

Server Load Balancer

Best Practices

Best Practices

How to use guaranteed-performance instances?

1. What are guaranteed-performance instances?
2. How are guaranteed-performance instances charged?
3. What is the price of each capacity?
4. How to choose capacities for guaranteed-performance instances?
5. Can I modify the capacity after the instance is created?
6. When will Alibaba Cloud start to charge guaranteed-performance instances?
7. After Alibaba Cloud starts to charge capacity fee on guaranteed-performance instances, will it charge additional fees on shared-performance instances?
8. Why sometimes guaranteed-performance instances cannot reach the performance limit defined in the capacity?
9. Why sometimes the performance of a low-capacity guaranteed-performance instance is worse than that of a shared-performance instance?
10. Can I still buy shared-performance instances?
11. Will intranet SLB instances be charged for capacity fee?

1. What are guaranteed-performance instances?

A guaranteed-performance instance provides guaranteed performance metrics (performance SLA) and is opposite to a shared-performance instance. For a shared-performance instance, the performance metrics are not guaranteed and the resources are shared by all instances.

All instances are shared-performance instances before Alibaba launches guaranteed-performance instances. You can view the instance type on the console. Hover your mouse pointer to the green icon of the target instance to view the performance metrics, as shown in the following figure.

Server Load Balancer									
ID/Name	Zone	IP Address(All)	Status	Network(All)	Port/Health Check	Backend Server	Instance Specification	Bandwidth Billing Method(All)	Billing Method(All)
lb- (None)	cn-hangzhou- f(Hangzhou) cn-hangzhou- e(Slave)	47.98.17.89(Public IP)	Running	Classic Network	Not Configured	Not Configured	Guaranteed-Performance Instance slb.s1.small	Max Connection:5000 CPS:3000 QPS:1000	Pay-As-You-Go 8-01-29 02:50 ended

The following are three key performance metrics for guaranteed-performance instances:

Max Connection

The maximum number of connections to a SLB instance. When the maximum number of connections reaches the limits of the capacity, the new connection will be dropped.

Connection Per Second (CPS)

The rate at which a new connection is established per second. When the CPS reaches the limits of the capacity, the new connection will be dropped.

Query Per Second (QPS)

The number of HTTP/HTTPS requests that can be processed per second. When the QPS reaches the limits of the capacity, the new connection will be dropped.

This metrics is only available for Layer-7 Server Load Balancer.

Alibaba Cloud Server Load Balancer provides the following capacities for guaranteed-performance instances:

Note: If you want to use a larger capacity, contact your customer manager.

Capacity		Max Connection	CPS	QPS
Capacity 1	Small I (slb.s1.small)	5,000	3,000	1,000
Capacity 2	Standard I (slb.s2.small)	50,000	5,000	5,000
Capacity 3	Standard II (slb.s2.medium)	100,000	10,000	10,000
Capacity 4	Higher I (slb.s3.small)	200,000	20,000	20,000
Capacity 5	Higher II (slb.s3.medium)	500,000	50,000	30,000
Capacity 6	Super I (slb.s3.large)	1,000,000	100,000	50,000

2. How are guaranteed-performance instances billed?

Guaranteed-performance instances are billed as follows:

Total fee (per instance) = instance fee + traffic fee + capacity fee

The corresponding capacity fee is billed for each guaranteed-performance instance no matter the network type of the instance, and is billed based on the actual usage depending on the capacity selected. If the actual performance metrics of an instance occurs between two capacities, the capacity fee is charged at the higher capacity fee. For more information about the SLB billing, see [Billing](#).

The corresponding capacity fee is billed for each guaranteed-performance instance no matter the network type of the instance, and is billed based on the actual usage depending on the capacity selected. If the actual performance metrics of an instance occurs between two capacities, the capacity fee is charged at the higher capacity fee.

For example, if you purchase the **Super I (slb.s3.large)** capacity, and the actual usage of your instance in an hour is as follow:

Max Connection	CPS	QPS
90,000	4,000	11,000

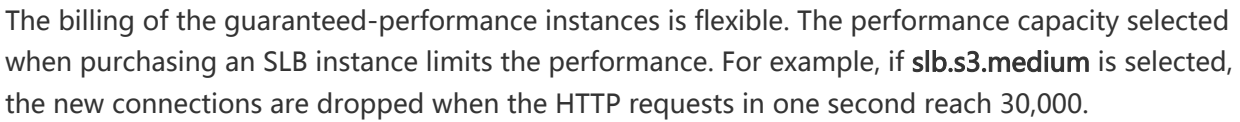
From the perspective of Max Connection, the actual metrics 90,000 occurs between the limit 50,000 defined in the **Standard I (slb.s2.small)** capacity and the limit 100,000 defined in the **Standard II (slb.s2.medium)** capacity. Therefore, the capacity of the Max Connection metrics in this hour is **Standard II (slb.s2.medium)**.

From the perspective of CPS, the actual metrics 4,000 occurs between the limit 3,000 defined in the **Small I (slb.s1.small)** capacity and the limit 5,000 defined in the **Standard I (slb.s2.small)** capacity. Therefore, the capacity of the CPS metrics in this hour is **Standard I (slb.s2.small)**.

From the perspective of QPS, the actual metrics 11,000 occurs between the limit 10,000 defined in the **Standard II (slb.s2.medium)** capacity and the limit 20,000 defined in the **Higher I (slb.s3.small)** capacity. Therefore, the capacity of the QPS metrics in this hour is **Higher I (slb.s3.small)**.

Comparing these three metrics, the capacity of the QPS metrics is highest, therefore, the capacity fee of the instance in this hour is charged at the price of the **Higher I (slb.s3.small)** capacity.

The following figure is an example showing how the capacity fee is billed for an SLB instance in the first three hours:



The following table lists the capacity price of each capacity. For the price on the SLB instance and traffic, see [Billing](#).

4

hi na N or th 5 (H u h e ha ot e)						
C hi na N or th 1 (Q in g d a o)						
C hi na N or th 2 (B eij in g)						
C hi na Ea st 2 (S ha n g ha i)						
C hi na S o ut h						

1 (S h e n z h e n)						
As ia Pa cif ic SE 1 (S in g a p o r e) As ia Pa cif ic SE 3 (K ua la Lu m p ur) As ia Pa cif ic SE 5 (J ak ar ta) As ia Pa cif ic S	Capacity 1	Small I (slb.s1.sma ll)	5000	3000	1000	Free of charge
	Capacity 2	Standard I (slb.s2.sma ll)	50000	5000	5000	0.06
	Capacity 3	Standard II (slb.s2.med ium)	100000	10000	10000	0.12
	Capacity 4	Higher I (slb.s3.sma ll)	200000	20000	20000	0.24
	Capacity 5	Higher II (slb.s3.med ium)	500000	50000	30000	0.37
	Capacity 6	Super I (slb.s3.larg e)	1000000 ⁶	100000	50000	0.61

O U 1 (M u m b a i)						
U S W e s t 1 (S i l i c o n V a l l e y)						
U S E a s t 1 (V i r g i n i a)						
H o n g K o n g						

Capacity fees of guaranteed-performance instances in the international regions can enjoy an 83% discount.

4. How to select a capacity?

Because the capacity fee is billed based on the actual usage, we recommend that you select the largest capacity (slb.s3.large). This guarantees the business flexibility (flexibility) and will not cause

extra costs. If your traffic does not reach the largest capacity, you can select a more reasonable capacity, such as `slb.s3.medium`.

5. Can I modify the capacity after the instance is created?

Yes.

You can change the capacity at any time and the change takes effect immediately. For more information, see [Change the configuration](#).

Note:

After you change a shared-performance instance to a guaranteed-performance instance, you cannot change it back.

If you change a shared-performance instance to a guaranteed-performance instance, a brief disconnection of service may occur for 10 to 30 seconds.

The IP of the SLB instance will not be changed after you changing the instance type or the capacity.

6. When will be the guaranteed-performance instances charged?

Alibaba Cloud launched the guaranteed-performance instances in May 2017, and will charge the capacity fee on guaranteed-performance instances from April 1, 2018.

Pay attention to official announcement, messages, and emails for notification.

7. After Alibaba Cloud starts to charge capacity fee on guaranteed-performance instances, will extra fees be charged on shared-performance instances?

No.

The billing of the original shared-performance instances is the same if you do not change it to a performance-guaranteed instance. However, if you change the shared-performance instance to the guaranteed-performance one, the capacity fee will be charged from April 1st, 2018.

8. Why sometimes guaranteed-performance instances cannot reach the performance limit as defined in the capacity?

It applies to the cask theory.

Guaranteed-performance instances do not guarantee that the three metrics can reach the capacity limits at the same time. The limitation is triggered as long as a metric first reaches the limitation defined in the capacity.

For example, you have purchased a guaranteed-performance instance of the **Higher I (slb.s3.small)** capacity. When the QPS of the instance reaches 20,000 but the number of maximum connections does not reach 200,000, the new connections are still dropped because the QPS has reached the limitation.

9. Why sometimes the performance of a smaller capacity guaranteed-performance instance is less efficient than that of a shared-performance instance?

For a shared-performance instance, all the resources are shared. Its performance cannot be guaranteed when the traffic load is high. However, a guaranteed-performance instance can guarantee the performance at any time. Because the capacity fee is billed based on the actual usage, we recommend that you select the highest capacity.

10. Can I still buy shared-performance instances?

Yes.

However, shared-performance instances will be phased out in the future. Please pay attention to the official announcement.

11. Will intranet SLB instances be charged for capacity fee?

If the intranet SLB instance is a shared-performance instance, no capacity fee is charged.

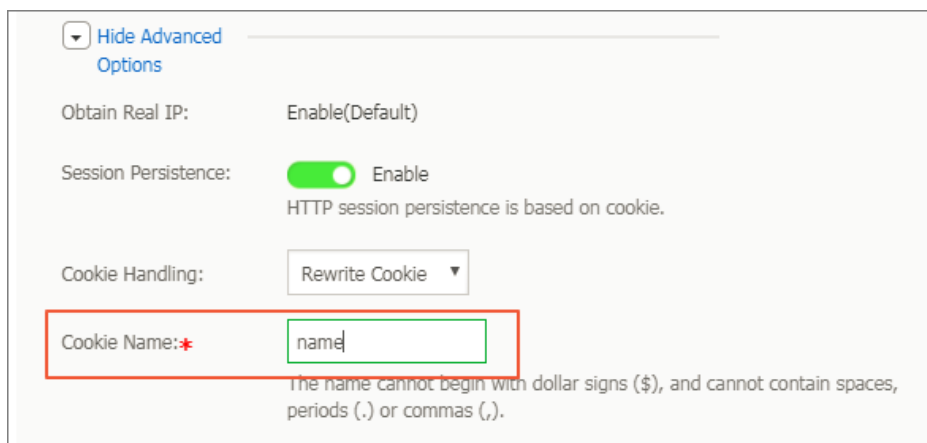
If the intranet SLB instance is a guaranteed-performance instance, corresponding capacity fee is charged, and no other fees are charged.

Configure cookie in the backend server

Server Load Balancer provides session persistence function. With session persistence enabled, Server Load Balancer can distribute requests from the same client to the same backend server during the session period.

For layer-4 listeners, session persistence is based on the IP address. The listener of Server Load Balancer forwards requests from the same IP address to the same backend server.

For layer-7 listeners, session persistence is based on cookies. If you choose the **Rewrite Cookie** method, you can set the **Cookie Name** as name, and set the key of vip.a.com 's cookie as name on the backend server.



The screenshot shows the 'Advanced Options' section of the Server Load Balancer console. The 'Session Persistence' toggle is set to 'Enable', with a note that 'HTTP session persistence is based on cookie.' The 'Cookie Handling' dropdown is set to 'Rewrite Cookie'. The 'Cookie Name' field is highlighted with a red box and contains the text 'name'. A red asterisk indicates a required field. Below the field, a note states: 'The name cannot begin with dollar signs (\$), and cannot contain spaces, periods (.) or commas (,).'.

Follow the instructions in this section to set cookies on a backend server.

Apache

Open the httpd.conf file and make sure that the following line is not commented.

```
LoadModule usertrack_module modules/mod_usertrack.so
```

Add the following configurations in the VirtualHost file.

```
CookieName name
CookieExpires "1 days"
CookieStyle Cookie
CookieTracking on
```

Nginx

Configure the configuration file as follows.

```
server {  
    listen 8080;  
    server_name wqwq.example.com;  
    location / {  
        add_header Set-Cookie name=xxxx;  
        root html;  
        index index.html index.htm;  
    }  
}
```

Lighttpd

Configure the configuration file as follows.

```
server.modules = ( "mod_setenv" )  
$HTTP["host"] == "test.example.com" {  
    server.document-root = "/var/www/html/"  
    setenv.add-response-header = ( "Set-Cookie" => "name=XXXXXX" )  
}
```

Use Open API to configure Server Load Balancer

In this tutorial, the request parameters are included in the request URL, and the URL does not include common parameters. For more information, see [API overview](#).

Note: To increase readability, the parameter values of the request URL in this example are not URL-encoded.

Prerequisites

You have created 2 ECS instances and granted access to their SSH and Web ports.

Procedure

Call `CreateLoadBalancer` interface to create a Server Load Balancer instance.

Request:

`https://slb.aliyuncs.com/?Action=CreateLoadBalancer&RegionId=cn-hangzhou-dg-a01`

Response:

```
{
  "RequestId": "3DE96B24-E2AB-4DFA-9910-1AADD60E13A5",
  "LoadBalancerId": "LoadBalancerId",
  "Address": "SLBIPAddress"
}
```

Call `CreateLoadBalancerHttpListener` interface to create a HTTP listener, of which the port is 80, for the Server Load Balancer instance.

Request:

`https://slb.aliyuncs.com/?Action=CreateLoadBalancerHttpListener&LoadBalancerId=LoadBalancerId&ListenerPort=80&BackendServerPort=80&ListenerStatus=active`

Call `SetLoadBalancerStatus` interface to active the Server Load Balancer instance.

Request:

`https://slb.aliyuncs.com/?Action=SetLoadBalancerStatus&LoadBalancerId=LoadBalancerId&LoadBalancerStatus=active`

Call `AddBackendServers` interface to add an ECS instance to backend servers.

Request:

`https://slb.aliyuncs.com/?Action=AddBackendServers&LoadBalancerId=LoadBalancerId&BackendServers=[{"ServerId":"ECS1InstanceId"}]`

Response:

```
{
  "RequestId": "FA2F2172-63F2-409D-927C-86BD1D536F13",
  "LoadBalancerId": "LoadBalancerId",
  "BackendServers": {
    "BackendServer": [
```

```
{
  "ServerId" : "ECS1InstanceId",
  "Weight" : 100
}
]
```

Call AddBackendServers interface again to add an ECS instance to backend servers.

Request:

[https://slb.aliyuncs.com/?Action=AddBackendServers&LoadBalancerId=LoadBalancerId&BackendServers=\[{"ServerId":"ECS2InstanceId"}\]](https://slb.aliyuncs.com/?Action=AddBackendServers&LoadBalancerId=LoadBalancerId&BackendServers=[{)

Response:

```
{
  "RequestId" : "C61FAD0A-2E87-4D0C-80B0-95AB758FCA70",
  "LoadBalancerId" : "LoadBalancerId",
  "BackendServers" : {
    "BackendServer" : [
      {
        "ServerId" : "ECS1InstanceId",
        "Weight" : 100
      },
      {
        "ServerId" : "ECS2InstanceId",
        "Weight" : 100
      }
    ]
  }
}
```

Call DescribeLoadBalancerAttribute interface to view the configuration of the Server Load Balancer instance.

Request:

<https://slb.aliyuncs.com/?Action=DescribeLoadBalancerAttribute&LoadBalancerId=LoadBalancerId>

Response:

```
{
  "RequestId" : "4747E9AE-ADFD-412D-B523-C1CBD45A2154",
  "LoadBalancerId" : "LoadBalancerId",
  "Address" : "SLBIPAddress",
  "IsPublicAddress" : "true",
  "ListenerPorts" : {
```

```
"ListenerPort" : [
  80
],
},
"BackendServers" : {
  "BackendServer" : [
    {
      "ServerId" : "ECS1InstanceId",
      "Weight" : 100
    },
    {
      "ServerId" : "ECS2InstanceId",
      "Weight" : 100
    }
  ]
}
}
```

Use your browser to access the IP address of the Server Load Balancer instance to verify whether the service is working.

Remove backend ECS

Directly removing backend ECS instances from a Server Load Balancer instance may cause service interruption. We recommend setting the weight of an ECS instance to zero first, and then remove it when no traffic is distributed to it.

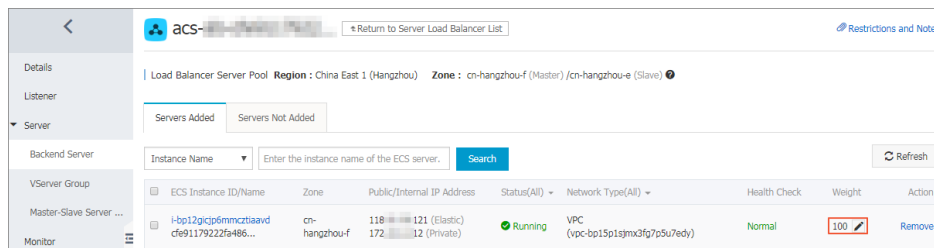
Log on Server Load Balancer console.

Choose a region and then click the ID of the target Server Load Balancer instance.

In the left-side navigation pane, click **Server > Backend Server**.

If the ECS instance is added to a server group, click **VServer Group** or **Master-Slave Server Group** accordingly.

Hover the mouse pointer to the weight of the target ECS instance and then set the value to 0.



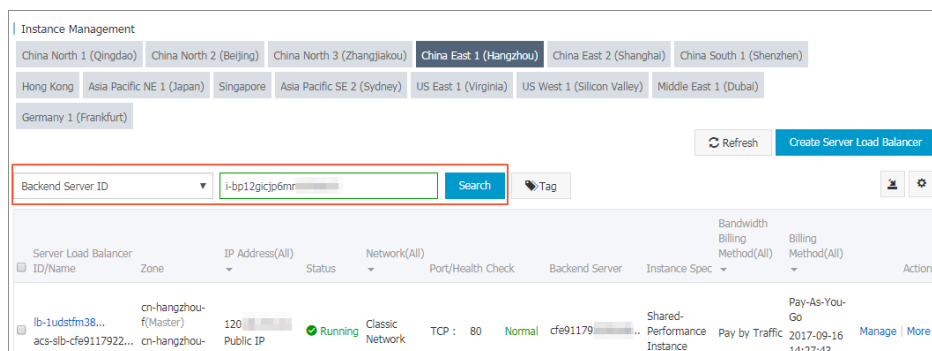
When no traffic is distributed to the ECS instance, click **Remove** to remove it from the backend server pool.

Troubleshoot

If there are ongoing service requests sent to the ECS instance after removing it from the backend server pool, check the following:

Whether the ECS instance is added to backend server pools of other Server Load Balancer instances.

You can use the ECS instance ID to filter Server Load Balancer instances that the ECS instance is added to.



Log on to the ECS instance, run the netstat command to check whether the ECS instance is deployed with public services.

Windows: Run `netstat -ano` to view all open ports on the instance.

Linux: Run this command to view all open ports on the instance or use other parameters of the netstat command.


```

~# netstat -ano
Active Internet connections (servers and established)
Proto Recv-Q Send-Q Local Address           Foreign Address         State       Timer
tcp        0      0 0.0.0.0:22              0.0.0.0:*               LISTEN      off (0.00/0/0)
tcp        0      0 0.0.0.0:111             0.0.0.0:*               LISTEN      off (0.00/0/0)
tcp        0      0 172.16.17.1:42285       0.0.0.0:*               ESTABLISHED off (0.00/0/0)
tcp        0 428 172.16.17.1:22          172.16.17.1:44832       ESTABLISHED on (0.16/0/0)
tcp6       0      0 :::111                  :::*                    LISTEN      off (0.00/0/0)
udp        0      0 0.0.0.0:42947           0.0.0.0:*               off (0.00/0/0)
udp        0      0 0.0.0.0:68              0.0.0.0:*               off (0.00/0/0)
udp        0      0 0.0.0.0:111             0.0.0.0:*               off (0.00/0/0)
udp        0      0 0.0.0.0:627             0.0.0.0:*               off (0.00/0/0)
udp        0      0 172.16.17.1:123         0.0.0.0:*               off (0.00/0/0)
udp        0      0 127.0.0.1:123           0.0.0.0:*               off (0.00/0/0)
udp        0      0 0.0.0.0:123             0.0.0.0:*               off (0.00/0/0)
udp6       0      0 :::111                  :::*                    off (0.00/0/0)
udp6       0      0 :::627                  :::*                    off (0.00/0/0)
udp6       0      0 :::123                  :::*                    off (0.00/0/0)
udp6       0      0 :::1275                 :::*                    off (0.00/0/0)
Active UNIX domain sockets (servers and established)
Proto RefCnt Flags       Type       State       I-Node  Path
unix    2      [ ]         DGRAM     7689       /run/systemd/shutdown
unix    7      [ ]         DGRAM     7691       /run/systemd/journal/dev-log
unix    2      [ AGC ]     SOCKSEQ   LISTENING  7695       /run/udev/control

```

Obtain the real IP address of the client

Introduction to the function of obtaining IP address

Alibaba Cloud Server Load Balancer provides the function of obtaining the real IP address of the client and this function is enabled by default.

For the Layer-4 load balancing service (TCP protocol), listeners distribute client requests to backend ECS servers without modifying the request headers. Therefore, you can obtain the real IP address from the backend ECS servers without additional configurations.

For the Layer-7 load balancing service (HTTP/HTTPS protocol), you have to configure the application servers, and then use the X-Forwarded-For header to obtain the real IP addresses of the clients.

Note: For the HTTPS load balancing service, the SSL certificates are configured in front-end listeners, the backend still uses the HTTP protocol. Therefore, the configurations on application servers are the same for HTTP and HTTPS protocols.

Configure web applications

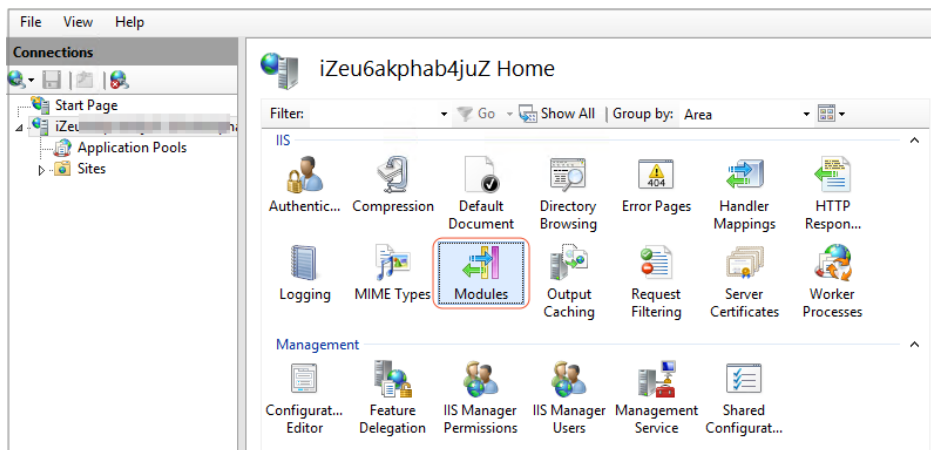
This section introduces some common methods used to configure web applications.

Configure IIS7/IIS8

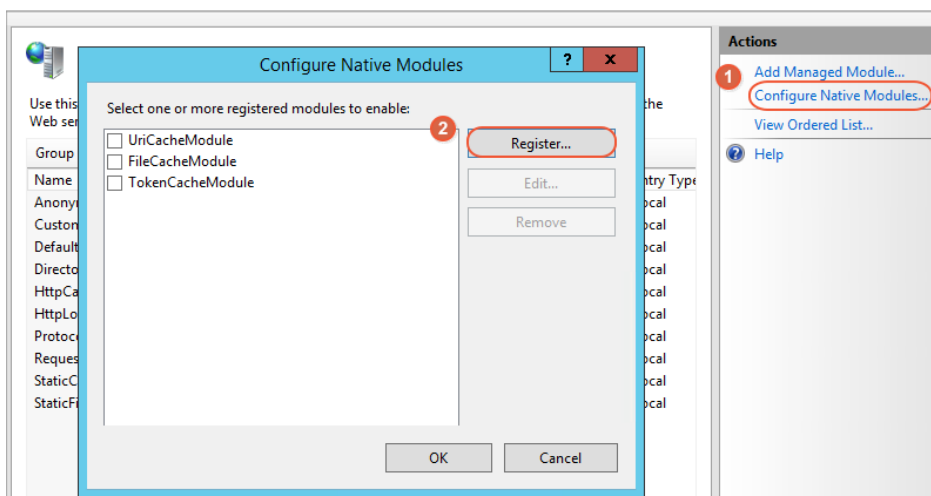
Download and extract the F5XForwardedFor.

Copy the F5XFFHttpModule.dll and F5XFFHttpModule.ini files from the extracted folder to a folder, such as C:\F5XForwardedFor\. Make sure that the IIS process has the write permission to this folder.

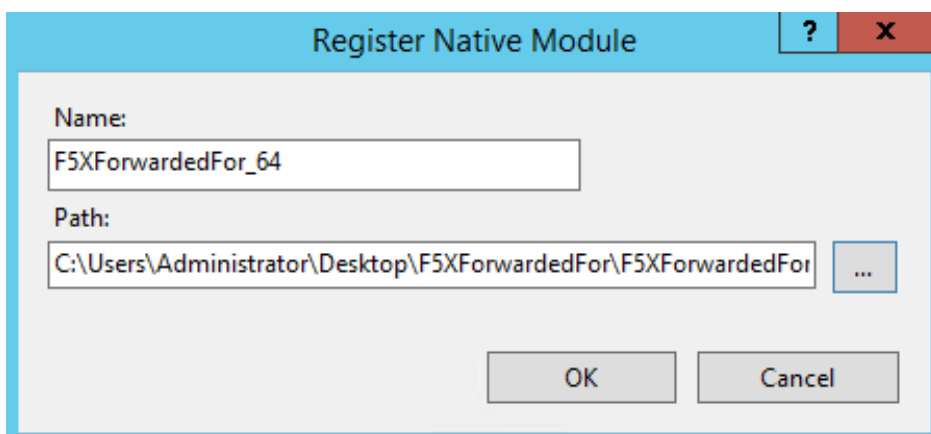
Open the IIS Manager, and then double-click the **Modules** function.



Click **Configure Native Modules**, and then click **Register**.

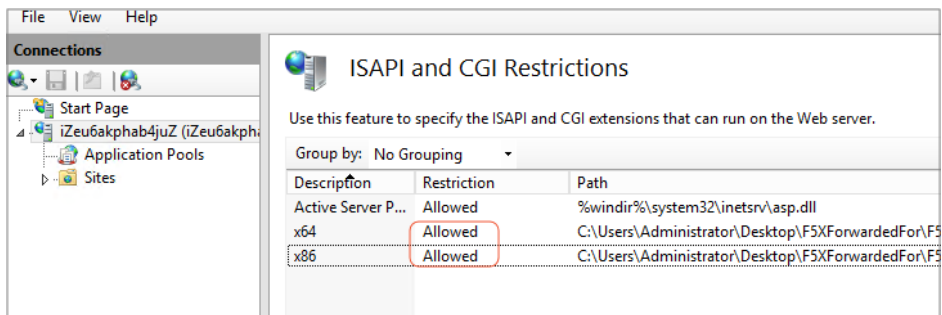


Add the copied the .dll file.



Add the ISAPI and CGI restrictions for the .dll file and set the restriction to **Allowed**.

Make sure that you have installed the ISAPI and CGI applications.



Restart the IIS Manager.

Configure Apache

Run the following command to install the mod_rpaf module.

```
wget http://stderr.net/apache/rpaf/download/mod_rpaf-0.6.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
```

Open the `/alidata/server/httpd/conf/httpd.conf` file and add the following information at the end of the content.

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

`RPAFproxy_ips`: the IP address is not the IP address of the Server Load Balancer instance. Check the Apache log to find the IP address, usually both the two IP addresses are entered.

Run the following command to restart the Apache server.

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

Configure Nginx

Run the following command to install `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`
```

Run the following command to open the nginx.conf file.

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

Find the following content and add the required information after it.

```
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;
```

The information to be added:

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

set_real_ip_from IP: the IP address is not the IP address of the Server Load Balancer instance. Check the Nginx log to find the IP address, usually both the two IP addresses are entered.

Run the following command to restart the Nginx server.

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

How to forward same-domain requests to different servers

In this case, we use four ECSs deployed with Nginx servers as the example to demonstrate how to configure forwarding rules specified by domain name and URL, so as to fulfill traffic forwarding as shown in the following table.

Frontend request	Forward traffic to
www.aaa.com/tom	Server SLB_tom1 and server SBL_tom2
www.aaa.com/jerry	Server SLB_jerry1 and server SBL_jerry2

Instance ID/Name	Zone	IP Address	Status(All)	Network Type(All)
i-bp1huan9m5ia3jvcmk0cmg7 SLB_jerry1	China East 1 Zone F	47.96.179.121(Elastic IP Address) 172.16.19.13(Private IP Address)	Running	VPC
i-bp1x52etca9hmgpku2py SLB_jerry2	China East 1 Zone F	47.96.172.148(Elastic IP Address) 172.16.33.32(Private IP Address)	Running	VPC
i-bp138ue0e0k0k0dc7rea SLB_tom1	China East 1 Zone F	114.62.125.14(Elastic IP Address) 172.16.19.20(Private IP Address)	Running	VPC
i-bp138ue0e0k0k0dc7rea SLB_tom2	China East 1 Zone F	47.96.169.115(Elastic IP Address) 172.16.30.30(Private IP Address)	Running	VPC

Procedure

Create an Internet-facing SLB instance.

For details, see [Create a server load balancer](#).

Resolve the domain name into the public IP of the SLB instance by using DNS.

For convenience, the public IP of the SLB instance is bound to domain name www.aaa.com in the host file in this case.

Create two VServer groups.

Locate the newly created instance in the Server Load Balancer console and click the instance ID to go to the **Instance Details** page.

In the left-side navigation pane, click **Server > VServer Group**.

Click **Create VServer Group**.

In the dialog box that appears, select the backend servers to be added and set ports and weights for them respectively. The ports for ECSs in the VServer group can be different.

In this case, enter **TOM** as the server group name, add server SLB_tom1 and server SBL_tom2 into the group, set the port number to 80, and keep the default weight value (100).

Notice: The network type of current server load balancer is VPC, instance type is Intranet IP. This VServer group can only add a VPC ECS.

*Group Name:

*Server Network Type: ☐ Classic Network ☒ VPC

Instance Na Enter the name of the ECS instan

Available Servers		
ECS Instance ID/Name	IP Address	Zone
i-bp1han9m6x3ecmk37mqz SLB_jerry1	47.96.175.121 (EIP) 172.16.33.33 (Private)	cn-hangzhou-f VPC
i-bp162wb1ayjhmp6u2py SLB_jerry2	47.96.172.148 (EIP) 172.16.33.32 (Private)	cn-hangzhou-f VPC
i-bp18us9p89pilgbc7yaa SLB_tom1	116.62.128.54 (EIP) 172.16.33.30 (Private)	cn-hangzhou-f VPC
i-bp1e9bjg74lm9ofxddd SLB_tom2	116.62.158.112 (EIP) 172.16.33.29 (Private)	cn-hangzhou-f VPC

Selected Servers Add(2/20)				
ECS Instance ID/Name	IP Address	Zone	Port	Weight
i-bp18us9p89pilg... SLB_tom1	116.62.128.54 (EIP) 172.16.33.30 (Private)	cn-hangzhou-f VPC	80	100
i-bp1e9bjg74lm9o... SLB_tom2	116.62.158.112 (EIP) 172.16.33.29 (Private)	cn-hangzhou-f VPC	80	100

Note: Already added an ECS instance in a VPC (ID: vpc-bp1w92wjrgz01fm6pubd8). Only ECS

Repeat the preceding steps to add another VServer group named JERRY, which includes server SLB_jerry1 and server SBL_jerry2.

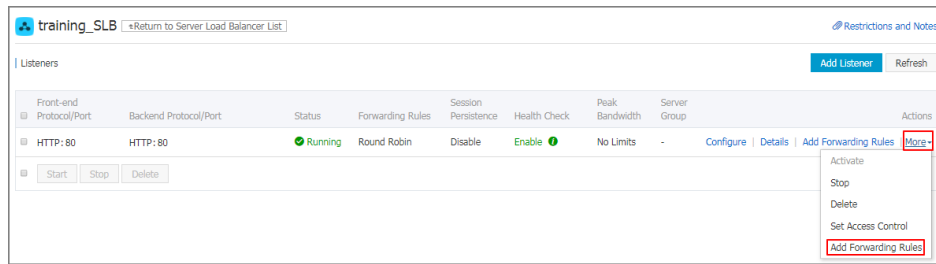
Add a listener.

In the left-side navigation pane, click **Listeners**, and click **Add Listener**.

Configure the listener. In this case, the listener is configured as follows:

- i. Frontend protocol [Port]: HTTP: 80
- ii. Backend protocol [Port]: HTTP: 80
- iii. Scheduling algorithm: Round-robin.
- iv. Keep the default values for other configuration items.

On the **Listeners** page, click **More > Add Forwarding Rules**.



On the **Forwarding rules** page, click **Add Forwarding Rules**.

Configure three forwarding rules.

Add Forwarding Rules

Rule Name	Domain Name	URL	VServer Group	Actions
rule1	www.aaa.com	/jerry	JERRY	Delete
rule2	www.aaa.com	/tom	TOM	Delete
rule3	www.aaa.com		TOM	Delete

Add Forwarding Rule

* Domain name rule:
The domain name can contain letters a-z, numbers 0-9, hyphens (-), and periods (.), and wildcard characters. The following two domain name formats are supported:
- Standard domain name: www.test.com
- Wildcard domain name: *.test.com. wildcard (*) must be the first character in the format of (*.)

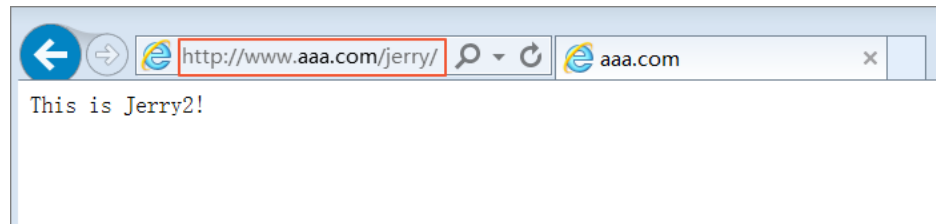
* URL rule:
URLs must be 2-80 characters in length. Only letters a-z, numbers 0-9, and characters '-', '/', '?', '%', '#', and '&' are allowed. URLs must be started with the character '/', but cannot be '/' alone.

* At least one domain name rule or URL rule is required.

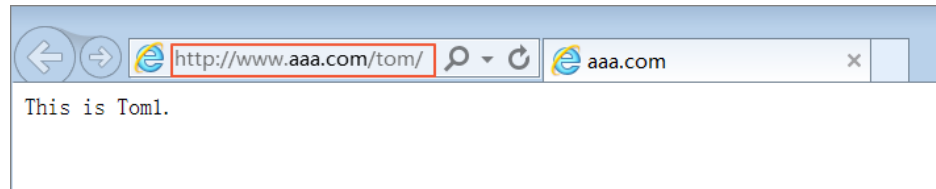
Confirm
Cancel

Test:

Enter `www.aaa.com/jerry` in the browser and the following result is returned.



Enter `www.aaa.com/tom` in the browser and the following result is returned.



Enter `www.aaa.com` in the browser and the following result is returned.

