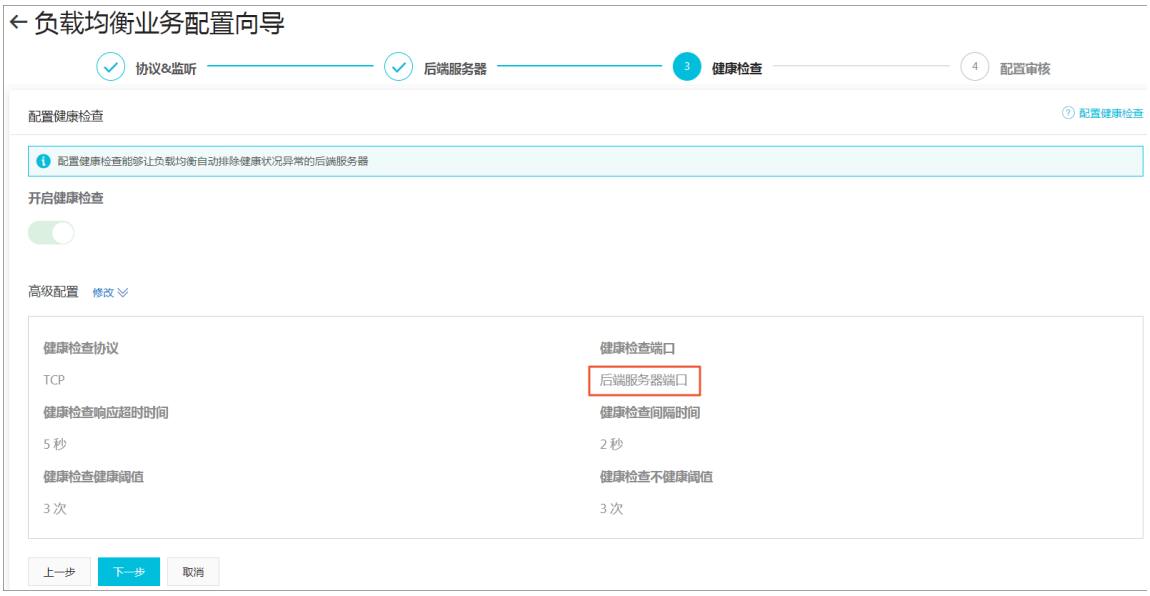
1. 确保后端服务器上没有针对100.64.0.0/10地址段进行任何形式的屏蔽，包括iptables或其他任何第三方防火墙/安全策略软件。

   负载均衡SLB通过100.64.0.0/10内部保留地址段中的IP地址与后端服务器通信，如被屏蔽则会导致健康检查异常，负载均衡无法正常工作。

2. 执行telnet命令，探测后端服务器。

a) 登录负载均衡控制台，查看健康检查配置。

其中，健康检查端口默认使用后端服务器端口，也可以手动设置端口。此处示例使用后端服务器端口，端口号为80。



b) 执行如下命令，尝试连接健康检查端口，负载均衡上配置的健康检查端口要与后端服务器上的监听的端口保持一致。

```
telnet 172.17.58.131 80
```

此处*172.17.58.131*为后端服务器的内网IP地址，*80*为健康检查端口，如保持默认健康检查端口设置，则使用后端服务器的端口，请根据实际情况配置。

- 正常情况下，会返回类似`Connected to xxx.xxx.xxx.xxx`信息，表示后端服务器上指定端口处于正常工作（监听）状态，此时健康检查是正常的，如下图所示。

- 异常示例：假设负载均衡上的监听配置保持不变，但是停止后端服务器上的80端口监听进程，执行telnet命令后，系统提示无法连接到该主机，连接被拒绝，表示80端口监听的进程不再工作，此时健康检查会出现异常，如下图所示。

3. （可选） 四层监听支持HTTP方式健康检查，如果使用HTTP方式进行健康检查，请参见七层监听（HTTP/HTTPS）健康检查异常排查进行排查。

7 如何排查七层监听（HTTP/HTTPS）健康检查异常

健康检查用于探测您的后端服务器是否处于正常工作状态，当健康检查出现异常时，通常说明您的后端服务器出现了异常，但也有可能是您的健康检查配置不正确导致，本文主要介绍对七层监听（HTTP/HTTPS）健康检查异常进行排查的详细步骤。

操作步骤

1. 确保后端服务器上没有针对100.64.0.0/10地址段进行任何形式的屏蔽，包括iptables或其他任何第三方防火墙/安全策略软件。

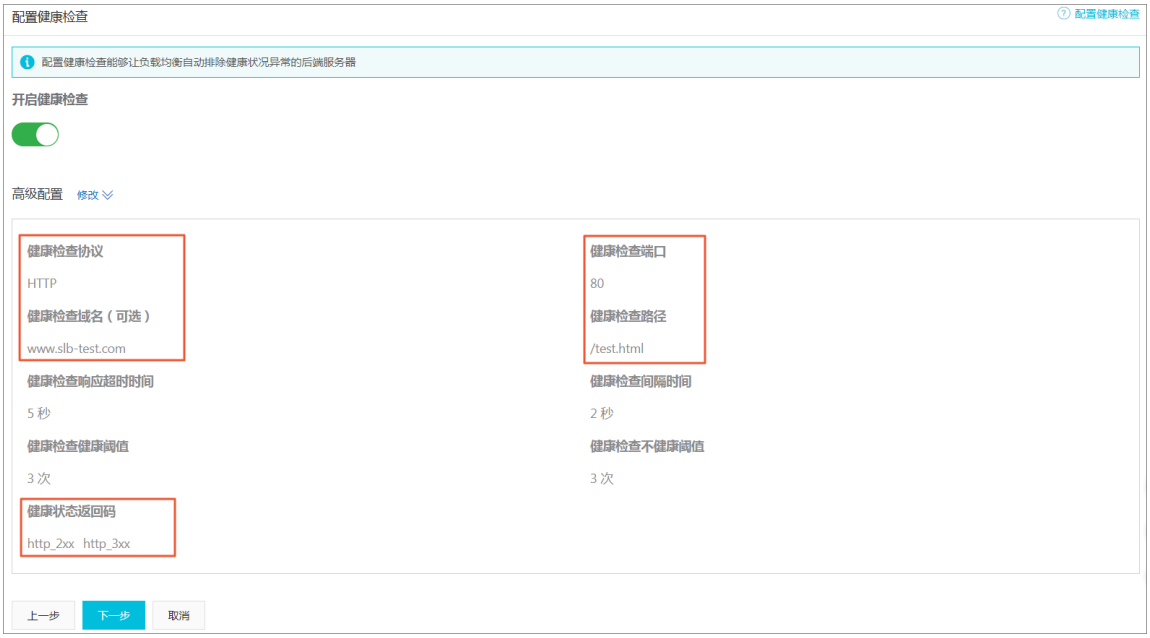
   负载均衡SLB通过100.64.0.0/10内部保留地址段中的IP地址与后端服务器通信，如被屏蔽则会导致健康检查异常，负载均衡无法正常工作。

2. 从后端服务器本地发起访问，确保后端服务器上的HTTP服务正常工作。

   a) 登录负载均衡控制台，在监听实例详情页中，查看健康检查配置。

   本次示例使用HTTP监听，出现健康检查异常的后端服务器内网IP为10.0.0.2，其他健康检查配置信息如下：

   - 健康检查端口：80
   - 健康检查域名：`www.slb-test.com`
   - 健康检查路径：`/test.html`



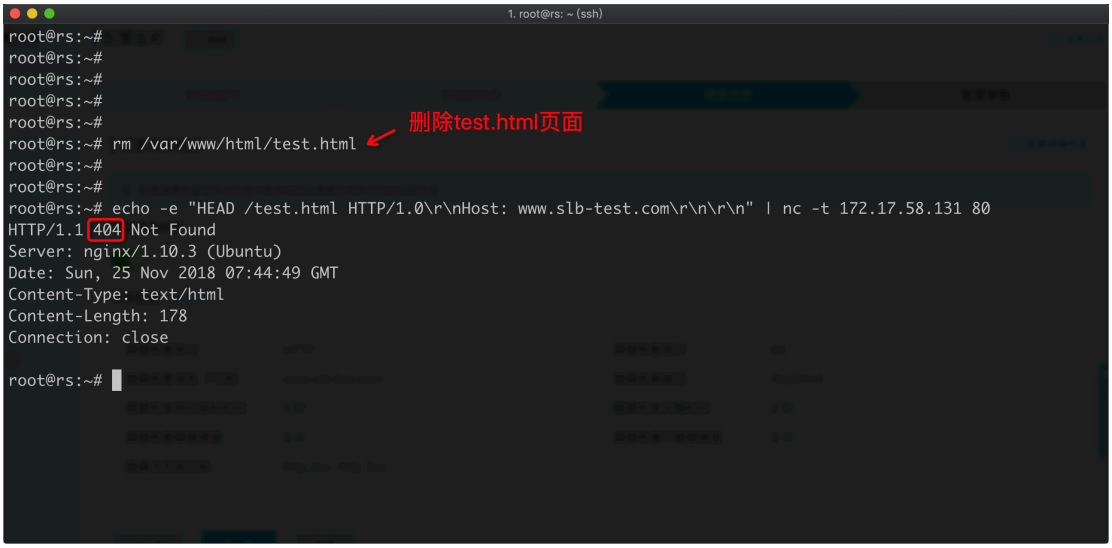
   b) 以Linux系统为例，执行nc或curl命令对后端服务器上的HTTP服务进行探测，健康检查路径、健康检查端口和健康检查域名配置必须与后端服务器上配置保持一致，否则会产生健康检查异常。

   此处使用nc命令为例，请根据实际情况配置健康检查路径、健康检查域名、健康检查内网地址和健康检查端口：

```
echo -e "HEAD /test.html HTTP/1.0\r\nHost: www.slb-test.com\r\n\r
\n" | nc -t 172.17.58.131 80
```

   - 正常情况下，返回`200`或其他`2xx/3xx`返回码，如下图所示。

- 异常示例：假设负载均衡上的监听配置保持不变，但是删除后端服务器上/test.html页面，执行nc命令后，得到404错误码，该错误码与负载均衡SLB监听中设置的2xx或者3xx状态码不符，此时会出现健康检查异常结果，如下图所示。



8 如何进行压力测试

压力测试性能概述

四层负载均衡采用开源软件LVS（Linux Virtual Server）+ Keepalived的方式实现负载均衡，七层负载均衡由Tengine实现。其中四层监听经过LVS后直接到达后端服务器，而七层监听经过LVS后，还需要再经过Tengine，最后到达后端服务器。七层比四层多了一个处理环节，因此，七层性能没有四层性能好。

如果您使用七层监听进行压力测试，发现压测性能比较低。挂了两台ECS的七层负载均衡监听性能还不如挂了一台ECS的四层负载均衡监听性能，除了七层本身的性能比四层低外，以下情况也可能会造成七层压测性能低：

- 客户端端口不足 。

  在进行压力测试时，客户端端口不足会导致建立连接失败。负载均衡会默认抹除TCP连接的timestamp属性，Linux协议栈的tw_reuse(time_wait 状态连接复用)无法生效，time_wait状态连接堆积导致客户端端口不足。

  解决方法：客户端使用长连接代替短连接。使用RST报文断开连接，即socket设置SO_LINGER属性。
- 后端服务器accept队列满 。

  后端服务器accept队列满，导致后端服务器不回复syn_ack报文，客户端超时。

  解决方法：默认net.core.somaxconn的值为128，执行`sysctl -w net.core.somaxconn=1024`命令更改net.core.somaxconn的值，并重启后端服务器上的应用。
- 后端服务器连接过多。

  由于架构设计的原因，使用七层负载均衡时，用户长连接经过Tengine后变成短连接，可能导致后端服务器连接过多，从而表现为压测性能低。
- 后端服务器依赖的应用成为瓶颈。

  请求经过负载均衡到达后端服务器后，后端服务器本身负载正常，但由于所有的后端服务器上的应用又依赖其它应用，例如数据库，当数据库成为瓶颈时，也会引起性能降低。
- 后端服务器的健康检查状态异常。

  在压测时，容易忽略后端服务器的健康检查状态，如果有后端服务器健康检查失败或者健康检查状态经常跳跃（好到坏，又从坏到好，反复变化），也会导致压测性能低。

压力测试建议

在进行压力测试时，请注意如下配置：

- 压测负载均衡转发能力建议使用短连接 。

  一般来说压测除了验证会话保持和均衡性等功能外，主要想验证负载均衡的转发能力，因此使用短连接比较合适，用于测试负载均衡和后端服务器的处理能力。使用短连接测试时，需要注意客户端端口不足的问题。
- 压测负载均衡吞吐量建议使用长连接，用于测试带宽上限或特殊业务。

  压测工具的超时时间建议设置为一个较小值，如5秒。超时时间太大的话，测试结果会体现在平均响应时间加长，不利于判断压测水位是否已到达。超时时间调小，测试结果会体现在成功率上，便于快速判断压测水位。
- 后端服务器提供一个静态网页用于压测，以避免应用逻辑带来的损耗。
- 压测时，监听配置建议如下：
  - 不开启会话保持功能，否则压力会集中在个别后端服务器。
  - 关闭健康检查功能，减少健康检查对后端服务器的访问请求。
  - 性能测试服务的5000并发规格能够提供5个及5个以上的公网IP。

压力测试工具建议

不建议您使用Apache ab作为压力测试工具。

Apache ab在大量并发场景下存在3s、6s、9s阶梯式停顿的现象。Apache ab会通过判断content length来确定请求是否成功，而负载均衡挂载多台后端服务器时，返回的content length会不一致，导致测试结果有误。

建议使用阿里云PTS。

可以设置足够高的并发，PTS会分配来自全国各地的公网IP，压力来源足够分散，并且可以在PTS中集成云监控，实时查看端到端的全部性能数据。

使用PTS简单压测示例

创建一个负载均实例，添加两台ECS实例作为后端服务器，分别创建一个TCP监听和HTTP监听，后端端口设置为80。ECS服务器的配置为CPU 1核，内存512M使用CentOS 6.3 64位的操作系统。

1. **安装Apache Web Server提供Web服务。**

```
yum install -y httpd
```

2. **初始化默认首页index.html。**

```
echo "testvm" > /var/www/html/index.html
```

3. **启动HTTP服务。**

```
service httpd start
```

4. **访问本地的80端口，确认Web服务可用。**

```
curl localhost
```

5. **在PTS中创建测试场景，开始压力测试。**

9 如何排查500/502/504错误

配置负载均衡之后，访问网站出现500 Internal Server Error、502 Bad Gateway和504 Gateway Timeout等错误，有可能由多种原因导致，例如运营商拦截、客户端异常行为导致云盾封堵、负载均衡配置错误、健康检查失败或者后端ECS Web应用访问问题。

本文档列举了此类问题的可能原因、解决方案以及排查步骤。

1. 可能原因以及解决方案
   - 源站域名没有备案或者域名没有在高防或者安全网络中配置七层转发规则
   - 客户端源IP地址被云盾拦截
   - 后端ECS安全防护软件阻挡
   - 后端ECS Linux内核参数配置错误
   - 后端ECS性能瓶颈
   - 健康检查失败导致负载均衡报502错误
   - 健康检查正常但Web应用报502错误
   - HTTP头部过大
2. 排查步骤
3. 提交工单

可能原因以及解决方案

1. 源站域名没有备案或者域名没有在高防或者安全网络中配置七层转发规则。

   解决方案：请将域名备案。如果负载均衡在高防或者安全网络中，请配置对应的域名规则。

2. 客户端源IP地址被云盾拦截。

   测试其他ISP运营商的客户端是否有相同问题，如果仅仅是某个固定运营商网络的客户端访问有问题，一般是运营商封堵导致。

   解决方案：通过提交工单反馈给阿里云售后技术支持，抓包确认是否有封堵行为。如果有，请联系运营商解决该问题。

3. 后端ECS安全防护软件阻挡。

   100.64.0.0/10（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）是负载均衡服务器IP段，主要用于健康检查和转发请求。例如安装安全软件或者开启系统内部防火墙，可以将此IP加入白名单，避免出现500或502错误。

   解决方案：配置杀毒、防火墙软件白名单，或者卸载此类软件快速测试。

4. 后端ECS Linux内核参数配置错误。

   对于后端ECS为Linux系统，改成TCP模式时需要注意关闭系统内核参数中rp_filter相关设置。

   解决方案：将系统配置文件*/etc/sysctl.conf*的以下三个配置的值设为0，然后执行`sysctl -p`。

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

5. 后端ECS性能瓶颈。

   例如CPU高，外网带宽跑满均可能导致访问异常。

   解决方案：检查后端ECS性能，解决性能瓶颈问题，如果是整体系统容量不够，可以通过扩容后端ECS 的数量消除问题。

6. 健康检查失败导致负载均衡出现502错误 。

   健康检查失败，请参见健康检查异常排查进行排查。

   此外，未开启负载均衡的健康检查，同时服务器中Web服务无法正常处理HTTP请求，比如Web服务未运行，也会出现502错误。

7. 健康检查正常但Web应用报502错误 。

   502 Bad Gateway错误提示表明负载均衡可以将来自客户端的请求转发到后端服务器中，但是服务器中Web应用处理异常抛出该提示，所以排错的方向是针对服务器中Web应用的配置以及运行情况进行分析。例如Web应用处理HTTP请求的时间超过了负载均衡的timeout时间。

   在七层HTTP模式下，后端对PHP请求的处理时间超过proxy_read_timeout设置的60秒，此时会出现负载均衡抛出的504 Gateway Time-out。对于四层监听，超时时间为900秒。

   解决方案：确保Web服务以及依赖正常运行，检查PHP请求处理情况，优化后端PHP请求处理。下面以Nginx+php-fpm为例进行分析说明：

   a. 处理PHP请求的进程数达到上限。

      当前服务器中PHP请求总数已经达到了php-fpm中max_children设置的上限，如果后续有新的PHP请求到达服务器中，这种情况下通常502与504的错误码会随机出现：

      - 如果已有的请求被处理完成，新请求被继续处理，一切正常。
      - 如果已有的PHP请求处理较慢，新的PHP一直处于等待状态，直至超过Nginx的fastcgi_read_timeout的值，就会出现504 Gateway timeout的错误。
      - 如果已有的PHP请求处理较慢，新的PHP处于等待状态，超过了Nginx的request_terminate_timeout的值，就会出现502 Bad Gateway的错误。

   b. PHP脚本执行时间处理超时，即如果php-fpm处理PHP脚本的时长超过了Nginx中request_terminate_timeout设置的值，就会出现502 Bad Gateway的错误，同时在Nginx日志中可以查看到如下错误记录：

      ```
      [error] 1760#0: *251777 recv() failed (104: Connection reset by
      peer) while reading response header from upstream, client: xxx.xxx
      .xxx.xxx, server: localhost, request: “GET /timeoutmore.php HTTP/1
      .1”, upstream: “fastcgi://127.0.0.1:9000”
      ```

   c. 健康检查针对的是静态页面，实际处理动态请求的进程异常，比如php-fpm未启动运行。

8. HTTP头部过大。

   Head头信息过大可能导致负载均衡无法正确处理相关数据，进而引发502错误。

   解决方案：减少通过Head头传递的数据量或者换成TCP监听。

9. 业务访问逻辑问题。

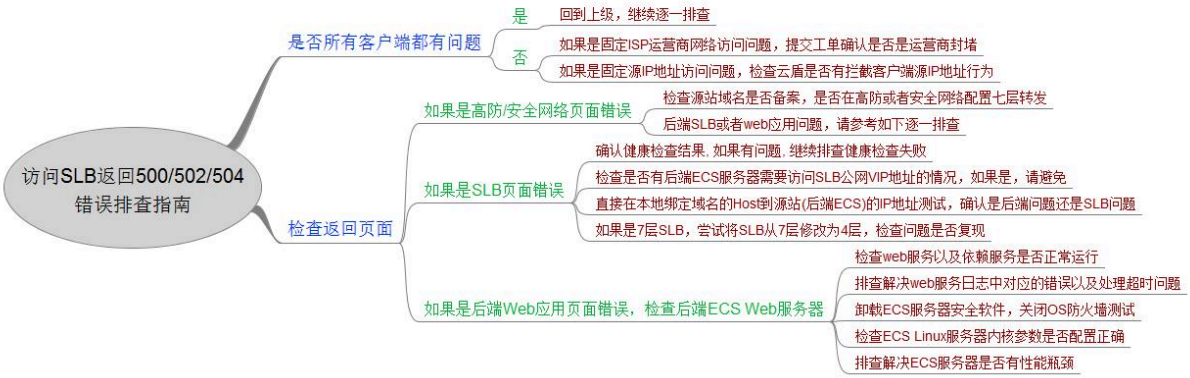
   确保不存在负载均衡后端ECS实例在服务器内部通过负载均衡公网IP地址访问SLB的情况。该情况下，后端业务服务器通过负载均衡地址访问自身所监控的端口后，根据负载均衡调度策略的不

同，可能会将相应的请求调度到自身服务器上。导致出现自己访问自己的情况，造成死循环，进而导致相应的请求出现500或502错误。

解决方案：确保负载均衡场景应用正确，避免后端ECS服务器需要访问负载均衡公网IP地址的情况。

排查步骤

- 检查500/502/504错误截图，判断是负载均衡问题，高防/安全网络配置问题，还是后端ECS配置问题。
- 如果有高防/安全网络，请确认高防/安全网络的七层转发配置正确。
- 请确认是所有客户端都有问题，还仅仅是部分客户端有问题。如果仅仅是部分客户端问题，排查该客户端是否被云盾阻挡，或者负载均衡域名或者IP是否被ISP运营商拦截。
- 检查负载均衡状态，是否有后端ECS健康检查失败的情况，如果有健康检查失败，解决健康检查失败问题。
- 在客户端用hosts文件将负载均衡的服务地址绑定到后端服务器的IP地址上，确认是否是后端问题。如果5XX错误间断发生，很可能是后端某一台ECS服务器的配置问题。
- 尝试将七层负载均衡切换为四层负载均衡，查看问题是否会复现。
- 检查后端ECS服务器是否存在CPU、内存、磁盘或网络等性能瓶颈。
- 如果确认是后端服务器问题，请检查后端ECS Web服务器日志是否有相关错误，Web服务是否正常运行，确认Web访问逻辑是否有问题，卸载服务器上杀毒软件重启测试。
- 检查后端ECS Linux操作系统的TCP内核参数是否配置正确。



提交工单

请根据上述排查步骤中的指导逐条排查，详细记录排查测试结果。提交工单时，请您提供上述信息以便售后支持尽快协助您解决问题。

如果问题还未解决，请联系售后技术支持。

10 负载均衡实例计费FAQ

包含以下常见问题：

1. 负载均衡实例如何计费？

参考按量计费

2. 负载均衡是否对入流量计费？

负载均衡目前只对出流量计费，入流量不计费。关于负载均衡网络流量路径的信息，参考网络流量说明。

3. 健康检查产生的流量是否会被计费？

不会。负载均衡健康检查产生的流量不会计入购买的实例流量费用中。

4. ECS加入负载均衡后端服务器资源池是否影响其计费？

不会。不论您配置在负载均衡实例后端的ECS采用何种计费方式，都不会因为其与负载均衡的关联而产生计费规则上的变化。因为负载均衡和ECS是根据您的使用情况分别计量计费并结算的。

5. 攻击流量是否会被计费？

负载均衡目前和云盾结合提供防护功能。从攻击开始达到清洗或黑洞阈值，到云盾开始清洗或黑洞，这期间有秒级的延时。因此这期间对攻击包的响应会产生一定的费用。这样的攻击本身也消耗了负载均衡的带宽资源。

6. 负载均衡实例的所有后端ECS都停止，或者没有挂载ECS，是否会被计费？

负载均衡实例还是会计费，具体说明如下：

- 按流量计费

  按使用流量计费只有在实例停止、被释放、或无任何访问的情况下才不会产生流量费用。

  负载均衡是放置于ECS之前的负载均衡服务设备，通过服务地址的方式提供服务。负载均衡实例的所有后端ECS停止，但负载均衡实例本身服务并未停止，当有请求发生时，入流量还是会到负载均衡的服务地址，负载均衡健康检查发现后端没有可用的ECS，会进行响应。

  对于四层负载均衡服务，响应的仅是三次握手的包。对于七层负载均衡服务，由于负载均衡七层服务是通过Tengine提供的，因此响应的是一个Tengine的503错误页。如果不停地有访问进来，负载均衡不停地响应，这些响应流量会被计费。

  对于没有挂载ECS的负载均衡实例也是这样的情况。因此，为避免这种情况下被计费，您不使用负载均衡实例的时候，可以停止这个负载均衡实例。

7. 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。

8. 性能保障型实例的流量费和实例费与共享型一样吗?

  一样。

9. 收取规格费以后，共享型实例也会额外收取费用吗?

  不会。

  原有的共享型实例（如果您不将其变更成性能保障型）将继续保持为性能共享型实例，不收取规格费。

  但如果您将其变更成性能保障型，将从4月1日起收取规格费。

10. 负载均衡监控数据与实际账单数据为什么不同?

- 负载均衡控制台监控指标中展示的实例流量数据是按照一分钟一个采集粒度，由负载均衡系统采集后上报云监控系统，再由云监控系统计算出每15分钟所有采集点的平均值；而负载均衡实例的账单计量数据是按照同样粒度的采集频度采集数据，在一个消费帐期后，以1个小时的累加值向账单计量系统上报并结算费用的。

  账单数据是每一分钟的累加值，而监控数据是每15分钟平均值的累加值。所以，这两组数据在实际计算生成方式上是有区别的，两种数据不具备对比性。
- 负载均衡系统从采集数据到向云监控上报数据，然后云监控对数据进行平均值计算，最后展示在监控控制台的整个过程中，不可避免的存在一定的延迟。虽然这个延迟很小，我们也会尽量保证

数据的实时性，但这种延迟也会导致其数据本身与账单计量数据存在一定程度的差异。用作计费的账单计量数据是可以容忍最多三小时延迟的，比如在01:00-02:00产生的账单计量数据，正常情况下会在03:00之前由负载均衡上报账单计量系统并进行计费，但是系统允许该上报时间最晚于05:00之前完成并计费。所以，从数据对实时性的要求不同来看，这两组数据也是不具备可比性的。

- 从监控和账单计量的产品定义上也是有区别的。监控的目的是为了通过一种手段来观察被监控实例的运行状态是否会出现异常，从而针对性地采取一定措施来解决由于异常导致的问题。账单计量的目的是根据实例的实际资源消耗情况来进行计费。所以，从账单核算的出发点来看应该是以账单计量系统生产的数据来作为计费的判断依据而不应该以监控数据作为计费的判断依据。

11. 为什么HTTPS协议实际产生的流量会比账单流量多一些?

HTTPS协议会使用一些流量用于协议握手，因此其实际产生的流量会多于账单流量。

11 负载均衡实例FAQ

包含以下常见问题：

1. 共享型实例可以变更成保障型实例么？

可以。

变更以后将从4月1起收取规格费，且不能再变更回性能共享型，即性能保障型实例无法变更为共享型。

2. 共享型实例性能规格如何？

性能共享型实例不提供性能保障，没有可提供的规格。

3. 性能保障型实例有很多规格，怎么选？

- 后付费实例（按量付费），可以选择最大规格，因为后付费实例的规格费按照实际使用量收取，闲时免费。

4. 是否可以调整性能保障型实例的规格？

可以。

- 按量付费的性能保障型实例的规格可以升配也可以降配，详情参见后付费实例变配。

> 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。
> - 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，再进行变配。

5. 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。

12 性能保障型实例FAQ

负载均衡性能保障型实例提供了可保障的性能指标，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。

包含以下常见问题：

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

如果需要更大规格，请联系您的客户经理申请。

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。

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。

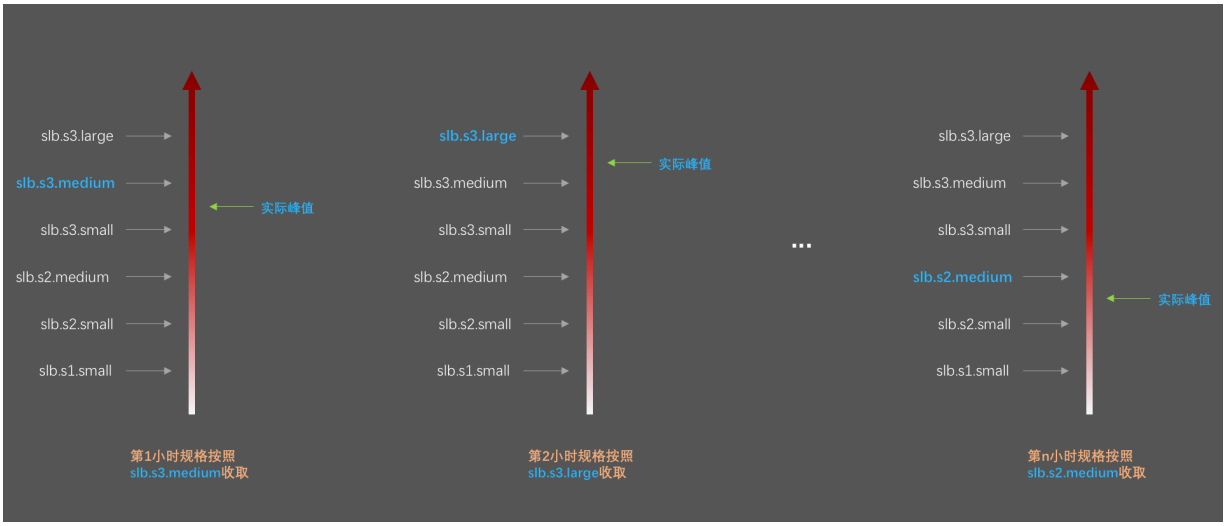
例如，您选择了超强型I (slb.s3.large)规格（最大连接数1,000,000；CPS 500,000；QPS 50,000）。您的实例在某个小时内各项指标产生的实际峰值如下：

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正常收取。更多详细信息，参考按量计费。

地域	规格	最大连接数	每秒新建连接数 (CPS)	每秒查询数(QPS)	规格费（美元/小时）
华东1（杭州） 华北3（张家口） 华北5（呼和浩特） 华北1（青岛） 华北2（北京） 华东2（上海） 华南1（深圳）	规格1：简约型I (slb.s1.small)	5000	3000	1000	免费
	规格2：标准型I (slb.s2.small)	50,000	5,000	5,000	0.05
	规格3：标准型II (slb.s2.medium)	100,000	10,000	10,000	0.10
	规格4：高阶型I (slb.s3.small)	200,000	20,000	20,000	0.20
	规格5：高阶型II (slb.s3.medium)	500,000	50,000	30,000	0.31
	规格6：超强型I (slb.s3.large)	1,000,000	100,000	50,000	0.51
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） 香港	规格1：简约型I (slb.s1.small)	5,000	3,000	1,000	免费
	规格2：标准型I (slb.s2.small)	50,000	5,000	5,000	0.06
	规格3：标准型II (slb.s2.medium)	100,000	10,000	10,000	0.12
	规格4：高阶型I (slb.s3.small)	200,000	20,000	20,000	0.24
	规格5：高阶型II (slb.s3.medium)	500,000	50,000	30,000	0.37
	规格6：超强型I (slb.s3.large)	1,000,000	100,000	50,000	0.61

4. 如何选择性能保障型实例?

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)。

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。

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。
- 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。

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。

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。

13 负载均衡服务FAQ

包含以下问题：

1. 负载均衡是否支持端口跳转?

支持。

详情参见配置监听转发（redirect）。

2. 禁用公网网卡是否影响负载均衡服务?

如果ECS有公网IP，禁用公网网卡就会影响负载均衡服务。

因为在有公网网卡的情况下，默认路由会走公网，如禁用就无法回包从而影响负载均衡服务。建议不要禁止，如一定要这么做，需要修改默认路由为私网才会不影响服务。但需要考虑业务是否对公网有依赖，如通过公网访问RDS等。

3. 为什么每个连接达不到带宽峰值?

因为负载均衡系统通过集群部署的方式为负载均衡实例提供服务，所有外部的访问请求都将平均分散到这些负载均衡系统服务器上进行转发。所以，设定的带宽峰值将被平均设定在多台系统服务器上。

单个连接下载的流量上限计算方法为：单个连接下载峰值=设置的负载均衡总带宽/(N-1)。N为流量转发分组个数，当前值为4。比如您在控制台上设置的是10Mb带宽上限，那么单个客户端可下载的最大流量为10/(4-1)=3.33Mb。

基于负载均衡的实现原理，建议在配置单个监听的带宽峰值时根据您实际的业务情况并结合其实现方式来设定一个较为合理的值，从而确保您业务的正常对外服务不会受到影响和限制。

4. 负载均衡各监听连接超时时间是多少?

- TCP监听： 900秒
- UDP监听： 90秒
- HTTP监听： 60秒
- HTTPS监听： 60秒

5. 为什么负载均衡服务地址会连接访问超时?

从服务端分析，以下情况会导致服务地址链接访问超时:

- 服务地址被安全防护

  如流量黑洞和清洗，WAF防护（WAF的特点是建连后向客户端和服务器集群双向发送RST报文）。

- 客户端端口不足

  尤其容易发生在压测的时候，客户端端口不足会导致建立连接失败，负载均衡默认会抹除TCP连接的timestamp属性，Linux协议栈的tw_reuse(time_wait状态连接复用)无法生效，time_wait状态连接堆积导致客户端端口不足。

  解决方法：客户端使用长连接代替短连接。使用RST报文断开连接（socket设置SO_LINGER属性），而不是发FIN包这种方式断开。

- 后端服务器accept队列满

  后端服务器accept队列满，导致后端服务器不回复syn_ack报文，客户端超时。

  解决方法：默认的`net.core.somaxconn`的值为128，执行`sysctl -w net.core.somaxconn=1024`更改它的值，并重启后端服务器上的应用。

- 从四层负载均衡后端服务器访问该四层负载均衡的服务地址

  四层负载均衡，在该负载均衡的后端服务器上去访问该负载均衡的服务地址会导致连接失败，常见的场景是后端应用使用URL拼接的方式跳转访问。

- 对连接超时的RST处理不当

  负载均衡上建立TCP连接后，如果900s未活动，则会向客户端和服务器双向发送RST断开连接，有的应用对RST异常处理不当，可能会对已关闭的连接再次发送数据导致应用超时。

6. 为什么有时候会话保持失败?

- 查看是否在监听配置中已经开启了会话保持功能。

- HTTP/HTTPS监听在后端服务器返回 4xx 响应码的报文中无法插入会话保持所需cookie。

  解决方案：改用TCP监听，因为TCP监听是以源客户端的IP来做会话保持的，另外后端ECS上也可以插入cookie，并增加cookie的判断来多重保障。
- 302重定向会改变会话保持中的SERVERID字串。

  负载均衡植入cookie时，如果后端ECS中有回复302重定向的报文，将改变会话保持中的SERVERID字串，导致会话保持失效。

  排查方法：在浏览器端捕抓请求与响应的回复，或用抓包软件抓包后分析是否存在302的响应报文，对比前后报文的cookie中的SERVERID字串是否不同了。

  解决方案：改用TCP监听，因为TCP监听是以源客户端的IP来做会话保持的，另外后端ECS上也可以插入cookie，并增加cookie的判断来多重保障。
- 会话保持时间设置过小，会话保持时间过小也会导致会话保持失败。

7. 如何查看会话保持字串?

可以在浏览器中用F12查看回应报文中是否含有SERVERID字串或用户指定的关键字，或者运行 `curl www.xxx.com -c /tmp/cookie123` 保存一下cookie，再用 `curl www.xxx.com -b /tmp/cookie123` 访问。

8. 如何使用Linux curl测试负载均衡会话保持?

1. 创建测试页面。

   在负载均衡所有后端ECS中创建测试页面，如下图所示页面中能显示本机内网IP。内网IP用于判断相应请求被指派到的物理服务器。通过观察该IP的一致性，来判断负载均衡会话保持的有效性。

2. Linux下curl测试。

   假设负载均衡服务IP地址是 1.1.1.1，创建的测试页面URL为： http://1.1.1.1/check.jsp

   a. 登录用来测试的Linux服务器。

   b. 执行以下命令负载均衡服务器cookie值。

   ```
   curl -c test.cookie http://1.1.1.1/check.jsp
   ```

   说明:

阿里云负载均衡会话保持默认模式是植入cookie，而curl测试默认是不会保存和发送cookie的。所以必须先保存相应的cookie，用于cookie测试。否则，curl测试结果是随机的，会误认为负载均衡会话保持无效。

c. 执行以下命令持续测试。

```
for ((a=1;a<=30;a++)); 2="" do="" curl="" -b="" 1.cookie=""
            check.jsp="">/dev/null | grep '10.170.*';sleep 1; done
、
```

说明:

a<=30是重复测试次数，可以按需修改；grep ‘10.170.*’ 是筛选显示的IP信息，根据后端ECS内网IP情况进行相应修改；

d. 观察上述测试返回的IP，如果是同一台ECS内网IP，则证明负载均衡会话保持有效；反之则证明负载均衡会话保持有问题。

9. 一个请求通过负载均衡到达后端服务器，如果客户端在未收到后端服务器的回复前主动断开和负载均衡的连接，负载均衡会同时断开和后端服务器的连接么?

负载均衡在读写过程中不会断开与后端服务器的连接。

14 后端服务器FAQ

包含以下常见问题：

1. 负载均衡运行中是否可调整ECS数量?

可以。

您可以在任意时刻增加或减少负载均衡的后端ECS实例数量并且支持不同ECS实例之间的切换。但是为了保证您对外服务的稳定，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保证负载均衡后端至少有一台正常运行的ECS实例。

2. 后端ECS实例的操作系统是否可以不同?

可以。

负载均衡本身不会限制后端ECS实例使用哪种操作系统，只要确保后端ECS实例中的应用服务部署相同且数据一致即可。但建议使用相同的操作系统，以便您日后的管理维护。

3. 可以使用不同地域的ECS实例作为后端服务器么?

不可以。

负载均衡不支持跨地域部署。负载均衡实例的地域和后端ECS实例的地域必须相同。

4. 为什么有100开头的IP在频繁访问ECS实例?

负载均衡系统除了会通过系统服务器的内网IP将来自外部的访问请求转到后端ECS实例之外，还会对ECS实例进行健康检查和可用性监控，这些访问的来源都是由负载均衡系统发起的。

负载均衡系统的地址段为100.64.0.0/10（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），所以会有很多100开头的IP地址访问ECS实例。

为了确保您对外服务的可用性，确保对上述地址的访问配置了放行规则。

5. ECS实例上没有配置压缩，为什么从负载均衡返回的响应却被压缩了？

可能是客户端浏览器端支持压缩。您可以在控制台上创建监听时关闭Gzip压缩功能，或改用TCP监听。

6. ECS实例使用了HTTP1.0是否支持chunked transfer传输编码？

支持。

7. 为什么负载均衡后端ECS实例频繁收到UA为KeepAliveClient的请求？

问题现象：

负载均衡后端的ECS实例即使在没有用户访问时也会频繁收到GET请求，来源的IP是阿里云的内网IP，User-Agent显示为KeepAliveClient。

问题原因：

监听协议选择的是TCP，而健康检查选择了HTTP协议。TCP监听下使用HTTP协议进行健康检查时，默认使用GET方法请求。

解决方案：

建议您将监听协议和健康检查协议统一设置为同一个协议。

15 健康检查FAQ

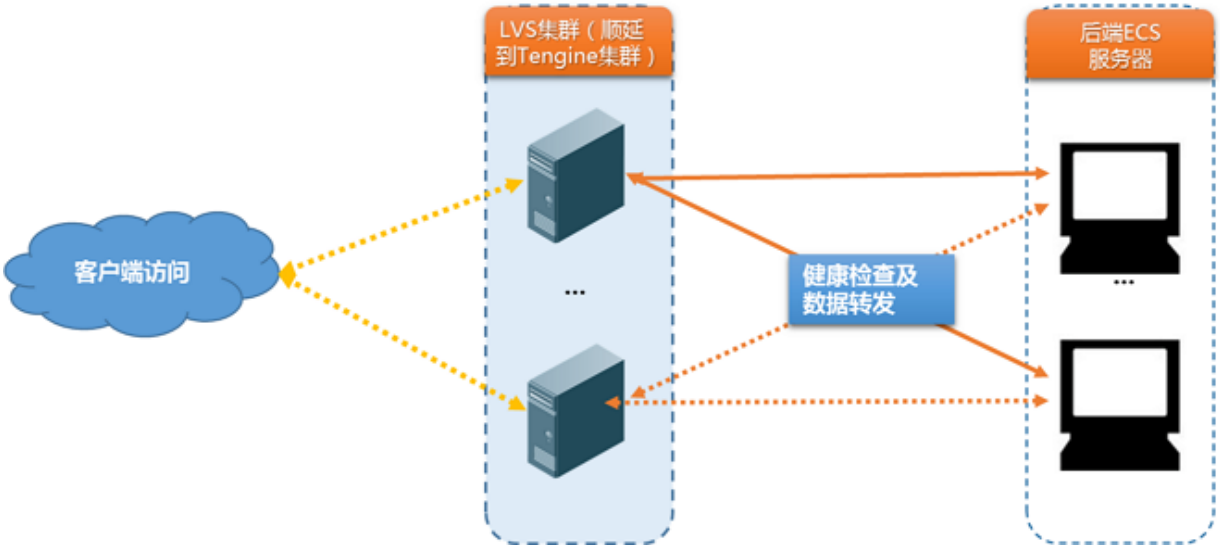
包含以下常见问题：

1. 健康检查的原理是什么？

负载均衡通过健康检查来探测后端ECS实例的可用性。开启健康检查功能后，当后端某个ECS实例健康检查出现异常时，来自客户端的新请求将不会再被转发到该ECS实例，直到该ECS实例恢复正常。

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）。

更多详细信息，参考负载均衡健康检查原理。



2. 推荐的健康检查配置是什么?

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。更多详细信息参见健康检查配置。

以下是TCP/HTTP/HTTPS监听建议使用的健康检查配置。

配置	推荐值
响应超时时间	5秒
健康检查间隔	2秒
不健康阈值	3

以下是UDP监听建议使用的健康检查配置。

配置	推荐值
响应超时时间	10秒
健康检查间隔	5秒
不健康阈值	3
健康阈值	3

说明:

此配置有利于用户服务及应用状态的尽快收敛。如果您有更高要求，可以适当地降低响应超时时间值，但必须先保证服务在正常状态下的处理时间小于这个值。

3. 是否可以关闭健康检查?

您只可能关闭HTTP/HTTPS监听的健康检查，不能关闭TCP/UDP监听的健康检查。具体操作，参考关闭健康检查。

说明:

如果关闭健康检查，当后端某个ECS实例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实例上，造成部分业务不可访问。建议您不要关闭健康检查。

4. TCP监听如何选择健康检查方式?

TCP监听支持HTTP和TCP两种健康检查方式:

- TCP协议健康检查通过发送SYN握手报文，检测服务器端口是否存活。
- HTTP协议健康检查通过发送HEAD/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。

TCP健康检查方式对服务器的性能资源消耗相对要少一些。如果您对后端服务器的负载高度敏感，则选择TCP健康检查；如果负载不是很高，则选择HTTP健康检查。

5. ECS实例权重设置为零对健康检查有什么影响?

该状态下，负载均衡不会再将流量转发给该ECS实例，且四层监听的后端服务器健康检查会显示异常（七层监听不会显示异常）。

将负载均衡后端ECS实例的权重置为零，相当于将该ECS实例移出负载均衡。一般是在ECS实例进行重启、配置调整等主动运维时将其权重设置为零。

6. HTTP监听向后端ECS实例执行健康检查使用的方法是什么?

HEAD方法。

如果后端ECS实例的服务关闭HEAD方法，会导致健康检查失败。建议在ECS实例上用HEAD方法访问自己IP地址进行测试:

```
curl -v -0 -I -H "Host:" -X HEAD http://IP:port
```

7. HTTP监听向后端ECS实例执行健康检查的IP地址是什么?

负载均衡健康检查使用的地址段是100.64.0.0/10（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）。如果后端ECS实例启用了iptables等访问控制，需要在内网网卡对100.64.0.0/10（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）做访问放行。

8. 为什么健康检查监控频率与web日志记录不一致?

负载均衡健康检查服务也是集群方式的，这样可以避免单点故障。负载均衡的代理分布到很多节点上，因此看到的健康检查日志访问频率和控制台设置的频率不一致，这是正常现象。

9. 健康检查是否会消耗系统资源?

HTTP模式的健康检查对后端ECS实例的资源消耗不大。

10. 负载均衡因后端数据库故障导致健康检查失败，如何处理?

问题现象：

ECS实例内配置了两个网站，www.test.com是静态网站，app.test.com是动态网站，都配置了负载均衡。后端数据库服务异常，导致访问www.test.com静态网站出现502错误。

问题原因：

负载均衡健康检查配置的检查域名是app.test.com，RDS或者自建数据库故障导致app.test.com访问异常，所以健康检查失败。

解决方案：

将负载均衡健康检查域名配置为www.test.com即可。

11. 负载均衡服务TCP端口健康检查成功，为什么在后端业务日志中出现网络连接异常信息?

问题现象：

负载均衡后端配置TCP服务端口后，后端业务日志中频繁出现类似如下网络连接异常错误信息。经进抓包分析，发现相关请求来自负载均衡服务器，同时负载均衡主动向服务器发送了RST数据包。

```
java.io.IOException: Connection reset by peer
        at sun.nio.ch.FileDispatcherImpl.read0(Native Method)
        at sun.nio.ch.SocketDispatcher.read(SocketDispatcher.java:39)
        at sun.nio.ch.IOUtil.readIntoNativeBuffer(IOUtil.java:223)
        at sun.nio.ch.IOUtil.read(IOUtil.java:192)
        at sun.nio.ch.SocketChannelImpl.read(SocketChannelImpl.java:379)
        at io.netty.buffer.UnpooledUnsafeDirectByteBuf.setBytes(UnpooledUnsafeDirectByteBuf.java:446)
        at io.netty.buffer.AbstractByteBuf.writeBytes(AbstractByteBuf.java:871)
        at io.netty.channel.socket.nio.NioSocketChannel.doReadBytes(NioSocketChannel.java:225)
        at io.netty.channel.nio.AbstractNioByteChannel$NioByteUnsafe.read(AbstractNioByteChannel.java:115)
        at io.netty.channel.nio.NioEventLoop.processSelectedKey(NioEventLoop.java:507)
        at io.netty.channel.nio.NioEventLoop.processSelectedKeysOptimized(NioEventLoop.java:464)
        at io.netty.channel.nio.NioEventLoop.processSelectedKeys(NioEventLoop.java:378)
        at io.netty.channel.nio.NioEventLoop.run(NioEventLoop.java:350)
        at io.netty.util.concurrent.SingleThreadEventExecutor$2.run(SingleThreadEventExecutor.java:116)
        at java.lang.Thread.run(Thread.java:745)
```

问题原因：

该问题和负载均衡的健康检查机制有关。

由于TCP对上层业务状态无感知，同时，为了降低负载均衡健康检查成本和对后端业务的冲击，当前负载均衡针对TCP协议服务端口的健康检查只会做简单的TCP三次握手，而后直接发送RST包断开TCP连接。数据交互流程如下：

1. 负载均衡服务器向后端负载均衡服务端口发送SYN请求包；
2. 后端服务器收到请求后，如果端口状态正常，则按照正常的TCP机制返回相应的SYN+ACK应答包；
3. 负载均衡服务器成功收到后端服务端口应答后，则认为端口监听是正常的，判定健康检查成功；
4. 负载均衡服务器向相应TCP服务端口直接发送RST包主动关闭连接，结束本次健康检查操作，且没有继续发送业务数据。

如上所述，由于健康检查成功后，负载均衡服务器直接发送TCP RST包中断了连接，并没有做进一步的业务数据交互，导致上层业务（比如Java连接池等）认为相应的连接是异常的，所以会出现`Connection reset by peer`等错误信息。

解决方案：

- 更换TCP协议为HTTP协议。
- 在业务层面，对来自SLB服务器IP地址段的相关请求做日志过滤，忽略相关错误信息。

12. 为什么业务本身没有异常但是健康检查显示异常？

问题现象：

负载均衡HTTP方式的健康检查始终失败，但测试`curl -I`得到的状态码是正常的。

```
echo -e 'HEAD /test.html HTTP/1.0\r\n\r\n' | nc -t 192.168.0.1 80
```

问题原因：

如果返回的状态与控制台配置的正常状态码不一致，则判定健康检查异常。如果您配置的正常状态码为http_2xx，则所有返回的非HTTP 2xx状态码均被认为是健康检查失败。

Tengine/Nginx配置会发现curl没有问题，但是echo测试会匹配到默认站点，导致测试文件*test.html*返回404错误，如下图所示。

解决方案：

- 修改主配置文件，将默认站点注释掉。
- 在健康检查配置中添加检查域名。

```
[root@iZ28s03z94sZ home]# echo -e "HEAD /test.html HTTP/1.0\r\n\r\n" |nc -t 10.161.93.136 80
HTTP/1.1 404 Not Found
Server: Tengine/2.1.0
Date: Mon, 16 Feb 2015 07:29:32 GMT
Content-Type: text/html
Content-Length: 585
Connection: close

[root@iZ28s03z94sZ home]# curl -I http://10.161.93.136/test.html
HTTP/1.1 200 OK
Server: Tengine/2.1.0
Date: Mon, 16 Feb 2015 07:29:41 GMT
Content-Type: text/html
Content-Length: 5
Last-Modified: Mon, 16 Feb 2015 07:27:00 GMT
Connection: keep-alive
ETag: "54e19bc4-5"
Accept-Ranges: bytes
```

16 七层监听（HTTPS/HTTP）FAQ

包含以下常见问题：

1. 为什么请求经过七层负载均衡转发后，后端服务器的响应头中的某些参数会被删除?

为了实现会话保持，负载均衡会修改后端服务器响应头中的Date、Server、X-Pad和X-Accel-Redirect等参数值。

解决方案：

- 在自定义的报文头部中加入一个前缀，如xl-server或xl-date，以避开负载均衡的处理。
- 将七层HTTP监听改为四层TCP监听。

2. 为什么在HTTP请求的头部增加了Transfer-Encoding: chunked字段?

将域名解析到七层负载均衡的服务地址后，从本地主机访问域名时发现在HTTP请求的头部增加了一个Transfer-Encoding: chunked字段，但是从本地主机直接访问后端服务器时是没有这个字段的。

由于七层负载均衡基于Tengine反向代理实现。Transfer-Encoding字段表示Web服务器如何对响应消息体编码，比如Transfer-Encoding: chunked表示Web服务器对响应消息体做了分块传输。

说明:

在四层负载均衡服务中，负载均衡仅转发流量，不存在该字段。

3. 为什么HTTP监听访问正常但HTTPS监听打开网址不加载样式?

现象：

分别创建HTTP和HTTPS监听，两个监听使用同样的后端服务器。以HTTP方式访问监听端口对应的网站时，网站正常显示，但使用HTTPS监听访问时，网站排版显示错乱。

原因：

负载均衡默认是不会屏蔽JS文件加载传输的，可能原因：

- 证书和浏览器安全级别不兼容导致。
- 证书是非正规第三方证书需要联系证书发布者检查证书问题。

解决方案：

1. 打开网站时，按照浏览器提示加载脚本。
2. 在客户端中添加对应证书。

4. HTTPS监听使用什么端口?

HTTPS监听对端口无特殊要求，建议您使用443端口。

5. 负载均衡支持哪些类型的证书?

支持上传PEM格式的服务器证书和CA证书。

服务器证书需要上传证书内容和私钥；CA证书只需要上传证书内容。

6. 负载均衡是否支持keytool创建的证书?

支持。

但在上传证书前，您需要将证书转换为PEM格式，详情参见转换证书格式。

7. 可以使用PKCS#12（PFX）格式的证书么?

可以。

但在上传证书前，您需要将证书转换为PEM格式，详情参见转换证书格式。

8. 添加证书时，为什么会出现KeyEncryption的错误?

该错误由于私钥内容有误导致。关于私钥格式说明，参见证书要求。

9. 负载均衡HTTPS支持哪些SSL协议版本?

TLSv1、TLSv1.1以及TLSv1.2。

10. HTTPS session ticket的保持时间是多久?

HTTPS session ticket保持时间为300秒。

11. 可以上传包含DH PARAMETERS字段的证书吗?

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，即不支持将PEM证书文件中含BEGIN DH PARAMETERS字段的证书上传。

12. HTTPS监听是否支持SNI?

SNI（Server Name Indication）是为了解决一个服务器使用多个域名和证书的SSL/TLS扩展，负载均衡HTTPS监听支持SNI功能，具体请参见配置教程。

13. HTTP/HTTPS监听访问后端服务器的HTTP协议版本是什么?

- 客户端请求的协议为HTTP/1.1或者HTTP2/0版本时，七层监听访问后端服务器的HTTP协议版本是HTTP/1.1
- 客户端请求的协议为除HTTP/1.1和HTTP2/0以外其他版本时，七层监听访问后端服务器的HTTP协议版本是HTTP/1.0

14 后端服务器能否获取客户端访问HTTP/HTTPS监听的协议版本?

可以。

15. HTTP/HTTPS连接的超时时间是如何规定的?

- HTTP长连接的请求数量限定是最多连续发送100个请求，超过限定将关闭这条连接。
- HTTP长连接两个HTTP/HTTPS请求之间的超时时间为15秒（存在误差1-2秒），超过后会关闭TCP连接，如果用户有长连接使用需求请尽量保持在13秒之内发送一个心跳请求。
- 负载均衡与后端一台ECS实例TCP三次握手完成过程的超时时间为5秒，超时后选择下一台ECS实例；查询访问日志的upstream响应时间可以定位。
- 负载均衡等待一台ECS实例回复请求的响应时间是60秒，超过后一般会返回504响应码或408响应码给客户端；查询访问日志的upstream响应时间可以定位。
- HTTPS session重用超时间为300秒，超过后同一客户端需要重新进行完整的SSL握手过程。

17 WS/WSS协议支持FAQ

什么是WS/WSS?

WebSocket (WS)是HTML5一种新的协议。它实现了浏览器与服务器全双工通信，能更好地节省服务器资源和带宽并达到实时通讯。WebSocket建立在TCP之上，同HTTP一样通过TCP来传输数据，但是它和HTTP最大不同是：

WebSocket是一种双向通信协议，在建立连接后，WebSocket服务器和Browser/Client Agent都能主动的向对方发送或接收数据，就像Socket一样；WebSocket需要类似TCP的客户端和服务器端通过握手连接，连接成功后才能相互通信。

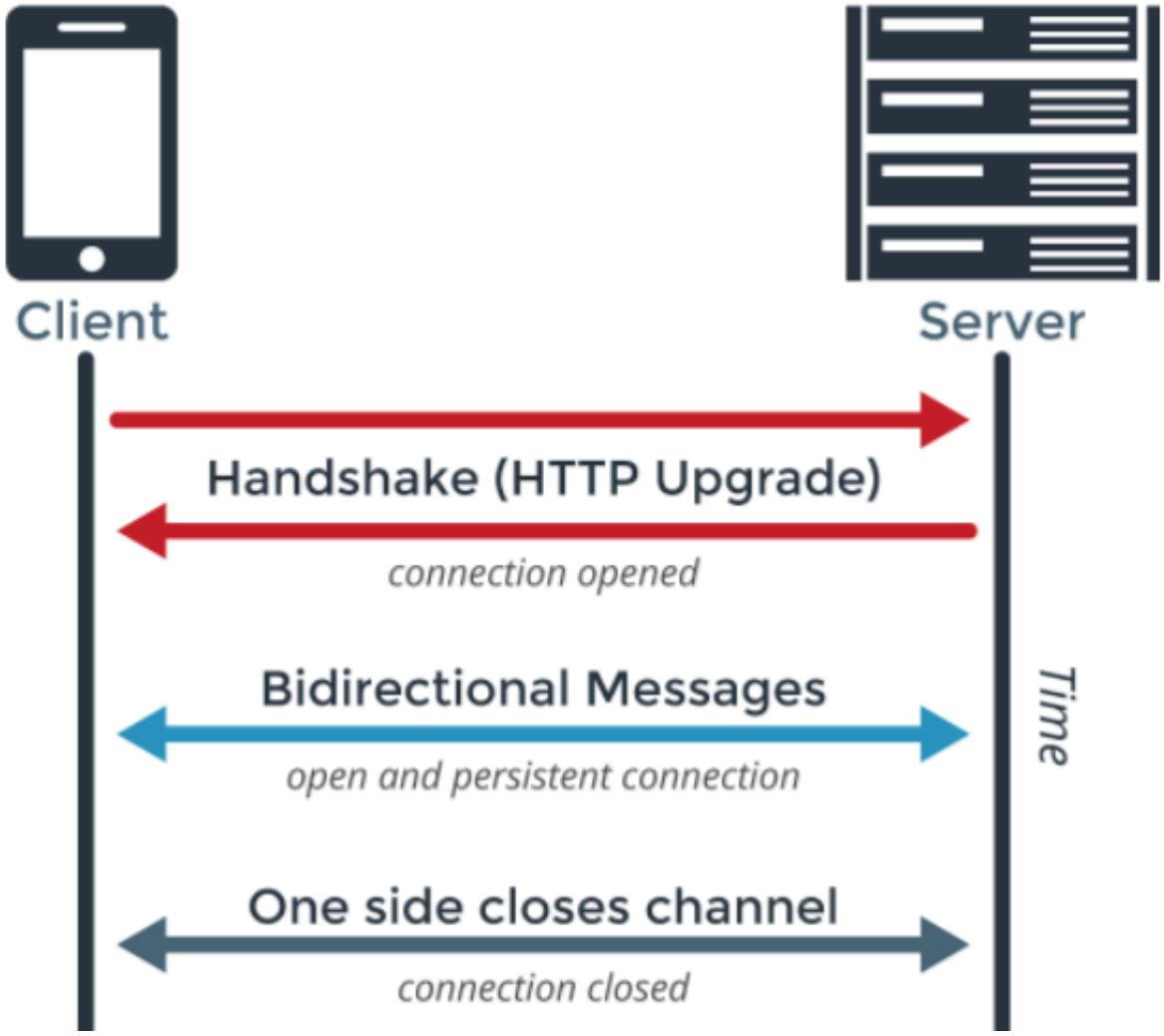
WSS（Web Socket Secure）是WebSocket的加密版本。

为何使用WS/WSS?

随着互联网的蓬勃发展，各种类型的Web应用层出不穷，很多应用要求服务端有能力进行实时推送能力（比如直播间聊天室），以往很多网站为了实现推送技术，所用的技术都是轮询。轮询是在特定的的时间间隔（如每1秒），由浏览器对服务器发出HTTP请求，然后由服务器返回最新的数据给客户端的浏览器。这种传统的模式带来很明显的缺点，即浏览器需要不断地向服务器发出请求，然而HTTP请求可能包含较长的头部，其中真正有效的数据可能只是很小的一部分，显然这样会浪费很多的带宽资源。

在这种情况下，HTML5定义了WebSocket协议，能更好地节省服务器资源和带宽，并且能够更实时地进行通讯。WebSocket实现了浏览器与服务器全双工(full-duplex)通信，允许服务器主动发送信息给客户端。

WebSocket协议的交互过程如下图所示。



如何在阿里云负载均衡上启用WS/WSS支持?

无需配置，当选用HTTP监听时，默认支持无加密版本WebSocket协议（WS协议）；当选择HTTPS监听时，默认支持加密版本的WebSocket协议（WSS协议）。

说明:
需要将实例升级为性能保障型实例。详细参见如何使用负载均衡性能保障型实例。

支持的地域

全部地域都已开放WS/WSS支持。

限制

WSS/WS协议支持的约束如下:

- 负载均衡与ECS后端服务的连接采用HTTP/1.1，建议后端服务器采用支持HTTP/1.1的Web Server。

- 若负载均衡与后端服务超过60秒无消息交互，会主动断开连接，如需要维持连接一直不中断，需要主动实现保活机制，每60秒内进行一次报文交互。

收费政策

WSS/WS协议支持不额外收取费用。