

Elastic Load Balance

User Guide

Issue 29
Date 2024-03-15



Copyright © Huawei Cloud Computing Technologies Co., Ltd. 2024. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Cloud Computing Technologies Co., Ltd.

Trademarks and Permissions



HUAWEI and other Huawei trademarks are the property of Huawei Technologies Co., Ltd.

All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei Cloud and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Contents

1 Load Balancer	1
1.1 Overview	1
1.2 Preparations for Creating a Load Balancer	3
1.3 Creating a Dedicated Load Balancer	7
1.4 Creating a Shared Load Balancer	14
1.5 Enabling Guaranteed Performance for a Shared Load Balancer	18
1.6 Configuring Modification Protection for Load Balancers	19
1.7 Modifying the Bandwidth	20
1.8 Changing the Specifications of a Dedicated Load Balancer	20
1.9 Changing the Billing Mode or Bandwidth Billing Option	22
1.10 Changing an IP Address	23
1.11 Binding an IP Address to or Unbinding an IP Address from a Load Balancer	24
1.12 Adding to or Removing from an IPv6 Shared Bandwidth	27
1.13 Exporting the Load Balancer List	28
1.14 Deleting a Load Balancer	28
2 Listener	30
2.1 Overview	30
2.2 Protocols and Ports	31
2.3 Adding a TCP Listener	33
2.4 Adding a UDP Listener	45
2.5 Adding an HTTP Listener	56
2.6 Adding an HTTPS Listener	72
2.7 Adding a UDP Listener (with a QUIC Backend Server Group Associated)	90
2.8 Configuring Modification Protection for a Listener	91
2.9 Configuring Timeout Durations	92
2.10 Modifying or Deleting a Listener	95
2.11 Transfer Client IP Address (Dedicated Load Balancers)	96
2.12 Transfer Client IP Address (Shared Load Balancers)	97
3 Advanced Features of HTTP/HTTPS Listeners	99
3.1 Forwarding Policy (Shared Load Balancers)	99
3.2 Forwarding Policy (Dedicated Load Balancers)	104
3.3 Advanced Forwarding (Dedicated Load Balancers)	107

3.3.1 Advanced Forwarding.....	107
3.3.2 Managing an Advanced Forwarding Policy.....	115
3.4 Mutual Authentication.....	117
3.5 HTTP/2.....	123
3.6 HTTP Redirection to HTTPS.....	124
3.7 HTTP Headers.....	126
3.8 SNI Certificate.....	128
4 Backend Server Group.....	131
4.1 Overview.....	131
4.2 Key Features.....	134
4.2.1 Health Check.....	134
4.2.2 Load Balancing Algorithms.....	140
4.2.3 Sticky Session.....	146
4.2.4 Forwarding Mode (Dedicated Load Balancers).....	148
4.2.5 Slow Start (Dedicated Load Balancers).....	149
4.3 Creating a Backend Server Group (Dedicated Load Balancers).....	149
4.4 Creating a Backend Server Group (Shared Load Balancers).....	157
4.5 Modifying a Backend Server Group.....	163
4.5.1 Overview.....	163
4.5.2 Modifying Health Check Settings.....	164
4.5.3 Changing the Load Balancing Algorithm.....	168
4.5.4 Modifying Sticky Session Settings.....	168
4.5.5 Modifying Slow Start Settings (Dedicated Load Balancers).....	169
4.6 Changing a Backend Server Group.....	171
4.7 Viewing a Backend Server Group.....	172
4.8 Deleting a Backend Server Group.....	172
5 Backend Server (Dedicated Load Balancers).....	174
5.1 Overview.....	174
5.2 Security Group Rules.....	176
5.3 Managing Backend Servers.....	179
5.3.1 Adding Backend Servers.....	179
5.3.2 Viewing Backend Servers.....	180
5.3.3 Removing Backend Servers.....	180
5.3.4 Changing Backend Server Weights/Ports.....	181
5.4 IP Addresses as Backend Servers.....	182
5.4.1 Overview.....	182
5.4.2 Enabling IP as a Backend.....	183
5.4.3 Adding IP Addresses as Backend Servers.....	184
5.4.4 Viewing Backend Servers.....	185
5.4.5 Removing Backend Servers.....	185
5.4.6 Changing Backend Server Weights/Ports.....	186
5.5 Supplementary Network Interfaces.....	187

5.5.1 Adding Supplementary Network Interfaces.....	187
5.5.2 Viewing Supplementary Network Interfaces.....	188
5.5.3 Removing Supplementary Network Interfaces.....	189
5.5.4 Changing the Weights/Ports of Supplementary Network Interfaces.....	189
6 Backend Server (Shared Load Balancers).....	191
6.1 Overview.....	191
6.2 Security Group Rules.....	192
6.3 Managing Backend Servers.....	195
6.3.1 Adding Backend Servers.....	195
6.3.2 Viewing Backend Servers.....	196
6.3.3 Removing Backend Servers.....	196
6.3.4 Changing Backend Server Weights.....	197
7 Certificate.....	199
7.1 Introduction to Certificates.....	199
7.2 Certificate and Private Key Format.....	200
7.3 Converting Certificate Formats.....	201
7.4 Adding a Certificate.....	202
7.5 Deleting a Certificate.....	205
7.6 Replacing the Certificate Bound to a Listener.....	206
7.7 Replacing the Certificate Bound to Different Listeners.....	207
7.8 Querying Listeners by Certificate.....	208
8 Access Control.....	209
8.1 Access Control.....	209
8.2 Managing IP Address Groups.....	211
8.2.1 Creating an IP Address Group.....	212
8.2.2 Viewing the Details of an IP Address Group.....	213
8.2.3 Managing IP Addresses in an IP Address Group.....	214
8.2.4 Deleting an IP Address Group.....	216
9 TLS Security Policy.....	217
10 Tag.....	227
11 Access Logging.....	230
12 Protection for Mission-Critical Operations.....	241
13 Self-service Troubleshooting.....	245
13.1 Overview.....	245
13.2 Troubleshooting an Unhealthy Backend Server.....	245
13.3 Other Issues.....	250
14 Monitoring.....	251
14.1 Monitoring Metrics.....	251
14.2 Setting an Alarm Rule.....	265

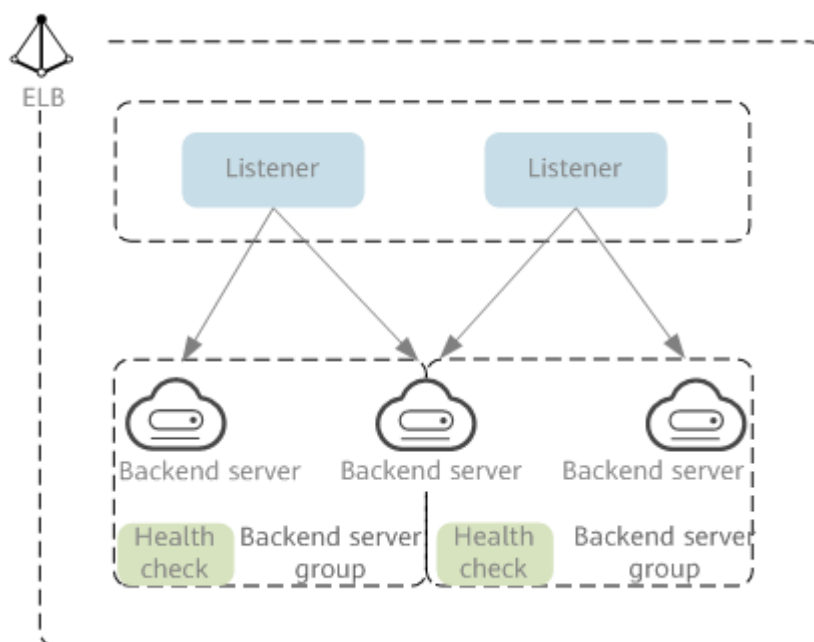
14.2.1 Creating an Alarm Rule.....	265
14.2.2 Modifying an Alarm Rule.....	266
14.3 Viewing Metrics.....	267
15 Auditing.....	269
15.1 Key Operations Recorded by CTS.....	269
15.2 Viewing Traces.....	270
16 Permissions Management.....	273
16.1 Creating a User and Granting Permissions.....	273
16.2 Creating a Custom Policy.....	274
17 Quotas.....	277
18 Appendix.....	279
18.1 Configuring the TOA Module.....	279
19 Change History.....	286

1 Load Balancer

1.1 Overview

A load balancer distributes incoming traffic across multiple backend servers. Before using a load balancer, you need to add at least one listener to it and associate one backend server with it.

Figure 1-1 ELB components



Network Type

Load balancers can work on both public and private network.

- Load balancers on the public network route requests over the Internet. Each load balancer has an EIP bound so that it can receive requests from clients on the Internet and routes the requests across backend servers.

Application scenario

- A load balancer is used as a single point of contact for clients when a group of servers provide services over the Internet.
- Fault tolerance and fault recovery are necessary.
- Load balancers on a private network route requests within a VPC.
This type of load balancers has only private IP addresses and can be accessed only in the VPC. They receive requests from clients in a VPC and route the requests across backend servers in the same VPC.

Application scenario

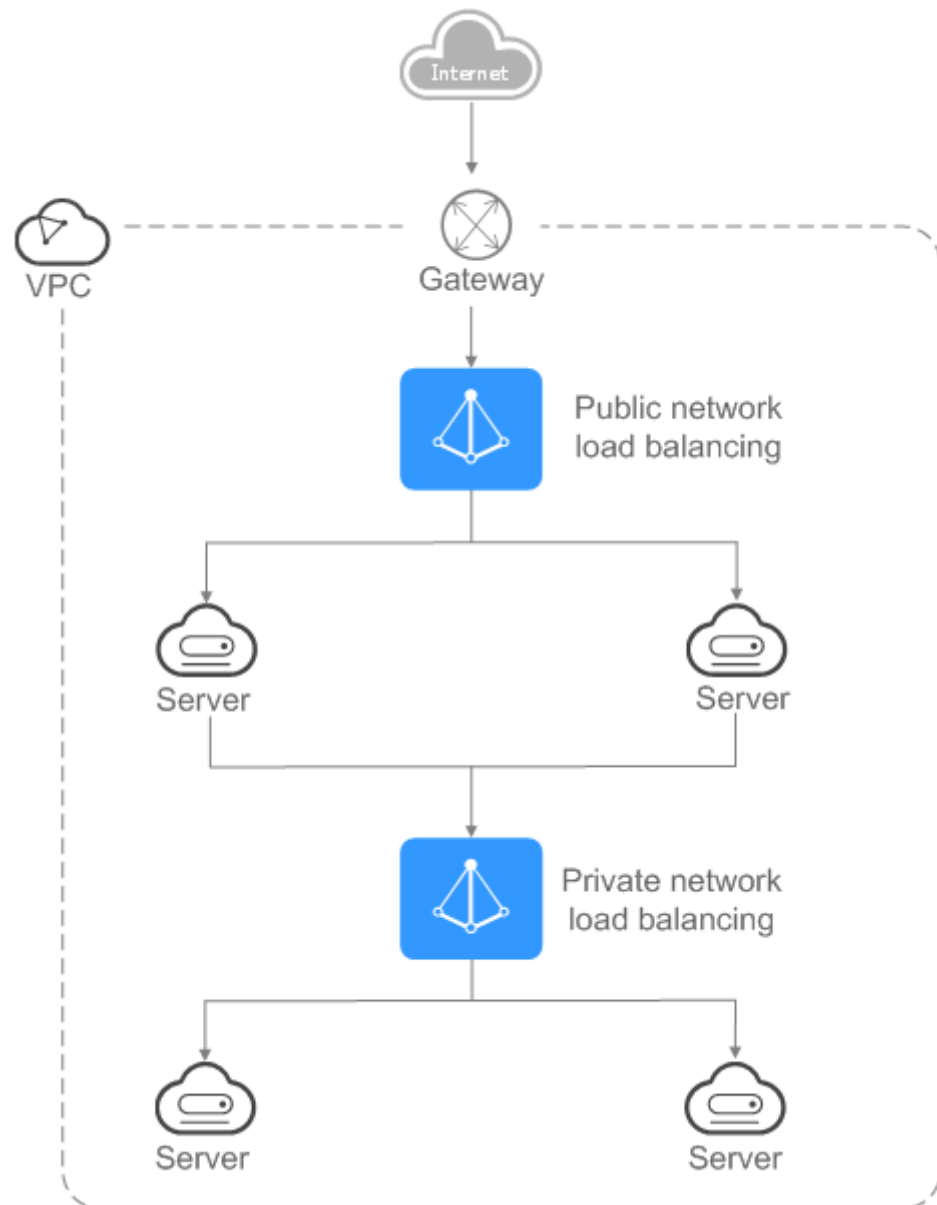
Both clients and backend servers are in the same VPC as the load balancer.

- There are multiple backend servers, and requests need to be evenly distributed across these servers.
- Fault tolerance and fault recovery are necessary.
- You do not want IP addresses of your physical devices to be exposed.

Load balancing on both public and private networks

Suppose that you have deployed both web servers and database servers. The web servers are accessible from users on the Internet, while the database servers can be accessed only on the private network. In this case, you can create two load balancers, one for the web servers and one for the database servers. The load balancer on the public network receives requests over the Internet and routes the requests to the web servers. Then, the load balancer on the private network forwards the requests to database servers.

Figure 1-2 Load balancing on both public and private networks



1.2 Preparations for Creating a Load Balancer

Before creating a load balancer, you must plan its region, network, protocol, and backend servers.

Region

When you select a region, note the following:

- The region must be close to your users to reduce network latency and improve the download speed.
- Shared load balancers cannot distribute traffic across regions. When creating a load balancer, select the same region as the backend servers.

- You can associate backend servers across regions or in a different VPC with a dedicated load balancer in either of the following ways:
 - If the backend servers are in different VPCs, you can use Cloud Connect to connect the VPCs across regions. For details, see the [Cloud Connect User Guide](#).
 - To add backend servers in a different VPC or an on-premises data center, you need to enable **IP as a Backend** for the load balancer. For details, see [Configuring Hybrid Load Balancing](#).

AZ

Dedicated load balancers can be deployed across AZs. If you select multiple AZs, a load balancer is created in each selected AZ.

Load balancers in these AZs work in active-active or multi-active mode and requests are distributed by the nearest load balancer in the same AZ.

To reduce network latency and improve access speed, you are suggested to deploy your load balancer in the AZ where backend servers are running.

If disaster recovery is required, create load balancers based on the scenario:

- **One load balancer in multiple AZs (disaster recovery at the AZ level)**

If the number of requests does not exceed what the largest specifications (large II) can handle, you can create a load balancer and select multiple AZs. In this way, if the load balancer in a single AZ is abnormal, the load balancer in other AZs can route the traffic, and disaster recovery can be implemented among multiple AZs.
- **Multiple load balancers and each load balancer in multiple AZs (disaster recovery at both the load balancer and AZ level)**

If the number of requests exceeds what the largest specifications (large II) can handle, you can create multiple load balancers and select multiple AZs for each load balancer. In this way, if a single load balancer is abnormal, other load balancers can distribute the traffic, and disaster recovery can be implemented among multiple load balancers and AZs.

 NOTE

- If requests are from the Internet, the load balancer in each AZ you select routes the requests based on source IP addresses. If you deploy a load balancer in two AZs, the requests the load balancers can handle will be doubled.
- For requests from a private network:
 - If clients are in the AZ you selected when you created the load balancer, requests are distributed by the load balancer in this AZ. If the load balancer is unavailable, requests are distributed by the load balancer in another AZ you select.
If the load balancer is available but the connections that the load balancer needs to handle exceed the amount defined in the specifications, service may be interrupted. To address this issue, you need upgrade specifications. You can monitor traffic usage on private network by AZ.
 - If clients are in an AZ that is not selected when you create the load balancer, requests are distributed by the load balancer in each AZ you select based on source IP addresses.
- If requests are from a Direct Connect connection, the load balancer in the same AZ as the Direct Connect connection routes the requests. If the load balancer in this AZ is unavailable, requests are distributed by the load balancer in another AZ.
- If clients are in a VPC that is different from where the load balancer works, the load balancer in the AZ where the original VPC subnet resides routes the requests. If the load balancer in this AZ is unavailable, requests are distributed by the load balancer in another AZ.

Network Type

Dedicated load balancers support IPv4 public network, IPv4 private network, and IPv6 network.

- If you select the public IPv4 network, the load balancer will have an IPv4 EIP bound to route requests over the Internet.
- If you select the private IPv4 network, a private IPv4 address will be assigned to the load balancer to route requests within a VPC.
- If you select the IPv6 network, the load balancer will have an IPv6 address, which allows the load balancer to route requests within a VPC. If you add the IPv6 address to a shared bandwidth, the load balancer can also process requests over the Internet.

Shared load balancers can work in both public and private networks.

- To route requests over the Internet, you need to bind an EIP to the load balancer. The load balancer also has a private IP address and can route requests in a VPC.
- To route requests in a VPC, bind only a private IP address to the load balancer.

Specifications

Dedicated load balancers provide a broad range of specifications to meet your requirements in different scenarios. Specifications for network load balancing are suitable for TCP or UDP requests, while specifications for application load balancing are broadly used to handle HTTP or HTTPS requests. Select appropriate specifications based on your traffic volume and service requirements. For details, see [Specifications of Dedicated Load Balancers](#).

The following are some principles for you to select the specifications:

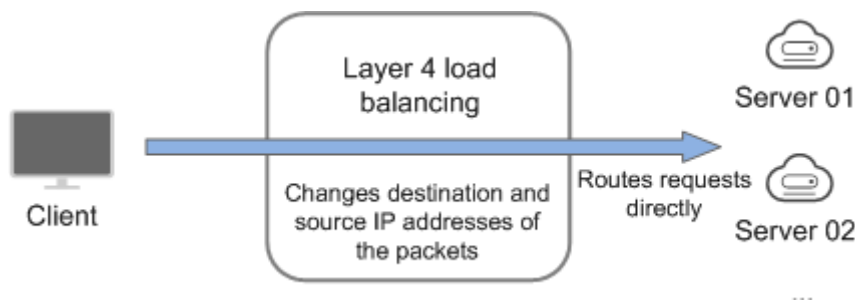
- For TCP or UDP load balancing, pay attention to the number of concurrent persistent connections, and consider Maximum Concurrent Connections as a key metric. Estimate the maximum number of concurrent connections that a load balancer can handle in the actual service scenario and select the corresponding specification.
- For HTTP or HTTPS load balancing, focus more on queries per second (QPS), which determines the service throughput of an application system. Estimate the QPS that a load balancer can handle in the actual service scenario and select the corresponding specification.
- Use the monitoring data from Cloud Eye to analyze the peak traffic, trend and regularity of the traffic to select the specifications more accurately.

Protocol

ELB provides load balancing at both Layer 4 and Layer 7.

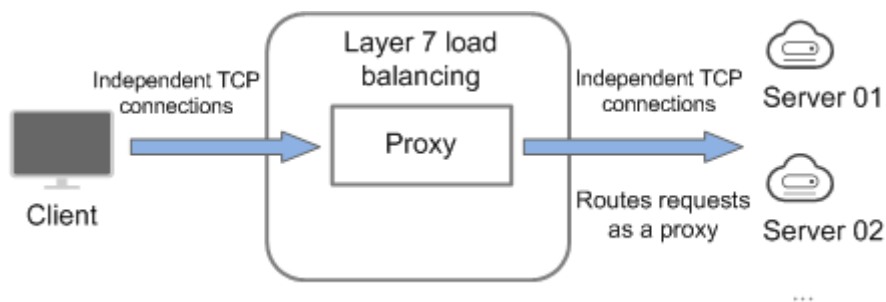
- If you choose TCP or UDP, the load balancer routes requests directly to backend servers. In this process, the destination IP address in the packets is changed to the IP address of the backend server, and the source IP address to the private IP address of the load balancer. A connection is established after a three-way handshake between the client and the backend server, and the load balancer only forwards the data.

Figure 1-3 Layer-4 load balancing



- Load balancing at Layer 7 is also called "content exchange". After the load balancer receives a request, it works as a proxy of backend servers to establish a connection (three-way handshake) with the client and then determines to which backend server the request is to be routed based on the fields in the HTTP/HTTPS request header and the load balancing algorithm you selected when you add the listener.

Figure 1-4 Layer-7 load balancing



Backend Servers

Before you use ELB, you need to create cloud servers, deploy required applications on them, and add the cloud servers to one or more backend server groups. When you create ECSs or BMSs, note the following:

- Cloud servers must be in the same region as the load balancer.
- Cloud servers that run the same OS are recommended so that you can manage them more easily.

1.3 Creating a Dedicated Load Balancer

Scenarios

You have prepared everything required for creating a load balancer. For details, see [Preparations for Creating a Load Balancer](#).

Constraints

- After a load balancer is created, the VPC cannot be changed. If you want to change the VPC, create a load balancer and select a different VPC.
- To ping the IP address of a load balancer, you need to add a listener to it.

Procedure



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, click Buy Elastic Load Balancer. Complete the basic configurations based on [Table 1-1](#).

Table 1-1 Parameters for configuring the basic information

Parameter	Description	Example Value
Type	Specifies the type of the load balancer. The type cannot be changed after the load balancer is created. For details about the differences, see Differences Between Dedicated and Shared Load Balancers .	Dedicated
Billing Mode	Specifies the billing mode of the dedicated load balancer. You are charged for how long you use each load balancer.	Pay-per-use

Parameter	Description	Example Value
Region	Specifies the desired region. Resources in different regions cannot communicate with each other over internal networks. For lower network latency and faster access to resources, select the nearest region.	-
AZ	<p>Specifies the AZ of the load balancer. You can deploy a load balancer in multiple AZs for high availability. If an AZ becomes faulty or unavailable, the load balancers in other AZs can route requests to backend servers to ensure service continuity and improve application reliability.</p> <p>If you deploy a load balancer in multiple AZs, its performance such as the number of new connections and the number of concurrent connections will multiply. For example, if you deploy a dedicated load balancer in two AZs, it can handle up to 40 million concurrent connections.</p> <p>For details about AZ planning, see AZ.</p>	-
Specifications	<p>Select Elastic or Fixed if pay-per-use is chosen as the billing mode.</p> <ul style="list-style-type: none">• Elastic specifications work well for fluctuating traffic, and you will be charged for how many LCUs you use.• Fixed specifications are suitable for stable traffic, and you will be charged for the specifications you select. <p>Select either Application load balancing (HTTP/HTTPS) or Network load balancing (TCP/UDP) or both, and then select the desired specification. You can select only one specification for Application load balancing (HTTP/HTTPS) and Network load balancing (TCP/UDP), respectively. Select the desired specifications based on your service size by referring to Specifications of Dedicated Load Balancers.</p> <p>NOTE For details about the regions where elastic specification is available, see Function Overview.</p>	Medium II

Parameter	Description	Example Value
Name	Specifies the load balancer name.	elb-test
Enterprise Project	Specifies an enterprise project by which cloud resources and members are centrally managed.	default
Description	Provides supplementary information about the load balancer.	-
Tag	<p>Identifies load balancers so that they can be easily found. A tag consists of a tag key and a tag value. The tag key marks a tag, and the tag value specifies specific tag content. For details about the naming specifications, see Table 1-2.</p> <p>A maximum of 10 tags can be added.</p> <p>NOTE If your organization has configured tag policies for ELB, add tags to load balancers based on the tag policies. If you add a tag that does not comply with the tag policies, load balancers may fail to be created. Contact your organization administrator to learn more about tag policies.</p>	<ul style="list-style-type: none">• Key: elb_key1• Value: elb-01

Table 1-2 Tag naming rules

Item	Requirement	Example Value
Tag key	<ul style="list-style-type: none">• Cannot be empty.• Must be unique for the same load balancer.• Can contain a maximum of 36 characters.• Only letters, digits, underscores (_), hyphens (-), at signs (@), and Chinese characters are allowed.	elb_key1
Tag value	<ul style="list-style-type: none">• Can contain a maximum of 43 characters.• Only letters, digits, underscores (_), hyphens (-), at signs (@), and Chinese characters are allowed.	elb-01

5. Configure the network parameters based on [Table 1-3](#).

Table 1-3 Parameters for network configurations

Parameter	Description	Example Value
IP as a Backend	<p>Specifies whether to associate backend servers that are not in the VPC of the load balancer. After this function is enabled, you can associate the backend servers with the load balancer by using their IP addresses.</p> <p>NOTE</p> <ul style="list-style-type: none">To use this function, configure correct VPC routes to ensure requests can be routed to backend servers.If you enable this function, more IP addresses in the subnet will be occupied. Ensure that the selected subnet has sufficient IP addresses. After you select a subnet, you can view the number of IP addresses required by the load balancer in the <code>infotip</code>.	-
Network Type	<p>Specifies the network where the load balancer works. You can select one or more network types.</p> <ul style="list-style-type: none">Public IPv4 network: The load balancer routes requests from the clients to backend servers over the Internet.Private IPv4 network: The load balancer routes requests from the clients to backend servers in a VPC.IPv6 network: An IPv6 address will be assigned to the load balancer to route requests from IPv6 clients. <p>NOTE</p> <p>If you do not select any of the options, the load balancer cannot communicate with the clients after it is created. When you are using ELB or testing network connectivity, ensure that the load balancer has a public or private IP address bound.</p>	Public IPv4 network
VPC	<p>Specifies the VPC where the load balancer works.</p> <p>Select an existing VPC or create a new one.</p> <p>For more information about VPC, see the Virtual Private Cloud User Guide.</p>	vpc-test

Parameter	Description	Example Value
Frontend Subnet	<p>Specifies the subnet where the load balancer will work.</p> <p>The system assigns IP addresses to load balancers for receiving requests based on the configured network type.</p> <ul style="list-style-type: none">• IPv4 private network: assigns IPv4 private addresses.• IPv6 network: assigns IPv6 private or public addresses. <p>NOTE If you select IPv6 network for Network Type and the selected VPC does not have any subnet that supports IPv6, enable IPv6 for the subnets or create a subnet that supports IPv6. For details, see the Virtual Private Cloud User Guide.</p>	subnet-test
Backend Subnet	<p>The load balancer uses the IP addresses in the backend subnet to forward requests to the backend servers.</p> <ul style="list-style-type: none">• Select Subnet of the load balancer by default.• Select an existing subnet in the VPC where the load balancer works.• Add a new subnet <p>NOTE</p> <ul style="list-style-type: none">• The number of IP addresses required depend on the specifications, number of AZs, and IP as a backend function you have configured when you create the load balancer. The actual number of occupied IP addresses depends on that displayed on the console.• An application load balancer requires 8 to 30 additional IP addresses in the backend subnet for traffic forwarding. The actual number of required IP addresses depends on the ELB cluster size. If load balancers are deployed in the same cluster and work in the same backend subnet, they share the same IP addresses to save resources.	Subnet of the load balancer
Private IPv4 network configuration		

Parameter	Description	Example Value
IPv4 Address	<p>Specifies how you want the IPv4 address to be assigned.</p> <ul style="list-style-type: none">• Automatically assign IP address: The system automatically assigns an IPv4 address to the load balancer.• Manually specify IP address: Manually specify an IPv4 address to the load balancer. <p>NOTE Network ACL rules configured for the backend subnet of the load balancer will not restrict the traffic from the clients to the load balancer. If these rules are configured, the clients can directly access the load balancer. To control access to the load balancer, configure access control for all listeners added to the load balancer.</p> <p>For details, see Access Control.</p>	Automatically assign IP address
IPv6 network configuration		
IPv6 Address	<p>Specifies how you want the IPv6 address to be assigned.</p> <p>NOTE Network ACL rules configured for the backend subnet of the load balancer will not restrict the traffic from the clients to the load balancer. If network ACL rules are configured, the clients can directly access the load balancer. To control access to the load balancer, configure access control for all listeners added to the load balancer.</p> <p>For details, see Access Control.</p>	Automatically assign IP address
Shared Bandwidth	<p>Specifies the shared bandwidth that the IPv6 address will be added to.</p> <p>You can choose not to select a shared bandwidth, select an existing shared bandwidth, or assign a shared bandwidth.</p>	Skip
Public IPv4 network configuration		
EIP	<p>This parameter is mandatory when Network Type is set to IPv4 public network.</p> <ul style="list-style-type: none">• New EIP: The system will assign a new EIP to the load balancer.• Use existing: Select an existing IP address.	-

Parameter	Description	Example Value
EIP Type	<p>Specifies the link type (BGP) when a new EIP is used.</p> <ul style="list-style-type: none">• Static BGP: When changes occur on a network using static BGP, carriers cannot adjust network configurations in real time to ensure optimal user experience.• Dynamic BGP: When changes occur on a network using dynamic BGP, routing protocols provide automatic, real-time optimization of network configurations, ensuring network stability and optimal user experience.	Dynamic BGP
Billed By	<p>Specifies how the bandwidth will be billed.</p> <p>You can select Bandwidth, Traffic, or Shared Bandwidth.</p> <ul style="list-style-type: none">• Bandwidth: You specify the maximum bandwidth and pay for the amount of time you use the bandwidth.• Traffic: You specify a maximum bandwidth and pay for the outbound traffic you use.• Shared Bandwidth	Shared Bandwidth
Bandwidth	Specifies the maximum bandwidth.	100 Mbit/s

6. Click **Next**.
7. Confirm the configuration and submit your request.

Popular Questions

- Can ELB Work in a Different AZ from Backend Servers?
Yes. ELB can route requests to backend servers in an AZ that is different from where the load balancer is deployed. To reduce network latency and improve access speed, you are suggested to deploy your load balancer in the AZ where backend servers are running.
- Can I Change the Specifications of an Existing Load Balancer?
Yes. For details, see [Changing the Specifications of a Dedicated Load Balancer](#).

1.4 Creating a Shared Load Balancer

Scenarios

You have prepared everything required for creating a load balancer. For details, see [Preparations for Creating a Load Balancer](#).

Load balancers receive requests from clients and route them to backend servers, which answer to these requests over the private network.

Constraints

- After a load balancer is created, the VPC cannot be changed. If you want to change the VPC, create a load balancer and select a different VPC.
- To ping the IP address of a load balancer, you need to add a listener and associate a backend server to it.

Procedure



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, click **Buy Elastic Load Balancer**. Configure the parameters based on [Table 1-4](#).

Table 1-4 Parameters for configuring the basic information

Parameter	Description	Example Value
Type	Specifies the type of the load balancer. The type cannot be changed after the load balancer is created. For details about the differences, see Differences Between Dedicated and Shared Load Balancers .	Shared
Billing Mode	Specifies the billing mode of the shared load balancer. You are charged for how long you use each load balancer.	Pay-per-use
Region	Specifies the desired region. Resources in different regions cannot communicate with each other over internal networks. For lower network latency and faster access to resources, select the nearest region.	-

Parameter	Description	Example Value
Name	Specifies the load balancer name.	elb-test
Enterprise Project	Specifies an enterprise project by which cloud resources and members are centrally managed.	default
Description	Provides supplementary information about the load balancer.	-
Tag	<p>Identifies load balancers so that they can be easily found. A tag consists of a tag key and a tag value. The tag key marks a tag, and the tag value specifies specific tag content. For details about the naming specifications, see Table 1-5.</p> <p>A maximum of 10 tags can be added.</p> <p>NOTE If your organization has configured tag policies for ELB, add tags to load balancers based on the tag policies. If you add a tag that does not comply with the tag policies, load balancers may fail to be created. Contact your organization administrator to learn more about tag policies.</p>	<ul style="list-style-type: none">• Key: elb_key1• Value: elb-01

Table 1-5 Tag naming rules

Item	Requirement	Example Value
Tag key	<ul style="list-style-type: none">• Cannot be empty.• Must be unique for the same load balancer.• Can contain a maximum of 36 characters.• Only letters, digits, underscores (_), hyphens (-), at signs (@), and Chinese characters are allowed.	elb_key1
Tag value	<ul style="list-style-type: none">• Can contain a maximum of 43 characters.• Only letters, digits, underscores (_), hyphens (-), at signs (@), and Chinese characters are allowed.	elb-01

5. Configure the network parameters based on [Table 1-6](#).

Table 1-6 Parameters for network configurations

Parameter	Description	Example Value
Network Type	<p>Specifies the network type of a load balancer. You can select either of the following:</p> <ul style="list-style-type: none">• Public network: The load balancer routes requests from the clients to backend servers over the Internet.• Private network: A private network load balancer routes requests from the clients to backend servers in the same VPC.	Public IPv4 network
VPC	<p>Specifies the VPC where the load balancer will work.</p> <p>Select an existing VPC or create a new one.</p> <p>For more information about VPC, see the Virtual Private Cloud User Guide.</p>	-
Frontend Subnet	<p>Specifies the subnet where the load balancer will work. Shared load balancers support private IPv4 network by default.</p> <p>The system assigns IPv4 private addresses in this subnet to load balancers.</p>	-
IPv4 Address	<p>Specifies how you want the IPv4 address to be assigned.</p> <ul style="list-style-type: none">• Automatically assign IP address: The system automatically assigns an IPv4 address to the load balancer.• Manually specify IP address: Manually specify an IPv4 address to the load balancer. <p>NOTE</p> <p>Network ACL rules configured for the backend subnet of the load balancer will not restrict the traffic from the clients to the load balancer. If these rules are configured, the clients can directly access the load balancer. To control access to a load balancer, configure access control for all listeners added to the load balancer.</p> <p>For details, see Access Control.</p>	Automatically assign IP address

Parameter	Description	Example Value
Guaranteed Performance	Specifies whether to enable the guaranteed performance option. This function allows your load balancers to handle up to 50,000 concurrent connections, 5,000 new connections per second, and 5,000 queries per second.	-
EIP	Specifies the public IP address that will be bound to the load balancer for receiving and forwarding requests over the Internet. You can use an existing EIP or apply for a new one. The following options are available: <ul style="list-style-type: none">• New EIP: The system will automatically assign an EIP.• Use existing: Select an existing EIP.	New EIP
EIP Type	Specifies the link type (BGP) when a new EIP is used. <ul style="list-style-type: none">• Static BGP: When changes occur on a network using static BGP, carriers cannot adjust network configurations in real time to ensure optimal user experience.• Dynamic BGP: When changes occur on a network using dynamic BGP, routing protocols provide automatic, real-time optimization of network configurations, ensuring network stability and optimal user experience.	Dynamic BGP
Billed By	Specifies how the bandwidth will be billed. You can select Bandwidth , Traffic , or Shared Bandwidth . <ul style="list-style-type: none">• Bandwidth: You specify the maximum bandwidth and pay for the amount of time you use the bandwidth.• Traffic: You specify a maximum bandwidth and pay for the total traffic you use.	Traffic
Bandwidth	Specifies the maximum bandwidth when a new EIP is used, in Mbit/s.	10 Mbit/s

6. Click **Next**.

7. Confirm the configuration and submit your request.

1.5 Enabling Guaranteed Performance for a Shared Load Balancer

Scenarios

Guaranteed performance allows shared load balancers to handle up to 50,000 concurrent connections, 5,000 new connections per second, and 5,000 queries per second. It provides you with more stable and reliable load balancing capabilities in case of traffic surge.

If your shared load balancers were created after February 10, 2023, guaranteed performance were enabled for them by default.

If your shared load balancers were created before created before February 10, 2023, perform the following operations to enable guaranteed performance.

Notes

- Guaranteed performance cannot be disabled once enabled.
- After guaranteed performance is enabled, shared load balancers will be charged on pay-per-use basis. For details about product prices, see [Product Pricing Details](#).

Enabling Guaranteed Performance



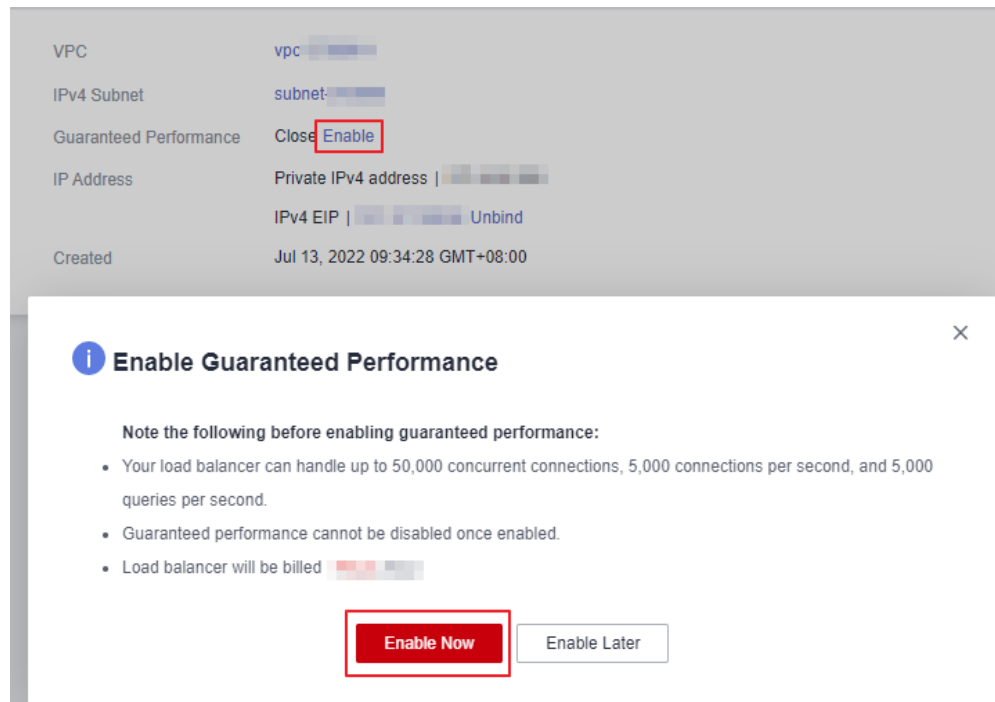
1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Click the target shared load balancer to enter the **Summary** page.
5. Click **Enable**.
6. Click **Enable Now** to enable guaranteed performance.



Figure 1-5 Enabling guaranteed performance

1.6 Configuring Modification Protection for Load Balancers

Scenario

You can enable modification protection for load balancers to prevent them from being modified or deleted by accident.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. On the **Summary** tab page, click **Configure** next to **Modification Protection**.
6. In the **Configure Modification Protection** dialog box, enable **Modification Protection**.
Fill in the reason if needed.
7. Click **OK**.

NOTE

Disable **Modification Protection** if you want to modify or delete a load balancer.

1.7 Modifying the Bandwidth

Scenario



If you set the **Network Type** of a load balancer to **Public IPv4 network** or **IPv6 network**, the load balancer can route requests over the Internet and you can modify the bandwidth used by the EIP bound to the load balancer as required.

NOTE

- When changing bandwidth, you need to change the specifications of the dedicated load balancer to avoid speed limit due to insufficient bandwidth.
- The bandwidth of the EIP bound to the load balancer is the limit for traffic required by the clients to access the load balancer.

Modifying the Bandwidth

When you modify the bandwidth, traffic routing will not be interrupted.

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer and click **More** in the **Operation** column.
5. Dedicated load balancers: Click **Modify IPv4 Bandwidth** or **Modify IPv6 Bandwidth**.
Shared load balancers: Click **Modify IPv4 Bandwidth**.
6. In the **New Configuration** area, modify the billing option and bandwidth and click **Next**.
You can select the bandwidth defined by the system or customize the bandwidth. The bandwidth ranges from 1 Mbit/s to 2,000 Mbit/s.
7. Confirm the modified bandwidth and click **Pay Now**.

NOTE

After you change the billing option and bandwidth, the price will be recalculated accordingly.

1.8 Changing the Specifications of a Dedicated Load Balancer

Scenario

You can change the specifications of a dedicated load balancer on the console:

- Change the elastic specifications to fixed specifications, or the other way round.
- Change an application load balancer to a network load balancer, or the other way round.
- Upgrade or downgrade the specifications, for example, upgrade small I to medium I, or downgrade large I to medium I.

 **NOTE**

- For details about the regions where elastic specification is available, see [Function Overview](#).
- You can only change the specifications of dedicated load balancers.

[Table 1-7](#) describe the supported specifications change options.

Table 1-7 Supported change options for a pay-per-use load balancer

Billing Mode	Specifications	Change to Elastic	Change to Fixed	Adding Load Balancing Type	Removing Load Balancing Type	Upgrading Specifications	Downgrading Specifications
Pay-per-use	Elastic	N/A	√	√	√	N/A	N/A
	Fixed	√	N/A	√	√	√	√


 **NOTE**


- Upgrading specifications does not interrupt your services.
- Downgrading specifications will temporarily interrupt services.
 - Network load balancing (TCP/UDP): New connections may not be able to be established.
 - Application load balancing (HTTP/HTTPS): New connections may not be able to be established and some persistent connections may be interrupted.

Constraints

- Retain at least one load balancing type: application or network.
- Before removing a load balancing type, you must delete the:
 - HTTP or HTTPS listeners added to an application load balancer.
 - TCP or UDP listeners added to a network load balancer.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.

3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer whose specifications you want to modify, click **More** in the **Operation** column, and select **Change Specifications**.
5. Select the new specifications and click **Next**.
6. Confirm the information and click **Submit**.

 **NOTE**

For yearly/monthly dedicated load balancers, confirm the order, select a payment method, and click **Confirm Payment**.

Popular Questions

- Can I Change the Type of a Load Balancer?
Yes, you can change an application load balancer to a network load balancer, or the other way round.
- Does Changing Specifications Affect Services?
Upgrading specifications does not affect your services, while downgrading specifications temporarily interrupt your services.

1.9 Changing the Billing Mode or Bandwidth Billing Option



Changing the Bandwidth Billing Option

Scenarios

For public network load balancers, you can change their billing options (billed by bandwidth or traffic) as required.

After you change the billing option, the price will be recalculated accordingly.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the **New Configuration** area, change the billing option and click **Next**.
5. Confirm the new billing option and click **Submit**.

1.10 Changing an IP Address

Scenarios

You can change the private IPv4 address and IPv6 address bound to a load balancer.

- You can change the private IPv4 address into another IPv4 IP address in the current subnet or other subnets.
- You can change the IPv6 address into another IPv6 IP address in other subnets.



NOTE

You can only change the IP address bound to a dedicated load balancer.



Constraints

To change the IPv6 address, ensure that the VPC where the load balancer works has subnets with IPv6 enabled.

Changing a Private IPv4 Address

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer whose IP address you want to change, and click **More > Change Private IPv4 Address** in the **Operation** column.
5. In the **Change Private IPv4 Address** dialog box, select the subnet where the IP address resides and specify the IP address.
 - To use an IP address from another subnet, select **Automatically assign IPv4 address**. The system automatically assigns an IPv4 address for your load balancer.
 - To use another IP address from the current subnet, specify an IP address.
6. Click **OK**.

Changing an IPv6 Address

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.

4. On the **Load Balancers** page, locate the load balancer whose IP address you want to change, and click **More > Change IPv6 Address** in the **Operation** column.
5. In the **Change IPv6 Address** dialog box, select a different subnet where the IP address resides and specify the IP address.
The system will automatically assign an IPv6 address to the load balancer from the subnet you select.
6. Click **OK**.

1.11 Binding an IP Address to or Unbinding an IP Address from a Load Balancer

Scenarios



You can bind an IP address to a load balancer or unbind the IP address from a load balancer based on service requirements.

- An IPv6 address, IPv4 EIP, and private IPv4 address can be bound to or unbound from a dedicated load balancer.
- Only an IPv4 EIP can be bound to or unbound from a shared load balancer.

NOTE

- Load balancers without IPv4 EIPs cannot route requests over the public IPv4 network.
- Load balancers without private IPv4 addresses cannot route requests over the private IPv4 network.
- After an IPv6 address is unbound, the load balancer cannot route requests over the IPv6 network.



Binding an IPv4 EIP

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer to which you want to bind an IPv4 EIP and click **More > Bind IPv4 EIP** in the **Operation** column.
5. In the **Bind IPv4 EIP** dialog box, select the EIP you want to bind to the load balancer.
6. Click **OK**.

Binding a Private IPv4 Address



Only dedicated load balancers support this function.

1. Log in to the management console.



2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer to which you want to bind a private IPv4 address and click **More > Bind Private IPv4 Address** in the **Operation** column.
5. In the **Bind Private IPv4 Address** dialog box, select the subnet where the IP address resides and specify the IP address.
 - By default, the system automatically assigns an IP address. To manually specify an IP address, deselect **Automatically assign IP address** and enter the IP address.
 - Ensure that the entered IP address belongs to the selected subnet and is not in use.
6. Click **OK**.

Binding an IPv6 Address

Only dedicated load balancers can have IPv6 addresses bound.

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer to which you want to bind an IPv6 address and click **More > Bind IPv6 Address** in the **Operation** column.
5. In the **Bind IPv6 Address** dialog box, select the subnet where the IP address resides.
6. Click **OK**.

Unbinding an IPv4 EIP



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer from which you want to unbind the IPv4 EIP and click **More > Unbind IPv4 EIP** in the **Operation** column.
5. In the displayed dialog box, confirm the IPv4 EIP that you want to unbind and click **Yes**.

 **NOTE**

After the IPv4 EIP is unbound, the load balancer cannot route requests over the Internet.

Unbinding a Private IPv4 Address

Only dedicated load balancers support this function.



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer from which you want to unbind the private IPv4 address and click **More > Unbind Private IPv4 Address** in the **Operation** column.
5. In the displayed dialog box, confirm the private IPv4 address that you want to unbind and click **Yes**.

 **NOTE**

After the private IPv4 address is unbound, the load balancer cannot route requests over the private IPv4 network.

Unbinding an IPv6 Address

Only dedicated load balancers can have IPv6 addresses bound.

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer from which you want to unbind the IPv6 address and click **More > Unbind IPv6 Address** in the **Operation** column.
5. In the displayed dialog box, confirm the IPv6 address that you want to unbind and click **Yes**.

 **NOTE**



After an IPv6 address is unbound, the load balancer cannot route requests over the IPv6 network.

1.12 Adding to or Removing from an IPv6 Shared Bandwidth



Scenarios

After you bind an IPv6 address to a dedicated load balancer, you can add the load balancer to a shared bandwidth to enable it to route requests over the Internet. After you are finished with the shared bandwidth, you can remove the load balancer from the shared bandwidth, so that it can only route requests within a VPC.

Adding to an IPv6 Shared Bandwidth

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer that you want to add to a shared bandwidth, and click **More > Add to IPv6 Shared Bandwidth** in the **Operation** column.
5. In the **Add to IPv6 Shared Bandwidth** dialog box, select the shared bandwidth to which you want to add the dedicated load balancer.
If no shared bandwidths are available, buy one as prompted.
6. Click **OK**.

Removing from an IPv6 Shared Bandwidth

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer that you want to remove from a shared bandwidth, and click **More > Remove from IPv6 Shared Bandwidth** in the **Operation** column.
5. In the displayed dialog box, confirm the shared bandwidth you want to remove.

NOTE

After the shared bandwidth is removed, the load balancer cannot route requests over the Internet.



6. Click **Yes**.

1.13 Exporting the Load Balancer List

Scenarios

You can export the load balancer list for backup.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the upper left corner of the load balancer list, click **Export**.

1.14 Deleting a Load Balancer

Scenarios

You can delete a load balancer if you do not need it any longer.

 **CAUTION**

A deleted load balancer cannot be recovered.

After a public network load balancer is deleted, its EIP will not be released and can be used by other resources.



Prerequisites

Delete the resources configured for the load balancer in the following sequence:

1. Delete all the forwarding policies added to HTTP and HTTPS listeners of the load balancer.
2. Delete the redirect created for each HTTP listener of the load balancer.
3. Remove all the backend servers from the backend server groups associated with each listener of the load balancer.
4. Delete all the listeners added to the load balancer.
5. Delete all backend server groups associated with each listener of the load balancer.

Deleting a Load Balancer

1. Log in to the management console.

2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the target load balancer and choose **More > Delete** in the **Operation** column.
A confirmation dialog box is displayed. Select **Release the EIP** as required.
5. Click **Yes**.

2 Listener

2.1 Overview

You need to add at least one listener after you have created a load balancer. This listener receives requests from clients and routes requests to backend servers using the protocol, port, and load balancing algorithm you select.

Supported Protocols

ELB provides load balancing at both Layer 4 and Layer 7.

Select TCP or UDP for load balancing at Layer 4 and HTTP or HTTPS at Layer 7.

Table 2-1 Protocols supported by ELB

Protocol		Description	Application Scenario
Layer 4	TCP	<ul style="list-style-type: none">• Source IP address-based sticky sessions• Fast data transfer	<ul style="list-style-type: none">• Scenarios that require high reliability and data accuracy, such as file transfer, email, and remote login• Web applications that receive a large number of concurrent requests and require high performance
Layer 4	UDP	<ul style="list-style-type: none">• Low reliability• Fast data transfer	Scenarios that require quick response, such as video chat, gaming, and real-time financial quotations
Layer 7	HTTP	<ul style="list-style-type: none">• Cookie-based sticky sessions• X-Forward-For request header	Web applications where data content needs to be identified, such as mobile games

Protocol		Description	Application Scenario
Layer 7	HTTPS	<ul style="list-style-type: none">• An extension of HTTP for encrypted data transmission to prevent unauthorized access• Encryption and decryption performed on load balancers• Multiple versions of encryption protocols and cipher suites	Web applications that require encrypted transmission

2.2 Protocols and Ports

Frontend Protocols and Ports

Frontend protocols and ports are used by load balancers to receive requests from clients. Load balancers use TCP or UDP at Layer 4, and HTTP or HTTPS at Layer 7. Select a protocol and a port that best suit your requirements.

 **NOTE**

The selected frontend protocols and entered ports cannot be changed. If you want to change them, create another listener.

Table 2-2 Frontend protocols and ports

Protocol	Port
TCP	There are some restrictions when you select the protocols and port numbers. <ul style="list-style-type: none">For each load balancer, UDP can use the same ports as other protocols, but these other protocols must have unique ports. For example, if you have a UDP listener that uses port 88, you can add a TCP, HTTP, or HTTPS listener that also uses port 88. However, if you already have an HTTP listener that uses port 443, you cannot add an HTTPS or TCP listener that uses the same port.The port numbers of the same protocol must be unique. For example, if you have a TCP listener that uses port 80, you cannot add another TCP listener that uses the same port. The port number ranges from 1 to 65535. The following are some commonly-used protocols and port numbers: TCP/80 HTTPS/443
UDP	
HTTP	
HTTPS	

Backend Protocols and Ports

Backend protocols and ports are used by backend servers to receive requests from load balancers. If Windows servers have Internet Information Services (IIS) installed, the default backend protocol and port are HTTP and 80.

Table 2-3 Backend protocols and ports

Protocol	Port
TCP	Backend servers can use the same ports. The port number ranges from 1 to 65535. The following are some commonly-used protocols and port numbers: TCP/80 HTTP/80 HTTPS/443
UDP	
QUIC	
HTTP	
HTTPS	

2.3 Adding a TCP Listener

Scenarios

You can add a TCP listener, if high reliability and high accuracy are required but slow speed is acceptable, for example, during file transfer, email sending and receiving, and remote login.

Constraints

- If the listener protocol is TCP, the protocol of the backend server group is TCP by default and cannot be changed.
- If you only select application load balancing (HTTP/HTTPS) for your dedicated load balancer, you cannot add a TCP listener to this load balancer.

Adding a TCP Listener to a Dedicated Load Balancer



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Under **Listeners**, click **Add Listener**. Configure the parameters based on [Table 2-4](#).

Table 2-4 Parameters for configuring a listener

Parameter	Description	Example Value
Name	Specifies the listener name.	listener-pnqy
Frontend Protocol	Specifies the protocol that will be used by the load balancer to receive requests from clients.	TCP
Frontend Port	Specifies the port that will be used by the load balancer to receive requests from clients. The port number ranges from 1 to 65535.	80

Parameter	Description	Example Value
Access Control	Specifies how access to the listener is controlled. For details, see Access Control . The following options are available: <ul style="list-style-type: none">• All IP addresses• Blacklist• Whitelist	Blacklist
IP Address Group	Specifies the IP address group associated with a whitelist or blacklist. If there is no IP address group, create one first. For more information, see Creating an IP Address Group .	ipGroup-b2
Transfer Client IP Address	Specifies whether to transmit IP addresses of the clients to backend servers. This function is enabled for dedicated load balancers by default and cannot be disabled.	N/A
Advanced Settings		
Idle Timeout	Specifies the length of time for a connection to keep alive, in seconds. If no request is received within this period, the load balancer closes the connection and establishes a new one with the client when the next request arrives. The idle timeout duration ranges from 10 to 4000 .	300
Description	Provides supplementary information about the listener. You can enter a maximum of 255 characters.	N/A

6. Click **Next: Configure Request Routing Policy** to configure the backend server group. For details about how to configure a backend server group, see [Table 2-5](#).

Table 2-5 Parameters for configuring a backend server group


Parameter	Description	Example Value
Backend Server Group	Specifies a group of servers with the same features to receive requests from the load balancer. Two options are available: <ul style="list-style-type: none">• Create new• Use existing NOTE The backend protocol of the backend server group must match the frontend protocol. For example, if the frontend protocol is TCP, the backend protocol must be TCP.	Create new
Backend Server Group Name	Specifies the name of the backend server group.	server_group
Backend Protocol	Specifies the protocol used by backend servers to receive requests. The backend protocol is TCP by default and cannot be changed.	TCP

Parameter	Description	Example Value
Load Balancing Algorithm	<p>Specifies the algorithm used by the load balancer to distribute traffic. The following options are available:</p> <ul style="list-style-type: none">● Weighted round robin: Requests are routed to different servers based on their weights, which indicate server processing performance. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests.● Weighted least connections: In addition to the number of active connections established with each backend server, each server is assigned a weight based on their processing capability. Requests are routed to the server with the lowest connections-to-weight ratio.● Source IP hash: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all backend servers are numbered. The generated key is used to allocate the client to a particular server. This allows requests from different clients to be routed based on source IP addresses and ensures that a client is directed to the same server that it was using previously. <p>NOTE</p> <ul style="list-style-type: none">● Choose an appropriate algorithm based on your requirements for better traffic distribution.● For Weighted round robin or Weighted least connections, no requests will be routed to a server with a weight of 0.	Weighted round robin

Parameter	Description	Example Value
Sticky Session	<p>Specifies whether to enable sticky sessions. If you enable sticky sessions, all requests from a client during one session are sent to the same backend server.</p> <p>This parameter is optional and can be enabled if you have selected Weighted round robin or Weighted least connections for Load Balancing Algorithm.</p>	N/A
Sticky Session Type	<p>Specifies the type of sticky sessions.</p> <p>Source IP address is the only choice available when TCP or UDP is used as the frontend protocol.</p> <p>Source IP address: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all the endpoints are numbered. The system allocates the client to a particular endpoint based on the generated key. Requests from the same IP address are forwarded to the same backend server for processing.</p>	Source IP address
Stickiness Duration (min)	<p>Specifies the minutes that sticky sessions are maintained. You can enable sticky sessions only if you select Weighted round robin or Weighted least connections for Load Balancing Algorithm.</p> <ul style="list-style-type: none">• Stickiness duration at Layer 4: 1 to 60• Stickiness duration at Layer 7: 1 to 1440	20
Description	<p>Provides supplementary information about the backend server group.</p> <p>You can enter a maximum of 255 characters.</p>	N/A

7. Click **Next: Add Backend Server**. Add backend servers and configure the health check for the backend server group. For details about how to add backend servers, see [Overview](#). For the parameters required for configuring a health check, see [Table 2-6](#).

Table 2-6 Parameters for configuring a health check

Parameter	Description	Example Value
Health Check	Specifies whether to enable health checks. If the health check is enabled, click  next to Advanced Settings to set health check parameters.	N/A
Health Check Protocol	Specifies the protocol that will be used by the load balancer to check the health of backend servers. If the backend protocol is TCP, the health check protocol can be TCP, HTTP, or HTTPS.	HTTP
Domain Name	Specifies the domain name that will be used for health checks. This parameter is mandatory if the health check protocol is HTTP or HTTPS. <ul style="list-style-type: none">You can use the private IP address of the backend server as the domain name.You can also specify a domain name that consists of at least two labels separated by periods (.). Use only letters, digits, and hyphens (-). Do not start or end strings with a hyphen. Max total: 100 characters. Max label: 63 characters.	www.elb.com
Health Check Port	Specifies the port that will be used by the load balancer to check the health of backend servers. The port number ranges from 1 to 65535. NOTE By default, the service port on each backend server is used. You can also specify a port for health checks.	80

Parameter	Description	Example Value
Path	Specifies the health check URL, which is the destination on backend servers for health checks. This parameter is mandatory if the health check protocol is HTTP or HTTPS. The path can contain 1 to 80 characters and must start with a slash (/). The path can contain letters, digits, hyphens (-), slashes (/), periods (.), question marks (?), number signs (#), percent signs (%), ampersands (&), and extended character sets _;~!. () *[]@\$^:'!,+	/index.html
Interval (s)	Specifies the interval for sending health check requests, in seconds. The interval ranges from 1 to 50.	5
Timeout (s)	Specifies the maximum time required for waiting for a response from the health check, in seconds. The timeout duration ranges from 1 to 50.	3
Maximum Retries	Specifies the maximum number of health check retries. The value ranges from 1 to 10.	3

8. Click **Next: Confirm**.
9. Confirm the configuration and click **Submit**.

Adding a TCP Listener to a Shared Load Balancer



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Under **Listeners**, click **Add Listener**. Configure the parameters based on [Table 2-7](#).

Table 2-7 Parameters for configuring a listener

Parameter	Description	Example Value
Name	Specifies the listener name.	listener-pnqy
Frontend Protocol	Specifies the protocol that will be used by the load balancer to receive requests from clients.	TCP
Frontend Port	Specifies the port that will be used by the load balancer to receive requests from clients. The port number ranges from 1 to 65535.	80
Access Control	Specifies how access to the listener is controlled. For details, see Access Control . The following options are available: <ul style="list-style-type: none">• All IP addresses• Blacklist• Whitelist	Whitelist
IP Address Group	Specifies the IP address group associated with a whitelist or blacklist. If there is no IP address group, create one first. For more information, see Creating an IP Address Group .	ipGroup-b2
Transfer Client IP Address	Specifies whether to transmit IP addresses of the clients to backend servers. This parameter is available when the listener protocol is TCP or UDP.	N/A
Advanced Settings		
Idle Timeout	Specifies the length of time for a connection to keep alive, in seconds. If no request is received within this period, the load balancer closes the connection and establishes a new one with the client when the next request arrives. The idle timeout duration ranges from 10 to 4000 .	300
Description	Provides supplementary information about the listener. You can enter a maximum of 255 characters.	N/A

6. Click **Next: Configure Request Routing Policy**. [Table 2-8](#) describes the parameters for configuring a backend server group.

Table 2-8 Parameters for configuring a backend server group


Parameter	Description	Example Value
Backend Server Group	Specifies a group of servers with the same features to receive requests from the load balancer. Two options are available: <ul style="list-style-type: none">• Create new• Use existing NOTE The backend protocol of the backend server group must match the frontend protocol. For example, if the frontend protocol is TCP, the backend protocol must be TCP.	Create new
Backend Server Group Name	Specifies the name of the backend server group.	server_group-sq4v
Backend Protocol	Specifies the protocol used by backend servers to receive requests. The backend protocol is TCP by default and cannot be changed.	TCP

Parameter	Description	Example Value
Load Balancing Algorithm	<p>Specifies the algorithm used by the load balancer to distribute traffic. The following options are available:</p> <ul style="list-style-type: none">• Weighted round robin: Requests are routed to different servers based on their weights, which indicate server processing performance. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests.• Weighted least connections: This algorithm is designed based on the least connections algorithm that uses the number of active connections to each backend server to make its load balancing decision. In addition to the number of connections, each server is assigned a weight based on its capacity. Requests are routed to the server with the lowest connections-to-weight ratio.• Source IP hash: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all backend servers are numbered. The generated key allocates the client to a particular server. This allows requests from different clients to be routed based on source IP addresses and ensures that a client is directed to the same server that it was using previously. <p>NOTE</p> <ul style="list-style-type: none">• Choose an appropriate algorithm based on your requirements for better traffic distribution.• For Weighted round robin or Weighted least connections, no requests will be routed to a server with a weight of 0.	Weighted round robin

Parameter	Description	Example Value
Sticky Session	<p>Specifies whether to enable sticky sessions. If you enable sticky sessions, all requests from a client during one session are sent to the same backend server.</p> <p>NOTE You can enable sticky sessions only if you have selected Weighted round robin or Weighted least connections for Load Balancing Algorithm.</p>	N/A
Sticky Session Type	<p>Specifies the type of sticky sessions. Source IP address is the only choice available when TCP or UDP is used as the frontend protocol.</p> <p>Source IP address: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all the endpoints are numbered. The system allocates the client to a particular endpoint based on the generated key. Requests from the same IP address are forwarded to the same backend server for processing.</p>	Source IP address
Stickiness Duration (min)	<p>Specifies the minutes that sticky sessions are maintained. You can enable sticky sessions only if you select Weighted round robin or Weighted least connections for Load Balancing Algorithm.</p> <ul style="list-style-type: none">• Stickiness duration at Layer 4: 1 to 60• Stickiness duration at Layer 7: 1 to 1440	20
Description	<p>Provides supplementary information about the backend server group.</p> <p>You can enter a maximum of 255 characters.</p>	N/A

7. Click **Next: Add Backend Server**. Add backend servers and configure the health check for the backend server group. For details about how to add backend servers, see [Overview](#). For the parameters required for configuring a health check, see [Table 2-9](#).

Table 2-9 Parameters for configuring a health check

Parameter	Description	Example Value
Health Check	Specifies whether to enable health checks. If the health check is enabled, click  next to Advanced Settings to set health check parameters.	N/A
Health Check Protocol	Specifies the protocol that will be used by the load balancer to check the health of backend servers. There are two options: TCP and HTTP.	HTTP
Domain Name	Specifies the domain name that will be used for health checks. This parameter is mandatory if the health check protocol is HTTP. <ul style="list-style-type: none">You can use the private IP address of the backend server as the domain name.You can also specify a domain name that consists of at least two labels separated by periods (.). Use only letters, digits, and hyphens (-). Do not start or end strings with a hyphen. Max total: 100 characters. Max label: 63 characters.	www.elb.com
Health Check Port	Specifies the port that will be used by the load balancer to check the health of backend servers. The port number ranges from 1 to 65535. NOTE By default, the service port on each backend server is used. You can also specify a port for health checks.	80

Parameter	Description	Example Value
Path	Specifies the health check URL, which is the destination on backend servers for health checks. This parameter is mandatory if the health check protocol is HTTP. The path can contain 1 to 80 characters and must start with a slash (/). The path can contain letters, digits, hyphens (-), slashes (/), periods (.), question marks (?), percent signs (%), ampersands (&), and underscores (_).	/index.html
Interval (s)	Specifies the interval for sending health check requests, in seconds. The interval ranges from 1 to 50 .	5
Timeout (s)	Specifies the maximum time required for waiting for a response from the health check, in seconds. The timeout duration ranges from 1 to 50 .	3
Maximum Retries	Specifies the maximum number of health check retries. The value ranges from 1 to 10 .	3

8. Click **Next: Confirm**.
9. Confirm the configuration and click **Submit**.

2.4 Adding a UDP Listener

Scenarios

UDP listeners are suitable for scenarios that focus more on timeliness than reliability, such as video chat, gaming, and real-time quotation in the financial market.

Constraints

- UDP listeners do not support fragmentation.
- The port of UDP listeners cannot be 4789.
- UDP packets can have any size less than 1,500 bytes. The packets will be discarded if they are too big. You need to modify the configuration files of the applications based on the maximum transmission unit (MTU) value.
- Dedicated load balancers: The backend protocol can be UDP or QUIC if the listener protocol is UDP.

- Shared load balancers: If the listener protocol is UDP, the protocol of the backend server group is UDP by default and cannot be changed.
- If you only select application load balancing (HTTP/HTTPS) for your dedicated load balancer, you cannot add a UDP listener to this load balancer.

Adding a UDP Listener to a Dedicated Load Balancer



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Under **Listeners**, click **Add Listener**. Configure the parameters based on [Table 2-10](#).

Table 2-10 Parameters for configuring a listener

Parameter	Description	Example Value
Name	Specifies the listener name.	listener
Frontend Protocol	Specifies the protocol that will be used by the load balancer to receive requests from clients.	UDP
Frontend Port	Specifies the port that will be used by the load balancer to receive requests from clients. The port number ranges from 1 to 65535.	80
Access Control	Specifies how access to the listener is controlled. For details, see Access Control . The following options are available: <ul style="list-style-type: none">• All IP addresses• Blacklist• Whitelist	Blacklist
IP Address Group	Specifies the IP address group associated with a whitelist or blacklist. If there is no IP address group, create one first. For more information, see Creating an IP Address Group .	ipGroup

Parameter	Description	Example Value
Transfer Client IP Address	Specifies whether to transmit IP addresses of the clients to backend servers. This function is enabled for dedicated load balancers by default and cannot be disabled.	N/A
Advanced Settings		
Idle Timeout	Specifies the length of time for a connection to keep alive, in seconds. If no request is received within this period, the load balancer closes the connection and establishes a new one with the client when the next request arrives. The idle timeout duration ranges from 10 to 4000 .	300
Description	Provides supplementary information about the listener. You can enter a maximum of 255 characters.	N/A

6. Click **Next: Configure Request Routing Policy** to configure the backend server group. [Table 2-11](#) describes the parameters for configuring a backend server group.

Table 2-11 Parameters for configuring a backend server group

Parameter	Description	Example Value
Backend Server Group	Specifies a group of servers with the same features to receive requests from the load balancer. Two options are available: <ul style="list-style-type: none">• Create new• Use existing NOTE The backend protocol of the backend server group must match the frontend protocol. For example, if the frontend protocol is TCP, the backend protocol must be TCP.	Create new
Backend Server Group Name	Specifies the name of the backend server group.	server_group


Parameter	Description	Example Value
Backend Protocol	Specifies the protocol that will be used by backend servers to receive requests. The backend protocol can be UDP or QUIC.	UDP

Parameter	Description	Example Value
Load Balancing Algorithm	<p>Specifies the algorithm that will be used by the load balancer to distribute traffic. The following options are available:</p> <ul style="list-style-type: none">• Weighted round robin: Requests are routed to different servers based on their weights, which indicate server processing performance. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests.• Weighted least connections: In addition to the number of active connections established with each backend server, each server is assigned a weight based on their processing capability. Requests are routed to the server with the lowest connections-to-weight ratio.• Source IP hash: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all backend servers are numbered. The generated key is used to allocate the client to a particular server. This allows requests from different clients to be routed based on source IP addresses and ensures that a client is directed to the same server that it was using previously. <p>NOTE</p> <ul style="list-style-type: none">• Choose an appropriate algorithm based on your requirements for better traffic distribution.• For Weighted round robin or Weighted least connections, no requests will be routed to a server with a weight of 0.	Weighted round robin

Parameter	Description	Example Value
Sticky Session	<p>Specifies whether to enable sticky sessions. If you enable sticky sessions, all requests from a client during one session are sent to the same backend server.</p> <p>This parameter is optional and can be enabled if you have selected Weighted round robin or Weighted least connections for Load Balancing Algorithm.</p>	N/A
Sticky Session Type	<p>Specifies the type of sticky sessions. Source IP address is the only choice available when TCP or UDP is used as the frontend protocol.</p> <p>Source IP address: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all the endpoints are numbered. The system allocates the client to a particular endpoint based on the generated key. Requests from the same IP address are forwarded to the same backend server for processing.</p>	Source IP address
Stickiness Duration (min)	<p>Specifies the minutes that sticky sessions are maintained. You can enable sticky sessions only if you select Weighted round robin or Weighted least connections for Load Balancing Algorithm.</p> <ul style="list-style-type: none">• Stickiness duration at Layer 4: 1 to 60• Stickiness duration at Layer 7: 1 to 1440	20
Description	<p>Provides supplementary information about the backend server group.</p> <p>You can enter a maximum of 255 characters.</p>	N/A


7. Click **Next: Add Backend Server**. Add backend servers and configure the health check for the backend server group. For details about how to add backend servers, see [Overview](#). For the parameters required for configuring a health check, see [Table 2-12](#).

Table 2-12 Parameters for configuring a health check

Parameter	Description	Example Value
Health Check	Specifies whether to enable health checks. If the health check is enabled, click  next to Advanced Settings to set health check parameters.	N/A
Health Check Protocol	Specifies the protocol that will be used by the load balancer to check the health of backend servers. If the backend protocol is UDP, the health check protocol is UDP and cannot be changed.	UDP
Health Check Port	Specifies the port that will be used by the load balancer to check the health of backend servers. The port number ranges from 1 to 65535. NOTE By default, the service port on each backend server is used. You can also specify a port for health checks.	80
Interval (s)	Specifies the interval for sending health check requests, in seconds. The interval ranges from 1 to 50.	5
Timeout (s)	Specifies the maximum time required for waiting for a response from the health check, in seconds. The timeout duration ranges from 1 to 50.	3
Maximum Retries	Specifies the maximum number of health check retries. The value ranges from 1 to 10.	3

8. Click **Next: Confirm**.
9. Confirm the configuration and click **Submit**.

Adding a UDP Listener to a Shared Load Balancer

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.


3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Under **Listeners**, click **Add Listener**. Configure the parameters based on [Table 2-13](#).

Table 2-13 Parameters for configuring a listener

Parameter	Description	Example Value
Name	Specifies the listener name.	listener
Frontend Protocol	Specifies the protocol that will be used by the load balancer to receive requests from clients.	UDP
Frontend Port	Specifies the port that will be used by the load balancer to receive requests from clients. The port number ranges from 1 to 65535.	80
Access Control	Specifies how access to the listener is controlled. For details, see Access Control . The following options are available: <ul style="list-style-type: none">• All IP addresses• Blacklist• Whitelist	Whitelist
IP Address Group	Specifies the IP address group associated with a whitelist or blacklist. If there is no IP address group, create one first. For more information, see Creating an IP Address Group .	ipGroup
Transfer Client IP Address	Specifies whether to transmit IP addresses of the clients to backend servers.	N/A
Advanced Settings		
Description	Provides supplementary information about the listener. You can enter a maximum of 255 characters.	N/A

6. Click **Next: Configure Request Routing Policy**. [Table 2-14](#) describes the parameters for configuring a backend server group.

Table 2-14 Parameters for configuring a backend server group


Parameter	Description	Example Value
Backend Server Group	Specifies a group of servers with the same features to receive requests from the load balancer. Two options are available: <ul style="list-style-type: none">• Create new• Use existing NOTE The backend protocol of the backend server group must match the frontend protocol. For example, if the frontend protocol is TCP, the backend protocol must be TCP.	Create new
Backend Server Group Name	Specifies the name of the backend server group.	server_group
Backend Protocol	Specifies the protocol that will be used by backend servers to receive requests. The backend protocol is UDP by default and cannot be changed.	UDP

Parameter	Description	Example Value
Load Balancing Algorithm	<p>Specifies the algorithm used by the load balancer to distribute traffic. The following options are available:</p> <ul style="list-style-type: none">● Weighted round robin: Requests are routed to different servers based on their weights, which indicate server processing performance. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests.● Weighted least connections: In addition to the number of active connections established with each backend server, each server is assigned a weight based on their processing capability. Requests are routed to the server with the lowest connections-to-weight ratio.● Source IP hash: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all backend servers are numbered. The generated key is used to allocate the client to a particular server. This allows requests from different clients to be routed based on source IP addresses and ensures that a client is directed to the same server that it was using previously. <p>NOTE</p> <ul style="list-style-type: none">● Choose an appropriate algorithm based on your requirements for better traffic distribution.● For Weighted round robin or Weighted least connections, no requests will be routed to a server with a weight of 0.	Weighted round robin

Parameter	Description	Example Value
Sticky Session	<p>Specifies whether to enable sticky sessions. If you enable sticky sessions, all requests from a client during one session are sent to the same backend server.</p> <p>NOTE You can enable sticky sessions only if you have selected Weighted round robin or Weighted least connections for Load Balancing Algorithm.</p>	N/A
Sticky Session Type	<p>Specifies the type of sticky sessions. Source IP address is the only choice available when TCP or UDP is used as the frontend protocol.</p> <p>Source IP address: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all the endpoints are numbered. The system allocates the client to a particular endpoint based on the generated key. Requests from the same IP address are forwarded to the same backend server for processing.</p>	Source IP address
Stickiness Duration (min)	<p>Specifies the minutes that sticky sessions are maintained. You can enable sticky sessions only if you select Weighted round robin or Weighted least connections for Load Balancing Algorithm.</p> <ul style="list-style-type: none">• Stickiness duration at Layer 4: 1 to 60• Stickiness duration at Layer 7: 1 to 1440	20
Description	<p>Provides supplementary information about the backend server group.</p> <p>You can enter a maximum of 255 characters.</p>	N/A

7. Click **Next: Add Backend Server**. Add backend servers and configure health check for the backend server group. For details about how to add backend servers, see [Overview](#). For the parameters required for configuring a health check, see [Table 2-15](#).

Table 2-15 Parameters for configuring a health check

Parameter	Description	Example Value
Health Check	Specifies whether to enable health checks. If the health check is enabled, click  next to Advanced Settings to set health check parameters.	N/A
Health Check Protocol	Specifies the protocol that will be used by the load balancer to check the health of backend servers. The health check protocol is UDP by default and cannot be changed.	UDP
Health Check Port	Specifies the port that will be used by the load balancer to check the health of backend servers. The port number ranges from 1 to 65535. NOTE By default, the service port on each backend server is used. You can also specify a port for health checks.	80
Interval (s)	Specifies the interval for sending health check requests, in seconds. The interval ranges from 1 to 50 .	5
Timeout (s)	Specifies the maximum time required for waiting for a response from the health check, in seconds. The timeout duration ranges from 1 to 50 .	3
Maximum Retries	Specifies the maximum number of health check retries. The value ranges from 1 to 10 .	3

8. Click **Next: Confirm**.
9. Confirm the configuration and click **Submit**.

2.5 Adding an HTTP Listener

Scenarios

HTTP listeners are suitable for applications that require identifying the data content, such as web applications and small mobile games.

Constraints

- If the listener protocol is HTTP, the protocol of the backend server group is HTTP by default and cannot be changed.
- If you only select network load balancing (TCP/UDP) for your dedicated load balancer, you cannot add an HTTP listener to this load balancer.

Adding an HTTP Listener to a Dedicated Load Balancer



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Under **Listeners**, click **Add Listener**. Configure the parameters based on [Table 2-16](#).

Table 2-16 Parameters for configuring a listener

Parameter	Description	Example Value
Name	Specifies the listener name.	listener
Frontend Protocol	Specifies the protocol that will be used by the load balancer to receive requests from clients.	HTTP
Frontend Port	Specifies the port that will be used by the load balancer to receive requests from clients. The port number ranges from 1 to 65535.	80
Redirect	Specifies whether to enable redirection. If you have both HTTPS and HTTP listeners, you can use this function to redirect the requests from the HTTP listener to the HTTPS listener to ensure security.	N/A
Redirected To	Specifies the HTTPS listener to which requests are redirected if Redirect is enabled.	listener_HTTPS_443

Parameter	Description	Example Value
Access Control	Specifies how access to the listener is controlled. For details, see Access Control . The following options are available: <ul style="list-style-type: none">• All IP addresses• Blacklist• Whitelist	Blacklist
IP Address Group	Specifies the IP address group associated with a whitelist or blacklist. If there is no IP address group, create one first. For more information, see Creating an IP Address Group .	ipGroup
Transfer Client IP Address	Specifies whether to transmit IP addresses of the clients to backend servers. This function is enabled for dedicated load balancers by default and cannot be disabled.	Enabled
Advanced Forwarding	Specifies whether to enable the advanced forwarding policy. You can add advanced forwarding policies to HTTP or HTTPS listeners to forward requests to different backend server groups based on HTTP request method, HTTP header, query string, or CIDR block in addition to domain names and URLs.	Enabled
Advanced Settings		
Idle Timeout	Specifies the length of time for a connection to keep alive, in seconds. If no request is received within this period, the load balancer closes the connection and establishes a new one with the client when the next request arrives. The idle timeout duration ranges from 0 to 4000 .	60

Parameter	Description	Example Value
Request Timeout	Specifies the length of time (in seconds) after which the load balancer closes the connection if the load balancer does not receive a request from the client. The request timeout duration ranges from 1 to 300 .	60
Response Timeout	Specifies the length of time (in seconds) after which the load balancer sends a 504 Gateway Timeout error to the client if the load balancer receives no response from the backend server after routing a request to the backend server and receives no response after attempting to route the same request to other backend servers. The response timeout duration ranges from 1 to 300 . NOTE If you have enabled sticky sessions and the backend server does not respond within the response timeout duration, the load balancer returns 504 Gateway Timeout to the clients.	60
Description	Provides supplementary information about the listener. You can enter a maximum of 255 characters.	N/A

6. Click **Next: Configure Request Routing Policy**.
 - a. You are advised to select an existing backend server group.
 - b. You can also click **Create new** to create a backend server group and configure parameters as described in [Table 2-17](#).

Table 2-17 Parameters for configuring a backend server group

Parameter	Description	Example Value
Backend Server Group	<p>Specifies a group of servers with the same features to receive requests from the load balancer. Two options are available:</p> <ul style="list-style-type: none">• Create new• Use existing <p>NOTE The backend protocol of the backend server group must match the frontend protocol. For example, if the frontend protocol is TCP, the backend protocol must be TCP.</p>	Create new
Backend Server Group Name	Specifies the name of the backend server group.	server_group
Backend Protocol	<p>Specifies the protocol that will be used by backend servers to receive requests.</p> <p>The backend protocol is HTTP by default and cannot be changed.</p>	HTTP


Parameter	Description	Example Value
Load Balancing Algorithm	<p>Specifies the algorithm that will be used by the load balancer to distribute traffic. The following options are available:</p> <ul style="list-style-type: none">• Weighted round robin: Requests are routed to different servers based on their weights, which indicate server processing performance. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests.• Weighted least connections: In addition to the number of active connections established with each backend server, each server is assigned a weight based on their processing capability. Requests are routed to the server with the lowest connections-to-weight ratio.• Source IP hash: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all backend servers are numbered. The generated key is used to allocate the client to a particular server. This allows requests from different clients to be routed based on source IP addresses and ensures that a client is directed to the same server that it was using previously. <p>NOTE</p> <ul style="list-style-type: none">• Choose an appropriate algorithm based on your requirements for better traffic distribution.• For Weighted round robin or Weighted least connections, no requests will be routed to a server with a weight of 0.	Weighted round robin

Parameter	Description	Example Value
Sticky Session	Specifies whether to enable sticky sessions. If you enable sticky sessions, all requests from a client during one session are sent to the same backend server. This parameter is optional and can be enabled if you have selected Weighted round robin or Weighted least connections for Load Balancing Algorithm .	N/A
Sticky Session Type	Specifies the type of sticky sessions for HTTP and HTTPS listeners. <ul style="list-style-type: none">• Load balancer cookie: The load balancer generates a cookie after receiving a request from the client. All subsequent requests with the same cookie are then routed to the same backend server. NOTE	Load balancer cookie
Stickiness Duration (min)	Specifies the minutes that sticky sessions are maintained. You can enable sticky sessions only if you select Weighted round robin or Weighted least connections for Load Balancing Algorithm . <ul style="list-style-type: none">• Stickiness duration at Layer 4: 1 to 60• Stickiness duration at Layer 7: 1 to 1440	20
Slow Start	Specifies whether to enable slow start, which is disabled by default. After you enable slow start, the load balancer linearly increases the proportion of requests to send to backend servers in this mode. When the slow start duration elapses, the load balancer sends full share of requests to backend servers and exits the slow start mode. For details, see Slow Start (Dedicated Load Balancers) .	N/A

Parameter	Description	Example Value
Slow Start Duration	Specifies the slow start duration if Slow Start is enabled. The duration ranges from 30 to 1200 , in seconds, and the default value is 30 .	30
Description	Provides supplementary information about the backend server group. You can enter a maximum of 255 characters.	N/A

7. Click **Next: Add Backend Server**. Add backend servers and configure health check for the backend server group. For details about how to add backend servers, see [Overview](#). For the parameters required for configuring a health check, see [Table 2-18](#).

Table 2-18 Parameters for configuring a health check

Parameter	Description	Example Value
Health Check	Specifies whether to enable health checks. If the health check is enabled, click  next to Advanced Settings to set health check parameters.	N/A
Health Check Protocol	Specifies the protocol that will be used by the load balancer to check the health of backend servers. If the backend protocol is HTTP or HTTPS, the health check protocol can be TCP, HTTP, or HTTPS.	HTTP

Parameter	Description	Example Value
Domain Name	<p>Specifies the domain name that will be used for health checks. This parameter is mandatory if the health check protocol is HTTP or HTTPS.</p> <ul style="list-style-type: none">You can use the private IP address of the backend server as the domain name.You can also specify a domain name that consists of at least two labels separated by periods (.). Use only letters, digits, and hyphens (-). Do not start or end strings with a hyphen. Max total: 100 characters. Max label: 63 characters.	www.elb.com
Health Check Port	<p>Specifies the port that will be used by the load balancer to check the health of backend servers. The port number ranges from 1 to 65535.</p> <p>NOTE By default, the service port on each backend server is used. You can also specify a port for health checks.</p>	80
Path	<p>Specifies the health check URL, which is the destination on backend servers for health checks. This parameter is mandatory if the health check protocol is HTTP or HTTPS. The path can contain 1 to 80 characters and must start with a slash (/).</p> <p>The path can contain letters, digits, hyphens (-), slashes (/), periods (.), question marks (?), number signs (#), percent signs (%), ampersands (&), and extended character sets <code>_~! . () *[]@\$^:' , +</code></p>	/index.html
Interval (s)	<p>Specifies the interval for sending health check requests, in seconds. The interval ranges from 1 to 50.</p>	5

Parameter	Description	Example Value
Timeout (s)	Specifies the maximum time required for waiting for a response from the health check, in seconds. The timeout duration ranges from 1 to 50 .	3
Maximum Retries	Specifies the maximum number of health check retries. The value ranges from 1 to 10 .	3

8. Click **Next: Confirm**.
9. Confirm the configuration and click **Submit**.

Adding an HTTP Listener to a Shared Load Balancer



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Under **Listeners**, click **Add Listener**. Configure the parameters based on [Table 2-19](#).

Table 2-19 Parameters for configuring a listener

Parameter	Description	Example Value
Name	Specifies the listener name.	listener
Frontend Protocol	Specifies the protocol that will be used by the load balancer to receive requests from clients.	HTTP
Frontend Port	Specifies the port that will be used by the load balancer to receive requests from clients. The port number ranges from 1 to 65535.	80

Parameter	Description	Example Value
Redirect	Specifies whether to enable redirection. Redirects requests to an HTTPS listener when HTTP is used as the frontend protocol. If you have both HTTPS and HTTP listeners, you can use this function to redirect the requests from the HTTP listener to the HTTPS listener to ensure security.	N/A
Redirected To	Specifies the HTTPS listener to which requests are redirected.	listener-9ecd (HTTPS/443)
Advanced Settings		
Access Control	Specifies how access to the listener is controlled. For details, see Access Control . The following options are available: <ul style="list-style-type: none">• All IP addresses• Blacklist• Whitelist	Whitelist
IP Address Group	Specifies the IP address group associated with a whitelist or blacklist. If there is no IP address group, create one first. For more information, see Creating an IP Address Group .	ipGroup-b2
Idle Timeout	Specifies the length of time for a connection to keep alive, in seconds. If no request is received within this period, the load balancer closes the connection and establishes a new one with the client when the next request arrives. The idle timeout duration ranges from 0 to 4000 .	60
Request Timeout	Specifies the length of time (in seconds) after which the load balancer closes the connection if the load balancer does not receive a request from the client. The request timeout duration ranges from 1 to 300 .	60

Parameter	Description	Example Value
Response Timeout	<p>A load balancer sends a request to a backend server. If the backend server does not respond within the timeout period, the load balancer sends the request to another backend server. If the backend server does not respond during the retry, the load balancer returns error code HTTP 504 to the client.</p> <p>The request timeout duration ranges from 1 to 300.</p> <p>NOTE If you have enabled sticky sessions and the backend server does not respond within the response timeout duration, the load balancer returns 504 Gateway Timeout to the clients.</p>	60
Description	<p>Provides supplementary information about the listener.</p> <p>You can enter a maximum of 255 characters.</p>	N/A

6. Click **Next: Configure Request Routing Policy**. [Table 2-20](#) describes the parameters for configuring a backend server group.

Table 2-20 Parameters for configuring a backend server group

Parameter	Description	Example Value
Backend Server Group	<p>Specifies a group of servers with the same features to receive requests from the load balancer. Two options are available:</p> <ul style="list-style-type: none">• Create new• Use existing <p>NOTE The backend protocol of the backend server group must match the frontend protocol. For example, if the frontend protocol is TCP, the backend protocol must be TCP.</p>	Create new
Backend Server Group Name	<p>Specifies the name of the backend server group.</p>	server_group

Parameter	Description	Example Value
Backend Protocol	Specifies the protocol that will be used by backend servers to receive requests. The backend protocol is HTTP by default and cannot be changed.	HTTP


Parameter	Description	Example Value
Load Balancing Algorithm	<p>Specifies the algorithm that will be used by the load balancer to distribute traffic. The following options are available:</p> <ul style="list-style-type: none">• Weighted round robin: Requests are routed to different servers based on their weights, which indicate server processing performance. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests.• Weighted least connections: In addition to the number of active connections established with each backend server, each server is assigned a weight based on their processing capability. Requests are routed to the server with the lowest connections-to-weight ratio.• Source IP hash: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all backend servers are numbered. The generated key is used to allocate the client to a particular server. This allows requests from different clients to be routed based on source IP addresses and ensures that a client is directed to the same server that it was using previously. <p>NOTE</p> <ul style="list-style-type: none">• Choose an appropriate algorithm based on your requirements for better traffic distribution.• For Weighted round robin or Weighted least connections, no requests will be routed to a server with a weight of 0.	Weighted round robin

Parameter	Description	Example Value
Sticky Session	<p>Specifies whether to enable sticky sessions. If you enable sticky sessions, all requests from a client during one session are sent to the same backend server.</p> <p>NOTE You can enable sticky sessions only if you have selected Weighted round robin for Load Balancing Algorithm.</p>	N/A
Sticky Session Type	<p>Specifies the type of sticky sessions for HTTP and HTTPS listeners.</p> <ul style="list-style-type: none">• Load balancer cookie: The load balancer generates a cookie after receiving a request from the client. All subsequent requests with the same cookie are then routed to the same backend server.• Application cookie: The application deployed on the backend server generates a cookie after receiving the first request from the client. All requests with the same cookie generated by backend application are then routed to the same backend server.	Load balancer cookie
Cookie Name	<p>Specifies the cookie name. If you select Application cookie, enter a cookie name.</p>	cookieName-qsps
Stickiness Duration (min)	<p>Specifies the minutes that sticky sessions are maintained. You can enable sticky sessions only if you select Weighted round robin for Load Balancing Algorithm.</p> <ul style="list-style-type: none">• Stickiness duration at Layer 4: 1 to 60• Stickiness duration at Layer 7: 1 to 1440	20
Description	<p>Provides supplementary information about the backend server group.</p> <p>You can enter a maximum of 255 characters.</p>	N/A

7. Click **Next: Add Backend Server**. Add backend servers and configure health check for the backend server group. For details about how to add backend

servers, see [Overview](#). For the parameters required for configuring a health check, see [Table 2-21](#).

Table 2-21 Parameters for configuring a health check

Parameter	Description	Example Value
Health Check	Specifies whether to enable health checks. If the health check is enabled, click  next to Advanced Settings to set health check parameters.	N/A
Health Check Protocol	Specifies the protocol that will be used by the load balancer to check the health of backend servers. There are two options: TCP and HTTP.	HTTP
Domain Name	Specifies the domain name that will be used for health checks. This parameter is mandatory if the health check protocol is HTTP. <ul style="list-style-type: none">You can use the private IP address of the backend server as the domain name.You can also specify a domain name that consists of at least two labels separated by periods (.). Use only letters, digits, and hyphens (-). Do not start or end strings with a hyphen. Max total: 100 characters. Max label: 63 characters.	www.elb.com
Health Check Port	Specifies the port that will be used by the load balancer to check the health of backend servers. The port number ranges from 1 to 65535. NOTE By default, the service port on each backend server is used. You can also specify a port for health checks.	80

Parameter	Description	Example Value
Path	Specifies the health check URL, which is the destination on backend servers for health checks. This parameter is mandatory if the health check protocol is HTTP. The path can contain 1 to 80 characters and must start with a slash (/). The path can contain letters, digits, hyphens (-), slashes (/), periods (.), question marks (?), percent signs (%), ampersands (&), and underscores (_).	/index.html
Interval (s)	Specifies the interval for sending health check requests, in seconds. The interval ranges from 1 to 50 .	5
Timeout (s)	Specifies the maximum time required for waiting for a response from the health check, in seconds. The timeout duration ranges from 1 to 50 .	3
Maximum Retries	Specifies the maximum number of health check retries. The value ranges from 1 to 10 .	3

8. Click **Next: Confirm**.
9. Confirm the configuration and click **Submit**.

2.6 Adding an HTTPS Listener

Scenarios

HTTPS listeners are best suited for applications that require encrypted transmission. Load balancers decrypt HTTPS requests before routing them to backend servers, which then send the processed requests back to load balancers for encryption before they are sent to clients.

When you add an HTTPS listener, ensure that the subnet of the load balancer has sufficient IP addresses. If the IP addresses are insufficient, add more subnets on the summary page of the load balancer. After you select a subnet, ensure that ACL rules are not configured for this subnet. If rules are configured, request packets may not be allowed.

Constraints

- Dedicated load balancers: If the listener protocol is HTTPS, the protocol of the backend server group can be HTTP or HTTPS.

- Shared load balancers: If the listener protocol is HTTPS, the protocol of the backend server group is HTTP by default and cannot be changed.
- If you only select network load balancing (TCP/UDP) for your dedicated load balancer, you cannot add an HTTPS listener to this load balancer.

Adding an HTTPS Listener to a Dedicated Load Balancer



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Under **Listeners**, click **Add Listener**. Configure the parameters based on [Table 2-22](#).

Table 2-22 Parameters for configuring a listener

Parameter	Description	Example Value
Name	Specifies the listener name.	listener-pnqy
Frontend Protocol	Specifies the protocol that will be used by the load balancer to receive requests from clients.	HTTPS
Frontend Port	Specifies the port that will be used by the load balancer to receive requests from clients. The port number ranges from 1 to 65535.	80
SSL Authentication	Specifies whether how you want the clients and backend servers to be authenticated. There are two options: One-way authentication or Mutual authentication . <ul style="list-style-type: none">• If only server authentication is required, select One-way authentication.• If you want the clients and the load balancer to authenticate each other, select Mutual authentication. Only authenticated clients will be allowed to access the load balancer.	One-way authentication

Parameter	Description	Example Value
Server Certificate	<p>Specifies the certificate that will be used by the backend server to authenticate the client when HTTPS is used as the frontend protocol.</p> <p>Both the certificate and private key are required.</p> <p>For details, see Adding a Certificate.</p>	N/A
CA Certificate	<p>Specifies the certificate that will be used by the backend server to authenticate the client when SSL Authentication is set to Mutual authentication.</p> <p>A CA certificate is issued by a certificate authority (CA) and used to verify the certificate issuer. If HTTPS mutual authentication is required, HTTPS connections can be established only when the client provides a certificate issued by a specific CA.</p> <p>For details, see Adding a Certificate.</p>	N/A
Enable SNI	<p>Specifies whether to enable SNI when HTTPS is used as the frontend protocol.</p> <p>SNI is an extension to TLS and is used when a server uses multiple domain names and certificates.</p> <p>This allows the client to submit the domain name information while sending an SSL handshake request. After the load balancer receives the request, the load balancer queries the corresponding certificate based on the domain name and returns it to the client. If no certificate is found, the load balancer will return the default certificate. For details, see SNI Certificate.</p>	N/A

Parameter	Description	Example Value
SNI Certificate	<p>Specifies the certificate associated with the domain name when the frontend protocol is HTTPS and SNI is enabled.</p> <p>Select an existing certificate or create one.</p> <p>For details, see Adding a Certificate.</p>	N/A
Access Control	<p>Specifies how access to the listener is controlled. For details, see Access Control. The following options are available:</p> <ul style="list-style-type: none">• All IP addresses• Blacklist• Whitelist	Whitelist
IP Address Group	<p>Specifies the IP address group associated with a whitelist or blacklist. If there is no IP address group, create one first. For more information, see Creating an IP Address Group.</p>	ipGroup-b2
Transfer Client IP Address	<p>Specifies whether to transmit IP addresses of the clients to backend servers.</p> <p>This function is enabled for dedicated load balancers by default and cannot be disabled.</p>	Enabled
Advanced Forwarding	<p>Specifies whether to enable the advanced forwarding policy. You can add advanced forwarding policies to HTTP or HTTPS listeners to forward requests to different backend server groups based on HTTP request method, HTTP header, query string, or CIDR block in addition to domain names and URLs.</p>	Enabled
Advanced Settings		
Security Policy	<p>Specifies the security policy you can use if you select HTTPS as the frontend protocol. For more information, see TLS Security Policy.</p>	TLS-1-0

Parameter	Description	Example Value
HTTP/2	Specifies whether you want to use HTTP/2 if you select HTTPS for Frontend Protocol . For details, see HTTP/2 .	N/A
Transfer Load Balancer EIP	Specifies whether to store the EIP bound to the load balancer in the X-Forwarded-ELB-IP header field and pass this field to backend servers.	N/A
Idle Timeout	Specifies the length of time for a connection to keep alive, in seconds. If no request is received within this period, the load balancer closes the connection and establishes a new one with the client when the next request arrives. The idle timeout duration ranges from 0 to 4000 .	60
Request Timeout	Specifies the length of time (in seconds) after which the load balancer closes the connection if the load balancer does not receive a request from the client. The request timeout duration ranges from 1 to 300 .	60
Response Timeout	Specifies the length of time (in seconds) after which the load balancer sends a 504 Gateway Timeout error to the client if the load balancer receives no response from the backend server after routing a request to the backend server and receives no response after attempting to route the same request to other backend servers. The request timeout duration ranges from 1 to 300 . NOTE If you have enabled sticky sessions and the backend server does not respond within the response timeout duration, the load balancer returns 504 Gateway Timeout to the clients.	60

Parameter	Description	Example Value
Description	Provides supplementary information about the listener. You can enter a maximum of 255 characters.	N/A

6. Click **Next: Configure Request Routing Policy**.
 - a. You are advised to select an existing backend server group.
 - b. You can also click **Create new** to create a backend server group and configure parameters as described in [Table 2-23](#).

Table 2-23 Parameters for configuring a backend server group

Parameter	Description	Example Value
Backend Server Group	Specifies a group of servers with the same features to receive requests from the load balancer. Two options are available: <ul style="list-style-type: none">• Create new• Use existing NOTE The backend protocol of the backend server group must match the frontend protocol. For example, if the frontend protocol is TCP, the backend protocol must be TCP.	Create new
Backend Server Group Name	Specifies the name of the backend server group.	server_group-sq4v
Backend Protocol	Specifies the protocol that will be used by backend servers to receive requests. If the frontend protocol is HTTPS, the backend protocol can be HTTP, or HTTPS.	HTTP


Parameter	Description	Example Value
Load Balancing Algorithm	<p>Specifies the algorithm that will be used by the load balancer to distribute traffic. The following options are available:</p> <ul style="list-style-type: none">• Weighted round robin: Requests are routed to different servers based on their weights, which indicate server processing performance. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests.• Weighted least connections: In addition to the number of active connections established with each backend server, each server is assigned a weight based on their processing capability. Requests are routed to the server with the lowest connections-to-weight ratio.• Source IP hash: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all backend servers are numbered. The generated key is used to allocate the client to a particular server. This allows requests from different clients to be routed based on source IP addresses and ensures that a client is directed to the same server that it was using previously. <p>NOTE</p> <ul style="list-style-type: none">• Choose an appropriate algorithm based on your requirements for better traffic distribution.• For Weighted round robin or Weighted least connections, no requests will be routed to a server with a weight of 0.	Weighted round robin

Parameter	Description	Example Value
Sticky Session	<p>Specifies whether to enable sticky sessions. If you enable sticky sessions, all requests from a client during one session are sent to the same backend server.</p> <p>This parameter is optional and can be enabled if you have selected Weighted round robin or Weighted least connections for Load Balancing Algorithm.</p>	N/A
Sticky Session Type	<p>Specifies the type of sticky sessions for HTTP and HTTPS listeners.</p> <ul style="list-style-type: none">• Load balancer cookie: The load balancer generates a cookie after receiving a request from the client. All subsequent requests with the same cookie are then routed to the same backend server. <p>NOTE</p>	Load balancer cookie
Stickiness Duration (min)	<p>Specifies the minutes that sticky sessions are maintained. You can enable sticky sessions only if you select Weighted round robin for Load Balancing Algorithm.</p> <ul style="list-style-type: none">• Stickiness duration at Layer 4: 1 to 60• Stickiness duration at Layer 7: 1 to 1440	20
Slow Start	<p>Specifies whether to enable slow start, which is disabled by default.</p> <p>After you enable slow start, the load balancer linearly increases the proportion of requests to send to backend servers in this mode. When the slow start duration elapses, the load balancer sends full share of requests to backend servers and exits the slow start mode.</p> <p>For details, see Slow Start (Dedicated Load Balancers).</p>	N/A

Parameter	Description	Example Value
Slow Start Duration	Specifies how long the slow start will last. The duration ranges from 30 to 1200 , in seconds, and the default value is 30 .	30
Description	Provides supplementary information about the backend server group. You can enter a maximum of 255 characters.	N/A

7. Click **Next: Add Backend Server**. Add backend servers and configure health check for the backend server group. For details about how to add backend servers, see [Overview](#). For the parameters required for configuring a health check, see [Table 2-24](#).

Table 2-24 Parameters for configuring a health check

Parameter	Description	Example Value
Health Check	Specifies whether to enable health checks. If the health check is enabled, click  next to Advanced Settings to set health check parameters.	N/A
Health Check Protocol	Specifies the protocol that will be used by the load balancer to check the health of backend servers. If the backend protocol is HTTP or HTTPS, the health check protocol can be TCP, HTTP, or HTTPS.	HTTP

Parameter	Description	Example Value
Domain Name	<p>Specifies the domain name that will be used for health checks. This parameter is mandatory if the health check protocol is HTTP or HTTPS.</p> <ul style="list-style-type: none">You can use the private IP address of the backend server as the domain name.You can also specify a domain name that consists of at least two labels separated by periods (.). Use only letters, digits, and hyphens (-). Do not start or end strings with a hyphen. Max total: 100 characters. Max label: 63 characters.	www.elb.com
Health Check Port	<p>Specifies the port that will be used by the load balancer to check the health of backend servers. The port number ranges from 1 to 65535.</p> <p>NOTE By default, the service port on each backend server is used. You can also specify a port for health checks.</p>	80
Path	<p>Specifies the health check URL, which is the destination on backend servers for health checks. This parameter is mandatory if the health check protocol is HTTP or HTTPS. The path can contain 1 to 80 characters and must start with a slash (/).</p> <p>The path can contain letters, digits, hyphens (-), slashes (/), periods (.), question marks (?), number signs (#), percent signs (%), ampersands (&), and extended character sets <code>_~! . () *[]@\$^:' , +</code></p>	/index.html
Interval (s)	<p>Specifies the interval for sending health check requests, in seconds. The interval ranges from 1 to 50.</p>	5

Parameter	Description	Example Value
Timeout (s)	Specifies the maximum time required for waiting for a response from the health check, in seconds. The timeout duration ranges from 1 to 50 .	3
Maximum Retries	Specifies the maximum number of health check retries. The value ranges from 1 to 10 .	3

8. Click **Next: Confirm**.
9. Confirm the configuration and click **Submit**.

Adding an HTTPS Listener to a Shared Load Balancer



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Under **Listeners**, click **Add Listener**. Configure the parameters based on [Table 2-25](#).

Table 2-25 Parameters for configuring a listener

Parameter	Description	Example Value
Name	Specifies the listener name.	listener-pnqy
Frontend Protocol	Specifies the protocol that will be used by the load balancer to receive requests from clients.	HTTPS
Frontend Port	Specifies the port that will be used by the load balancer to receive requests from clients. The port number ranges from 1 to 65535.	80

Parameter	Description	Example Value
SSL Authentication	<p>Specifies whether how you want the clients and backend servers to be authenticated.</p> <p>There are two options: One-way authentication or Mutual authentication.</p> <ul style="list-style-type: none"> • If only server authentication is required, select One-way authentication. • If you want the clients and the load balancer to authenticate each other, select Mutual authentication. Only authenticated clients will be allowed to access the load balancer. 	One-way authentication
CA Certificate	<p>Specifies the certificate that allows the clients and backend servers to mutually authenticate each other.</p> <p>For details, see Adding a Certificate.</p>	N/A
Server Certificate	<p>Specifies the certificate used by the server to authenticate the client when HTTPS is used as the frontend protocol.</p> <p>Both the certificate and private key are required.</p> <p>For details, see Adding a Certificate.</p>	N/A

Parameter	Description	Example Value
Enable SNI	<p>Specifies whether to enable SNI when HTTPS is used as the frontend protocol.</p> <p>SNI is an extension to TLS and is used when a server uses multiple domain names and certificates.</p> <p>This allows the client to submit the domain name information while sending an SSL handshake request. After the load balancer receives the request, the load balancer queries the corresponding certificate based on the domain name and returns it to the client. If no certificate is found, the load balancer will return the default certificate. For details, see SNI Certificate.</p>	N/A
SNI Certificate	<p>Specifies the certificate associated with the domain name when the frontend protocol is HTTPS and SNI is enabled.</p> <p>Select an existing certificate or create one.</p> <p>For details, see Adding a Certificate.</p>	N/A
Advanced Settings		
Access Control	<p>Specifies how access to the listener is controlled. For details, see Access Control. The following options are available:</p> <ul style="list-style-type: none">• All IP addresses• Blacklist• Whitelist	Whitelist
IP Address Group	<p>Specifies the IP address group associated with a whitelist or blacklist. If there is no IP address group, create one first. For more information, see Creating an IP Address Group.</p>	ipGroup-b2

Parameter	Description	Example Value
HTTP/2	Specifies whether you want to use HTTP/2 if you select HTTPS for Frontend Protocol . For details, see HTTP/2 .	N/A
Security Policy	Specifies the security policy you can use if you select HTTPS as the frontend protocol. There are four options. For more information, see TLS Security Policy .	TLS-1-2
Idle Timeout	Specifies the length of time for a connection to keep alive, in seconds. If no request is received within this period, the load balancer closes the connection and establishes a new one with the client when the next request arrives. The idle timeout duration ranges from 0 to 4000 .	60
Request Timeout	Specifies the length of time (in seconds) after which the load balancer closes the connection if the load balancer does not receive a request from the client. The request timeout duration ranges from 1 to 300 .	60
Response Timeout	Specifies the length of time (in seconds) after which the load balancer sends a 504 Gateway Timeout error to the client if the load balancer receives no response from the backend server after routing a request to the backend server and receives no response after attempting to route the same request to other backend servers. The request timeout duration ranges from 1 to 300 . NOTE If you have enabled sticky sessions and the backend server does not respond within the response timeout duration, the load balancer returns 504 Gateway Timeout to the clients.	60

Parameter	Description	Example Value
Description	Provides supplementary information about the listener. You can enter a maximum of 255 characters.	N/A

6. Click **Next: Configure Request Routing Policy**. [Table 2-26](#) describes the parameters for configuring a backend server group.

Table 2-26 Parameters for configuring a backend server group

Parameter	Description	Example Value
Backend Server Group	Specifies a group of servers with the same features to receive requests from the load balancer. Two options are available: <ul style="list-style-type: none">• Create new• Use existing NOTE To associate an existing backend server group, ensure that it is not in use. The backend protocol of the backend server group must match the frontend protocol. For example, if the frontend protocol is TCP, the backend protocol must be TCP.	Create new
Backend Server Group Name	Specifies the name of the backend server group.	server_group-sq4v
Backend Protocol	Specifies the protocol used by backend servers to receive requests. The backend protocol is HTTP by default and cannot be changed.	HTTP


Parameter	Description	Example Value
Load Balancing Algorithm	<p>Specifies the algorithm used by the load balancer to distribute traffic. The following options are available:</p> <ul style="list-style-type: none">● Weighted round robin: Requests are routed to different servers based on their weights, which indicate server processing performance. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests.● Weighted least connections: In addition to the number of active connections established with each backend server, each server is assigned a weight based on their processing capability. Requests are routed to the server with the lowest connections-to-weight ratio.● Source IP hash: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hash key, and all backend servers are numbered. The generated key is used to allocate the client to a particular server. This allows requests from different clients to be routed based on source IP addresses and ensures that a client is directed to the same server that it was using previously. <p>NOTE</p> <ul style="list-style-type: none">● Choose an appropriate algorithm based on your requirements for better traffic distribution.● For Weighted round robin or Weighted least connections, no requests will be routed to a server with a weight of 0.	Weighted round robin

Parameter	Description	Example Value
Sticky Session	Specifies whether to enable sticky sessions. If you enable sticky sessions, all requests from a client during one session are sent to the same backend server. NOTE You can enable sticky sessions only if you have selected Weighted round robin for Load Balancing Algorithm .	N/A
Sticky Session Type	Specifies the type of sticky sessions for HTTP and HTTPS listeners. <ul style="list-style-type: none">• Load balancer cookie: The load balancer generates a cookie after receiving a request from the client. All subsequent requests with the same cookie are then routed to the same backend server.• Application cookie: The application deployed on the backend server generates a cookie after receiving the first request from the client. All requests with the same cookie generated by backend application are then routed to the same backend server.	Load balancer cookie
Cookie Name	Specifies the cookie name. If you select Application cookie , enter a cookie name.	cookieName-qsps
Stickiness Duration (min)	Specifies the minutes that sticky sessions are maintained. You can enable sticky sessions only if you select Weighted round robin for Load Balancing Algorithm . <ul style="list-style-type: none">• Stickiness duration at Layer 4: 1 to 60• Stickiness duration at Layer 7: 1 to 1440	20
Description	Provides supplementary information about the backend server group. You can enter a maximum of 255 characters.	N/A

7. Click **Next: Add Backend Server**. Add backend servers and configure health check for the backend server group. For details about how to add backend

servers, see [Overview](#). For the parameters required for configuring a health check, see [Table 2-27](#).

Table 2-27 Parameters for configuring a health check

Parameter	Description	Example Value
Health Check	Specifies whether to enable health checks. If the health check is enabled, click  next to Advanced Settings to set health check parameters.	N/A
Health Check Protocol	Specifies the protocol that will be used by the load balancer to check the health of backend servers. There are two options: TCP and HTTP.	HTTP
Domain Name	Specifies the domain name that will be used for health checks. This parameter is mandatory if the health check protocol is HTTP. <ul style="list-style-type: none">You can use the private IP address of the backend server as the domain name.You can also specify a domain name that consists of at least two labels separated by periods (.). Use only letters, digits, and hyphens (-). Do not start or end strings with a hyphen. Max total: 100 characters. Max label: 63 characters.	www.elb.com
Health Check Port	Specifies the port that will be used by the load balancer to check the health of backend servers. The port number ranges from 1 to 65535. NOTE By default, the service port on each backend server is used. You can also specify a port for health checks.	80

Parameter	Description	Example Value
Path	Specifies the health check URL, which is the destination on backend servers for health checks. This parameter is mandatory if the health check protocol is HTTP. The path can contain 1 to 80 characters and must start with a slash (/). The path can contain letters, digits, hyphens (-), slashes (/), periods (.), question marks (?), percent signs (%), ampersands (&), and underscores (_).	/index.html
Interval (s)	Specifies the interval for sending health check requests, in seconds. The interval ranges from 1 to 50 .	5
Timeout (s)	Specifies the maximum time required for waiting for a response from the health check, in seconds. The timeout duration ranges from 1 to 50 .	3
Maximum Retries	Specifies the maximum number of health check retries. The value ranges from 1 to 10 .	3

- Click **Next: Confirm**.
- Confirm the configuration and click **Submit**.

2.7 Adding a UDP Listener (with a QUIC Backend Server Group Associated)

Scenarios

If you use UDP as the frontend protocol, you can select QUIC as the backend protocol and select the connection ID to route requests with the same connection ID to the same backend server. QUIC has the advantages of low latency, high reliability, and no head-of-line blocking (HOL blocking), and is very suitable for the mobile Internet. No new connections need to be established when you switch between a Wi-Fi and a mobile data network.



NOTE

- QUIC versions include Q043, Q046, and Q050.
- UDP listeners using QUIC as backend protocol do not support fragmentation.

Constraints and Limitations

- Only dedicated load balancers support the QUIC protocol.
- You can add only UDP listeners if you want to use QUIC as the backend protocol.

Adding a UDP Listener with a QUIC Backend Server Group Associated

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
Select **Network load balancing (TCP/UDP)** and select a specification for the load balancer.
5. Under **Listeners**, click **Add Listener**.
6. In the **Configure Listener** step, set **Frontend Protocol** to **UDP**, configure other parameters based on the site requirements, and click **Next: Configure Request Routing Policy**.
7. In the **Configure Routing Policy** step, set **Backend Protocol** to **QUIC** and configure other parameters as required.
8. Configure the parameters and click **Submit**.

Follow-Up Operations

After you add a listener, associate backend servers with the listener by performing the operations in [Overview](#).

2.8 Configuring Modification Protection for a Listener

Scenario

You can enable modification protection for a listener to prevent it from being modified or deleted.



Constraints

If you enable modification protection for a listener, you cannot:

- Modify the basic information and forwarding policy of the listener.
- Change the default backend server group.
- Delete the listener and its load balancer.

Procedure

1. Log in to the management console.

2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click its name.
6. On the **Summary** tab page, click **Configure** next to **Modification Protection**.
7. In the **Configure Modification Protection** dialog box, enable **Modification Protection**.

 **NOTE**

Disable **Modification Protection** if you want to modify or delete a listener.

2.9 Configuring Timeout Durations

Scenarios

You can configure timeout durations (idle timeout, request timeout, and response timeout) for your listeners to meet varied demands. For example, if the size of a request from an HTTP or HTTPS client is large, you can increase the request timeout duration to ensure that the request can be successfully routed.

For shared load balancers, you can only change the timeout durations of TCP, HTTP, and HTTPS listeners, but cannot change the timeout durations of UDP listeners.

For dedicated load balancers, you can change the timeout durations of TCP, UDP, HTTP, and HTTPS listeners.

Figure 2-1 Timeout durations at Layer 7

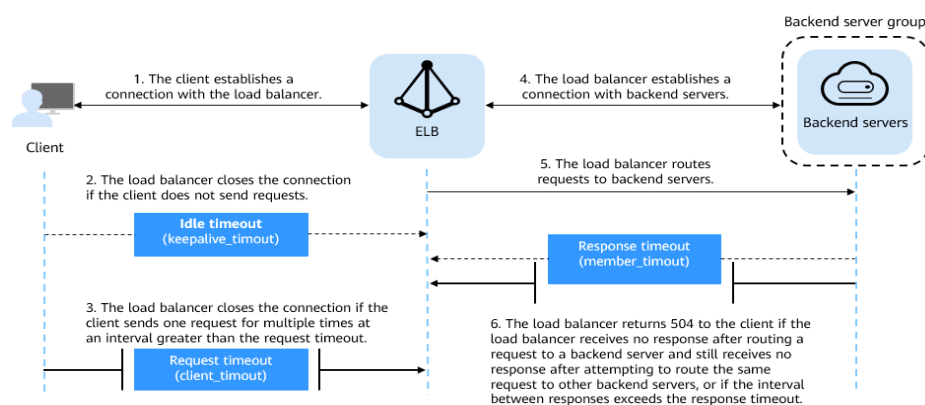


Figure 2-2 Timeout durations at Layer 4

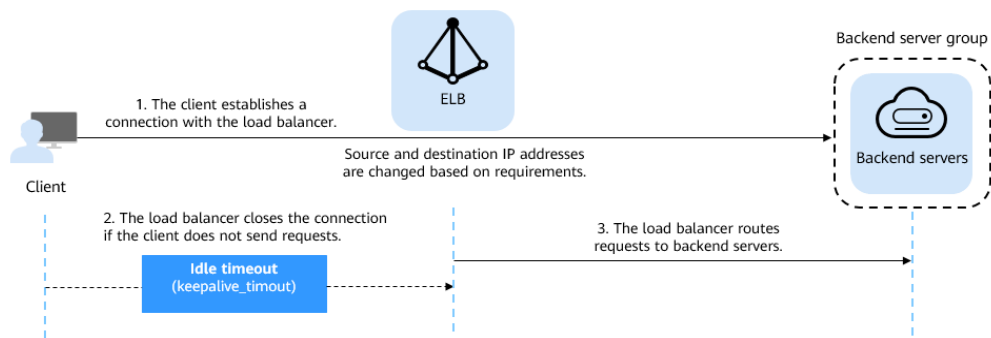


Table 2-28 Timeout durations



Protocol	Type	Description	Value Range	Default Timeout Duration
TCP	Idle Timeout	Duration for a connection to be kept alive. If no request is received within this period, the load balancer closes the connection and establishes a new one with the client when the next request arrives.	10–4000s	300s
UDP	Idle Timeout		10–4000s	Shared load balancers: 10s Dedicated load balancers: 300s
HTTP/HTTPS	Idle Timeout	Duration after which the load balancer closes the connection with the client if the load balancer does not receive a request from the client.	10–4000s	60s
	Request Timeout		10–300s	60s

Protocol	Type	Description	Value Range	Default Timeout Duration
	Response Timeout	<p>Duration after which the load balancer sends a 504 Gateway Timeout error to the client if the load balancer receives no response after routing a request to a backend server and receives no response after attempting to route the same request to other backend servers.</p> <p>NOTE If sticky sessions are enabled and the backend server does not respond within the response timeout duration, the load balancer returns the 504 error code without attempting to route the same request to other backend servers.</p>	1–300s	60s

Constraints

If modification protection is enabled for a listener, its timeout durations cannot be modified.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click the name of the listener.
6. On the **Summary** tab page, click **Edit** on the top right.
7. In the **Edit** dialog box, expand **Advanced Settings**.
8. Configure **Idle Timeout (s)**, **Request Timeout (s)**, or **Response Timeout (s)** as you need.
9. Click **OK**.

2.10 Modifying or Deleting a Listener

Scenarios

You can modify a listener as needed or delete a listener if you no longer need it. Deleted listeners cannot be recovered.



NOTE

Frontend Protocol/Port and **Backend Protocol** cannot be modified after you have configured them. If you want to modify the protocol or port of the listener, add another listener to the load balancer.



Constraints

If modification protection is enabled for a listener, the listener cannot be deleted or modified.

Modifying a listener

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Modify the listener in either of the following ways:
 - On the **Listeners** page, locate the listener, and click **Edit** in the **Operation** column.
 - Click the name of the target listener. On the **Summary** tab page, click **Edit** on the top right corner.
6. On the **Edit** dialog box, modify parameters, and click **OK**.

Deleting a listener

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click **Delete** in the **Operation** column.
6. In the displayed dialog box, enter **DELETE**.
7. Click **OK**.

2.11 Transfer Client IP Address (Dedicated Load Balancers)

Transfer Client IP Address

If you enable **Transfer Client IP Address**, your load balancer will use the IP address of the client to access the backend server.

[Table 2-29](#) lists whether you can enable or disable the transfer client IP address function.

Table 2-29 Transfer client IP address

Listener Type	Enabling Transfer Client IP Address	Disabling Transfer Client IP Address
TCP and UDP	Enabled by default	×
HTTP and HTTPS	Enabled by default	×

Constraints

- If **Transfer Client IP Address** is enabled, a server cannot serve as both a backend server and a client.
If the client and the backend server are using the same server and the **Transfer Client IP Address** option is enabled, the backend server will think the packet is sent by itself but not from the client and will not return a response packet to the load balancer. As a result, the return traffic will be interrupted.
- After this function is enabled, unidirectional download or push traffic may be interrupted when backend servers are being migrated. After the backend servers are migrated, retransmit the packets to restore the traffic.
- If you add IP addresses as backend servers, the source IP addresses of the clients cannot be passed to these servers. Install the **TOA module** to obtain source IP addresses.

Alternatives for Obtaining the IP Address of the Client

You can obtain the IP address of a client in one of the ways listed in [Table 2-30](#).

Table 2-30 Alternatives

Listener Type	Alternatives
TCP and UDP	Configuring the TOA Module
HTTP and HTTPS	Layer 7 Load Balancing

2.12 Transfer Client IP Address (Shared Load Balancers)

Scenario

Generally, load balancers use IP addresses in 100.125.0.0/16 to communicate with backend servers. If you want a load balancer to communicate with backend servers using real IP addresses of the clients, you can enable **Transfer Client IP Address** to pass the IP addresses of the clients to backend servers.

[Table 2-31](#) lists whether you can enable or disable the transfer client IP address function.



Table 2-31 Transfer client IP address

Listener Type	Enabling Transfer Client IP Address	Disabling Transfer Client IP Address
TCP and UDP	√	√
HTTP and HTTPS	Enabled by default	×

Constraints

- When you enable or disable the function, if the listener has backend servers associated, traffic to this listener will be interrupted for about 10 seconds. The interruption duration is twice the health check interval configured for the backend server group.
- If **Transfer Client IP Address** is enabled, a server cannot serve as both a backend server and a client. If the client and the backend server are using the same server and the **Transfer Client IP Address** option is enabled, the backend server will think the packet is sent by itself but not from the client and will not return a response packet to the load balancer. As a result, the return traffic will be interrupted.
- If a backend server has been associated with the listener and health checks are enabled, enabling this function will check the health of the backend server, and traffic to this server will be interrupted for one or two health check intervals.
- After this function is enabled, unidirectional download or push traffic may be interrupted when backend servers are being migrated. After the backend servers are migrated, retransmit the packets to restore the traffic.

Enabling Transfer Client IP Address



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.

4. Locate the load balancer and click its name.
5. You can use either of the following methods to enable the function:
 - On the **Listeners** page, locate the listener, and click **Edit** in the **Operation** column.
 - Click the name of the target listener. On the **Summary** tab page, click **Edit** on the top right corner.
6. In the displayed dialog box, enable **Transfer Client IP Address**.
7. Confirm the configurations and click **OK**.

 **NOTE**

After **Transfer Client IP Address** is enabled, configure security groups, network ACLs, and OS and software security policies so that IP addresses of the clients can access these backend servers.

Disabling Transfer Client IP Address

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. You can use either of the following methods to disable the function:
 - On the **Listeners** page, locate the listener, and click **Edit** in the **Operation** column.
 - Click the name of the target listener. On the **Summary** tab page, click **Edit** on the top right corner.
6. In the displayed dialog box, disable **Transfer Client IP Address**.
7. Confirm the configurations and click **OK**.

Alternatives for Obtaining the IP Address of the Client

You can obtain the IP address of a client in one of the ways listed in [Table 2-32](#).

Table 2-32 Alternatives

Listener Type	Alternatives
TCP and UDP	Configuring the TOA Module
HTTP and HTTPS	Layer 7 Load Balancing

3 Advanced Features of HTTP/HTTPS Listeners

3.1 Forwarding Policy (Shared Load Balancers)

Scenarios

You can add forwarding policies to HTTP or HTTPS listeners to forward requests to different backend server groups based on domain names or URLs.

This is suited for applications that are deployed on multiple backend servers and provide multiple types of services such as videos, images, audios, and texts.

A forwarding policy consists of a forwarding rule and an action.

- There are two types of forwarding rules: domain name and URL.
- HTTP listeners can forward requests to a backend server group and redirect requests to another listener.
- HTTPS listeners can forward requests to a backend server group.

How Requests Are Matched

- After you add a forwarding policy, the load balancer forwards requests based on the specified domain name or URL:
 - If the domain name or URL in a request matches that specified in the forwarding policy, the request is forwarded to the backend server group you select or create when you add the forwarding policy.
 - If the domain name or URL in a request does not match that specified in the forwarding policy, the request is forwarded to the default backend server group of the listener.
- Matching priority:
 - Forwarding policy priorities are independent of each other regardless of domain names. If a forwarding rule uses both domain names and URLs, requests are matched based on domain names first.

- If the forwarding rule is a URL, the priorities follow the order of exact match, prefix match, and regular expression match. If the matching types are the same, the longer the URL length, the higher the priority.

Table 3-1 Example forwarding policies

Request	Forwarding Policy	Forwarding Rule	Specified Value
www.elb.com/ test	1	URL	/test
	2	Domain name	www.elb.com

 **NOTE**

In this example, request **www.elb.com/test** matches both forwarding policies 1 and 2, but is routed based on forwarding policy 2.


Constraints and Limitations

- Forwarding policies can be added only to HTTP and HTTPS listeners.
- Forwarding policies must be unique.
- A maximum of 100 forwarding policies can be configured for a listener. If the number of forwarding policies exceeds the quota, the excess forwarding policies will not be applied.
- When you add a forwarding policy, note the following:
 - Each URL path must exist on the backend server. If the path does not exist, the backend server will return 404 Not Found.
 - In the regular expression match, the characters are matched sequentially, and matching ends when any rule is successfully matched. Matching rules cannot overlap with each other.
 - A URL path cannot be configured for two forwarding policies.
 - A domain name cannot exceed 100 characters.

 **CAUTION**

If you add a forwarding policy that is the same as an existing one by calling APIs, there will be a conflict. Even if you delete the existing forwarding policy, the new forwarding policy is still faulty. Delete the newly-added forwarding policy and add a different one.

Adding a Forwarding Policy

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.


3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. On the **Listeners** tab page, add a forwarding policy in either of the following ways:
 - On the **Listeners** page, locate the listener, and click **Add/Edit Forwarding Policy** in the **Forwarding Policies** column.
 - Locate the target listener, click its name, and click **Forwarding Policies**.
6. Click **Add Forwarding Policy**. Configure the parameters based on [Table 3-2](#).
7. After the configuration is complete, click **Save**.

Table 3-2 Forwarding policy parameters

Parameter		Description	Example Value
Forwarding Rule	Domain name	Specifies the domain name used for forwarding requests. The domain name in the request must exactly match that in the forwarding policy. You need to specify either a domain name or URL.	www.test.com
	URL	Specifies the URL used for forwarding requests. There are three URL matching rules: <ul style="list-style-type: none">• Exact match The request URL must exactly match that specified in the forwarding policy.• Prefix match The requested URL starts with the specified URL string.• Regular expression match The requested URL matches the specified URL string based on the regular expression.	/login.php
Action	Forward to a backend server group	If the request matches the configured forwarding rule, the request is forwarded to the specified backend server group.	Forward to a backend server group

Parameter	Description	Example Value
Redirect to another listener	<p>If the request matches the configured forwarding rule, the request is redirected to the specified HTTPS listener.</p> <p>This action can be configured only for HTTP listeners.</p> <p>NOTE</p> <p>If you select Redirect to another listener and create a redirect for the current listener, this listener will redirect the requests to the specified HTTPS listener, but access control configured for the listener will still take effect.</p> <p>For example, if you configure a redirect for an HTTP listener, HTTP requests to access a web page will be redirected to the HTTPS listener you select and handled by the backend servers associated with the HTTPS listener. As a result, the clients access the web page over HTTPS. The configuration of the HTTP listener will become invalid.</p>	N/A
Backend Server Group	<p>Select a backend server group that will receive requests from the load balancer.</p> <p>This parameter is mandatory when you set Action to Forward to a backend server group.</p>	N/A
Listener	<p>Select an HTTPS listener that will receive requests redirected from the current HTTP listener.</p> <p>This parameter is mandatory when Action is set to Redirect to another listener.</p>	N/A

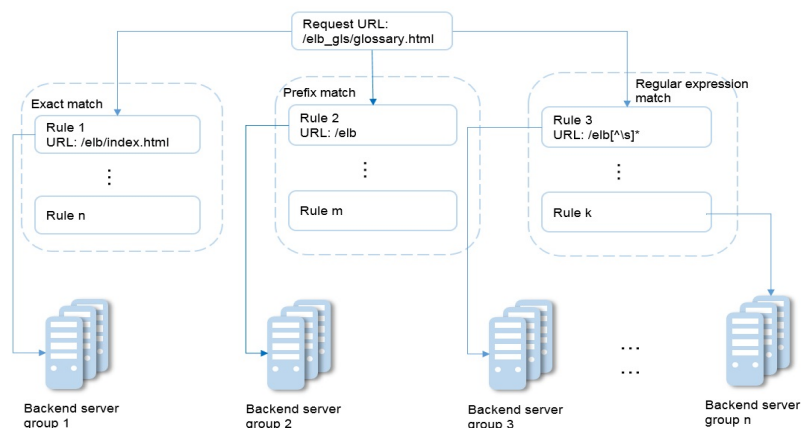
URL Matching Example

The following table lists how a URL is matched, and [Figure 3-1](#) shows how a request is forwarded to a backend server group.

Table 3-3 URL matching



URL Matching Rule	URL	URL in the Forwarding Policy			
		/elb/index.html	/elb	/elb[^\s]*	/index.html
N/A	N/A	/elb/index.html	/elb	/elb[^\s]*	/index.html
Exact match	/elb/index.html	√	N/A	N/A	N/A
Prefix match		√	√	N/A	N/A
Regular expression match		√	N/A	√	N/A

Figure 3-1 Request forwarding





In this figure, the system first searches for an exact match of the requested URL (`/elb_gls/glossary.html`). If there is no exact match, the system searches for a prefix match. If a match is found, the request is forwarded to backend server group 2 even if a regular expression match is also found, because the prefix match has a higher priority.

Modifying a Forwarding Policy

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.

5. Click **Listeners**, locate the listener, and click its name.
6. On the **Forwarding Policies** tab page, select the forwarding policy, and click **Edit**.
7. Modify the parameters and click **Save**.

Deleting a Forwarding Policy

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click its name.
6. On the **Forwarding Policies** tab page, select the forwarding policy, and click **Delete** on the top right.
7. In the displayed dialog box, click **Yes**.

3.2 Forwarding Policy (Dedicated Load Balancers)

Overview

You can add forwarding policies to HTTP or HTTPS listeners to forward requests to different backend server groups based on domain names or URLs.

A forwarding policy consists of one or more forwarding rules and an action. For details, see [Table 3-4](#).

Table 3-4 Rules and actions supported by a forwarding policy

Policy Type	Forwarding Rules	Actions
Forwarding policy	Domain name and URL	Forward to another backend server group and Redirect to another listener (only for HTTP listeners)
Advanced forwarding policy	Domain name, URL, HTTP request method, HTTP header, query string, and CIDR block	The following actions are supported: forward to a backend server group, redirect to another listener, redirect to another URL, and return a specific response body.

NOTE

You can configure an advanced forwarding policy by referring to [Managing an Advanced Forwarding Policy](#).

How Requests Are Matched

- After you add a forwarding policy, the load balancer forwards requests based on the specified domain name or URL:
 - If the domain name or URL in a request matches that specified in the forwarding policy, the request is forwarded to the backend server group you create or select when you add the forwarding policy.
 - If the domain name or URL in a request does not match that specified in the forwarding policy, the request is forwarded to the default backend server group of the listener.
- Matching priority:
 - Forwarding policy priorities are independent of each other regardless of domain names. If a forwarding rule uses both domain names and URLs, requests are matched based on domain names first.
 - If the forwarding rule is a URL, the priorities follow the order of exact match, prefix match, and regular expression match. If the matching types are the same, the longer the URL length, the higher the priority.

Table 3-5 Example forwarding policies

Request	Forwarding policy	Forwarding Rule	Specified Value
www.elb.com/ test	1	URL	/test
	2	Domain name	www.elb.com

NOTE

In this example, request **www.elb.com/test** matches both forwarding policies 1 and 2, but is routed based on forwarding policy 2.

Notes and Constraints

- You can add forwarding policies to HTTP and HTTPS listeners.
- Forwarding policies must be unique.
- A maximum of 100 forwarding policies can be configured for a listener. If the number of forwarding policies exceeds the quota, the excess forwarding policies will not be applied.
- When you add a forwarding policy, note the following:
 - Each URL path must exist on the backend server. Otherwise, the backend server returns 404 when you access the backend server.
 - In the regular expression match, the rules are matched sequentially, and matching ends when any rule is successfully matched. Matching rules cannot overlap with each other.
 - A URL path cannot be configured for two forwarding policies.
 - A domain name cannot exceed 100 characters.

Adding a Forwarding Policy



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer and click its name.
5. On the **Listeners** tab page, add a forwarding policy in either of the following ways:
 - Click **Add/Edit Forwarding Policy** in the **Forwarding Policies** column.
 - Locate the target listener, click its name, and click **Forwarding Policies**.
6. Click **Add Forwarding Policy**. Configure the parameters based on [Table 3-6](#).

Table 3-6 Forwarding policy parameters

Parameter	Type	Description	Example Value
Forwarding Rule	Domain name	Specifies the domain name used for forwarding requests. The domain name in the request must exactly match that in the forwarding policy. You need to specify either a domain name or URL.	www.test.com
	URL	Specifies the URL used for forwarding requests. There are three URL matching rules: <ul style="list-style-type: none">• Exact match: The request URL must exactly match that specified in the forwarding policy.• Prefix match: The requested URL starts with the specified URL string.• Regular expression match: The URLs are matched using a regular expression.	/login.php
Action	Forward to a backend server group	If the request matches the configured forwarding rule, the request is forwarded to the specified backend server group.	-

Parameter	Type	Description	Example Value
	Redirect to another listener	<p>If the request matches the configured forwarding rule, the request is redirected to the specified HTTPS listener.</p> <p>This action can be configured only for HTTP listeners.</p> <p>NOTE If you select Redirect to another listener, the HTTP listener will redirect requests to the specified HTTPS listener, but access control configured for the HTTP listener still takes effect.</p>	-

7. Click **Save**.

3.3 Advanced Forwarding (Dedicated Load Balancers)

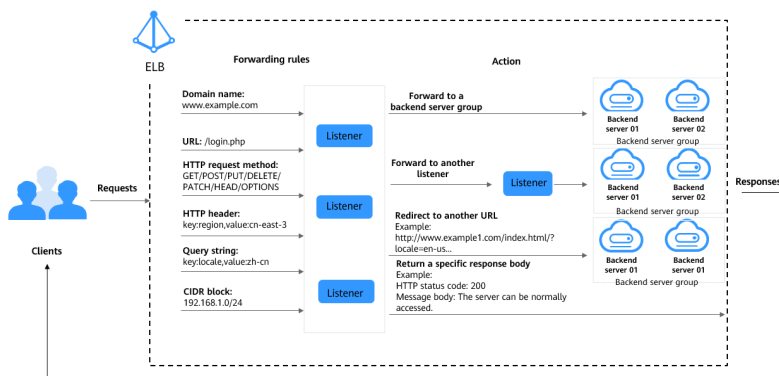
3.3.1 Advanced Forwarding

Overview

Advanced forwarding policies are available only for dedicated load balancers. If you have enabled **Advanced Forwarding**, you can add advanced forwarding policies to HTTP and HTTPS listeners of dedicated load balancers.

You can add advanced forwarding policies to HTTP or HTTPS listeners to forward requests to different backend server groups based on HTTP request method, HTTP header, query string, or CIDR block in addition to domain names and URLs. [Table 3-7](#) describes the rules and actions that you can configure for request forwarding.

Figure 3-2 How advanced forwarding works



The following describes how an advanced forwarding policy works:

- Step 1** The client sends a request to the load balancer.
- Step 2** The load balancer matches the request based on the forwarding rule you configure.
- Step 3** The load balancer forwards the request to the corresponding backend server or returns a fixed response to the client based on the action you configure.
- Step 4** The load balancer sends a response to the client.

----End

Table 3-7 Rules and actions supported by an advanced forwarding policy

Forwarding Policy	Description
Forwarding rule	There are six types of forwarding rules: domain name, URL, HTTP request method, HTTP header, query string, and CIDR block For details, see Forwarding Rule .
Action	The following actions are supported: forward to a backend server group, redirect to another listener, redirect to another URL, and return a specific response body. For details, see Action Types .

How Requests Are Matched

After you add an HTTP or HTTPS listener to a load balancer, a default forwarding policy is generated. This policy uses the protocol and port specified for the listener to match requests and forward the requests to the backend server group you specified when adding the listener.

The default forwarding policy has the lowest priority and is not included when you sort forwarding policies. It can be edited but cannot be deleted.

Each request is matched based on the forwarding policy priority (a smaller value indicates a higher priority). Once a forwarding policy is matched, the request is forwarded based on this forwarding policy.

- If the request is matched with any forwarding policy of the listener, it is forwarded based on this forwarding policy.
- If the request is not matched with any forwarding policy, it is forwarded based on the default forwarding policy.

Forwarding Rule

Advanced forwarding policies support the following types of forwarding rules: domain name, URL, HTTP request method, HTTP header, query string, and CIDR block (source IP addresses).

Table 3-8 Forwarding rules

Forwarding Rule	Description
Domain name	<ul style="list-style-type: none">• Description Route requests based on the domain name.<ul style="list-style-type: none">– You can configure multiple domain names in a forwarding policy. Each domain name contains at least two labels separated by periods (.). Max total: 100 characters. Max label: 63 characters.– Each label can contain letters, digits, hyphens (-), periods (.), and asterisks (*). A label must start with a letter, digit, or asterisk (*) and cannot end with a hyphen (-). An asterisk (*) must be used as the leftmost label if you want to configure a wildcard domain name.• Matching rules Exact match domains and wildcard domains are supported. <p>Example Request URL: <code>https://www.example.com/login.php?locale=en-us#videos</code> Domain name in the forwarding rule: <code>www.example.com</code></p>
URL	<ul style="list-style-type: none">• Description Route requests based on URLs. You can configure multiple URLs in a forwarding policy. A URL can contain letters, digits, and special characters <code>_~';@^-%#\$.*+?,=!: \()[]{}.</code> If the URL contains special characters such as question marks (?) or pound keys (#), escape the special characters before configuring the forwarding rule.• Matching rules<ul style="list-style-type: none">– Exact match: The request URL must exactly match that specified in the forwarding policy. The URL must start with a slash (/) and can use asterisks (*) and question marks (?) as wildcards.– Prefix match: The requested URL starts with the specified URL string. The URL must start with a slash (/) and can use asterisks (*) and question marks (?) as wildcard characters.– Regular expression match: The URLs are matched using a regular expression. <p>For more information about URL matching rules, see URL Matching.</p> <p>Example Request URL: <code>https://www.example.com/login.php?locale=en-us#videos</code> URL in the forwarding rule: <code>/login.php</code></p>

Forwarding Rule	Description
Query string	<p>Route requests based on the query string.</p> <p>A query string consists of a key and one or more values. You need to set the key and values separately.</p> <ul style="list-style-type: none">The key can contain only letters, digits, and special characters <code>!\$'()*+.,/;=?@^_-'</code>A key can have one or more values. The value can contain letters, digits, and special characters <code>!\$'()*+.,/;=?@^_-'</code>. Asterisks (*) and question marks (?) can be used as wildcard characters. <p>Example Request URL: <code>https://www.example.com/login.php?locale=en-us#videos</code> A query string needs to be configured for the forwarding rule: Key: <code>locale</code> Value: <code>en-us</code></p>
HTTP request method	<p>Route requests based on the HTTP method.</p> <ul style="list-style-type: none">You can configure multiple request methods in a forwarding policy.The following methods are available: GET, POST, PUT, DELETE, PATCH, HEAD, and OPTIONS. <p>Example GET</p>
HTTP header	<p>Route requests based on the HTTP header.</p> <p>An HTTP header consists of a key and one or more values. You need to configure the key and values separately.</p> <ul style="list-style-type: none">The key can contain only letters, digits, underscores (_), and hyphens (-).A key can have one or more values. The value can contain letters, digits, and special characters <code>!#\$%&'()*+.,\/:;<=>?@[^_`{ }~</code>. Asterisks (*) and question marks (?) can be used as wildcard characters. <p>Example Key: <code>Accept-Language</code> Value: <code>en-us</code></p>
CIDR block	<p>Route requests based on the source IP addresses from where requests originate.</p> <p>Example <code>192.168.1.0/24</code> or <code>2020:50::44/127</code></p>

Action Types

Advanced forwarding policies support the following actions: forward to a backend server group, redirect to another listener, redirect to another URL, and return a specific response body.

Table 3-9 Actions of an advanced forwarding policy

Action	Description
Forward to a backend server group	Requests are forwarded to the specified backend server group.
Redirect to another listener	<p>Requests are redirected to another listener, which then routes the requests to its associated backend server group.</p> <p>NOTE</p> <p>If you select Redirect to another listener and create a redirect for the listener, it will redirect the requests to the specified HTTPS listener, but access control configured for the listener will still take effect.</p> <p>For example, if you configure a redirect for an HTTP listener, HTTP requests to access a web page will be redirected to the HTTPS listener you select and handled by the backend servers associated with the HTTPS listener. As a result, the clients access the web page over HTTPS. The configuration of the HTTP listener will become invalid.</p>

Action	Description
Redirect to another URL	<p>Requests are redirected to the configured URL.</p> <p>When clients access website A, the load balancer returns 302 or any other 3xx status code and automatically redirects the clients to website B. You can custom the redirection URL that will be returned to the clients.</p> <p>Configure at least one of the following components:</p> <ul style="list-style-type: none">• Protocol: <code>\${protocol}</code>, HTTP, or HTTPS <code>\${protocol}</code>: retains the protocol of the request.• Domain Name: A domain name consists of at least two labels separated by periods (.). Each label can contain only letters, digits, hyphens (-), and periods (.), must start with a letter, digit, or asterisk (*), and cannot end with a hyphen (-). <code>\${host}</code>: retains the domain name of the request.• Port: ranges from 1 to 65535. <code>\${port}</code>: retains the port number of the request.• Path: A path can contain letters, digits, and special characters <code>_~';@^-%#&\$.*+?,=!: \\()[]{}</code> and must start with a slash (/). <code>\${path}</code>: retains the path of the request. <p>NOTE</p> <p>If you select regular expression match, the request path will be overwritten by the variables that match the regular expressions. For details, see URL Matching Based on Regular Expressions.</p> <ul style="list-style-type: none">• Query String: A query string can contain only letters, digits, and special characters <code>!\$'()*+,-./:;=?@&^_-'</code>. Ampersand (&) can only be used as separators.• HTTP Status Code: 301, 302, 303, 307, or 308 <p>Example URL for redirection: <code>http://www.example1.com/index.html?locale=en-us#videos</code> Protocol: HTTP Domain name: <code>www.example1.com</code> Port: 8081 Path: <code>/index.html</code> Query String: <code>locale=en-us</code> HTTP Status Code: 301</p>

Action	Description
Return a specific response body	<p>Load balancers return a fixed response to the clients. You can custom the status code and response body that load balancers directly return to the clients without the need to route the requests to backend servers.</p> <p>Configure the following components:</p> <ul style="list-style-type: none"> • HTTP Status Code: By default, 2xx, 4xx, and 5xx status codes are supported. • Content-Type: text/plain, text/css, text/html, application/javascript, or application/json • Message Body: This parameter is optional. The value is a string of 0 to 1,024 characters. <p>Example</p> <p>text/plain Sorry, the language is not supported.</p> <p>text/css <head><style type="text/css">div {background-color:red}#div {font-size:15px;color:red}</style></head></p> <p>text/html <form action="/" method="post" enctype="multipart/form-data"><input type="text" name="description" value="some text"><input type="file" name="myFile"><button type="submit">Submit</button></form></p> <p>application/javascript String.prototype.trim = function() {var reExtraSpace = /\s*(.*?)\s+\$/;return this.replace(reExtraSpace, "\$1")}</p> <p>application/json { "publicip": { "type": "5_bgp", "ip_version": 4}, "bandwidth": {"name": "bandwidth123", "size": 10, "share_type": "PER"}}</p> <p>NOTE Ensure that the response body does not contain carriage return characters. Otherwise, it cannot be saved.</p>

URL Matching

Table 3-10 shows how URLs configured in the forwarding policies match the URLs in the requests.

Table 3-10 URL matching examples

Request URL	Forwarding Policy	URL in the Forwarding Policy	Matching Mode	Forwarding Policy Priority	Destination Backend Server Group
/elb/abc.html	Forwarding policy 01	/elb/abc.html	Prefix match	1	Backend server group 01

Request URL	Forwarding Policy	URL in the Forwarding Policy	Matching Mode	Forwarding Policy Priority	Destination Backend Server Group
	Forwarding policy 02	/elb	Prefix match	2	Backend server group 02
/exa/index.html	Forwarding policy 03	/exa[^\s]*	Regular expression match	3	Backend server group 03
	Forwarding policy 04	/exa/index.html	Regular expression match	4	Backend server group 04
/mpl/index.html	Forwarding policy 05	/mpl/index.html	Exact match	5	Backend server group 05

URLs are matched as follows:

- When the request URL is /elb/abc.html, it matches both forwarding policy 01 and forwarding policy 02. However, the priority of forwarding policy 01 is higher than that of forwarding policy 02. Forwarding policy 01 is used, and requests are forwarded to backend server group 01.
- When the request URL is /exa/index.html, it matches both forwarding policy 03 and forwarding policy 04. However, the priority of forwarding policy 03 is higher than that of forwarding policy 04. Forwarding policy 03 is used, and requests are forwarded to backend server group 03.
- If the request URL is /mpl/index.html, it matches forwarding policy 05 exactly, and requests are forwarded to backend server group 05.

URL Matching Based on Regular Expressions

A path can contain letters, digits, and special characters `_~';@^-%#&$.*+?,=!:|\/()\[\]{}` and must start with a slash (/). `${path}` retains the path of the request.

If you select regular expression match, the request path will be overwritten by the variables that match the regular expressions.

How Request Paths Are Overwritten

1. URL matching: The client sends a request, and the request matches a regular expression in the forwarding rule. You can specify one or more regular expressions as the match conditions and set multiple capture groups represented by parentheses () for one regular expression.

2. Extraction and replacement: extracts the content from the capture groups.
3. Destination path: writes them to \$1, \$2, all the way to \$9 configured for the path.

Example

When a client requests to access `/test/ELB/elb/index`, which matches the regular expression `/test/(.*/)(.*/)index`, \$1 will be replaced by `ELB` and \$2 by `elb`, and then the request will be redirected to `/ELB/elb`.

Table 3-11 URL matching based on regular expressions

Matching Step		Description
Forwarding rule: URL	Regular expression match	<ul style="list-style-type: none">• Matching condition: <code>/test/(.*/)(.*/)index</code>• Request URL: <code>/test/ELB/elb/index</code>
Action: redirect to another URL	Path	<ul style="list-style-type: none">• Path: <code>/\$1/\$2</code>• Extracting content<ul style="list-style-type: none">\$1: <code>ELB</code>\$2: <code>elb</code>• Destination path: <code>/ELB/elb</code>

3.3.2 Managing an Advanced Forwarding Policy

Scenarios

You can add advanced forwarding policies to HTTP or HTTPS listeners of dedicated load balancers to route requests more specifically.



Each advanced forwarding policy consists of one or more forwarding rules and an action.

- Dedicated load balancers support the following types of forwarding rules: domain name, URL, HTTP request method, HTTP header, query string, and CIDR block (source IP addresses). For details, see [Forwarding Rule](#).
- The following actions are supported: forward to a backend server group, redirect to another listener, redirect to another URL, and return a specific response body. For details, see [Action Types](#).
- Multiple forwarding rules can be configured in a single forwarding policy.
- Forwarding policies can be sorted based on their priorities.



Constraints

- Advanced forwarding cannot be disabled once enabled.
- An advanced forwarding policy can contain a maximum of 10 conditions.

Enabling Advanced Forwarding



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer and click its name.
5. Click the **Listeners** tab and click the target listener.
6. On the **Summary** tab page, click **Enable** next to **Advanced Forwarding**.
7. Click **OK**.

Adding an Advanced Forwarding Policy



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer and click its name.
5. On the **Listeners** tab page, add a forwarding policy in either of the following ways:
 - Click **Add/Edit Forwarding Policy** in the **Forwarding Policies** column.
 - Locate the target listener, click its name, and click **Forwarding Policies**.
6. Click **Add Forwarding Policy** and configure the parameters based on [Table 3-8](#) and [Table 3-9](#).
7. Click **Save**.

Sorting Forwarding Policies

Multiple forwarding policies can be sorted to set their priorities.

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click its name.
6. On the **Forwarding Policies** tab page, click **Sort**.
7. Drag the forwarding policies to adjust their priorities.
8. Click **Save**.



Modifying a Forwarding Policy

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click its name.
6. On the **Forwarding Policies** tab page, select the forwarding policy, and click **Edit**.
7. Modify the parameters and click **Save**.

Deleting a Forwarding Policy

You can delete a forwarding policy if you no longer need it.

Deleted forwarding policies cannot be recovered.

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click its name.
6. On the **Forwarding Policies** tab page, select the forwarding policy and click **Delete** on the top right.
7. In the displayed dialog box, click **Yes**.

3.4 Mutual Authentication

Scenarios

In common HTTPS service scenarios, only the server certificate is required for authentication. For some mission-critical services, such as financial transactions, you need to deploy both the server certificate and the client certificate for mutual authentication.

Self-signed certificates are used as an example to describe how to configure mutual authentication. Self-signed certificates do not provide all the security properties provided by certificates signed by a CA. It is recommended that you purchase certificates from [SSL Certificate Manager \(SCM\)](#) or CAs.

Creating a CA Certificate Using OpenSSL

1. Log in to a Linux server with OpenSSL installed.
2. Create the **server** directory and switch to the directory:
mkdir ca
cd ca
3. Create the certificate configuration file **ca_cert.conf**. The file content is as follows:

```
[ req ]
distinguished_name = req_distinguished_name
prompt = no

[ req_distinguished_name ]
O = ELB
```

4. Create the CA certificate private key **ca.key**.
openssl genrsa -out ca.key 2048

Figure 3-3 Private key of the CA certificate

```
[root@elbv30003 ca]# openssl genrsa -out ca.key 2048
Generating RSA private key, 2048 bit long modulus (2 primes)
.....+++++
.....+++++
e is 65537 (0x010001)
[root@elbv30003 ca]#
```

5. Create the certificate signing request (CSR) file **ca.csr** for the CA certificate.
openssl req -out ca.csr -key ca.key -new -config ./ca_cert.conf
6. Create the self-signed CA certificate **ca.crt**.
openssl x509 -req -in ca.csr -out ca.crt -sha1 -days 5000 -signkey ca.key

Figure 3-4 Creating a self-signed CA certificate

```
[root@elbv30003 ca]# openssl x509 -req -in ca.csr -out ca.crt -sha1 -days 5000 -signkey ca.key
Signature ok
subject=O = ELB
Getting Private key
[root@elbv30003 ca]#
```

Issuing a Server Certificate Using the CA Certificate

The server certificate can be a CA signed certificate or a self-signed one. In the following steps, a self-signed certificate is used as an example to describe how to create a server certificate.

1. Log in to the server where the CA certificate is generated.
2. Create a directory at the same level as the directory of the CA certificate and switch to the directory.

```
mkdir server  
cd server
```

3. Create the certificate configuration file **server_cert.conf**. The file content is as follows:

```
[ req ]
distinguished_name = req_distinguished_name
```

```
prompt          = no
[ req_distinguished_name ]
O               = ELB
CN             = www.test.com
```

NOTE

Set the **CN** field to the domain name or IP address of the Linux server.

4. Create the server certificate private key **server.key**.
openssl genrsa -out server.key 2048
5. Create the CSR file **server.csr** for the server certificate.
openssl req -out server.csr -key server.key -new -config ./server_cert.conf
6. Use the CA certificate to issue the server certificate **server.crt**.
openssl x509 -req -in server.csr -out server.crt -sha1 -CAcreateserial -days 5000 -CA ../ca/ca.crt -CAkey ../ca/ca.key

Figure 3-5 Issuing a server certificate

```
[root@elbv30003 server]# openssl x509 -req -in server.csr -out server.crt -sha1 -CAcreateserial -days 5000 -CA ../ca/ca.crt -CAkey ../ca/ca.key
Signature ok
subject=O = ELB, CN = www.test.com
Getting CA Private Key
[root@elbv30003 server]#
```

Issuing a Client Certificate Using the CA Certificate

1. Log in to the server where the CA certificate is generated.
2. Create a directory at the same level as the directory of the CA certificate and switch to the directory.

mkdir client

cd client

3. Create the certificate configuration file **client_cert.conf**. The file content is as follows:

```
[ req ]
distinguished_name = req_distinguished_name
prompt            = no

[ req_distinguished_name ]
O               = ELB
CN             = www.test.com
```

NOTE

Set the **CN** field to the domain name or IP address of the Linux server.

4. Create the client certificate private key **client.key**.
openssl genrsa -out client.key 2048

Figure 3-6 Creating a client certificate private key

```
[root@elbv30003 client]# openssl genrsa -out client.key 2048
Generating RSA private key, 2048 bit long modulus (2 primes)
.....+++++
.....+++++
e is 65537 (0x010001)
[root@elbv30003 client]#
```

5. Create the CSR file **client.csr** for the client certificate.

```
openssl req -out client.csr -key client.key -new -config ./client_cert.conf
```

Figure 3-7 Creating a client certificate CSR file

```
[root@lbv30003 client]# openssl req -out client.csr -key client.key -new -config ./client_cert.conf
```

6. Use the CA certificate to issue the client certificate **client.crt**.

```
openssl x509 -req -in client.csr -out client.crt -sha1 -CAcreateserial -days 5000 -CA ../ca/ca.crt -CAkey ../ca/ca.key
```

Figure 3-8 Issuing a client certificate

```
[root@lbv30003 client]# openssl x509 -req -in client.csr -out client.crt -sha1 -CAcreateserial -days 5000 -CA ../ca/ca.crt -CAkey ../ca/ca.key
Signature ok
subject=D = ELB, CN = www.test.com
Getting CA Private Key
[root@lbv30003 client]#
```

7. Convert the client certificate to a **.p12** file that can be identified by the browser.

```
openssl pkcs12 -export -clcerts -in client.crt -inkey client.key -out client.p12
```

 **NOTE**

A password is required during command execution. Save this password, which will be required when you import the certificate using the browser.

Configuring the Server Certificate and Private Key

1. Log in to the load balancer management console.
2. In the navigation pane on the left, choose **Certificates**.
3. In the navigation pane on the left, choose **Certificates**. On the displayed page, click **Add Certificate**. In the **Add Certificate** dialog box, select **Server certificate**, copy the content of server certificate **server.crt** to the **Certificate Content** area and the content of private key file **server.key** to the **Private Key** area, and click **OK**.

 **NOTE**

Delete the last newline character before you copy the content.

 **NOTE**

The certificate and private key must be PEM-encoded.

Configuring the CA Certificate

- Step 1** Log in to the load balancer management console.
- Step 2** In the navigation pane on the left, choose **Certificates**.
- Step 3** Click **Add Certificate**. In the **Add Certificate** dialog box, select **CA certificate**, copy the content of CA certificate **ca.crt** created in [Creating a CA Certificate Using OpenSSL](#) to the **Certificate Content** area, and click **OK**.

 **NOTE**

Delete the last newline character before you copy the content.

 **NOTE**

The certificate must be PEM-encoded.

----End

Configuring Mutual Authentication

1. Log in to the load balancer management console.
2. Locate the load balancer and click its name. Under **Listeners**, click **Add Listener**. Select **HTTPS** for **Frontend Protocol** and **Mutual authentication** for **SSL Authentication**, and select a CA certificate and server certificate.

Add backend servers.

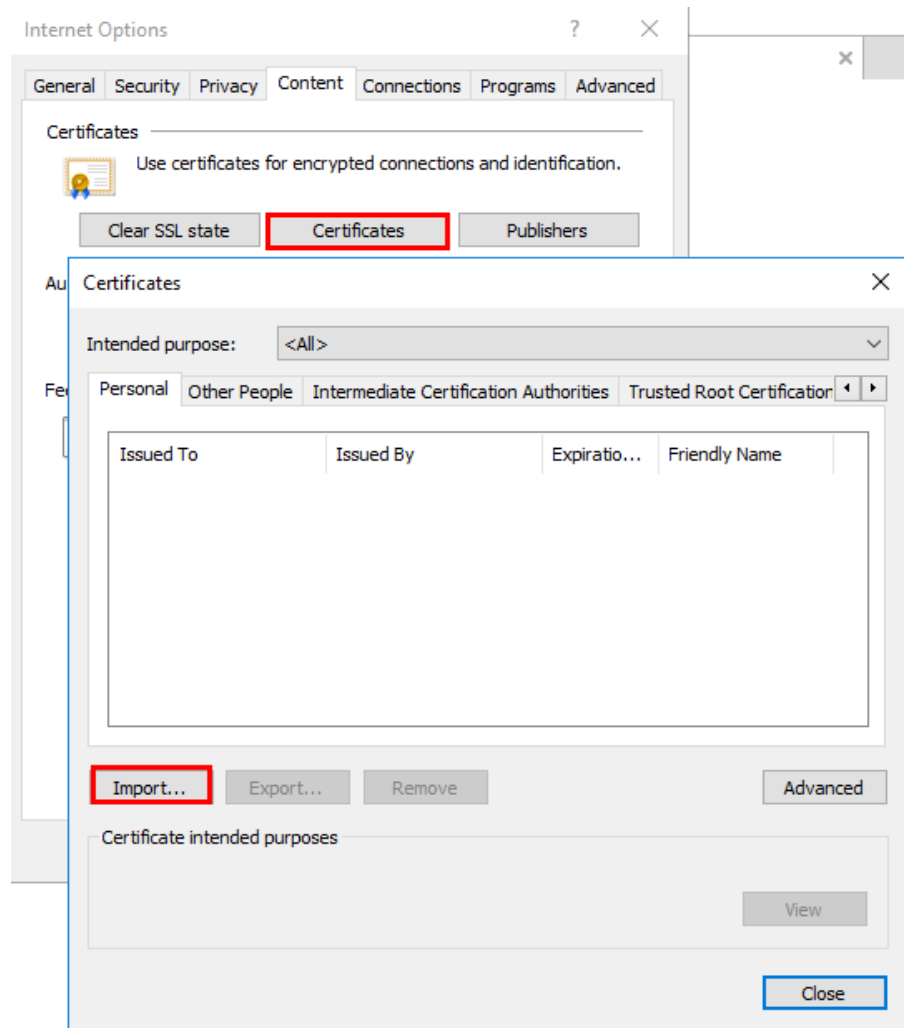
For detailed operations, see [Overview](#).

Importing and Testing the Client Certificate

Method 1: Using a browser

1. Import the client certificate using a browser (Internet Explorer 11 is used as an example).
 - a. Export **client.p12** from the Linux server.
 - b. Open the browser, choose **Settings > Internet Options** and click **Content**.
 - c. Click **Certificates** and then **Import** to import the **client.p12** certificate.

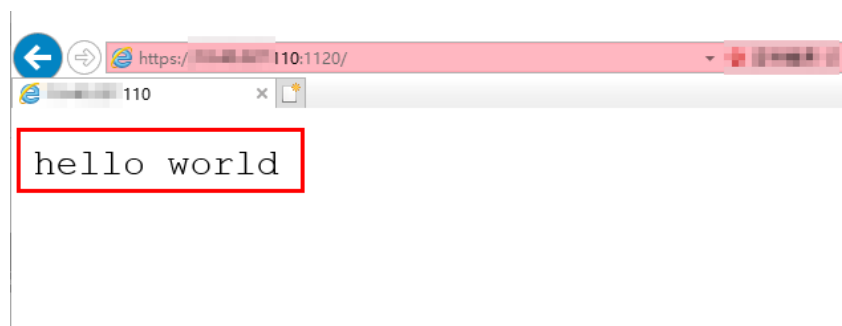
Figure 3-9 Importing the `client.p12` certificate



2. Verify the import.

Enter the access address in the address box of your browser. A window is displayed asking you to select the certificate. Select the client certificate and click **OK**. If the website can be accessed, the certificate is successfully imported.

Figure 3-10 Accessing the website



Method 2: Using cURL

1. Import the client certificate.
Copy client certificate **client.crt** and private key **client.key** to a new directory, for example, **/home/client_cert**.
2. Verify the import.
On the Shell screen, run the following command:

```
curl -k --cert /home/client_cert/client.crt --key /home/client_cert/client.key https://  
XXX.XXX.XXX.XXX:XXX/ -I
```


Ensure that the certificate address, private key address, IP address and listening port of the load balancer are correct. Replace **https://
XXX.XXX.XXX.XXX:XXX** with the actual IP address and port number. If the expected response code is returned, the certificate is successfully imported.

Figure 3-11 Example of a correct response code

```
[192.168.10.216 test]#curl -k --cert client.crt --key client.key https://192.168.10.16:4500 -I  
HTTP/1.1 200 OK  
Date: Fri, 25 Sep 2020 10:11:17 GMT  
Content-Type: application/octet-stream  
Connection: keep-alive  
Set-Cookie: name=d92f80b6-55e9-4b61-9c37-932ccd7b02f2; path=/; Expires=Sat, 26-Sep-20 10:11:19 GMT  
Server: elb
```

3.5 HTTP/2



Scenarios

Hypertext Transfer Protocol 2.0 (HTTP/2) is the next-generation HTTP protocol. HTTP/2 is used to secure connections between the load balancer and clients. You can enable HTTP/2 when you add HTTPS listeners. If you have already added an HTTPS listener, you can also enable this function.



Constraints

You can enable HTTP/2 only for HTTPS listeners.

Enabling HTTP/2 When Adding a Listener

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Under **Listeners**, click **Add Listener**.
6. In the **Add Listener** dialog box, set **Frontend Protocol** to **HTTPS**.
7. Expand **Advanced Settings** and enable HTTP/2.
8. Confirm the configurations and click **Submit**.

Enabling or Disabling HTTP/2 When Modifying a Listener

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click its name.
6. On the **Summary** tab page, click **Edit** on the top right.
7. In the **Edit** dialog box, expand **Advanced Settings** and enable or disable HTTP/2.
8. Click **OK**.

3.6 HTTP Redirection to HTTPS

Scenarios

HTTPS is an extension of HTTP. HTTPS encrypts data between a web server and a browser.

If you enable redirection, all HTTP requests to your website are transmitted over HTTPS connections to improve security.


CAUTION

- If the listener protocol is HTTP, only the GET or HEAD method can be used for redirection. If you create a redirect for an HTTP listener, the client browser will change POST or other methods to GET. If you want to use other methods rather than GET and HEAD, add an HTTPS listener.
 - HTTP requests are forwarded to the HTTPS listener as HTTPS requests, which are then routed to backend servers over HTTP.
 - If an HTTP listener is redirected to an HTTPS listener, no certificate can be deployed on the backend servers associated with the HTTPS listener. If certificates are deployed, HTTPS requests will not take effect.
-

Prerequisites

- You have added an HTTPS listener.
- You have added an HTTP listener.

Creating Redirection to HTTPS

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.


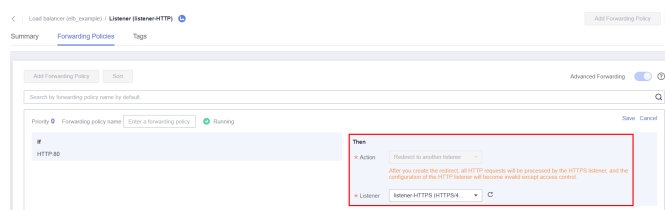
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Listeners**, locate the HTTP listener, and click its name.
6. On the **Forwarding Policies** tab page, click **Add Forwarding Policy**.

Table 3-12 Configuring parameters for redirection

Parameter	Setting
Action	Select Redirect to another listener .
Listener	Select the HTTPS listener to which requests are redirected.

7. After the forwarding policy is added, click **Save**.



Figure 3-12 Redirection to an HTTPS listener



 **NOTE**



- If requests to an HTTP listener are redirected, the listener will become invalid, but access control to the listener will still take effect.
- If you create a redirect for an HTTP listener, the backend server will return HTTP 301 Move Permanently to the clients.

Modifying Redirection to HTTPS

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Listeners**, locate the HTTP listener, and click its name.
6. On the **Forwarding Policies** tab page, locate the target forwarding policy and click **Edit**.
7. You can change the HTTPS listener to which requests are redirected as required.

8. Click **Save**.

Deleting Redirection to HTTPS

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click its name.
6. On the **Forwarding Policies** tab page, click **Delete** on the right of the target forwarding policy.
7. In the displayed dialog box, click **Yes**.

3.7 HTTP Headers

HTTP headers are a list of strings sent and received by both the client and server on every Hypertext Transfer Protocol (HTTP) request and response. This section describes HTTP headers supported by HTTP and HTTPS listeners.

Table 3-13 Transfer headers

Header	Feature	Description	Dedicated Load Balancers	Shared Load Balancers
X-Forwarded-ELB-IP	Transfer Load Balancer EIP	If this option is enabled, the EIP bound to the load balancer will be transmitted to backend servers through the X-Forwarded-ELB-IP header. The format is as follows (XX.XXX.XX.XXX indicates the EIP of the load balancer): X-Forwarded-ELB-IP: XX.XXX.XX.XXX	√	√
X-Forwarded-ELB-ID	Transfer Load Balancer ID	If this option is enabled, the load balancer ID will be transmitted to backend servers through the X-Forwarded-ELB-ID header.	√	×
X-Forwarded-Port	Transfer Listener Port Number	If this option is enabled, the port number used by the listener will be transmitted to backend servers through the X-Forwarded-Port header.	√	×

Header	Feature	Description	Dedicated Load Balancers	Shared Load Balancers
X-Forwarded-For-Port	Transfer Port Number in the Request	If this option is enabled, the port number used by the client will be transmitted to backend servers through the X-Forwarded-For-Port header.	√	×



Table 3-14 Rewrite headers

Header	Feature	Description	Dedicated Load Balancers	Shared Load Balancers
X-Forwarded-Host	Rewrite X-Forwarded-Host	<ul style="list-style-type: none">• If this option is enabled, the Host header of the client request will be rewritten into the X-Forwarded-Host header and transmitted to the backend servers.• If this option is disabled, the X-Forwarded-Host header of the client will be transmitted to the backend servers.	√	√
X-Forwarded-Proto	Rewrite X-Forwarded-Proto	<ul style="list-style-type: none">• If this option is enabled, the listener protocol will be rewritten into the X-Forwarded-Proto header field and transmitted to the backend servers.• If this option is disabled, the protocol used by the client will be transmitted to the backend servers through the X-Forwarded-Proto header.	√	×
X-Real-IP	Rewrite X-Real-IP	<ul style="list-style-type: none">• If this option is enabled, the source IP address of the client will be rewritten into the X-Real-IP header and transmitted to the backend servers.• If this option is disabled, the X-Real-IP header of the client will be transmitted to the backend servers.	√	×



 NOTE

- √ indicates the load balancer supports the header, whereas × indicates the load balancer does not support the header.

Enabling HTTP/HTTPS Headers

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. You can enable these header features in either of the following ways:
 - On the **Load Balancers** page, locate the load balancer and click its name. Under **Listeners**, click **Add Listener**.
 - On the **Load Balancers** page, locate the load balancer and click **Add Listener** in the **Operation** column.
5. On the **Configure Listener** page, expand **Advanced Settings** and enable the features as needed.
6. Configure the listener as prompted.
7. Confirm the configuration and click **Submit**.

Modifying HTTP Header Features

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer and click its name.
5. Click the **Listeners** tab, locate the target listener and click **Edit** in **Operation** column.
6. On the displayed page, expand **Advanced Settings** and enable or disable the features.
7. Click **OK**.

3.8 SNI Certificate

Scenarios

If you have an application that can be accessed through multiple domain names and each domain name uses a different certificate, you can enable Server Name Indication (SNI) when you add an HTTPS listener.

SNI, an extension to Transport Layer Security (TLS), enables a server to present multiple certificates on the same IP address and port number. SNI allows the client

to indicate the domain name of the website while sending an SSL handshake request. Once receiving the request, the load balancer queries the right certificate based on the hostname or domain name and returns the certificate to the client. If no certificate is found, the load balancer will return the default certificate.

You can enable SNI only when you add HTTPS listeners. Load balancers can have multiple SNI certificates bound.

Constraints

An HTTPS listener can have up to 30 SNI certificates. All the certificates can have up to 30 domain names.

NOTE

Listeners of a dedicated load balancer can have up to 50 SNI certificates. You can [submit a service ticket](#) to increase the quota.



Prerequisites

- You have added an HTTPS listener to the load balancer by performing the operations in [Adding an HTTPS Listener](#).
- You have created an SNI certificate by performing the operations in [Adding a Certificate](#).

NOTE

- You need to specify a domain name for an SNI certificate. The domain name must be the same as that in the certificate.
- A domain name can be used by both an ECC certificate and an RSA certificate. If there are two SNI certificates that use the same domain name, the ECC certificate is displayed preferentially.
- Domain names in an SNI certificate are matched as follows:
If the domain name of the certificate is *.test.com, a.test.com and b.test.com are supported, and a.b.test.com and c.d.test.com are not supported.
The domain name with the longest suffix is matched: If a certificate contains both *.b.test.com and *.test.com, a.b.test.com preferentially matches *.b.test.com.
- If a certificate has expired, you need to manually replace or delete it by following the instructions in [Replacing the Certificate Bound to a Listener](#).

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click its name.
6. On the **Summary** tab page, click **Edit** on the top right.
7. Enable SNI and select an SNI certificate.

8. Click **OK**.

4 Backend Server Group

4.1 Overview

Introduction

A backend server group is a logical collection of one or more backend servers to receive massive concurrent requests at the same time. A backend server can be an ECS, BMS, supplementary network interface, or IP address.

The following process describes how a backend server group forwards traffic:

1. A client sends a request to your application. The listeners added to your load balancer use the protocols and ports you have configured forward the request to the associated backend server group.
2. Healthy backend servers in the backend server group receive the request based on the load balancing algorithm, handle the request, and return a result to the client.
3. In this way, massive concurrent requests can be processed at the same time, improving the availability of your applications.

For dedicated load balancers, the backend server group type can be **Hybrid** or **IP as a backend server**. You can add an ECS, BMS, supplementary network interface, or IP address to a hybrid backend server group. If you set the type to **IP as a backend server**, you can only add IP addresses as backend servers.

Shared load balancers have only one type of backend server group, where you can only add cloud servers.

Figure 4-1 shows the architecture of different types of backend server groups.

Figure 4-1 Backend server group architecture

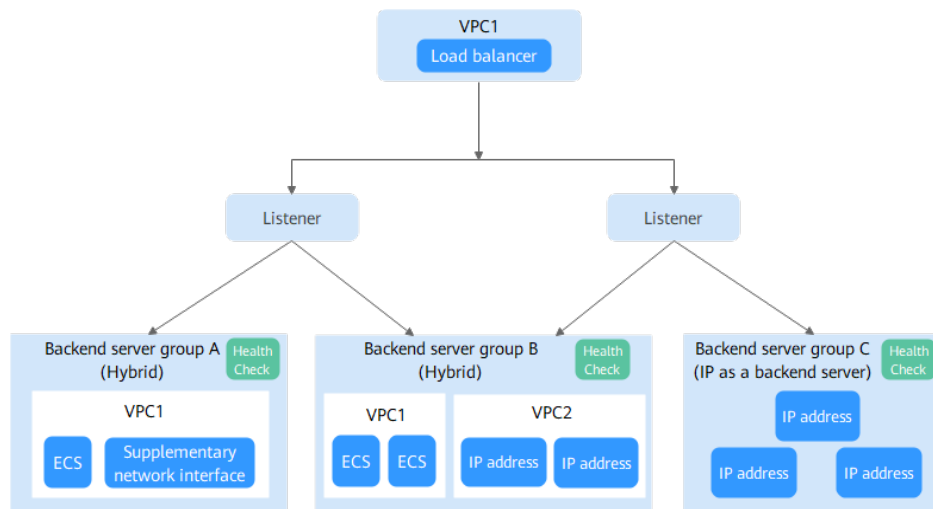


Table 4-1 Backend server group types

Backend Server Group Type	Backend Server Type	Example	Reference
Hybrid	<ul style="list-style-type: none"> ECSs, BMSs, or supplementary network interfaces that are in the same VPC as the load balancer Cloud servers in other VPCs or on-premises servers if IP as a backend is enabled for the load balancer 	<p>As shown in Figure 4-1:</p> <ul style="list-style-type: none"> In backend server group A, you can add servers or supplementary network interfaces in VPC1. In backend server group B, you can add IP addresses in VPC2 as backend servers. 	<ul style="list-style-type: none"> Adding Backend Servers Adding Supplementary Network Interfaces Adding IP Addresses as Backend Servers
IP as a backend server	Cloud servers in other VPCs or on-premises servers if IP as a backend is enabled for the load balancer	As shown in Figure 4-1 , IP addresses can be added to backend server group C as backend servers.	Adding IP Addresses as Backend Servers

Advantages

Backend server groups can bring the following benefits:

- **Reduced costs and easier management:** You can add or remove backend servers as traffic changes over the time. This can help avoid low resource utilization and makes it easy to manage backend servers.
- **Higher reliability:** Traffic is routed only to healthy backend servers in the backend server group.

Key Functions

You can configure the key functions listed in [Table 4-2](#) for each backend server group to ensure service stability.

Table 4-2 Key functions

Key Function	Description	Detail
Health Check	Specifies whether to enable the health check option. Health checks determine whether backend servers are healthy. If a backend server is detected unhealthy, it will not receive requests from the associated load balancer, improving your service reliability.	Health Check
Load Balancing Algorithm	The load balancer distributes traffic based on the load balancing algorithm you have configured for the backend server group.	Load Balancing Algorithms
Sticky Session	Specifies whether to enable the sticky session option. If you enable this option, all requests from a client during one session are sent to the same backend server.	Sticky Session
Slow Start	Specifies whether to enable slow start. After you enable it, the load balancer linearly increases the proportion of requests to backend servers in this mode. When the slow start duration elapses, the load balancer sends full share of requests to backend servers and exits the slow start mode. NOTE Slow start is only available for HTTP and HTTPS backend server groups of dedicated load balancers.	Slow Start (Dedicated Load Balancers)

Precautions for Creating a Backend Server Group

The backend protocol of the new backend server group must match the frontend protocol of the listener as described in [Table 4-3](#).

You can create a backend server group by referring to [Table 4-4](#).

Table 4-3 The frontend and backend protocol

Frontend Protocol	Backend Protocol
TCP	TCP
UDP	<ul style="list-style-type: none">• UDP• QUIC
HTTP	HTTP
HTTPS	<ul style="list-style-type: none">• HTTP• HTTPS

Table 4-4 Creating a backend server group

Load Balancer Type	Reference
Dedicated	Creating a Backend Server Group (Dedicated Load Balancers)
Shared	Creating a Backend Server Group (Shared Load Balancers)

4.2 Key Features

4.2.1 Health Check

ELB periodically sends requests to backend servers to check whether they can process requests. This process is called health check.

If a backend server is detected unhealthy, the load balancer will stop route requests to it. After the backend server recovers, the load balancer will resume routing requests to it.

If backend servers have to handle large number of requests, frequent health checks may overload the backend servers and cause them to respond slowly. To address this problem, you can prolong the health check interval or use TCP or UDP instead of HTTP. You can also disable health check. If you choose to disable health check, requests may be routed to unhealthy servers, and service interruptions may occur.

Health Check Protocol

You can configure health checks when configuring backend server groups. Generally, you can use the default setting or select a different health check protocol as you need.

If you want to modify health check settings, see details in [Modifying Health Check Settings](#).

Select a health check protocol that matches the backend protocol as described in [Table 4-5](#) and [Table 4-6](#).

Table 4-5 The backend protocol and health check protocols (dedicated load balancers)

Backend Protocol	Health Check Protocol
TCP	TCP, HTTP, or HTTPS
UDP	UDP
QUIC	UDP
HTTP	TCP, HTTP, or HTTPS
HTTPS	TCP, HTTP, or HTTPS

Table 4-6 The backend protocol and health check protocols (shared load balancers)

Backend Protocol	Health Check Protocol
TCP	TCP or HTTP
UDP	UDP
HTTP	TCP or HTTP
HTTPS	TCP or HTTP

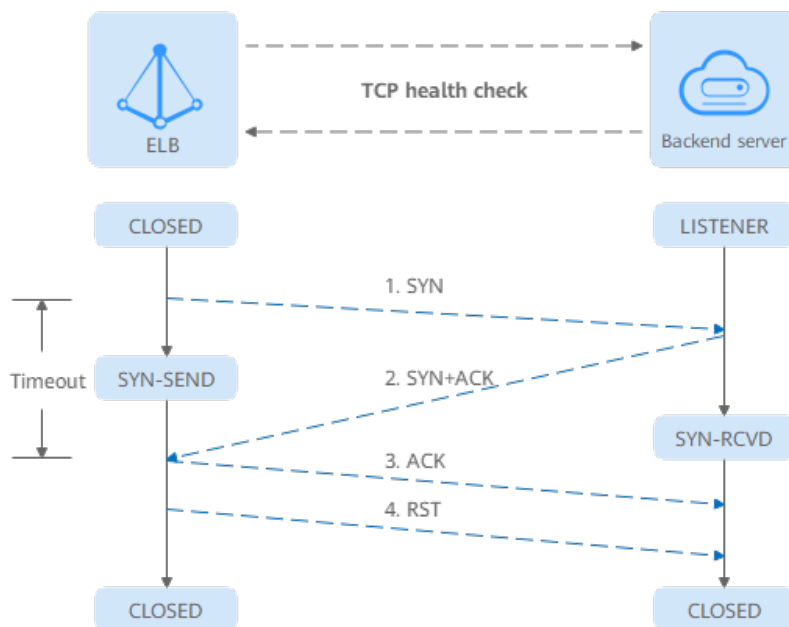
Health Check Source IP Address

- A dedicated load balancer uses the IP addresses in its backend subnet to send requests to backend servers and verify their health status. To perform health checks, ensure that the security group rules of the backend servers allow access from the backend subnet where the load balancer is running. For details, see [Security Group Rules](#).
- A shared load balancer uses an IP address in 100.125.0.0/16 to send requests to backend servers and verify their health status. To perform health checks, ensure that the security group rules of the backend server allow access from 100.125.0.0/16. For details, see [Security Group Rules](#).

TCP Health Check

For TCP, HTTP, and HTTPS backend protocols, you can use TCP to initiate three-way handshakes to obtain the statuses of backend servers.

Figure 4-2 TCP health check



The TCP health check process is as follows:

1. The load balancer sends a TCP SYN packet to the backend server (in the format of $\{Private\ IP\ address\}:\{Health\ check\ port\}$).
2. The backend server returns an SYN-ACK packet.
 - If the load balancer does not receive the SYN-ACK packet within the timeout duration, it declares that the backend server is unhealthy and sends an RST packet to the backend server to terminate the TCP connection.
 - If the load balancer receives the SYN-ACK packet from the backend server within the timeout duration, it sends an ACK packet to the backend server and declares that the backend server is healthy. After that, the load balancer sends an RST packet to the backend server to terminate the TCP connection.

NOTICE

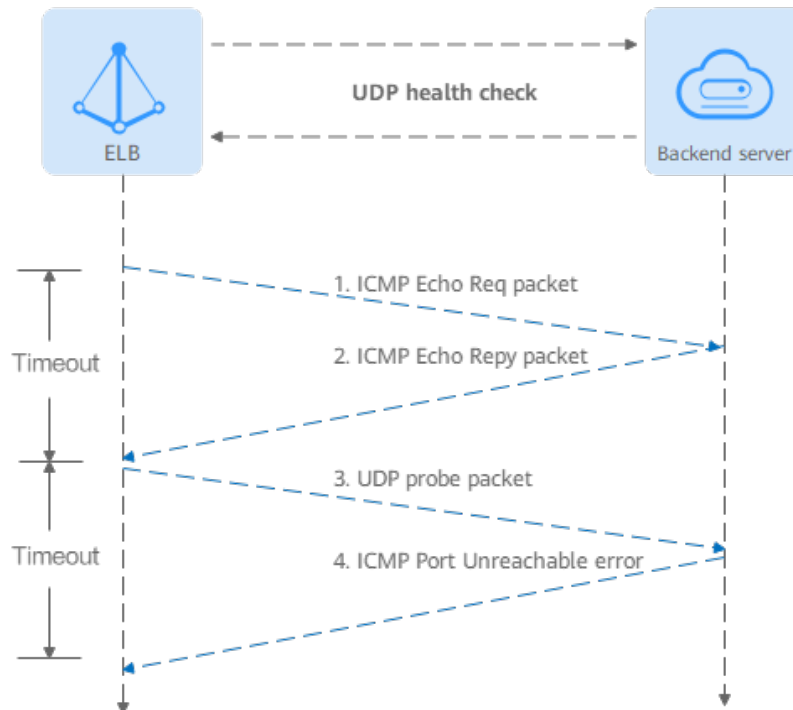
After a successful TCP three-way handshake, an RST packet will be sent to close the TCP connection. The application on the backend server may consider this packet a connection error and reply with a message, for example, "Connection reset by peer". To avoid this issue, take either of the following actions:

- Use [HTTP Health Check](#).
- Have the backend server ignore the connection error.

UDP Health Check

For UDP backend protocol, ELB sends ICMP and UDP probe packets to backend servers to check their health.

Figure 4-3 UDP health check



The UDP health check process is as follows:

1. The load balancer sends an ICMP Echo Request packet to the backend server.
 - If the load balancer does not receive an ICMP Echo Reply packet within the health check timeout duration, the backend server is declared unhealthy.
 - If the load balancer receives an ICMP Echo Reply packet within the timeout period, it sends a UDP probe packet to the backend server.
2. If the load balancer does not receive an ICMP Port Unreachable error within the health check timeout duration, it declares the backend server is healthy. If the load balancer receives an ICMP Port Unreachable error, the backend server is declared unhealthy.

HTTP Health Check

You can also configure HTTP health checks to obtain server statuses through HTTP GET requests if you select TCP, HTTP, or HTTPS as the backend protocol. [Figure 4-4](#) shows how an HTTP health check works.

Figure 4-4 HTTP health check



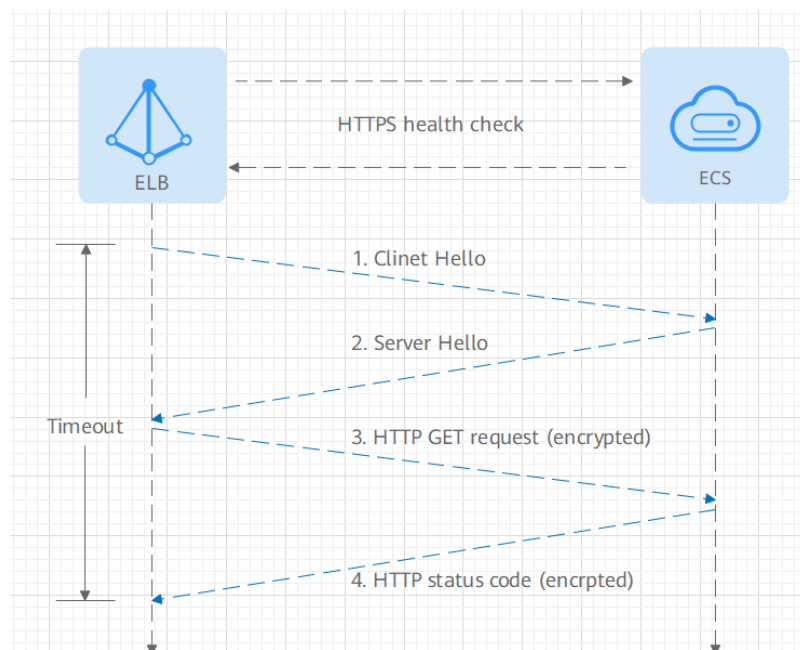
The HTTPS health check process is as follows:

1. The load balancer sends an HTTP GET request to the backend server (in format of $\{Private\ IP\ address\}:\{Health\ check\ port\}\{Health\ check\ path\}$). (You can specify a domain name when configuring a health check.)
2. The backend server returns an HTTP status code to ELB.
 - If the load balancer receives the status code within the health check timeout duration, it compares the status code with the preset one. If the status codes are the same, the backend server is declared healthy.
 - If the load balancer does not receive any response from the backend server within the health check timeout duration, it declares the backend server is unhealthy.

HTTPS Health Check

For TCP, HTTP, and HTTPS backend protocols, you can use HTTPS to establish an SSL connection over TLS handshakes to obtain the statuses of backend servers. [Figure 4-5](#) shows how an HTTPS health check works.

Figure 4-5 HTTPS health check



The HTTPS health check process is as follows:

1. The load balancer sends a Client Hello packet to establish an SSL connection with the backend server.
2. After receiving the Server Hello packet from the backend server, the load balancer sends an encrypted HTTP GET request to the backend server (in the format of *{Private IP address}:{Health check port}/{Health check path}*). (You can specify a domain name when configuring a health check.)
3. The backend server returns an HTTP status code to the load balancer.
 - If the load balancer receives the status code within the health check timeout duration, it compares the status code with the preset one. If the status codes are the same, the backend server is declared healthy.
 - If the load balancer does not receive any response from the backend server within the health check timeout duration, it declares the backend server is unhealthy.

Health Check Time Window

Health checks greatly improve service availability. However, if health checks are too frequent, service availability will be compromised. To avoid the impact, ELB declares a backend server healthy or unhealthy after several consecutive health checks.

The health check time window is determined by the factors in [Table 4-7](#):

Table 4-7 Factors affecting the health check time window

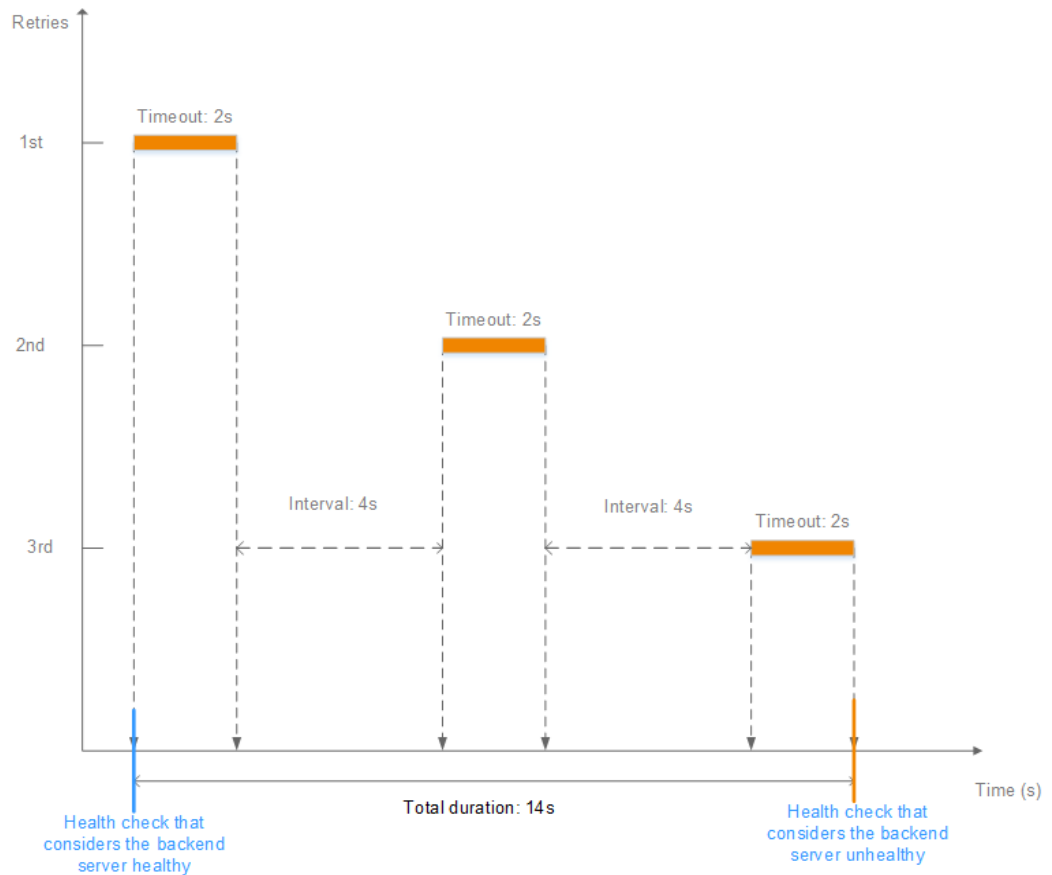
Factor	Description
Check Interval	How often health checks are performed.
Timeout Duration	How long the load balancer waits for the response from the backend server.
Health Check Threshold	The number of consecutive successful or failed health checks required for determining whether the backend server is healthy or unhealthy.

The following is a formula for you to calculate the health check time window:

- Time window for a backend server to be detected healthy = Timeout duration x Healthy threshold + Interval x (Healthy threshold - 1)
- Time window for a backend server to be detected unhealthy = Timeout duration x Unhealthy threshold + Interval x (Unhealthy threshold - 1)

As shown in [Figure 4-6](#), if the health check interval is 4s, the health check timeout duration is 2s, and unhealthy threshold is 3, the time window for a backend server to be considered unhealthy is calculated as follows: $2 \times 3 + 4 \times (3 - 1) = 14s$.

Figure 4-6 Health check timeout duration



Rectifying an Unhealthy Backend Server

If a backend server is detected unhealthy, see [How Do I Troubleshoot an Unhealthy Backend Server?](#)

4.2.2 Load Balancing Algorithms

Overview

Load balancers receive requests from clients and forward them to backend servers in one or more AZs. Each load balancer has at least a listener and a backend server. The load balancing algorithm you select when you create the backend server group determines how requests are distributed.

ELB supports the following load balancing algorithms: weighted round robin, weighted least connections, source IP hash, and connection ID.

You can select the load balancing algorithm that best suits your needs.

Table 4-8 Load balancing algorithms

Load Balancing Algorithm	Description
Weighted round robin	Routes requests to backend servers in sequence based on their weights.
Weighted least connections	Routes requests to backend servers with the smallest connections-to-weight ratio.
Consistent hashing <ul style="list-style-type: none"> Source IP hash Connection ID 	<p>Calculates the request fields using the consistent hashing algorithm to obtain a hash value and routes requests with the same hash value to the same backend server, even if the number of backend servers in the backend server group changes.</p> <ul style="list-style-type: none"> Source IP hash: Calculates the source IP address of each request and routes requests from the same source IP address to the same backend server. Connection ID: Calculates the QUIC connection ID and routes requests with the same ID to the same backend server.

Weighted Round Robin

Figure 4-7 shows an example of how requests are distributed using the weighted round robin algorithm. Two backend servers are in the same AZ and have the same weight, and each server receives the same proportion of requests.

Figure 4-7 Traffic distribution using the weighted round robin algorithm

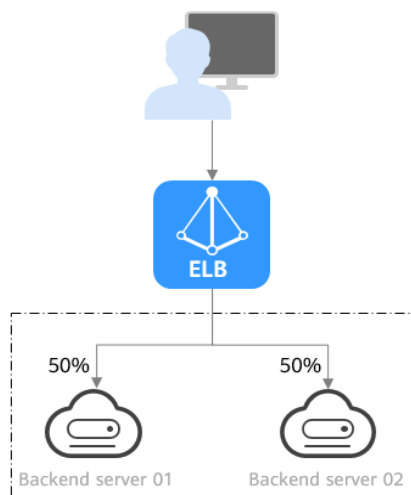


Table 4-9 Weighted round robin

Description	Requests are routed to backend servers based on their weights. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests.
When to Use	<p>This algorithm is typically used for short connections, such as HTTP connections.</p> <ul style="list-style-type: none"> Flexible load balancing: When you need more refined load balancing, you can set a weight for each backend server to specify the percentage of requests to each server. For example, you can set higher weights to backend servers with better performance so that they can process more requests. Dynamic load balancing: You can adjust the weight of each backend server in real time when the server performance or load fluctuates.
Disadvantages	<ul style="list-style-type: none"> You need to set a weight for each backend server. If you have a large number of backend servers or your services require frequent adjustments, setting weights would be time-consuming. If the weights are inappropriate, the requests processed by each server may be imbalanced. As a result, you may need to frequently adjust server weights.

Weighted Least Connections

Figure 4-8 shows an example of how requests are distributed using the weighted least connections algorithm. Two backend servers are in the same AZ and have the same weight, 100 connections have been established with backend server 01, and 50 connections have been established with backend server 02. New requests are preferentially routed to backend server 02.

Figure 4-8 Traffic distribution using the weighted least connections algorithm

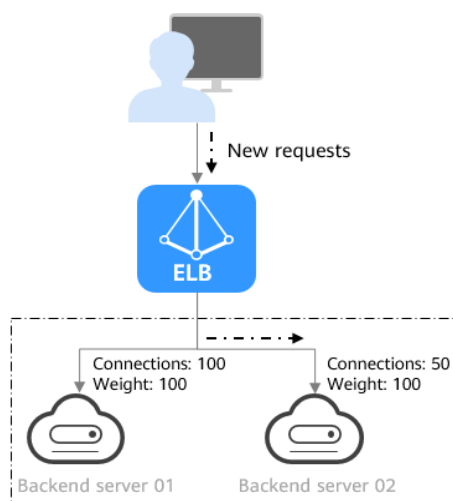


Table 4-10 Weighted least connections

Description	In addition to the number of active connections established with each backend server, each server is assigned a weight based on their processing capability. Requests are routed to the server with the lowest connections-to-weight ratio.
When to Use	<p>This algorithm is often used for persistent connections, such as connections to a database.</p> <ul style="list-style-type: none">• Flexible load balancing: Load balancers distribute requests based on the number of established connections and the weight of each backend server and route requests to the server with the lowest connections-to-weight ratio. This helps prevent servers from being underloaded or overloaded.• Dynamic load balancing: When the number of connections to and loads on backend servers change, you can use the weighted least connection algorithm to dynamically adjust the requests distributed to each server in real time.• Stable load balancing: You can use this algorithm to reduce the peak loads on each backend server and improve service stability and reliability.
Disadvantages	<ul style="list-style-type: none">• Complex calculation: The weighted least connections algorithm needs to calculate and compare the number of connections established with each backend server in real time before selecting a server to route requests.• Dependency on connections to backend servers: The algorithm routes requests based on the number of connections established with each backend server. If monitoring data is inaccurate or outdated, requests may not be distributed evenly across backend servers. The algorithm can only collect statistics on the connections between a given load balancer and a backend server, but cannot obtain the total number of connections to the backend server if it is associated with multiple load balancers.• Too much loads on new servers: If existing backend servers have to handle a large number of requests, new requests will be routed to new backend servers. This may deteriorate new servers or even cause them to fail.

Source IP Hash

Figure 4-9 shows an example of how requests are distributed using the source IP hash algorithm. Two backend servers are in the same AZ and have the same weight. If backend server 01 has processed a request from IP address A, the load balancer will route new requests from IP address A to backend server 01.

Figure 4-9 Traffic distribution using the source IP hash algorithm

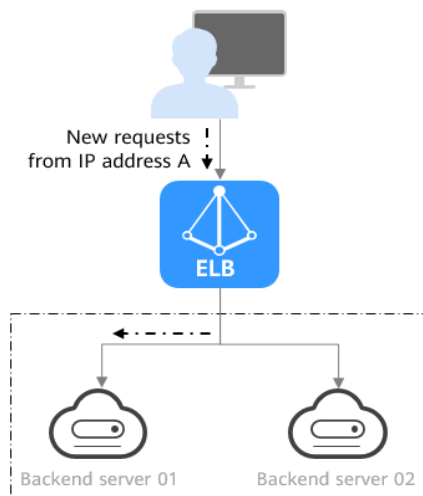


Table 4-11 Source IP hash

Description	The source IP hash algorithm calculates the source IP address of each request and routes requests from the same IP address to the same backend server.
When to Use	<p>This algorithm is often used for applications that need to maintain user sessions or state.</p> <ul style="list-style-type: none"> • Session persistence: Source IP hash ensures that requests with the same source IP address are distributed to the same backend server. • Data consistency: Requests with the same hash value are distributed to the same backend server. • Load balancing: In scenarios that have high requirements for load balancing, this algorithm can distribute requests to balance loads among servers.
Disadvantages	<ul style="list-style-type: none"> • Imbalanced loads across servers: This algorithm tries its best to ensure request consistency when backend servers are added or removed. If the number of backend servers decreases, some requests may be redistributed, causing imbalanced loads across servers. • Complex calculation: This algorithm calculates the hash values of requests based on hash factors. If servers are added or removed, some requests may be redistributed, making calculation more difficult.

Connection ID

Figure 4-10 shows an example of how requests are distributed using the connection ID algorithm. Two backend servers are in the same AZ and have the same weight. If backend server 01 has processed a request from client A, the load balancer will route new requests from client A to backend server 01.

Figure 4-10 Traffic distribution using the connection ID algorithm

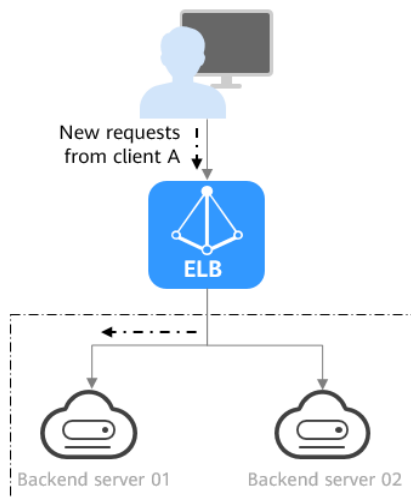


Table 4-12 Connection ID

Description	<p>The connection ID algorithm calculates the QUIC connection ID and routes requests with the same ID to the same backend server. A QUIC ID identifies a QUIC connection. This algorithm distributes requests by QUIC connection.</p> <p>You can use this algorithm to distribute requests only to QUIC backend server groups.</p>
When to Use	<p>This algorithm is typically used for QUIC requests.</p> <ul style="list-style-type: none"> • Session persistence: The connection ID algorithm ensures that requests with the same QUIC ID are distributed to the same backend server. • Data consistency: Requests with the same hash value are distributed to the same backend server. • Load balancing: In scenarios that have high requirements for load balancing, this algorithm can distribute requests to balance loads among servers.
Disadvantages	<ul style="list-style-type: none"> • Imbalanced loads across servers: This algorithm tries its best to ensure request consistency when backend servers are added or removed. If the number of backend servers decreases, some requests may be redistributed, causing imbalanced loads across servers. • Complex calculation: This algorithm calculates the hash values of requests based on hash factors. If servers are added or removed, some requests may be redistributed, making calculation more difficult.

4.2.3 Sticky Session

Sticky sessions ensure that requests from a client always get routed to the same backend server before a session elapses.

Here is an example that describes how sticky session works. Assume that you have logged in to a server. After a while, you send another request. If sticky sessions are not enabled, the request may be routed to another server, and you will be asked to log in again. If sticky sessions are enabled, all your requests are processed by the same server, and you do not need to repeatedly log in.

Differences Between Sticky Sessions at Layer 4 and Layer 7

The following table describes the differences of sticky sessions at Layer 4 at Layer 7.

Table 4-13 Sticky session comparison

OSI Layer	Listener Protocol	Sticky Session Type	Stickiness Duration	Scenarios Where Sticky Sessions Become Invalid
Layer 4	TCP or UDP	Source IP address: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hashing key, and all backend servers are numbered. The system allocates the client to a particular server based on the generated key. This allows requests from the same IP address are forwarded to the same backend server.	<ul style="list-style-type: none">• Default: 20 minutes• Maximum: 60 minutes• Range: 1 minute to 60 minutes	<ul style="list-style-type: none">• Source IP addresses of the clients change.• The session stickiness duration has been reached.

OSI Layer	Listener Protocol	Sticky Session Type	Stickiness Duration	Scenarios Where Sticky Sessions Become Invalid
Layer 7	HTTP or HTTPS	<ul style="list-style-type: none"> • Load balancer cookie: The load balancer generates a cookie after receiving a request from the client. All subsequent requests with the cookie are routed to the same backend server. The load balancer itself does not generate cookies. • Application cookie: The application deployed on the backend server generates a cookie after receiving the first request from the client. All subsequent requests with the same cookie are routed to the same backend server. 	<ul style="list-style-type: none"> • Default: 20 minutes • Maximum: 1,440 minutes • Range: 1 minute to 1,440 minutes 	<ul style="list-style-type: none"> • If requests sent by the clients do not contain a cookie, sticky sessions will not take effect. • Requests from the clients exceed the session stickiness duration.

 **NOTE**

- If you set **Load Balancing Algorithm** to **Source IP hash**, you do not need to manually enable and configure **Sticky Session**. Source IP hash allows requests from the same client to be directed to the same server.
- If you set **Load Balancing Algorithm** to **Weighted round robin** or **Weighted least connections**, you need to manually enable and configure **Sticky Session**.

Constraints and Limitations

- If you use **Cloud Connect connection**, **Direct Connect** or **VPN** to access ELB, you must select **Source IP hash** as the load balancing algorithm and disable sticky sessions for ELB.
- Dedicated load balancers support **Source IP address** and **Load balancer cookie**.
- Shared load balancers support three types of sticky session: **Source IP address**, **Load balancer cookie**, and **Application cookie**.

NOTE

- For HTTP and HTTPS listeners, enabling or disabling sticky sessions may cause few seconds of service interruption.
- If you enable sticky sessions, traffic to backend servers may be unbalanced. If this happens, disable sticky sessions and check the requests received by each backend server.

4.2.4 Forwarding Mode (Dedicated Load Balancers)

The load balancer routes the traffic across backend servers based on the forwarding mode. There are two options: **Load balancing** and **Active/Standby**.

NOTE

- This feature is only available for backend server groups that are bound to dedicated load balancers.

Table 4-14 Forwarding modes

Forwarding Mode	Description	When to Use
Load balancing	You can add multiple backend servers to a backend server group. And then the load balancer distributes requests across these backend servers based on the load balancing algorithm configured for this backend server group.	You want your load balancer to forward requests based on the forwarding policies configured for the listener.
Active/Standby	You must add two backend servers to the backend server group, one acting as the active server and the other as the standby server. Active/Standby forwarding requires at least one healthy backend server. The load balancer routes the traffic to the active server if it works normally. If the active server becomes unhealthy, the load balancer then routes the traffic to the standby server.	You need higher service availability.

4.2.5 Slow Start (Dedicated Load Balancers)

If you enable slow start, the load balancer linearly increases the proportion of requests to the new backend servers added to the backend server group. When the slow start duration elapses, the load balancer sends full share of requests to backend servers and exits the slow start mode. For details about how to set weights for backend servers, see [Backend Server Weights](#).

Slow start gives applications time to warm up and respond to requests with optimal performance.

NOTE

Slow start is only available for HTTP and HTTPS backend server groups of dedicated load balancers.

Backend servers will exit slow start in either of the following cases:

- The slow start duration elapses.
- Backend servers become unhealthy during the slow start duration.

Constraints

- Weighted round robin must be selected as the load balancing algorithm.
- Slow start takes effect only for new backend servers and does not take effect when the first backend server is added to a backend server group.
- After the slow start duration elapses, backend servers will not enter the slow start mode again.
- Slow start takes effect when health check is enabled and the backend servers are running normally.
- If health check is disabled, slow start takes effect immediately.

4.3 Creating a Backend Server Group (Dedicated Load Balancers)

Scenario

To route requests, you need to associate at least one backend server group to each listener.

NOTE

This section describes how you can create a backend server group for a dedicated load balancer.

You can create a backend server group for a load balancer in any of the ways described in [Table 4-15](#).

Table 4-15 Creating a backend server group

Scenario	Procedure
Creating a backend server group and associating it with a load balancer	Procedure
Creating a backend server group when adding a listener	You can add listeners using different protocols as required. For details, see Overview . References are as follows: <ul style="list-style-type: none">• Adding a TCP Listener• Adding a UDP Listener• Adding an HTTP Listener• Adding an HTTPS Listener
Changing the backend server group associated with the listener	Changing a Backend Server Group



Constraints

The backend protocol of the new backend server group must match the frontend protocol of the listener as described in [Table 4-16](#).

Table 4-16 The frontend and backend protocol

Frontend Protocol	Backend Protocol
TCP	TCP
UDP	<ul style="list-style-type: none">• UDP• QUIC
HTTP	HTTP
HTTPS	<ul style="list-style-type: none">• HTTP• HTTPS

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.

5. Click **Create Backend Server Group** in the upper right corner.
6. Configure the routing policy based on [Table 4-17](#).

Table 4-17 Parameters required for configuring a routing policy

Parameter	Description	Example Value
Load Balancer Type	Specifies the type of load balancers that can use the backend server group. Dedicated load balancers are recommended. The following parameters apply to exclusive load balancers.	-
Load Balancer	Specifies whether to associate a load balancer. You can associate an existing dedicated load balancer when you create a backend server group or associate one later. <ul style="list-style-type: none">• Associate later• Associate existing	Associate later
Forwarding Mode	Specifies the forwarding mode to distribute traffic. There are two options: Load balancing and Active/Standby . <ul style="list-style-type: none">• Load balancing: You can add one or more backend servers to the backend server group.• Active/Standby: You can add only two backend servers to the backend server group, one acting as the active server and the other as the standby server. If the active server is faulty, traffic is forwarded to the standby server, improving service reliability.	Load balancing

Parameter	Description	Example Value
Backend Server Group Type	<p>Specifies the type of the backend server group.</p> <ul style="list-style-type: none">• Hybrid: You can add ECSs and supplementary network interfaces as backend servers, or add IP addresses as servers when IP as a Backend is enabled. When you create a hybrid backend server group, you must specify a VPC and associate the backend server group with a load balancer in this VPC.• IP as a backend server: You can add IP addresses as backend servers only when you enable IP as a Backend.	Hybrid
Backend Server Group Name	<p>Specifies the name of the backend server group.</p>	server_group
VPC	<p>Specifies the VPC where the backend server group works. You can associate the backend server group with a load balancer in this VPC.</p> <p>This parameter is mandatory if you select Hybrid for Backend Server Group Type.</p> <p>You can select an existing VPC or create a new one.</p> <p>For more information about VPC, see the Virtual Private Cloud User Guide.</p>	vpc-test
Backend Protocol	<p>Specifies the protocol that backend servers in the backend server group use to receive requests from the listeners. The protocol varies depending on the forwarding mode:</p> <ul style="list-style-type: none">• Load balancing: HTTP, HTTPS, gRPC, TCP, UDP, TLC, and QUIC• Active/Standby: TCP, UDP, and QUIC	HTTP


Parameter	Description	Example Value
Load Balancing Algorithm	<p>Specifies the algorithm used by the load balancer to distribute traffic. The following options are available:</p> <ul style="list-style-type: none">• Weighted round robin: Requests are routed to different servers based on their weights. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests.• Weighted least connections: In addition to the number of connections, each server is assigned a weight based on its capacity. Requests are routed to the server with the lowest connections-to-weight ratio.• Source IP hash: Allows requests from different clients to be routed based on source IP addresses and ensures that requests from the same client are forwarded to the same server.• Connection ID: This algorithm is available when you have selected QUIC for Backend Protocol. This algorithm allows requests with different connection IDs to be routed to different backend servers and ensures that requests with the same connection ID are routed to the same backend server. <p>For more information about load balancing algorithms, see Load Balancing Algorithms.</p>	Weighted round robin
Sticky Session	<p>Specifies whether to enable sticky sessions if you have selected Weighted round robin or Weighted least connections for Load Balancing Algorithm.</p> <p>If you enable sticky sessions, all requests from the same client during one session are sent to the same backend server.</p> <p>For more information about sticky sessions, see Sticky Session.</p>	-

Parameter	Description	Example Value
Sticky Session Type	<p>Specifies the sticky session type. This parameter is mandatory if Sticky Session is enabled. You can select one of the following type:</p> <ul style="list-style-type: none">• Source IP address: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hashing key, and all backend servers are numbered. The system allocates the client to a particular server based on the generated key. This allows requests from the same IP address are forwarded to the same backend server.• Load balancer cookie: The load balancer generates a cookie after receiving a request from the client. All subsequent requests with the cookie are routed to the same backend server. <p>NOTE</p> <ul style="list-style-type: none">• Source IP address is available when you have selected TCP, QUIC, or UDP for Backend Protocol.• Load balancer cookie is available when you have selected HTTP or HTTPS for Backend Protocol.	Source IP address
Stickiness Duration (min)	<p>Specifies the minutes that sticky sessions are maintained. This parameter is mandatory if Sticky Session is enabled.</p> <ul style="list-style-type: none">• Sticky sessions at Layer 4: 1 to 60• Sticky sessions at Layer 7: 1 to 1440	20

Parameter	Description	Example Value
Slow Start	<p>Specifies whether to enable slow start. This parameter is optional if you have selected Weighted round robin for Load Balancing Algorithm.</p> <p>After you enable this option, the load balancer linearly increases the proportion of requests to backend servers in this mode.</p> <p>When the slow start duration elapses, the load balancer sends full share of requests to backend servers and exits the slow start mode.</p> <p>NOTE Slow start is only available for HTTP and HTTPS backend server groups of dedicated load balancers.</p> <p>For more information about the slow start, see Slow Start (Dedicated Load Balancers).</p>	-
Slow Start Duration (s)	<p>Specifies how long the slow start will last, in seconds.</p> <p>This parameter is mandatory if Slow Start is enabled.</p>	30
Description	Provides supplementary information about the backend server group.	-

- Click **Next** to add backend servers and configure health check.
Add cloud servers,, supplementary network interfaces, or IP addresses to this backend server group. For details, see [Overview](#).
Configure health check for the backend server group based on [Table 4-18](#). For more information about health checks, see [Health Check](#).

Table 4-18 Parameters required for configuring a health check

Parameter	Description	Example Value
Health Check	<p>Specifies whether to enable health checks.</p> <p>If the health check is enabled, click  next to Advanced Settings to set health check parameters.</p>	-

Parameter	Description	Example Value
Health Check Protocol	<p>Specifies the protocol that will be used by the load balancer to check the health of backend servers.</p> <ul style="list-style-type: none">• The backend protocol can be TCP, HTTP, or HTTPS.• If the protocol of the backend server group is UDP, the health check protocol is UDP by default.	HTTP
Domain Name	<p>Specifies the domain name that will be used for health checks.</p> <p>This parameter is mandatory if the health check protocol is HTTP or HTTPS.</p> <ul style="list-style-type: none">• You can use the private IP address of the backend server as the domain name.• You can also specify a domain name that consists of at least two labels separated by periods (.). Use only letters, digits, and hyphens (-). Do not start or end strings with a hyphen. Max total: 100 characters. Max label: 63 characters.	www.elb.com
Health Check Port	<p>Specifies the port that will be used by the load balancer to check the health of backend servers. The port number ranges from 1 to 65535.</p> <p>NOTE By default, the service port on each backend server is used. You can also specify a port for health checks.</p>	80

Parameter	Description	Example Value
Path	<p>Specifies the health check URL, which is the destination on backend servers for health checks. This parameter is mandatory if the health check protocol is HTTP or HTTPS.</p> <p>The path can contain 1 to 80 characters and must start with a slash (/).</p> <p>The path can contain letters, digits, hyphens (-), slashes (/), periods (.), question marks (?), number signs (#), percent signs (%), ampersands (&), and extended character sets _;~!. () *[]@\$^:'!,+</p>	/index.html
Interval (s)	<p>Specifies the maximum time between two consecutive health checks, in seconds.</p> <p>The interval ranges from 1 to 50.</p>	5
Timeout (s)	<p>Specifies the maximum time required for waiting for a response from the health check, in seconds. The interval ranges from 1 to 50.</p>	3
Maximum Retries	<p>Specifies the maximum number of health check retries. The value ranges from 1 to 10.</p>	3

- Click **Next**.
- Confirm the specifications and click **Create Now**.

Related Operations

You can associate the backend server group with the listener of a dedicated load balancer in either ways listed in [Table 4-15](#).

4.4 Creating a Backend Server Group (Shared Load Balancers)

Scenario

To route requests, you need to associate a backend server group to each listener.

 NOTE

This section describes how you can create a backend server group for shared load balancer. You can create a backend server group in the ways listed in [Table 4-19](#).

Table 4-19 Creating a backend server group

Scenario	Procedure
Creating a backend server group and associating it with a load balancer	Procedure
Creating a backend server group when adding a listener	You can add listeners using different protocols as required. For details, see Overview . References are as follows: <ul style="list-style-type: none">• Adding a TCP Listener• Adding a UDP Listener• Adding an HTTP Listener• Adding an HTTPS Listener
Changing the backend server group associated with the listener	Changing a Backend Server Group

Constraints

- The backend protocol of the new backend server group must match the frontend protocol of the listener as described in [Table 4-3](#).
- The backend server group of a shared load balancer can be associated with only one listener.

Procedure



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. Click **Create Backend Server Group** in the upper right corner.
6. Configure the routing policy based on [Table 4-20](#).

Table 4-20 Parameters required for configuring a routing policy


Parameter	Description	Example Value
Load Balancer Type	Specifies the type of load balancers that can use the backend server group.	Shared
Load Balancer	Specifies whether to associate a load balancer.	N/A
Backend Server Group Name	Specifies the name of the backend server group.	server_group
Backend Protocol	Specifies the protocol that backend servers in the backend server group use to receive requests from the listeners. The protocol varies depending on the forwarding mode: The options are HTTP, TCP, and UDP.	HTTP

Parameter	Description	Example Value
Load Balancing Algorithm	<p>Specifies the algorithm used by the load balancer to distribute traffic. The following options are available:</p> <ul style="list-style-type: none">• Weighted round robin: Requests are routed to different servers based on their weights. Backend servers with higher weights receive proportionately more requests, whereas equal-weighted servers receive the same number of requests.• Weighted least connections: In addition to the number of connections, each server is assigned a weight based on its capacity. Requests are routed to the server with the lowest connections-to-weight ratio.• Source IP hash: Allows requests from different clients to be routed based on source IP addresses and ensures that requests from the same client are forwarded to the same server.• Connection ID: This algorithm is available when you have selected QUIC for Backend Protocol. This algorithm allows requests with different connection IDs to be routed to different backend servers and ensures that requests with the same connection ID are routed to the same backend server. <p>For more information about load balancing algorithms, see Load Balancing Algorithms.</p>	Weighted round robin
Sticky Sessions	<p>Specifies whether to enable sticky sessions. If you enable sticky sessions, all requests from the same client during one session are sent to the same backend server.</p> <p>For more information about sticky sessions, see Sticky Session.</p>	N/A

Parameter	Description	Example Value
Sticky Session Type	<p>Specifies the type of sticky sessions. After the sticky session is enabled, you need to select a sticky session type:</p> <ul style="list-style-type: none">• Source IP address: The source IP address of each request is calculated using the consistent hashing algorithm to obtain a unique hashing key, and all backend servers are numbered. The system allocates the client to a particular server based on the generated key. This enables requests from different clients to be routed and ensures that a client is directed to the same server that it was using previously.• Load balancer cookie: The load balancer generates a cookie after receiving a request from the client. All subsequent requests with the cookie are routed to the same backend server.• Application cookie: The application deployed on the backend server generates a cookie after receiving the first request from the client. All subsequent requests with the same cookie are routed to the same backend server. <p>NOTE</p> <ul style="list-style-type: none">• Source IP address is available when you have selected TCP, UDP, or QUIC for Backend Protocol.• Load balancer cookie is available when you have selected HTTP or HTTPS for Backend Protocol.	Source IP address
Stickiness Duration (min)	<p>Specifies the time that sticky sessions are maintained, in minutes.</p> <ul style="list-style-type: none">• Sticky sessions at Layer 4: 1 to 60• Sticky sessions at Layer 7: 1 to 1440	20
Description	Provides supplementary information about the backend server group.	N/A

7. Click **Next** to add backend servers and configure health check based on [Table 4-21](#). For more information about health checks, see [Health Check](#).

Table 4-21 Parameters required for configuring a health check

Parameter	Description	Example Value
Health Check	Specifies whether to enable health checks. If the health check is enabled, click  next to Advanced Settings to set health check parameters.	N/A
Health Check Protocol	<ul style="list-style-type: none">• The health check protocol can be TCP or HTTP.• If the protocol of the backend server group is UDP, the health check protocol is UDP by default.	HTTP
Domain Name	Specifies the domain name that will be used for health checks. By default, the private IP address of each backend server is used. A domain name consists of at least two character strings separated by periods (.). The total length of a domain name cannot exceed 100 characters with each character string not exceeding 63 characters. Only letters, digits, and hyphens (-) are allowed. Strings cannot start or end with a hyphen.	www.elb.com
Health Check Port	Specifies the port that will be used by the load balancer to check the health of backend servers. The port number ranges from 1 to 65535 . NOTE By default, the service port on each backend server is used. You can also specify a port for health checks.	80

Parameter	Description	Example Value
Path	Specifies the health check URL, which is the destination on backend servers for health checks. The path can contain 1 to 80 characters and must start with a slash (/). The path can contain letters, digits, hyphens (-), slashes (/), periods (.), question marks (?), number signs (#), percent signs (%), ampersands (&).	/index.html
Interval (s)	Specifies the maximum time between two consecutive health checks, in seconds. The interval ranges from 1 to 50 .	5
Timeout (s)	Specifies the maximum time required for waiting for a response from the health check, in seconds. The interval ranges from 1 to 50 .	3
Maximum Retries	Specifies the maximum number of health check retries. The value ranges from 1 to 10 .	3

8. Click **Next**.
9. Confirm the specifications and click **Create Now**.

4.5 Modifying a Backend Server Group

4.5.1 Overview

After a backend server group is created, you can modify its health check settings and basic information.

Health Check

If backend servers have to handle large number of requests, frequent health checks may overload the backend servers and cause them to respond slowly. To address this problem, you can prolong the health check interval or use TCP or UDP instead of HTTP. You can also disable health check. If you choose to disable health check, requests may be routed to unhealthy servers, and service interruptions may occur.

For details about the health check, see [Health Check](#).

For details about how to modify health check settings, see [Modifying Health Check Settings](#).

Basic Information

You can modify the basic information of a backend server group listed in [Table 4-22](#).

Table 4-22 Basic information that can be modified

Parameter	Description
Name	Change the name by performing the operations in Changing the Load Balancing Algorithm .
Load Balancing Algorithm	Change the load balancing algorithm by performing the operations in Changing the Load Balancing Algorithm . For details about load balancing algorithms, see Load Balancing Algorithms .
Sticky Session	Enable or disable sticky session by performing the operations in Modifying Sticky Session Settings . For details about the sticky session function, see Sticky Session .
Slow Start	Enable or disable slow start by performing the operations in Modifying Slow Start Settings (Dedicated Load Balancers) . For details about the slow start function, see Slow Start (Dedicated Load Balancers) .
Description	Change the description of the backend server group by performing the operations in Changing the Load Balancing Algorithm .

4.5.2 Modifying Health Check Settings

Scenario

This section describes how you can modify the health check settings.

After the protocol is changed, the load balancer uses the new protocol to check the health of backend servers. The load balancer continues to route traffic to the backend servers after they are detected healthy.

Before the new configurations take effect, the load balancer may return the HTTP 503 error code to the clients.

NOTE

This section applies to dedicated and shared load balancers.

Constraints and Notes

- The health check protocol can be different from the backend protocol.
- To reduce the vCPU usage of the backend servers, it is recommended that you use TCP for health checks. If you want to use HTTP for health checks, you can use static files to return the health check results.
- If health check is enabled, security group rules must allow traffic from the health check port to the backend servers over the health check protocol.
 - Dedicated load balancers: For details, see [Security Group Rules](#).
 - Shared load balancers: For details, see [Security Group Rules](#).

NOTE

After you enable health check, the load balancer immediately checks the health of backend servers.

- If a backend server is detected healthy, the load balancer will start routing requests to it over new connections based on the configured loading balancing algorithms and weights.
- If a backend server is detected unhealthy, the load balancer will stop routing traffic to it.

Enabling Health Check



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** tab page, locate the backend server group.
6. On the **Summary** page, click **Health Check** on the right.
7. In the **Configure Health Check** dialog box, configure the parameters based on [Table 4-23](#).

Table 4-23 Parameters required for configuring health check


Parameter	Description	Example Value
Health Check	Specifies whether to enable health checks.	-
Health Check Protocol	<ul style="list-style-type: none">• The health check protocol can be TCP, HTTP, or HTTPS.• If the protocol of the backend server group is UDP, the health check protocol is UDP by default.	HTTP


Parameter	Description	Example Value
Domain Name	<p>Specifies the domain name that will be used for health checks. This parameter is mandatory if the health check protocol is HTTP or HTTPS.</p> <ul style="list-style-type: none">You can use the private IP address of the backend server as the domain name.You can also specify a domain name that consists of at least two labels separated by periods (.). Use only letters, digits, and hyphens (-). Do not start or end strings with a hyphen. Max total: 100 characters. Max label: 63 characters.	www.elb.com
Health Check Port	<p>Specifies the port that will be used by the load balancer to check the health of backend servers. The port number ranges from 1 to 65535.</p> <p>NOTE By default, the service port on each backend server is used. You can also specify a port for health checks.</p>	80

Parameter	Description	Example Value
Path	<p>Specifies the health check URL, which is the destination on backend servers for health checks. This parameter is mandatory if the health check protocol is HTTP or HTTPS. The path can contain 1 to 80 characters and must start with a slash (/).</p> <ul style="list-style-type: none">• If the backend server group is associated with a dedicated load balancer, the check path can contain letters, digits, hyphens (-), slashes (/), periods (.), question marks (?), number signs (#), percent signs (%), ampersands (&), and extended character sets <code>_;~!. () *[]@\$^:';,+</code>• If the backend server group is associated with a shared load balancer, the path can contain letters, digits, hyphens (-), slashes (/), periods (.), question marks (?), percent signs (%), ampersands (&), and extended character <code>_</code>	/index.html
Interval (s)	<p>Specifies the maximum time between two consecutive health checks, in seconds.</p> <p>The interval ranges from 1 to 50.</p>	5
Timeout (s)	<p>Specifies the maximum time required for waiting for a response from the health check, in seconds. The interval ranges from 1 to 50.</p>	3
Maximum Retries	<p>Specifies the maximum number of health check retries. The value ranges from 1 to 10.</p>	3

8. Click **OK**.

Disabling Health Check

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.

3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the target backend server group.
6. On the **Summary** page, click **Health Check** on the right.
7. In the **Configure Health Check** dialog box, disable health check.
8. Click **OK**.

4.5.3 Changing the Load Balancing Algorithm

Scenario



This section describes how you can change the load balancing algorithm.

For details about load balancing algorithms, see [Load Balancing Algorithms](#).

NOTE

This section applies to dedicated and shared load balancers.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, locate the target backend server group and click **Edit** in the **Operation** column.
6. In the **Modify Backend Server Group** dialog box, change the load balancing algorithm.
7. Click **OK**.

NOTE

The new load balancing algorithm takes effect immediately and will be used to route requests over new connections. However, the previous load balancing algorithm will still be used to route requests over established connections.

4.5.4 Modifying Sticky Session Settings



Scenario

This section describes how you can modify the sticky session settings.



 NOTE

- This section applies to dedicated and shared load balancers.
- You can also configure sticky sessions when adding a listener or creating a backend server group.

Enabling Sticky Session

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, locate the backend server group and click **Edit** in the **Operation** column.
6. In the **Modify Backend Server Group** dialog box, enable sticky session, select the sticky session type, and set the session stickiness duration.
7. Click **OK**.

Disabling Sticky Session

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, locate the backend server group and click **Edit** in the **Operation** column.
6. In the **Modify Backend Server Group** dialog box, disable sticky session.
7. Click **OK**.

4.5.5 Modifying Slow Start Settings (Dedicated Load Balancers)

Scenario

This section describes how you can modify the slow start settings.

For details, see [Slow Start \(Dedicated Load Balancers\)](#).

NOTE

- This section applies only to dedicated load balancers.
- You can also configure slow start when adding a listener or creating a backend server group.

Enabling Slow Start



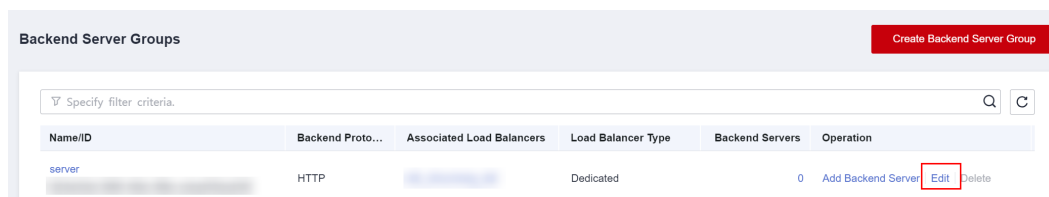


1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, locate the backend server group and click **Edit** in the **Operation** column.

Figure 4-11 Modifying a backend server group



6. In the **Modify Backend Server Group** dialog box, enable slow start and set the slow start duration.
The slow start duration ranges from 30 to 1200 in seconds. When the slow start duration elapses, the load balancer sends full share of requests to backend servers and exits the slow start mode.
7. Click **OK**.

Disabling slow start

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, locate the backend server group and click **Edit** in the **Operation** column.
6. In the **Modify Backend Server Group** dialog box, disable slow start.
7. Click **OK**.

4.6 Changing a Backend Server Group

Scenario

This section describes how you can change the default backend server group configured for a listener.



TCP or UDP listeners forward requests to the default backend server groups.

HTTP or HTTPS listeners forward requests based on the priorities of the forwarding policies. If you do not add a forwarding policy, the listener will route the requests to the default backend server group.

Constraints and Limitations

- The backend server group cannot be changed if redirection is enabled.
- The backend protocol of the backend server group must match the frontend protocol of the listener. For details, see [Table 4-3](#).

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the target load balancer and click its name.
5. On the **Listeners** tab, locate the target listener and click its name.
6. On the **Summary** page, click **Change Backend Server Group** on the right.
7. In the displayed dialog box, click the server group name box.
Select a backend server group from the drop-down list or create a group.
 - a. Click the name of the backend server group or enter the name in the search box to search for the target group.
 - b. Click **Create Backend Server Group**. After the backend server group is created, click the refresh icon.

NOTE

The backend protocol of the new backend server group must match the frontend protocol of the listener.

8. Click **OK**.



4.7 Viewing a Backend Server Group

Scenario

This section describes how you can view the following information about a backend server group:

- Basic information: the name, ID, and backend protocol
- Health check: whether health check is enabled and health check configurations
- Backend servers: servers that have been added to the backend server group

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. On the **Summary** tab page, view the basic information and health check settings.

4.8 Deleting a Backend Server Group

Scenario



This section describes how you can delete a backend server group.

Constraints and Limitations

- Before you delete a backend server group, you need to:
 - Disassociate it from the listener. For details, see [Changing a Backend Server Group](#).
 - Ensure the backend server group is not used by a forwarding policy of an HTTP or HTTPS listener.
- Remove all backend servers from the backend server group.

Procedure

1. Log in to the management console.

2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, locate the backend server group and click **Delete** in the **Operation** column.
6. In the displayed dialog box, click **Yes**.

5 Backend Server (Dedicated Load Balancers)

5.1 Overview

Backend servers receive and process requests from the associated load balancer.

If the incoming traffic increases, you can add more backend servers to ensure the stability and reliability of applications and eliminate SPOFs. If the incoming traffic decreases, you can remove some backend servers to reduce the cost.

If the load balancer is associated with an AS group, instances are automatically added to or removed from the load balancer.

Different types of backend servers can be added to different types of backend server groups as described in [Table 5-1](#).

Table 5-1 Backend server group and backend server types

Backend Server Group Type	Backend Server Types	Reference
Hybrid	<ul style="list-style-type: none">Cloud servers or supplementary network interfaces that are in the same VPC as the load balancer, if IP as a Backend is disabledIP addresses of servers in other VPCs or in your on-premises data center, if IP as a Backend is enabled <p>NOTE When you create a hybrid backend server group, you must specify a VPC and associate the backend server group with a load balancer in this VPC.</p>	<ul style="list-style-type: none">Adding Backend ServersAdding Supplementary Network InterfacesAdding IP Addresses as Backend Servers

Backend Server Group Type	Backend Server Types	Reference
IP as a backend server	IP addresses of cloud or on-premises servers NOTE IP as a Backend must have been enabled for the load balancer.	<ul style="list-style-type: none">Adding IP Addresses as Backend Servers

Precautions

- It is recommended that you select backend servers running the same OS for easier management and maintenance.
- The load balancer checks the health of each server added to the associated backend server group if you have configured health check for the backend server group. If the backend server responds normally, the load balancer will consider it healthy. If the backend server does not respond normally, the load balancer will periodically check its health until the backend server is considered healthy.
- If a backend server is stopped or restarted, connections established with the server will be disconnected, and data being transmitted over these connections will be lost. To avoid this from happening, configure the retry function on the clients to prevent data loss.
- If you enable sticky sessions, traffic to backend servers may be unbalanced. If this happens, disable sticky sessions and check the requests received by each backend server.

Constraints and Limitations

- A maximum of 500 backend servers can be added to a backend server group.
- Inbound security group rules must be configured to allow traffic over the port of each backend server and health check port. For details, see [Security Group Rules](#).
- If you select only network load balancing, a server cannot serve as both a backend server and a client.

Backend Server Weights

You need to set a weight for each backend server in a backend server group to receive requests. The higher the weight you have configured for a backend server, the more requests the backend server receives.

You can set an integer from **0** to **100**. If you set the weight of a backend server to **0**, new requests will not be routed to this server.

Three load balancing algorithms allow you to set weights to backend servers, as shown in the following table. For more information about load balancing algorithms, see [Load Balancing Algorithms](#).

Table 5-2 Server weights in different load balancing algorithms

Load Balancing Algorithm	Weight Setting
Weighted round robin	<ul style="list-style-type: none">• If none of the backend servers have a weight of 0, the load balancer routes requests to backend servers based on their weights. Backend servers with higher weights receive proportionately more requests.• If two backend servers have the same weights, they receive the same number of requests.
Weighted least connections	<ul style="list-style-type: none">• If none of the backend servers have a weight of 0, the load balancer calculates the load of each backend server using the formula (Overhead = Number of current connections/Backend server weight).• The load balancer routes requests to the backend server with the lowest overhead.
Source IP hash	<ul style="list-style-type: none">• If none of the backend servers have a weight of 0, requests from the same client are routed to the same backend server within a period of time.• If the weight of a backend server is 0, no requests are routed to this backend server.

5.2 Security Group Rules

Scenarios

To ensure normal communications between the load balancer and backend servers, you need to check the security group rules and network ACL rules configured for the backend servers.

- Security group rules must allow traffic from the backend subnet where the load balancer resides to the backend servers. (By default, the backend subnet of a load balancer is the same as the subnet where the load balancer resides.) For details about how to configure security group rules, see [Configuring Security Group Rules](#).
- Network ACL rules are optional for subnets. If network ACL rules are configured for the backend subnet of the load balancer, the network ACL rules must allow traffic from the backend subnet of the load balancer to the backend servers. For details about how to configure rules, see [Configuring Network ACL Rules](#).

NOTE

If the load balancer has a TCP or UDP listener and IP as a backend is disabled, security group rules and network ACL rules will not take effect.

You can use access control to limit which IP addresses are allowed to access the listener. Learn how to configure [Access Control](#).

Constraints and Limitations

- If health check is enabled for a backend server group, security group rules must allow traffic from the health check port over the health check protocol.
- If UDP is used for health check, there must be a rule that allows ICMP traffic to check the health of the backend servers.

Configuring Security Group Rules

If you have no VPCs when creating a server, the system automatically creates one for you. Default security group rules allow only communications among the servers in the VPC. To ensure that the load balancer can communicate with these servers over both the frontend port and health check port, configure inbound rules for security groups containing these servers.


1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Under **Compute**, click **Elastic Cloud Server**.
4. On the **Elastic Cloud Server** page, click the name of the ECS that has been added to a backend server group.
The page providing details about the ECS is displayed.
5. Click **Security Groups**, locate the security group, and view security group rules.
6. Click the ID of a security group rule or **Modify Security Group Rule**. The security group details page is displayed.
7. On the **Inbound Rules** tab page, click **Add Rule**. Configure an inbound rule based on [Table 5-3](#).

Table 5-3 Security group rules

Backend Protocol	Policy	Protocol & Port	Source IP Address
HTTP or HTTPS	Allow	Protocol: TCP Port: the port used by the backend server and health check port	Backend subnet of the load balancer
TCP	Allow	Protocol: TCP Port: health check port	
UDP	Allow	Protocol: UDP and ICMP Port: health check port	

 NOTE

- After a load balancer is created, do not change the subnet. If the subnet is changed, the IP addresses occupied by the load balancer will not be released, and traffic from the previous backend subnet is still need to be allowed to backend servers.
 - Traffic from the new backend subnet is also need to be allowed to backend servers.
8. Click **OK**.

Configuring Network ACL Rules



To control traffic in and out of a subnet, you can associate a network ACL with the subnet. Network ACL rules control access to subnets and add an additional layer of defense to your subnets.

The default network ACL rule denies all inbound and outbound traffic. You can configure an inbound rule to allow traffic from the backend subnet of the load balancer through the port of the backend server.

- If the load balancer is in the same subnet as the backend servers, network ACL rules will not take effect. In this case, the backend servers will be considered healthy and can be accessed by the clients.
- If the load balancer is not in the same subnet as the backend servers, network ACL rules will take effect. In this case, the backend servers will be considered unhealthy and cannot be accessed by the clients.

 NOTE

Network ACL rules configured for the backend subnet of the load balancer will not restrict the traffic from the clients to the load balancer. If network ACL rules are configured, the clients can directly access the load balancer. To control access to the load balancer, configure access control for all listeners added to the load balancer by performing the operations in [Access Control](#).

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Virtual Private Cloud**.
4. In the navigation pane on the left, choose **Access Control > Network ACLs**.
5. In the network ACL list, click the name of the network ACL to switch to the page showing its details.
6. On the **Inbound Rules** or **Outbound Rules** tab page, click **Add Rule** to add an inbound or outbound rule.
 - **Action:** Select **Allow**.
 - **Type:** Select the same type as the backend subnet of the load balancer.
 - **Protocol:** The protocol must be the same as the backend protocol.
 - **Source:** Set it to the backend subnet of the load balancer.
 - **Source Port Range:** Select a port range.

- **Destination:** Enter a destination address allowed in this direction. The default value is **0.0.0.0/0**, which indicates that traffic from all IP addresses is permitted.
 - **Destination Port Range:** Select a port range.
 - (Optional) **Description:** Describe the network ACL rule.
7. Click **OK**.

5.3 Managing Backend Servers

5.3.1 Adding Backend Servers

Scenario



When you use ELB to route traffic to backend servers, you need to ensure that at least one backend server is running properly and can receive requests from the associated load balancer.

If the incoming traffic increases, you can add more backend servers to ensure the stability and reliability of applications and eliminate SPOFs. If the incoming traffic decreases, you can remove some backend servers to reduce the cost.

Constraints and Limitations

- The cloud servers must be in the same VPC as the backend server group.

Procedure



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **Add** on the right.
7. You can search for backend servers using specified keywords. Select the backend servers you want to add and click **Next**.
8. Specify the weights and ports for the backend servers, and click **Finish**. Backend server ports can be set in batches.

5.3.2 Viewing Backend Servers

Scenario

You can view backend servers that have been added to a backend server group, including their status, private IP addresses, health check results, weights, and ports.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **Backend Servers**.
7. In the backend server list, view the backend servers.

5.3.3 Removing Backend Servers

Scenario

You can remove a backend server that is no longer needed from a backend server group.


Once a backend server is removed, it is disassociated from the load balancer and will no longer receive requests from the load balancer. The backend server still exists. You can add the backend server to the backend server group again when traffic increases or the reliability needs to be enhanced.


Notes

After the backend server is removed, requests are still routed to it. This is because a persistent connection is established between the load balancer and the backend server and requests are routed to this server until the TCP connection times out.

If no data is transmitted over this TCP connection after it times out, ELB disconnects the connection.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.

3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **Backend Servers**.
7. Select the backend servers you want to remove and click **Remove** above the backend server list.
8. In the displayed dialog box, click **Yes**.

5.3.4 Changing Backend Server Weights/Ports



Scenario

You can change the weights/ports configured for backend servers based on their capability to process requests.

Constraints and Limitations

- You can set an integer from **0** to **100**. If you set the weight of a backend server to **0**, new requests will not be routed to this server.
- The weights can only be specified when you select weighted round robin, weighted least connections, or source IP hash as the load balancing algorithm. For more information about load balancing algorithms, see [Backend Server Weights](#).

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **Backend Servers**.
7. Select the backend servers and click **Modify Port/Weight** up above the backend server list.
8. In the displayed dialog box, modify weights/ports as you need.
 - Modifying ports:
 - Changing the port of a single backend server: Set the port in the **Backend Port** column.

- Changing the ports of multiple backend servers: Set the ports next to **Batch Modify Ports** and click **OK**.
- Modifying weights:
 - Changing the weight of a single backend server: Set the weight in the **New Weight** column.
 - Changing the weights of multiple backend servers: Set the weight next to **Batch Modify Weights** and click **OK**.

NOTE

You can change the weights of multiple backend servers to **0** so that they will not receive requests from the load balancer.

9. Click **OK**.

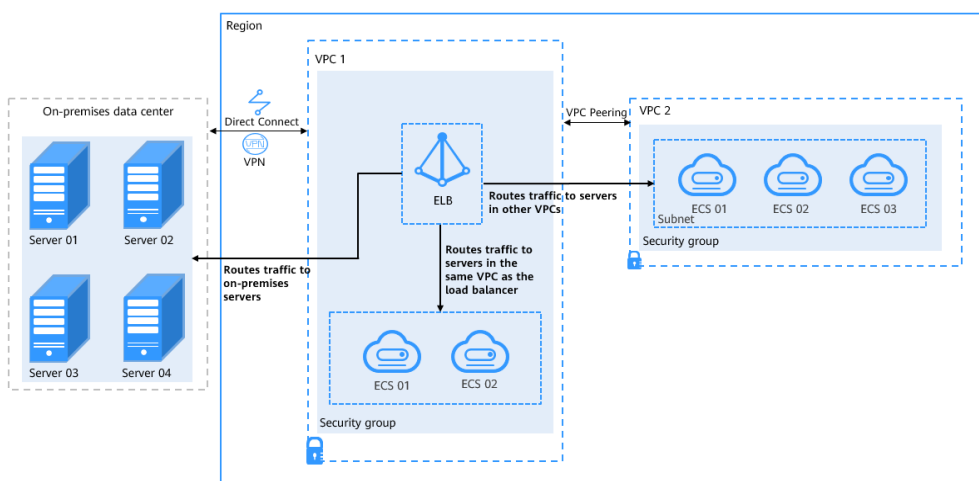
5.4 IP Addresses as Backend Servers

5.4.1 Overview

Dedicated load balancers support hybrid load balancing. You can add servers and supplementary network interfaces in the VPC where the load balancer is created, in a different VPC, or in an on-premises data center, by using private IP addresses of the servers to the backend server group of the load balancer.

In this way, incoming traffic can be flexibly distributed to cloud servers and on-premises servers.

Figure 5-1 Routing requests to cloud and on-premises servers



Constraints and Limitations

- IP as a backend cannot be disabled after it is enabled.
- Only private IPv4 addresses can be added as backend servers.
- A maximum of 50,000 concurrent connections can be established with a backend server that is added by using its IP address.

- If you add IP addresses as backend servers, the source IP addresses of the clients cannot be passed to these servers. Install the **TOA module** to obtain source IP addresses.

Scenario

After you enable IP as a backend, you can add backend servers by using their IP addresses. You need to get prepared for different scenarios as shown in **Table 5-4**.

Table 5-4 Adding IP addresses as backend servers

Where Servers Are Running	Preparations
In a different VPC from the load balancer	Set up a VPC peering connection between the VPC where the load balancer is running and the VPC where the servers are running. For details about how to set up a VPC peering connection, see the Virtual Private Cloud User Guide .
In the same VPC as the load balancer	Set up a VPC peering connection for the VPC where the load balancer and the servers are running, and then add routes for the VPC peering connection. For details, see Routing Traffic to Backend Servers in the Same VPC as the Load Balancer .
In on-premises data centers	Connect the on-premises data center to the VPC where the load balancer is running through Direct Connect or VPN. For details about how to connect on-premises data centers to the cloud, see the Direct Connect User Guide or Virtual Private Network User Guide .

5.4.2 Enabling IP as a Backend


Scenario


You can enable IP as a backend for an existing dedicated load balancer.

Constraints and Limitations

- IP as a backend cannot be disabled after it is enabled.
- The protocol of backend server groups can only be TCP, UDP, HTTP, or HTTPS.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.

3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. On the **Load Balancers** page, locate the load balancer and click its name.
5. On the **Summary** tab page, click **Enable** next to **IP as a Backend**.
6. Click **OK**.

5.4.3 Adding IP Addresses as Backend Servers

Scenario

If you enable IP as a backend, you can associate backend servers with the load balancer by using their IP addresses.

You need to get prepared for different scenarios as shown in [Table 5-5](#).



Table 5-5 Adding IP addresses as backend servers

Where Servers Are Running	Preparations
In a different VPC from the load balancer	Set up a VPC peering connection between the VPC where the load balancer is running and the VPC where the servers are running. For details about how to set up a VPC peering connection, see the Virtual Private Cloud User Guide .
In the same VPC as the load balancer	Set up a VPC peering connection for the VPC where the load balancer and the servers are running, and then add routes for the VPC peering connection. For details, see Routing Traffic to Backend Servers in the Same VPC as the Load Balancer .
In on-premises data centers	Connect the on-premises data center to the VPC where the load balancer is running through Direct Connect or VPN. For details about how to connect on-premises data centers to the cloud, see the Direct Connect User Guide or Virtual Private Network User Guide .

Constraints and Limitations

- If IP as a backend is not enabled when you create a load balancer, you can enable it on the **Summary** page of the load balancer.
- Only private IPv4 addresses can be added as backend servers.
- The backend subnet of the load balancer must have sufficient IP addresses (at least 16 IP addresses). Otherwise, backend servers cannot be added through IP addresses. If the IP addresses are insufficient, you can add more backend subnets on the **Summary** page of the load balancer.
- Security group rules of backend servers added through IP addresses must allow traffic from the backend subnet of the load balancer. If traffic is not allowed, health checks will fail.

Procedure



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **Add** on the **IP as Backend Servers** area.
7. Specify the IP addresses, ports, and weights for the backend servers.
8. Click **OK**.

5.4.4 Viewing Backend Servers

Scenario

You can view backend servers added to a backend server group, including their IP addresses, health check results, weights, and ports.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **IP as Backend Servers**.
7. In the backend server list, view the added backend servers.

5.4.5 Removing Backend Servers

Scenario



You can remove backend servers from a backend server group when you do not need them to process requests.

Notes

After the backend server is removed, requests are still routed to it. This is because a persistent connection is established between the load balancer and the backend server and requests are routed to this server until the TCP connection times out.

If no data is transmitted over this TCP connection after it times out, ELB disconnects the connection.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **IP as Backend Servers**.
7. Select the backend servers to be removed and click **Remove** above the backend server list.
8. In the displayed dialog box, click **Yes**.

5.4.6 Changing Backend Server Weights/Ports



Scenario


You can change the weights and ports specified for backend servers based on their capability to process requests.

Constraints and Limitations

- You can set an integer from **0** to **100**. If you set the weight of a backend server to **0**, new requests will not be routed to this server.
- The weights can only be specified when you select weighted round robin, weighted least connections, or source IP hash as the load balancing algorithm. For more information about load balancing algorithms, see [Backend Server Weights](#).

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.

4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
 5. On the **Backend Server Groups** page, click the name of the backend server group.
 6. Switch to the **Backend Servers** tab page and click **IP as Backend Servers**.
 7. Select the backend servers and click **Modify Port/Weight** up the backend server list.
 8. In the displayed dialog box, modify weights and ports as you need.
 - Modifying ports:
 - Changing the port of a single backend server: Set the port in the **Backend Port** column.
 - Changing the ports of multiple backend servers: Set the ports next to **Batch Modify Ports** and click **OK**.
 - Modifying weights:
 - Changing the weight of a single backend server: Set the weight in the **New Weight** column.
 - Changing the weights of multiple backend servers: Set the weight next to **Batch Modify Weights** and click **OK**.
-  **NOTE**
- You can change the weights of multiple backend servers to **0** so that they will not receive requests from the load balancer.
9. Click **OK**.

5.5 Supplementary Network Interfaces

5.5.1 Adding Supplementary Network Interfaces

Scenario

In addition to cloud servers and on-premises servers, you can add supplementary network interfaces to a backend server group.

Supplementary network interfaces allow you to configure more NICs than a cloud server would normally support. They can be attached to VLAN subinterfaces of elastic network interfaces.

For details on how to create a VPC, see the *Virtual Private Cloud User Guide*.



 **NOTE**

For regions where supplementary network interfaces are supported, see [Function Overview](#).

Constraints and Limitations

Supplementary network interfaces can only be added to a hybrid backend server group.

Procedure



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group that you want to add supplementary network interfaces to.
6. Switch to the **Backend Servers** tab page and click **Add** on the **Supplementary Network Interfaces** area.
You can search for supplementary network interfaces by ID, private IP address, network interface private IP address, subnet name, or subnet ID.
7. Specify the weights and ports for the supplementary network interfaces and click **Finish**.

5.5.2 Viewing Supplementary Network Interfaces

Scenario

You can view supplementary network interfaces that have been added to a backend server group.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **Supplementary Network Interfaces**.
7. View the added supplementary network interfaces.

5.5.3 Removing Supplementary Network Interfaces

Scenario



You can remove supplementary network interfaces from a backend server group if you do not need them to process requests.

Notes

After the backend server is removed, requests are still routed to it. This is because a persistent connection is established between the load balancer and the backend server and requests are routed to this server until the TCP connection times out.

If no data is transmitted over this TCP connection after it times out, ELB disconnects the connection.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **Supplementary Network Interfaces**.
7. Select the supplementary network interfaces and click **Remove** above the list.
8. In the displayed dialog box, click **Yes**.

5.5.4 Changing the Weights/Ports of Supplementary Network Interfaces




Scenario

You can change the weights specified for supplementary network interfaces based on their capability to process requests.

Constraints and Limitations

- You can set an integer from **0** to **100**. If you set the weight of a backend server to **0**, new requests will not be routed to this server.
- The weights can only be specified when you select weighted round robin, weighted least connections, or source IP hash as the load balancing algorithm. For more information about load balancing algorithms, see [Backend Server Weights](#).

Procedure

1. Log in to the management console.
 2. In the upper left corner of the page, click  and select the desired region and project.
 3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
 4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
 5. On the **Backend Server Groups** page, click the name of the backend server group.
 6. Switch to the **Backend Servers** tab page and click **Supplementary Network Interfaces**.
 7. Select the backend servers and click **Modify Weight** up the backend server list.
 8. In the displayed dialog box, modify weights and ports as you need.
 - Modifying ports:
 - Changing the port of a single backend server: Set the port in the **Backend Port** column.
 - Changing the ports of multiple backend servers: Set the ports next to **Batch Modify Ports** and click **OK**.
 - Modifying weights:
 - Changing the weight of a single backend server: Set the weight in the **Weight** column.
 - Changing the weights of multiple backend servers: Set the weight next to **Batch Modify Weights** and click **OK**.
-  **NOTE**
- You can change the weights of multiple backend servers to **0** so that they will not receive requests from the load balancer.
9. Click **OK**.

6 Backend Server (Shared Load Balancers)

6.1 Overview

Backend servers receive and process requests from the associated load balancer.

If the incoming traffic increases, you can add more backend servers to ensure the stability and reliability of applications and eliminating SPOFs. If the incoming traffic decreases, you can remove some backend servers to reduce the cost.

If the load balancer is associated with an AS group, instances are automatically added to or removed from the load balancer.

You can only add servers in the same VPC as the load balancer. For details, see [Adding Backend Servers](#).

Precautions

- It is recommended that you select backend servers running the same OS for easier management and maintenance.
- The load balancer checks the health of each server added to the associated backend server group if you have configured health check for the backend server group. If the backend server responds normally, the load balancer will consider it healthy. If the backend server does not respond normally, the load balancer will periodically check its health until the backend server is considered healthy.
- If a backend server is stopped or restarted, connections established with the server will be disconnected, and data being transmitted over these connections will be lost. To avoid this from happening, configure the retry function on the clients to prevent data loss.
- If you enable sticky sessions, traffic to backend servers may be unbalanced. If this happens, disable sticky sessions and check the requests received by each backend server.

Constraints and Limitations

- A maximum of 500 backend servers can be added to a backend server group.

- Inbound security group rules must be configured to allow traffic over the port of each backend server and health check port. For details, see [Security Group Rules](#).

Backend Server Weights

You need to set a weight for each backend server in a backend server group to receive requests. The higher the weight you have configured for a backend server, the more requests the backend server receives.

You can set an integer from **0** to **100**. If you set the weight of a backend server to **0**, new requests will not be routed to this server.

Three load balancing algorithms allow you to set weights to backend servers, as shown in the following table. For more information about load balancing algorithms, see [Load Balancing Algorithms](#).

Table 6-1 Server weights in different load balancing algorithms

Load Balancing Algorithm	Weight Setting
Weighted round robin	<ul style="list-style-type: none">• If none of the backend servers have a weight of 0, the load balancer routes requests to backend servers based on their weights. Backend servers with higher weights receive proportionately more requests.• If two backend servers have the same weights, they receive the same number of requests.
Weighted least connections	<ul style="list-style-type: none">• If none of the backend servers have a weight of 0, the load balancer calculates the load of each backend server using the formula (Overhead = Number of current connections/Backend server weight).• The load balancer routes requests to the backend server with the lowest overhead.
Source IP hash	<ul style="list-style-type: none">• If none of the backend servers have a weight of 0, requests from the same client are routed to the same backend server within a period of time.• If the weight of a backend server is 0, no requests are routed to this backend server.

6.2 Security Group Rules

To ensure normal communications between the load balancer and backend servers, you need to check the security group rules and network ACL rules configured for the backend servers.

When backend servers receive requests from the load balancer, source IP addresses are translated into those in 100.125.0.0/16.

- Security group rules must allow traffic from the 100.125.0.0/16 to backend servers. For details about how to configure security group rules, see [Configuring Security Group Rules](#).
- Network ACL rules are optional for subnets. If network ACL rules are configured for the backend subnet of the load balancer, the network ACL rules must allow traffic from the backend subnet of the load balancer to the backend servers. For details about how to configure these rules, see [Configuring Network ACL Rules](#).

NOTE

If **Transfer Client IP Address** is enabled for the TCP or UDP listeners, network ACL rules and security group rules will not take effect. You can use access control to limit which IP addresses are allowed to access the listener. Learn how to configure [access control](#).

Constraints and Limitations

- If health check is enabled for a backend server group, security group rules must allow traffic from the health check port over the health check protocol.
- If UDP is used for health check, there must be a rule that allows ICMP traffic. If there is no such rule, the health of the backend servers cannot be checked.

Configuring Security Group Rules

If you have no VPCs when creating a server, the system automatically creates one for you. Default security group rules allow only communications among the servers in the VPC. To ensure that the load balancer can communicate with these servers over both the frontend port and health check port, configure inbound rules for security groups containing these servers.


1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Under **Compute**, click **Elastic Cloud Server**.
4. On the **Elastic Cloud Server** page, click the name of the ECS that has been added to a backend server group.
The page providing details about the ECS is displayed.
5. Click **Security Groups**, locate the security group, and view security group rules.
6. Click the ID of a security group rule or **Modify Security Group Rule**. The security group details page is displayed.
7. On the **Inbound Rules** tab page, click **Add Rule**. Configure an inbound rule based on [Table 6-2](#).

Table 6-2 Security group rules

Backend Protocol	Policy	Protocol & Port	Source IP Address
HTTP	Allow	Protocol: TCP Port: the port used by the backend server and health check port	100.125.0.0/16
TCP	Allow	Protocol: TCP Port: health check port	100.125.0.0/16
UDP	Allow	Protocol: UDP and ICMP Port: health check port	100.125.0.0/16

8. Click **OK**.

Configuring Network ACL Rules

To control traffic in and out of a subnet, you can associate a network ACL with the subnet. Network ACL rules control access to subnets and add an additional layer of defense to your subnets. Default network ACL rules reject all inbound and outbound traffic. If the subnet of a load balancer or associated backend servers has a network ACL associated, the load balancer cannot receive traffic from the Internet or route traffic to backend servers, and backend servers cannot receive traffic from and respond to the load balancer.



Configure an inbound network ACL rule to permit access from 100.125.0.0/16.

ELB translates the public IP addresses used to access backend servers into private IP addresses in 100.125.0.0/16. You cannot configure rules to prevent public IP addresses from accessing backend servers.

NOTE

Network ACL rules configured for the backend subnet of the load balancer will not restrict the traffic from the clients to the load balancer. If these rules are configured, the clients can directly access the load balancer. To control access to the load balancer, configure access control for all listeners added to the load balancer

For details, see [Access Control](#).

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Virtual Private Cloud**.

4. In the navigation pane on the left, choose **Access Control > Network ACLs**.
5. In the network ACL list, click the name of the network ACL to switch to the page showing its details.
6. On the **Inbound Rules** or **Outbound Rules** tab page, click **Add Rule** to add an inbound or outbound rule.
 - **Action**: Select **Allow**.
 - **Protocol**: The protocol must be the same as the backend protocol.
 - **Source**: Set it to **100.125.0.0/16**.
 - **Source Port Range**: Select a port range.
 - **Destination**: Enter a destination address allowed in this direction. The default value is **0.0.0.0/0**, which indicates that traffic from all IP addresses is permitted.
 - **Destination Port Range**: Select a port range.
 - (Optional) **Description**: Describe the network ACL rule.
7. Click **OK**.

6.3 Managing Backend Servers

6.3.1 Adding Backend Servers

Scenario

You can add backend servers to a backend server group to process requests from clients.



When you use ELB to route requests, ensure that at least one backend server is running properly and can receive requests routed by the load balancer.

After a backend server is unbound from a load balancer, the backend server does not receive requests forwarded by the load balancer, but the backend server is disassociated from the load balancer. You can add the backend server to the backend server group again when traffic increases or the reliability needs to be enhanced.

Constraints and Limitations

Only servers in the same VPC as the load balancer can be added.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.



1. On the **Backend Server Groups** page, click the name of the backend server group.
2. Switch to the **Backend Servers** tab page and click **Add** on the right.
3. Search for backend servers using specified keywords.
4. Specify the weights and ports for the backend servers, and click **Finish**.
Backend server ports can be set in batches.

6.3.2 Viewing Backend Servers

Scenario

You can view backend servers that have been added to a backend server group, including their status, private IP addresses, health check results, weights, and ports.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **Backend Servers**.
7. In the backend server list, view the backend servers.

6.3.3 Removing Backend Servers

Scenario

You can remove a backend server that is no longer needed from a backend server group.



Once a backend server is removed, it is disassociated from the load balancer and will no longer receive requests from the load balancer. The backend server still exists. You can add the backend server to the backend server group again when traffic increases or the reliability needs to be enhanced.

Notes

After the backend server is removed, requests are still routed to it. This is because a persistent connection is established between the load balancer and the backend server and requests are routed to this server until the TCP connection times out.

If no data is transmitted over this TCP connection after it times out, ELB disconnects the connection.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **Backend Servers**.
7. Select the backend servers you want to remove and click **Remove** above the backend server list.
8. In the displayed dialog box, click **Yes**.

6.3.4 Changing Backend Server Weights



Scenarios

You can change the weights specified for backend servers based on their capability to process requests.

Constraints and Limitations

- You can set an integer from **0** to **100**. If you set the weight of a backend server to **0**, new requests will not be routed to this server.
- The weights can only be specified when you select weighted round robin, weighted least connections, or source IP hash as the load balancing algorithm. For more information about load balancing algorithms, see [Backend Server Weights](#).

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > Backend Server Groups**.
5. On the **Backend Server Groups** page, click the name of the backend server group.
6. Switch to the **Backend Servers** tab page and click **Backend Servers**.
7. Select the backend servers and click **Modify Weight** up above the backend server list.

8. In the displayed dialog box, modify weights as you need.
 - Changing the weight of a single backend server: Set the weight in the **Weight** column.
 - Changing the weights of multiple backend servers: Set the weight next to **Batch Modify Weights** and click **OK**.

 **NOTE**

You can change the weights of multiple backend servers to **0** so that they will not receive requests from the load balancer.

9. Click **OK**.

7 Certificate

7.1 Introduction to Certificates

ELB supports two types of certificates. If you need an HTTPS listener, you need to bind a server certificate to it. To enable mutual authentication, you also need to bind a CA certificate to the listener.

- **Server certificate:** used for SSL handshake negotiations if an HTTPS listener is used. Both the certificate content and private key are required.
- **CA certificate:** issued by a certificate authority (CA) and used to verify the certificate issuer. If HTTPS mutual authentication is required, HTTPS connections can be established only when the client provides a certificate issued by a specific CA.

 **NOTE**

SSL Certificate Manager (SCM) allows you to purchase a certificate from Huawei Cloud or upload your own certificates for easier management.

Precautions

- A certificate can be used by multiple load balancers but only needs to be uploaded to each load balancer once.
- You must specify a domain name for an SNI certificate. The domain name must be the same as that in the certificate. An SNI certificate can have multiple domain names.
- For each certificate type, a listener can have only one certificate by default, but a certificate can be bound to more than one listener. If SNI is enabled for the listener, multiple server certificates can be bound.
- Only original certificates are supported. That is to say, you cannot encrypt your certificates.
- You do not need to configure certificates for both shared load balancers and associated backend servers. If you configure a certificate for backend servers, HTTPS listeners cannot be added to the load balancer. In this case, you can add a TCP listener to transparently transmit HTTPS traffic to backend servers. This restriction does not apply to dedicated load balancers.

- You can use self-signed certificates. However, note that self-signed certificates pose security risks. Therefore, it is recommended that you use certificates issued by third parties.
- ELB supports certificates only in PEM format. If you have a certificate in any other format, you must convert it to a PEM-encoded certificate.
- If a certificate has expired, you need to manually replace or delete it.

7.2 Certificate and Private Key Format

Certificate Format

You can copy and paste the certificate body to create a certificate or directly upload the certificate.

A certificate issued by the Root CA is unique, and no additional certificates are required. The configured site is considered trustable by access devices such as a browser.

The body of the server and CA certificates must meet the following requirements:

- The content starts with -----BEGIN CERTIFICATE----- and ends with -----END CERTIFICATE-----.
- Each row contains 64 characters except the last row.
- There are no empty rows.

The following is an example:

```
-----BEGIN CERTIFICATE-----
MIIDljCCAougAwIBAgIJALV96mEtVF4EMA0GCSqGSIb3DQEBBQUAMGoxCzAJBgNV
BAYTAnh4MQswCQYDVQQLIEwJ4eDELMAkGA1UEBxMCeHgxGzAJBgNVBAAoTAnh4MQsw
CQYDVQQLLEwJ4eDELMAkGA1UEAxMCeHgxGjAYBgkqhkiG9w0BCQEWc3h4eEaxNjMu
Y29tMB4XDTE3MTE5MzAyMjYxM1oXDTEwMTExMjYxM1owajELMAkGA1UEBhMC
eHgxGzAJBgNVBAAgTAnh4MQswCQYDVQQLHEwJ4eDELMAkGA1UEChMCeHgxGzAJBgNV
BAcTAnh4MQswCQYDVQDEwJ4eDEaMBGCSqGSIb3DQEJARYLeHh4QDE2My5jb20w
gZ8wDQYJKoZIhvcNAQEBBQADgY0AMIGJAoGBAMU832iM+d3FILgTWmpZBUoYcIwV
cAAE7FsZ9LNEROyjJpyi256oypdBvGs9JAUBN5WaFk81UQx29wAyNixX+bKa0DB
WpUDqr84V1f9vdQc75v9WoujcnlKszpV6qePPC7igJpu4QOI362BrWzJCYQbg4
Uzo1KYBhLFxl0TovAgMBAAgJgc8wgcwWwHQYDVR0OBBYEFMbtvDyvE2KsRy9zPq/J
WojovG+WMIGcBgNVHSMegZQwgZGAFMbtvDyvE2KsRy9zPq/JWojovG+Ww6kbDBq
MQswCQYDVQQLGEwJ4eDELMAkGA1UECBMCeHgxGzAJBgNVBAAcTAnh4MQswCQYDVQK
EwJ4eDELMAkGA1UECXMceHgxGzAJBgNVBAMTAnh4MRowGAYJKoZIhvcNAQkBFgt4
eHhAMTYzLmNvbYIJALV96mEtVF4EMAwGA1UdEwQFMAMBAf8wDQYJKoZIhvcNAQEF
BQADgYEAA5SkC/1iwiALa2RU3YCxqZFEESZvQxikrDkDbFeoa6Tk49Fnb1f7FCW6
PTtY3HPWL5ygsMsSy0Fi3xp3jmulwzJhcQ3tcK5gC99HWp6Kw37RL8WoB8GWFUOQ
4tHLOjBixkZROPRhH+zMlrqUexv6fsb3NWKhnlfh1Mj5wQE4Ldo=
-----END CERTIFICATE-----
```

Private Key Format

When creating a server certificate, you also need to upload the private key of the certificate. You can copy and paste the private key content or directly upload the private key in the required format.

Private keys must be unencrypted and meet the following requirements:

- The value must be in PEM format.
 - The content must start with -----BEGIN RSA PRIVATE KEY----- and end with -----END RSA PRIVATE KEY-----.

- The content must start with -----BEGIN EC PRIVATE KEY----- and end with -----END EC PRIVATE KEY-----.
- There are no empty rows. Each row must contain 64 characters except the last row.

The following is an example:

```
-----BEGIN RSA PRIVATE KEY-----
MIICXQIBAAKBgQDFPN9ojPndxSC4E1pqWQVKGHCFFIXAAGBOxbGfSzXqzsoyacotu
eqMqXQbXrPSQFATeVmhzPNVEMdvcAMjYsV/mymtAwVqVA6q/OfdX/b3UHO+b/VqL
o3J5SrM86VeqnjzWu4oCSabuEDiN+tga1syQmEG4OFM6NSmAYSxcZdE6LwIDAQAB
AoGBAJvLzJCyIsCjCKHWL6onbSUTDtyFwPViD1QrVAtQYabF14g8CGUZG/9fgheu
TXPtTDcVU7cZdUArvgYW3I9F9IBb2lmF3a44xfiAKdDhzr4DK/vQhvHPuuTeZA41
r2zp8Cu+Bp40pSxmoAOK3B0/peZAka01Ju7c7ZChDWrxleHZAKEA/6dcaWHofG5
eW5YlBsms3f0m0GH38nRl7oxyCW6yMIDkFHURVMBKW1OhrcuGo8u0nTmi5IH9gRg
5bH8XcujlQJBAMWBQgzCHyoSeryD3TFieXIFzgDBw6Ve5hyMjUtjvgdVKoxRPvpO
kclc39QHP6Dm2wrXXHEej+9RILxBZCVQNbMCQQC42i+Ut0nHvPuXN/UkXzomDHde
h1ySsOAO4H+8Y6OSI87l3HUrByCQ7stX1z3L0HofjHqV9Koy9emGTFLEzSdAkB7
Ei6cUKKmtzkYe3rr+RcATEmwAw3tEJOHmrW5ErApVZKr2TzLMQZ7WZpIPzQRCYnY
ZZZLDuZWFFG3vW+wKKktAkAaQ5GNzbwKRLpXF1FZFuNF7erxypzstbUmU/31b7tS
i5LmxTGKL/xRYtZEHjya4lkkkg40q1MrUsglybFYMF2
-----END RSA PRIVATE KEY-----
```

7.3 Converting Certificate Formats

Scenarios

ELB supports certificates only in PEM format. If you have a certificate in any other format, you must convert it to a PEM-encoded certificate. There are some common methods for converting a certificate from any other format to PEM.

From DER to PEM

The DER format is usually used on a Java platform.

Run the following command to convert the certificate format:

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

Run the following command to convert the private key format:

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

From P7B to PEM

The P7B format is usually used by Windows Server and Tomcat.

Run the following command to convert the certificate format:

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

From PFX to PEM

The PFX format is usually used by Windows Server.

Run the following command to convert the certificate format:

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

Run the following command to convert the private key format:

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

7.4 Adding a Certificate

Scenarios

To enable authentication for securing data transmission over HTTPS, ELB allows you to bind certificates to HTTPS listeners of a load balancer.

- **Server certificate:** You can purchase a certificate from SSL Certificate Manager (SCM) or upload your own certificates.
- **CA certificate:** You can only upload your own CA certificates.

NOTE

If you want to use the same certificate in two regions, you need to create a certificate in each region.

Adding a Server Certificate



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Certificates**.
5. Click **Add Certificate** on the top right corner and set parameters by referring to [Table 7-1](#).

Table 7-1 Server certificate parameters

Parameter	Description	Example Value
Certificate Type	<p>Specifies the certificate type.</p> <ul style="list-style-type: none">• Server certificate: used for SSL handshake negotiations if an HTTPS listener is used. Both the certificate content and private key are required.• CA certificate: issued by a certificate authority (CA) and used to verify the certificate issuer. If HTTPS mutual authentication is required, HTTPS connections can be established only when the client provides a certificate issued by a specific CA.	Server certificate

Parameter	Description	Example Value
Source	<p>Specifies the source of a certificate. You can purchase a certificate from SCM or upload your own certificates.</p> <ul style="list-style-type: none">• SCM certificate: server certificate provided by SCM. You need to buy a certificate or upload your own certificate on the SCM console.• Your certificate: You need to upload the certificate content and private key of your own certificate on the ELB console. <p>NOTE You are advised to use SCM to manage your certificates.</p>	SCM certificate
Certificate	<p>This parameter is only available for SCM certificates.</p> <p>You can select certificates provided by SCM.</p>	-
Certificate Name	<p>Specifies the name of your certificate. This parameter is only available for your certificates.</p>	-
Enterprise Project	<p>Specifies an enterprise project by which cloud resources and members are centrally managed.</p>	default
Certificate Content	<p>Specifies the content of a certificate. This parameter is only available for your certificates.</p> <p>The content must be in PEM format. Click Upload and select a certificate. Ensure that your browser is the latest version.</p> <p>The format is as follows:</p> <pre>-----BEGIN CERTIFICATE----- Base64-encoded certificate -----END CERTIFICATE-----</pre>	-

Parameter	Description	Example Value
Private Key	<p>Specifies the private key of a certificate. This parameter is only available for your certificates.</p> <p>Click Upload and select a private key. Ensure that your browser is the latest version.</p> <p>The value must be an unencrypted private key. The private key must be in PEM format. The format is as follows:</p> <pre>-----BEGIN PRIVATE KEY----- [key] -----END PRIVATE KEY-----</pre>	-
Domain Name	<p>The domain name must be specified if the certificate is intended for SNI.</p> <p>Only one domain name can be specified for each certificate, and the domain name must be the same as that in the certificate.</p>	-
Description	(Optional) Provides supplementary information about the certificate.	-

Adding a CA Certificate



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Certificates**.
5. Click **Add Certificate** on the top right corner and set parameters by referring to [Table 7-2](#).

Table 7-2 CA certificate parameters

Parameter	Description	Example Value
Certificate Type	Specifies the certificate type. <ul style="list-style-type: none">• Server certificate: used for SSL handshake negotiations if an HTTPS listener is used. Both the certificate content and private key are required.• CA certificate: issued by a certificate authority (CA) and used to verify the certificate issuer. If HTTPS mutual authentication is required, HTTPS connections can be established only when the client provides a certificate issued by a specific CA.	CA certificate
Certificate Name	Specifies the name of the CA certificate.	-
Enterprise Project	Specifies an enterprise project by which cloud resources and members are centrally managed.	default
Certificate Content	The content must be in PEM format. Click Upload and select a certificate. Ensure that your browser is the latest version. The format is as follows: -----BEGIN CERTIFICATE----- Base64-encoded certificate -----END CERTIFICATE-----	-
Description	(Optional) Provides supplementary information about the certificate.	-

6. Click **OK**.

7.5 Deleting a Certificate



Scenarios

If a certificate is no longer needed, you can delete it on the ELB console.

Constraints

A certificate that has been bound to an HTTPS listener cannot be deleted. Disassociate the certificate from the listener first by referring to [Replacing a Certificate](#).

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Certificates**.
5. Locate the certificate and click **Delete** in the **Operation** column.
6. Click **Yes**.

7.6 Replacing the Certificate Bound to a Listener

Scenarios

You need to bind a certificate when you add an HTTPS listener to a load balancer. If the certificate used by a listener has expired or needs to be replaced due to other reasons, you can replace the certificate on the **Listeners** tab page.

If the certificate is also used by other services such as WAF, replace the certificate on all these services to prevent service unavailability.

NOTE

Replacing certificates and private keys does not affect your applications.



Prerequisites

You have created a certificate by following the instructions in [Adding a Certificate](#).

Binding a Certificate

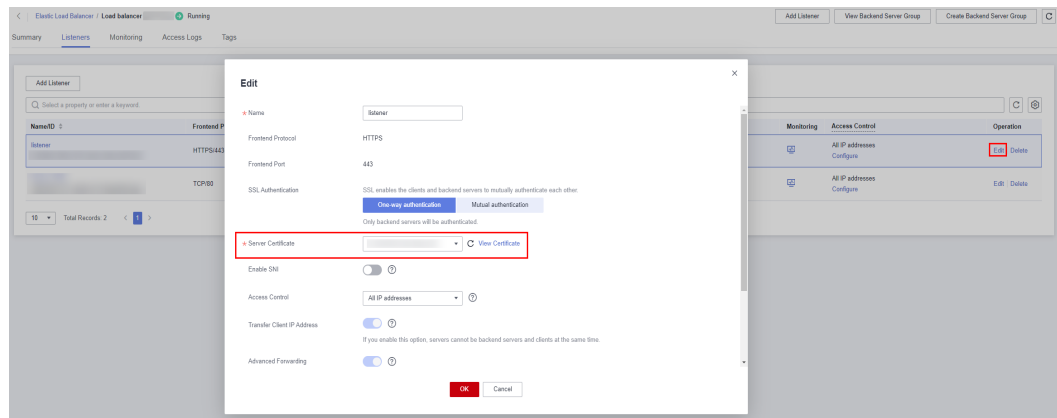
You can bind certificates when you add an HTTPS listener. For details, see [Adding an HTTPS Listener](#).

Replacing a Certificate

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click the **Listeners** tab, locate the listener, and click **Edit** in **Operation** column.
6. Select a server certificate.

7. Click **OK** in the **Edit** dialog box.

Figure 7-1 Replacing a certificate



7.7 Replacing the Certificate Bound to Different Listeners

Scenario

If the certificate that is bound to different listeners has expired or needs to be replaced due to other reasons, you can replace the certificate by modifying it on the **Certificates** page.



NOTE

Replacing the certificate and private keys does not affect your applications.

Constraints

- Only HTTPS listeners require certificates.
- The new certificate takes effect immediately. The previous certificate is used for established connections, and the new one is used for new connections.
- SSL Certificate Manager (SCM) allows you to purchase a certificate from Huawei Cloud or upload your own certificates for easier management.

Modifying a Certificate

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Certificates**.
5. Locate the certificate and click **Modify** in the **Operation** column.
6. Modify the parameters as required.



7. Confirm the information and click **OK**.

7.8 Querying Listeners by Certificate

Scenarios

You need to quickly view details of the listeners to which a certificate is bound.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Certificates**.
5. In the certificate list, click the listener name in the **Listener (Frontend Protocol/Port)** column to view its details.

If there are more than 5 listeners, no listener is displayed in the **Listener (Frontend Protocol/Port)** column. Click **View All**. On the displayed page, click **Listeners**, locate the listener, and click its name to view its details.

8 Access Control

8.1 Access Control

Access control allows you to add a whitelist or blacklist to specify IP addresses that are allowed or denied to access a listener. A whitelist allows specified IP addresses to access the listener, while a blacklist denies access from specified IP addresses.

NOTICE

- Adding the whitelist or blacklist may cause risks.
 - Once the whitelist is set, only the IP addresses specified in the whitelist can access the listener.
 - Once the blacklist is set, the IP addresses specified in the blacklist cannot access the listener.
- Whitelists and blacklists do not conflict with inbound security group rules. Whitelists define the IP addresses that are allowed to access the listeners, while blacklists specify IP addresses that are denied to access the listeners. Inbound security group rules control access to backend servers by specifying the protocol, ports, and IP addresses.
- Access control does not restrict the ping command. You can still ping backend servers from the restricted IP addresses.
 - To ping the IP address of a shared load balancer, you need to add a listener and associate a backend server to it.
 - To ping the IP address of a dedicated load balancer, you only need to add a listener to it.
- Access control policies only take effect for new connections, but not for connections that have been established. If a whitelist is configured for a listener but IP addresses that are not in the whitelist can access the backend server associated with the listener, one possible reason is that a persistent connection is established between the client and the backend server. To deny IP addresses that are not in the whitelist from accessing the listener, the persistent connection between the client and the backend server needs to be disconnected.

Configuring Access Control



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. You can configure access control for a listener in either of the following ways:
 - On the **Listeners** page, locate the listener and click **Configure** in the **Access Control** column.
 - Click the name of the target listener. On the **Summary** page, click **Configure** on the right of **Access Control**.
6. In the displayed **Configure Access Control** dialog box, configure parameters as shown in [Table 8-1](#).

Table 8-1 Parameter description

Parameter	Description	Example Value
Access Control	Specifies how access to the listener is controlled. Three options are available: <ul style="list-style-type: none">• All IP addresses: All IP addresses can access the listener.• Whitelist: Only IP addresses in the IP address group can access the listener.• Blacklist: IP addresses in the IP address group are not allowed to access the listener.	Blacklist
IP Address Group	Specifies the IP address group associated with a whitelist or blacklist. If there is no IP address group, create one first. For more information, see IP Address Group Overview .	ipGroup-b2
Access Control	If you have set Access Control to Whitelist or Blacklist , you can enable or disable access control. <ul style="list-style-type: none">• Only after you enable access control, the whitelist or blacklist takes effect.• If you disable access control, the whitelist or blacklist does not take effect.	N/A

7. Click **OK**.

8.2 Managing IP Address Groups

8.2.1 Creating an IP Address Group

IP Address Group Overview

An IP address group is a collection of IP addresses that you can use to manage IP addresses with the same security requirements or whose security requirements change frequently.

ELB allows you to use a whitelist or blacklist for access control. If you want to configure an **access control** policy, you must select an IP address group.

- **Whitelist:** Only IP addresses in the IP address group can access the listener. If the IP address group does not contain any IP address and you have selected whitelist for access control, no IP addresses can access the listener.
- **Blacklist:** IP addresses in the IP address group are denied to access the listener. If the IP address group does not contain any IP address and you have selected blacklist for access control, all IP addresses can access the listener.

Constraints

- By default, you can create a maximum of 50 IP address groups.
- An IP address group can be associated with a maximum of 50 listeners.

Procedure



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > IP Address Groups**.
5. On the displayed page, click **Create IP Address Group**.
6. Configure the parameters based on [Table 8-2](#).

Table 8-2 Parameters required for creating an IP address group

Parameter	Description	Example Value
Name	Specifies the name of the IP address group.	ipGroup-01
Enterprise Project	Specifies an enterprise project by which cloud resources and members are centrally managed. For details, see the Enterprise Management User Guide .	-

Parameter	Description	Example Value
IP Addresses	<p>Specifies IPv4 or IPv6 IP addresses or CIDR blocks that are added to the whitelist or blacklist for access control.</p> <ul style="list-style-type: none">Each line must contain an IP address or a CIDR block and end with a line break.Each IP address or CIDR block can include a description with a vertical bar () separated, for example, 192.168.10.10 ECS01. The description is 0 to 255 characters long and cannot contain angle brackets (<>).You can add a maximum of 300 IP addresses or CIDR blocks in each IP address group.	10.168.2.24 10.168.16.0/24
Description	Provides supplementary information about the IP address group.	-

7. Click **OK**.



8.2.2 Viewing the Details of an IP Address Group

Scenarios

This section describes how you can view information about an IP address group, including:

- Name, ID, and creation time
- IP addresses and CIDR blocks
- Associated listeners

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > IP Address Groups**.
5. On the **IP Address Groups** page, click the name of the target address group.
6. Viewing basic information about the IP address group.

- a. On the **IP Addresses** tab, view the IP addresses.
- b. On the **Associated Listeners** tab, view the listeners associated with the IP address group.

8.2.3 Managing IP Addresses in an IP Address Group

After an IP address group is created, you can manage the IP addresses in an IP address group as required:

- [Adding IP Addresses](#)
- [Changing IP Addresses](#)
- [Deleting an IP Address](#)



Constraints

The IP addresses can be in the following formats:

- Each line must contain an IP address or a CIDR block and end with a line break.
- Each IP address or CIDR block can include a description with a vertical bar (|) separated, for example, 192.168.10.10 | ECS01. The description is 0 to 255 characters long and cannot contain angle brackets (<>).
- You can add a maximum of 300 IP addresses or CIDR blocks in each IP address group.

Adding IP Addresses



After an IP address group is created, you can add IP addresses to an IP address group.

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > IP Address Groups**.
5. On the **IP Address Groups** page, click the name of the target address group.
6. In the lower part of the displayed page, choose **IP Addresses** tab and click **Add IP Addresses**.
7. On the **Add IP Addresses** page, add IP addresses.
8. Click **OK**.

Changing IP Addresses

You can perform the following steps to change all IP addresses in an IP address group:



1. Log in to the management console.

2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > IP Address Groups**.
5. On the **IP Address Groups** page, you can:
 - a. Modify the basic information and change IP addresses of an IP address group:
 - i. Locate the target address group, click **Modify** in the **Operation** column.
 - ii. You can modify the name and description of an IP address group, and change all its IP addresses.
 - iii. Click **OK**.
 - b. Only change IP addresses:
 - i. Click the name of the target IP address group.
 - ii. In the lower part of the displayed page, choose **IP Addresses** tab and click **Change IP Address**.
 - iii. Change IP addresses as you needed.
 - iv. Click **OK**.

Deleting an IP Address

If you want to delete IP addresses in batches from an IP address group, see [Changing IP Addresses](#).

To delete an IP address from an IP address group, perform the following operations:

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > IP Address Groups**.
5. On the **IP Address Groups** page, click the name of the target address group.
6. In the IP address list, locate the IP address you want to delete and click **Delete** in the **Operation** column.
A confirmation dialog box is displayed.
7. Confirm the information and click **Yes**.

8.2.4 Deleting an IP Address Group



Scenarios

If you no longer need an IP address group, you can delete it. This section describes how you can delete an IP address group.

Constraints

An IP address group that has been used for controlling access to a listener cannot be deleted. You can view the listeners associated with an IP address group by referring to [Viewing the Details of an IP Address Group](#). For details about how to disassociate an IP address group from a listener, see [Configuring Access Control](#).

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **Elastic Load Balance > IP Address Groups**.
5. On the **IP Address Groups** page, locate the IP address group, and click **Delete** in the **Operation** column.
6. Click **Yes**.

9 TLS Security Policy

Scenarios



When you add HTTPS listeners, you can select appropriate security policies to improve security. A security policy is a combination of TLS protocols of different versions and supported cipher suites.

- Dedicated load balancers: You can select the default security policy or create a custom policy. For details, see [Creating a Custom Security Policy](#).
- Shared load balancers: You can select the default security policy.

NOTE

Custom security policies can be created only in the CN-Hong Kong, AP-Bangkok, and AP-Singapore.

Adding a Security Policy

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Under **Listeners**, click **Add Listener**.
6. In the **Add Listener** dialog box, set **Frontend Protocol** to **HTTPS**.
7. Expand **Advanced Settings** and select a security policy.

[Table 9-1](#) shows the default security policies. Select a default security policy or create a custom security policy by referring to [Creating a Custom Security Policy](#).

Table 9-1 Default security policies

Security Policy	Description	TLS Versions	Cipher Suites
TLS-1-0	TLS 1.0, TLS 1.1, and TLS 1.2 and supported cipher suites (high compatibility and moderate security)	TLS 1.2 TLS 1.1 TLS 1.0	<ul style="list-style-type: none"> • ECDHE-RSA-AES256-GCM-SHA384 • ECDHE-RSA-AES128-GCM-SHA256 • ECDHE-ECDSA-AES256-GCM-SHA384 • ECDHE-ECDSA-AES128-GCM-SHA256 • AES128-GCM-SHA256 • AES256-GCM-SHA384 • ECDHE-ECDSA-AES128-SHA256 • ECDHE-RSA-AES128-SHA256 • AES128-SHA256
TLS-1-1	TLS 1.1 and TLS 1.2 and supported cipher suites (moderate compatibility and moderate security)	TLS 1.2 TLS 1.1	<ul style="list-style-type: none"> • AES256-SHA256 • ECDHE-ECDSA-AES256-SHA384 • ECDHE-RSA-AES256-SHA384 • ECDHE-ECDSA-AES128-SHA • ECDHE-RSA-AES128-SHA • ECDHE-RSA-AES256-SHA • ECDHE-ECDSA-AES256-SHA • AES128-SHA • AES256-SHA
TLS-1-2	TLS 1.2 and supported cipher suites (moderate compatibility and high security)	TLS 1.2	

Security Policy	Description	TLS Versions	Cipher Suites
<p>TLS-1-2-Strict</p>	<p>Strict TLS 1.2 and supported cipher suites (low compatibility and ultra-high security)</p>	<p>TLS 1.2</p>	<ul style="list-style-type: none"> ● ECDHE-RSA-AES256-GCM-SHA384 ● ECDHE-RSA-AES128-GCM-SHA256 ● ECDHE-ECDSA-AES256-GCM-SHA384 ● ECDHE-ECDSA-AES128-GCM-SHA256 ● AES128-GCM-SHA256 ● AES256-GCM-SHA384 ● ECDHE-ECDSA-AES128-SHA256 ● ECDHE-RSA-AES128-SHA256 ● AES128-SHA256 ● AES256-SHA256 ● ECDHE-ECDSA-AES256-SHA384 ● ECDHE-RSA-AES256-SHA384

Security Policy	Description	TLS Versions	Cipher Suites
TLS-1-0-WITH-1-3 (for dedicated load balancers)	TLS 1.0 and later, and supported cipher suites (ultra-high compatibility and low security)	TLS 1.3 TLS 1.2 TLS 1.1 TLS 1.0	<ul style="list-style-type: none"> • ECDHE-RSA-AES256-GCM-SHA384 • ECDHE-RSA-AES128-GCM-SHA256 • ECDHE-ECDSA-AES256-GCM-SHA384 • ECDHE-ECDSA-AES128-GCM-SHA256 • AES128-GCM-SHA256 • AES256-GCM-SHA384 • ECDHE-ECDSA-AES128-SHA256 • ECDHE-RSA-AES128-SHA256 • AES128-SHA256 • AES256-SHA256 • ECDHE-ECDSA-AES256-SHA384 • ECDHE-RSA-AES256-SHA384 • ECDHE-ECDSA-AES128-SHA • ECDHE-RSA-AES128-SHA • ECDHE-RSA-AES256-SHA • ECDHE-ECDSA-AES256-SHA • AES128-SHA • AES256-SHA • TLS_AES_128_GCM_SHA256 • TLS_AES_256_GCM_SHA384 • TLS_CHACHA20_POLY1305_SHA256 • TLS_AES_128_CCM_SHA256 • TLS_AES_128_CCM_8_SHA256

Security Policy	Description	TLS Versions	Cipher Suites
TLS-1-2-FS-WITH-1-3 (for dedicated load balancers)	TLS 1.2 and later, and supported forward secrecy cipher suites (high compatibility and ultra-high security)	TLS 1.3 TLS 1.2	<ul style="list-style-type: none">• ECDHE-RSA-AES256-GCM-SHA384• ECDHE-RSA-AES128-GCM-SHA256• ECDHE-ECDSA-AES256-GCM-SHA384• ECDHE-ECDSA-AES128-GCM-SHA256• ECDHE-ECDSA-AES128-SHA256• ECDHE-RSA-AES128-SHA256• ECDHE-ECDSA-AES256-SHA384• ECDHE-RSA-AES256-SHA384• TLS_AES_128_GCM_SHA256• TLS_AES_256_GCM_SHA384• TLS_CHACHA20_POLY1305_SHA256• TLS_AES_128_CCM_SHA256• TLS_AES_128_CCM_8_SHA256
TLS-1-2-FS (for dedicated load balancers)	TLS 1.2 and supported forward secrecy cipher suites (moderate compatibility and ultra-high security)	TLS 1.2	<ul style="list-style-type: none">• ECDHE-RSA-AES256-GCM-SHA384• ECDHE-RSA-AES128-GCM-SHA256• ECDHE-ECDSA-AES256-GCM-SHA384• ECDHE-ECDSA-AES128-GCM-SHA256• ECDHE-ECDSA-AES128-SHA256• ECDHE-RSA-AES128-SHA256• ECDHE-ECDSA-AES256-SHA384• ECDHE-RSA-AES256-SHA384

Security Policy	Description	TLS Versions	Cipher Suites
tls-1-2-strict-no-cbc (dedicated load balancers)	TLS 1.2 and supported cipher suites that exclude CBC encryption algorithm (low compatibility and ultra-high security)	TLS 1.2	<ul style="list-style-type: none"> • ECDHE-ECDSA-AES256-GCM-SHA384 • ECDHE-ECDSA-AES128-GCM-SHA256 • ECDHE-RSA-AES256-GCM-SHA384 • ECDHE-RSA-AES128-GCM-SHA256

 **NOTE**

- TLS-1-0-WITH-1-3, TLS-1-2-FS-WITH-1-3, TLS-1-2-FS, hybrid-policy-1-0, and tls-1-2-strict-no-cbc are available only for dedicated load balancers.
- The latest TLS version supported by dedicated load balancers is TLS 1.3, while the latest version supported by shared load balancers is TLS 1.2.
- This table lists the cipher suites supported by ELB. Generally, clients also support multiple cipher suites. In actual use, the intersection of the cipher suites supported by ELB and those supported by clients is used, and the cipher suites supported by ELB take precedence.

8. Click **OK**.

Differences Between Security Policies

Table 9-2 Differences between the security policies

Security Policy	TLS-1-0	TLS-1-1	TLS-1-2	TLS-1-2-Strict	TLS-1-0-WITH-1-3	TLS-1-2-FS-WITH-1-3	TLS-1-2-FS
TLS versions							
TLS 1.3	-	-	-	-	√	√	√
TLS 1.2	√	√	√	√	√	√	√
TLS 1.1	√	√	-	-	√	-	-

Security Policy	TLS-1-0	TLS-1-1	TLS-1-2	TLS-1-2-Strict	TLS-1-0-WITH-1-3	TLS-1-2-FS-WITH-1-3	TLS-1-2-FS
TLS 1.0	√	-	-	-	√	-	-
Cipher suite							
EDHE-RSA-AES128-GCM-SHA256	√	√	√	√	-	-	-
ECDHE-RSA-AES256-GCM-SHA384	√	√	√	√	√	√	√
ECDHE-RSA-AES128-SHA256	√	√	√	√	√	√	√
ECDHE-RSA-AES256-SHA384	√	√	√	√	√	√	√
AES128-GCM-SHA256	√	√	√	√	√	-	-
AES256-GCM-SHA384	√	√	√	√	√	-	-
AES128-SHA256	√	√	√	√	√	-	-
AES256-SHA256	√	√	√	√	√	-	-
ECDHE-RSA-AES128-SHA	√	√	√	-	√	-	-
ECDHE-RSA-AES256-SHA	√	√	√	-	√	-	-
AES128-SHA	√	√	√	-	√	-	-
AES256-SHA	√	√	√	-	√	-	-
ECDHE-ECDSA-AES128-GCM-SHA256	√	√	√	√	√	√	√
ECDHE-ECDSA-AES128-SHA256	√	√	√	√	√	√	√
ECDHE-ECDSA-AES128-SHA	√	√	√	-	√	-	-
ECDHE-ECDSA-AES256-GCM-SHA384	√	√	√	√	√	√	√
ECDHE-ECDSA-AES256-SHA384	√	√	√	√	√	√	√
ECDHE-ECDSA-AES256-SHA	√	√	√	-	√	-	-
ECDHE-RSA-AES128-GCM-SHA256	-	-	-	-	√	√	√

Security Policy	TLS-1-0	TLS-1-1	TLS-1-2	TLS-1-2-Strict	TLS-1-0-WITH-1-3	TLS-1-2-FS-WITH-1-3	TLS-1-2-FS
TLS_AES_256_GCM_SHA384	-	-	-	-	√	√	√
TLS_CHACHA20_POLY1305_SHA256	-	-	-	-	√	√	√
TLS_AES_128_GCM_SHA256	-	-	-	-	√	√	√
TLS_AES_128_CCM_8_SHA256	-	-	-	-	√	√	√
TLS_AES_128_CCM_SHA256	-	-	-	-	√	√	√

Creating a Custom Security Policy



1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **TLS Security Policies**.
5. On the displayed page, click **Create Custom Security Policy** in the upper right corner.
6. Configure the parameters based on [Table 9-3](#).

Table 9-3 Custom security policy parameters



Parameter	Description	Example Value
Name	Specifies the name of the custom security policy.	tls-test
TLS Version	Specifies the TLS version supported by the custom security policy. You can select multiple versions: <ul style="list-style-type: none"> • TLS 1.0 • TLS 1.1 • TLS 1.2 • TLS 1.3 	-

Parameter	Description	Example Value
Cipher Suite	Specifies the cipher suites that match the selected TLS versions.	-
Description	Provides supplementary information about the custom security policy.	-

7. Click **OK**.



Modifying a Custom Security Policy

You can modify a custom security policy as you need.

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **TLS Security Policies**.
5. On the **TLS Security Policies** page, click **Custom Security Policies**, locate the custom security policy, and click **Modify** in the **Operation** column.
6. In displayed dialog box, modify the custom security policy as described in [Table 9-3](#).
7. Click **OK**.



Deleting a Custom Security Policy

You can delete a custom security policy as you need.

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, choose **TLS Security Policies**.
5. On the **TLS Security Policies** page, click **Custom Security Policies**, locate the custom security policy, and click **Delete** in the **Operation** column.
6. Click **Yes**.

Changing a Security Policy

When you change a security policy, ensure that the security group containing backend servers allows traffic from 100.125.0.0/16 to backend servers and allows ICMP packets for UDP health checks. Otherwise, backend servers will be considered unhealthy, and routing will be affected.

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click its name.
6. On the **Summary** tab page, click **Edit** on the top right.
7. In the **Modify Listener** dialog box, expand **Advanced Settings** and change the security policy.
8. Click **OK**.

10 Tag



Scenarios

If you have a large number of cloud resources, you can add different tags to the resources to quickly identify them and use these tags to easily manage your resources.

If your organization has configured tag policies for ELB, add tags to load balancers based on the tag policies. If you add a tag that does not comply with the tag policies, load balancers may fail to be created. Contact your organization administrator to learn more about tag policies.

Adding a Tag to a Load Balancer

You can add a tag to a load balancer in the following method:



- Add a tag when you create a load balancer.
For detailed operations, see [Creating a Dedicated Load Balancer](#) and [Creating a Shared Load Balancer](#).
- Add a tag to an existing load balancer.
 - a. Log in to the management console.
 - b. In the upper left corner of the page, click  and select the desired region and project.
 - c. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
 - d. Locate the load balancer and click its name.
 - e. Under **Tags**, click **Add Tag**.
 - f. In the **Add Tag** dialog box, enter a tag key and value and click **OK**.

NOTE

- A maximum of 10 tags can be added to a load balancer.
- Each tag is a key-value pair, and the tag key is unique.

Adding a Tag to a Listener



To add a tag to an existing listener, perform the following steps:

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Listeners**, locate the listener, and click its name.
6. Under **Tags**, click **Add Tag**.
7. In the **Add Tag** dialog box, enter a tag key and value and click **OK**.

NOTE

- A maximum of 10 tags can be added to a listener.
- Each tag is a key-value pair, and the tag key is unique.

Modifying a Tag

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. Click **Tags**, select the tag to be edited, and click **Edit** in the **Operation** column. In the **Edit Tag** dialog box, change the tag value.



NOTE

The tag key cannot be changed.

6. Click **OK**.

The operations for modifying a listener tag are not detailed here. Refer to the operations of modifying a load balancer tag.

Deleting a Tag

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.

5. Click **Tags**, select the tag to be deleted, and click **Delete** in the **Operation** column.
6. In the displayed dialog box, click **Yes**.

The operations for deleting a listener tag are not detailed here. Refer to the operations of deleting a load balancer tag.

11 Access Logging

Scenarios

ELB logs HTTP and HTTPS requests received by load balancers, including the time when the request was sent, client IP address, request path, and server response. To enable access logging, you need to interconnect ELB with LTS and create a log group and a log stream on the LTS console.



Access logging is supported by HTTP/HTTPS listeners of both dedicated and shared load balancers.


NOTE

ELB displays operations data, such as access logs, on the LTS console. Do not transmit private or sensitive data through fields in access logs. Encrypt your sensitive data if necessary.

Configuring LTS


To view access logs, you first need to configure LTS by following the instructions in the [Log Tank Service User Guide](#).

1. Create a log group.
 - a. Log in to the management console.
 - b. In the upper left corner of the page, click  and select the desired region and project.
 - c. Click  in the upper left corner and **Management & Governance > Log Tank Service**.
 - d. In the navigation pane on the left, choose **Log Management**.
 - e. Click **Create Log Group**. In the displayed dialog box, enter a name for the log group.
Set **Log Retention Duration** as required.
 - f. Click **OK**.
2. Create a log stream.

- a. On the LTS console, click  on the left of a log group name.
- b. Click **Create Log Stream**. In the displayed dialog box, enter a name for the log stream.
- c. Select an enterprise project as required.
- d. Click **OK**.

Configuring Access Logging

Configure access logging on the ELB console.

1. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
2. Locate the load balancer and click its name.
3. Under **Access Logs**, click **Configure Access Logging**.
4. Enable access logging and select the log group and log stream you created.
5. Click **OK**.

Viewing Access Logs

After you enable access logging, you can obtain details about the requests sent to your load balancer.

There are two ways for you to view access logs.

- On the ELB console, click the name of the load balancer and click **Access Logs** to view logs.
- (Recommended) On the LTS console, click the name of the corresponding log topic. On the displayed page, click **Real-Time Logs**

The following is an example log. For details about the fields in the log, see [Table 11-1](#). The log format cannot be modified.

```
$msec $access_log_topic_id [$time_iso8601] $log_ver $remote_addr:$remote_port $status  
"$request_method $scheme://$host$router_request_uri $server_protocol" $request_length $bytes_sent  
$body_bytes_sent $request_time "$upstream_status" "$upstream_connect_time" "$upstream_header_time"  
"$upstream_response_time" "$upstream_addr" "$http_user_agent" "$http_referer" "$http_x_forwarded_for"  
$lb_name $listener_name $listener_id  
$pool_name "$member_name" $tenant_id $eip_address:$eip_port "$upstream_addr_priv" $certificate_id  
$ssl_protocol $ssl_cipher $sni_domain_name $tcpinfo_rtt $self_defined_header
```

Table 11-1 Parameter description

Parameter	Description	Description	Example Value
msec	Time in seconds with a millisecond resolution	Floating-point data	1530153091.868
access_log_topic_id	Log stream ID	UUID	04465dfa-640f-4567-8b58-45c9f8bbc23f

Parameter	Description	Description	Example Value
time_iso8601	Local time in the ISO 8601 standard format	-	2018-06-28T10:31:31+08:00
log_ver	Log format version	Fixed value: elb_01	elb_01
remote_addr: remote_port	IP address and port number of the client	Records the IP address and port of the client.	10.184.30.170:59605
status	HTTP status code	Records the request status code.	200
request_method scheme:// host request_uri server_protocol	<i>Request method Protocol://Host name: Request URI Request protocol</i>	<ul style="list-style-type: none">• request_method: request method• scheme: HTTP or HTTPS• host: host name, which can be a domain name or an IP address• request_uri: indicates the native URI initiated by the browser without any modification does not include the protocol and host name.	POST https://setting1.hicloud.com/AccountServer/!UserInfoMng/stAuth?Version=26400&Version=ID_SDK_2.6.4.300
request_length	Length of the request received from the client, including the header and body	Integer	295
bytes_sent	Number of bytes sent to the client	Integer	58470080
body_bytes_sent	Number of bytes sent to the client (excluding the response header)	Integer	58469792

Parameter	Description	Description	Example Value
request_time	Request processing time in seconds from the time when the load balancer receives the first request packet from the client to the time when the load balancer sends the response packet	Floating-point data	499.769
upstream_status	<p>Response status code returned by the backend server</p> <ul style="list-style-type: none"> When the load balancer attempts to retry a request, there will be multiple response status codes. If the request is not correctly routed to the backend server, a hyphen (-) is displayed as a null value for this field. 	HTTP status code returned by the backend server to the load balancer	200 or "-", 200", or "502, 502: 200", or "502:"

Parameter	Description	Description	Example Value
upstream_connect_time	<p>Time taken to establish a connection with the backend server, in seconds, with a millisecond resolution</p> <ul style="list-style-type: none">• When the load balancer attempts to retry a request, there will be multiple connection times.• If the request is not correctly routed to the backend server, a hyphen (-) is displayed as a null value for this field.	Floating-point data	0.008, "-", 0.008", "0.008, 0.005:0.004", or "0.008:"
upstream_header_time	<p>Time taken to receive the response header from the backend server, in seconds, with a millisecond resolution</p> <ul style="list-style-type: none">• When the load balancer attempts to retry a request, there will be multiple response times.• If the request is not correctly routed to the backend server, a hyphen (-) is displayed as a null value for this field.	Floating-point data	0.008, "-", 0.008", "0.008, 0.005:0.004", or "0.008:"

Parameter	Description	Description	Example Value
upstream_response_time	<p>Time taken to receive the response from the backend server, in seconds, with a millisecond resolution</p> <ul style="list-style-type: none">• When the load balancer attempts to retry a request, there will be multiple response times.• If the request is not correctly routed to the backend server, a hyphen (-) is displayed as a null value for this field.	Floating-point data	0.008, "-", 0.008", "0.008, 0.005:0.004", or "0.008:"
upstream_addr	<p>IP address and port number of the backend server. There may be multiple values separated by commas and spaces, and each value is in the format of <i>{IP address}:{Port number}</i> or <i>-</i>.</p> <p>This parameter is only available for dedicated load balancers.</p>	IP address and port number	-, or 192.168.1.2:8080
http_user_agent	<p>http_user_agent in the request header received by the load balancer, indicating the system model and browser information of the client</p>	Records the browser-related information.	Mozilla/5.0 (Windows NT 6.1; WOW64) AppleWebKit/537.36 (KHTML, like Gecko) Chrome/67.0.3396.99 Safari/537.36

Parameter	Description	Description	Example Value
http_referer	http_referer in the request header received by the load balancer, indicating the page link of the request	Request for a page link	http://10.154.197.90/
http_x_forwarded_for	http_x_forwarded_for in the request header received by the load balancer, indicating the IP address of the proxy server that the request passes through	IP address	10.154.197.90
lb_name	Load balancer name in the format of loadbalancer_Load balancer ID	String	loadbalancer_789424af-3fd2-4292-8c62-2a2dd7005175
listener_name	Listener name in the format of listener_Listener ID	String	listener_fde03b66-f960-440e-954a-0be8b2b75093
listener_id	Listener ID (This field can be ignored.)	String	-
pool_name	Backend server group name in the format of pool_backend server group ID	String	pool_066a5dc5-a3e4-4ea1-99f1-2a5716b681f6
member_name	Backend server name in the format of member_server ID (this field is not supported yet). There may be multiple values separated by commas and spaces, and each value is a member ID (member_id) or -.	String	member_47b07465-075a-4d2f-8ce9-0b9f39bff160 (There may be multiple values separated by commas and spaces, and each value is a member ID (member_id) or -.)
tenant_id	Tenant ID	String	04dd36f921000fe20f95c00bba986340

Parameter	Description	Description	Example Value
eip_address:eip_port	EIP of the load balancer and frontend port that were set when the listener was added	EIP of the load balancer and frontend port that were set when the listener was added	4.17.12.248:443
upstream_addr_priv	IP address and port number of the backend server. There may be multiple values separated by commas and spaces, and each value is in the format of <i>{IP address}:{Port number}</i> or <i>-</i> . This parameter is only available for dedicated load balancers.	IP address and port number	-, 192.168.1.2:8080 (There may be multiple values by commas and spaces, and each value is in the format of <i>{IP address}:{Port number}</i> or <i>-</i> .)
certificate_id	[HTTPS listener] Certificate ID used for establishing an SSL connection This field is not supported yet.	String	17b03b19-b2cc-454e-921b-4d187cce31dc
ssl_protocol	[HTTPS listener] Protocol used for establishing an SSL connection For a non-HTTPS listener, a hyphen (-) is displayed as a null value for this field.	String	TLS 1.2
ssl_cipher	[HTTPS listener] Cipher suite used for establishing an SSL connection For a non-HTTPS listener, a hyphen (-) is displayed as a null value for this field.	String	ECDHE-RSA-AES256-GCM-SHA384

Parameter	Description	Description	Example Value
sni_domain_name	[HTTPS listener] SNI domain name provided by the client during SSL handshake For a non-HTTPS listener, a hyphen (-) is displayed as a null value for this field.	String	www.test.com
tcpinfo_rtt	TCP Round Trip Time (RTT) between the load balancer and client in microseconds	Integer	39032
self_defined_header	This field is reserved. The default value is '-'.	String	-

Example Log

```
1644819836.370 eb11c5a9-93a7-4c48-80fc-03f61f638595 [2022-02-14T14:23:56+08:00] elb_01
192.168.1.1:888 200 "POST https://www.test.com/example /HTTP/1.1" 1411 251 3 0.011 "200" "0.000"
"0.011" "0.011" "100.64.0.129:8080" "okhttp/3.13.1" "-" "-"
loadbalancer_295a7eee-9999-46ed-9fad-32a62ff0a687 listener_20679192-8888-4e62-a814-a2f870f62148
3333fd44fe3b42cbaa1dc2c641994d90 pool_89547549-6666-446e-9dbc-e3a551034c46 "-"
f2bc165ad9b4483a9b17762da851bbbb 121.64.212.1:443 "10.1.1.2:8080" - TLSv1.2 ECDHE-RSA-AES256-
GCM-SHA384 www.test.com 56704 -
```

The following table describes the fields in the log.

Table 11-2 Fields in the log

Field	Example Value
msec	1644819836.370
access_log_topic_id	eb11c5a9-93a7-4c48-80fc-03f61f638595
time_iso8601	[2022-02-14T14:23:56+08:00]
log_ver	elb_01
remote_addr: remote_port	192.168.1.1:888
status	200
request_method scheme://host request_uri server_protocol	"POST https://www.test.com/ example/1 HTTP/1.1"
request_length	1411

Field	Example Value
bytes_sent	251
body_bytes_sent	3
request_time	0.011
upstream_status	"200"
upstream_connect_time	"0.000"
upstream_header_time	"0.011"
upstream_response_time	"0.011"
upstream_addr	"100.64.0.129:8080"
http_user_agent	"okhttp/3.13.1"
http_referer	"_"
http_x_forwarded_for	"_"
lb_name	loadbalancer_295a7eee-9999-46ed-9fad-32a62ff0a687
listener_name	listener_20679192-8888-4e62-a814-a2f870f62148
listener_id	3333fd44fe3b42cbaa1dc2c641994d90
pool_name	pool_89547549-6666-446e-9dbc-e3a551034c46
member_name	"_"
tenant_id	f2bc165ad9b4483a9b17762da851bbb
eip_address:eip_port	121.64.212.1:443
upstream_addr_priv	"10.1.1.2:8080"
certificate_id	-
ssl_protocol	TLSv1.2
ssl_cipher	ECDHE-RSA-AES256-GCM-SHA384
sni_domain_name	www.test.com
tcpinfo_rtt	56704
self_defined_header	-

Log analysis:



At 14:23:56 GMT+08:00 on Feb 14, 2022, the load balancer receives an HTTP/1.1 POST request from a client whose IP address and port number are 192.168.1.1 and 888, then routes the request to a backend server whose IP address and port number are 100.64.0.129 and 8080, and finally returns 200 OK to the client after receiving the status code from the backend server.

Analysis results:

The backend server responds to the request normally.

Configuring Log Transfer

If you want to analyze access logs later, transfer the logs to OBS or Data Ingestion Service (DIS) for storage.

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Click  in the upper left corner and **Management & Governance > Log Tank Service**.
4. In the navigation pane on the left, choose **Log Transfer**.
5. On the **Log Transfer** page, click **Configure Log Transfer** in the upper right corner.
 1. Configure the parameters. For details, see the [Log Tank Service User Guide](#).

12 Protection for Mission-Critical Operations

Scenarios

ELB supports sensitive operation protection. When you perform sensitive operations on the management console, you need to enter a credential that can prove your identity. You can perform corresponding operations only after your identity is authenticated. It is recommended that you enable operation protection to secure your account.

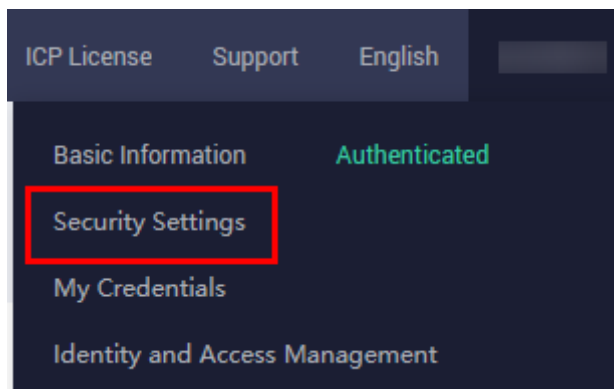
This function can be configured only by the administrator and takes effect for the resources in your account and the resources of users under your account. Common users have only the view permissions. To modify the permissions, contact the administrator.

Enabling Operation Protection

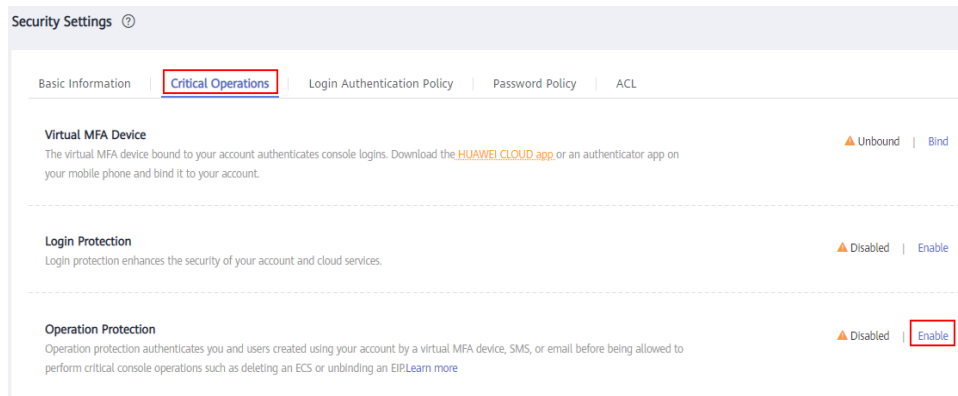
Operation protection is disabled by default. Perform the following operations to enable it:

1. Log in to the management console.
2. Move the cursor to the username in the upper right corner of the page and select **Security Settings** from the drop-down list.

Figure 12-1 Security Settings



3. On the **Security Settings** page, choose **Critical Operations > Operation Protection > Enable**.

Figure 12-2 Critical Operations

4. On the **Operation Protection** page, select **Enable** to enable operation protection.

When you or the IAM users under your account perform critical operations, for example, deleting ECS resources, you are required to enter a verification code based on the selected verification method.

NOTE

- When performing a critical operation, you will be asked to choose a verification method from email, SMS, and virtual MFA device.
 - If you have bound only a mobile number, only SMS verification is available.
 - If you have bound only an email address, only email verification is available.
 - If you have not bound an email address, mobile number, or virtual MFA device, you are required to bind one to continue with the critical operation.
- You can change the mobile number, email address, and virtual MFA device on the [Basic Information](#) page.

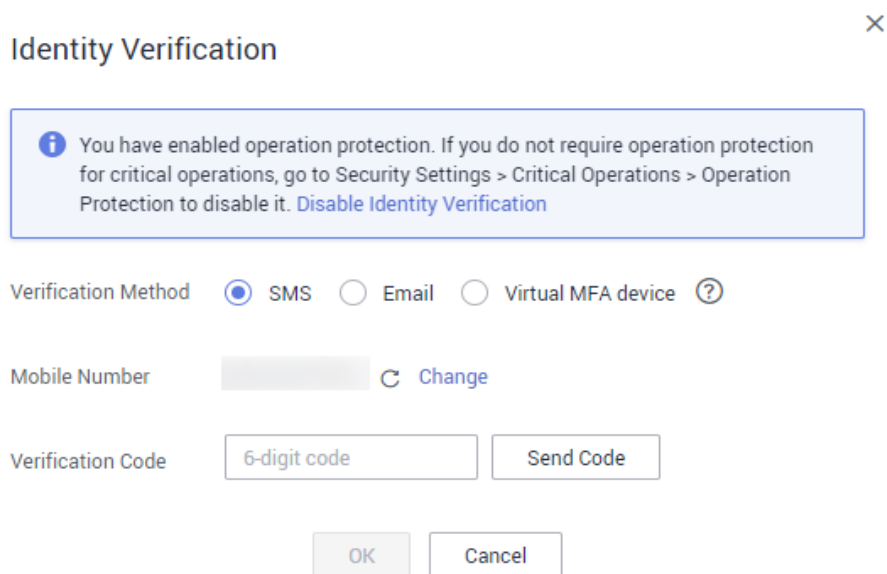
Verifying Operation Protection

After operation protection is enabled, when you perform a mission-critical operation, the system will verify your identity.

- If you have bound an email address, enter the email verification code.
- If you have bound a mobile number, enter the SMS verification code.
- If you have bound a virtual MFA device, enter a 6-digit dynamic verification code of the MFA device.

When you attempt to delete a load balancer, the following dialog box is displayed, and you need to select a verification method:

Figure 12-3 Identity Verification

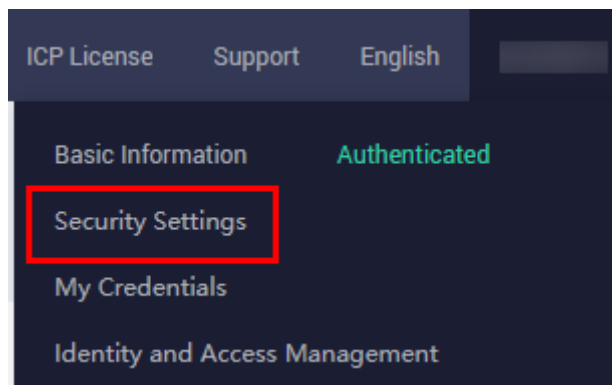


Disabling Operation Protection

Perform the following operations to disable operation protection.

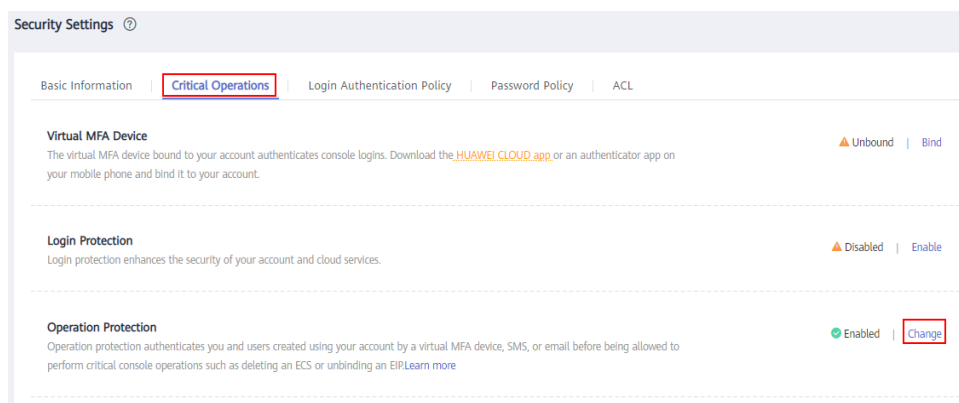
1. Log in to the management console.
2. Move the cursor to the username in the upper right corner of the page and select **Security Settings** from the drop-down list.

Figure 12-4 Security Settings



3. On the **Security Settings** page, choose **Critical Operations > Operation Protection > Change**.

Figure 12-5 Modifying operation protection settings



4. On the **Operation Protection** page, select **Disable** and click **OK**.

References

- [How Do I Bind a Virtual MFA Device?](#)
- [How Do I Obtain an MFA Verification Code?](#)

13 Self-service Troubleshooting

13.1 Overview

ELB self-service troubleshooting helps you detect and fix unhealthy backend servers in a timely manner. It also gets you familiar with billing and service features that you might be curious about. During the troubleshooting process, resource configurations will not be changed and services will work normally.

You may find the answers to the issues listed in [Table 13-1](#).

 **NOTE**

Self-service troubleshooting is available in AP-Bangkok.

Table 13-1 ELB self-service troubleshooting

Issue	Description
Troubleshooting an Unhealthy Backend Server	<ul style="list-style-type: none">• Checks the security group rules.• Checks the network ACL configurations.• Checks the health check ports.
ELB Billing	Describes how ELB is billed.
Differences Between Dedicated and Shared Load Balancers	Describes the advantages of each type of load balancer.

13.2 Troubleshooting an Unhealthy Backend Server

Scenarios

This section describes how you can use ELB self-service troubleshooting to detect and fix unhealthy backend servers in a timely manner.

 NOTE

Self-service troubleshooting is available in AP-Bangkok.

Prerequisites



Before troubleshooting an unhealthy backend server, make sure you have completed the following:

- [Creating a Dedicated Load Balancer](#)
- [Creating a Backend Server Group \(Dedicated Load Balancers\)](#)
- [Adding a TCP Listener](#)
- [Modifying Health Check Settings](#)

Constraints

- You can only troubleshoot an unhealthy backend server.
- The backend server must be associated with a listener.
- IP as backend servers does not support self-service troubleshooting.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. In the navigation pane on the left, click **Self-service Troubleshooting**.
5. On the **Elastic Load Balance** tab, click **Unhealthy backend servers**.
6. Select the load balancer that has unhealthy backend servers.
7. Select the unhealthy backend server you want to troubleshoot.
8. Click **Troubleshoot**. On the displayed page, view the troubleshooting progress and details.

View and rectify the faults in a timely manner as described in [Table 13-2](#).

Table 13-2 Health check items

Health Check Category	Health Check Item	Reason	Suggestion
Security group rule configurations	The protocol configured for the inbound rule	The inbound rules of the security group do not allow traffic over the health check protocol.	Change the security group rules by referring to the following: <ul style="list-style-type: none"> • Security Group Rules • Security Group Rules
	The source IP address configured for the inbound rule	The inbound rules of the security group do not allow traffic from the source IP address to the backend server.	
	The port configured for the inbound rule	The inbound rules of the security group do not allow traffic over the health check port.	
	The protocol configured for the outbound rule	The outbound rules of the security group do not allow traffic over the health check protocol.	
	The source IP address configured for the outbound rule	The outbound rules of the security group do not allow traffic from the source IP address to the backend server.	
	The port configured for the outbound rule	The outbound rules of the security group do not allow traffic over the health check port.	

Health Check Category	Health Check Item	Reason	Suggestion
Network ACL rule configurations	The protocol configured for the inbound rule	The inbound rules of the network ACL do not allow traffic over the health check protocol.	Change the network ACL rules by referring to the following: <ul style="list-style-type: none">• Security Group Rules• Security Group Rules
	The source IP address configured for the inbound rule	The inbound rules of the network ACL do not allow traffic from the source IP address to the backend server.	
	The source port configured for the inbound rule	The inbound rules of the network ACL do not allow traffic over all source ports.	
	The destination address configured for the inbound rule	The inbound rules of the network ACL do not allow traffic to the destination address.	
	The destination port configured for the inbound rule	The inbound rules of the network ACL do not allow traffic over the destination port.	
	The protocol configured for the outbound rule	The outbound rules of the network ACL do not allow traffic over the health check protocol.	

Health Check Category	Health Check Item	Reason	Suggestion
	The source IP address configured for the outbound rule	The outbound rules of the network ACL do not allow traffic from the source IP address to the backend server.	
	The source port configured for the outbound rule	The outbound rules of the network ACL do not allow traffic over the health check port.	
	The destination address configured for the outbound rule	The outbound rules of the network ACL do not allow traffic to the destination address.	
	The destination port configured for the outbound rule	The outbound rules of the network ACL do not allow traffic over all destination ports.	
Health check configurations	The port configured for the health check	The specified health check port is different from that used by the backend server.	Use the backend port as the health check port by referring to Modifying Health Check Settings .

 NOTE

- If all the check items are reported as normal, perform further checks as guided by [How Do I Troubleshoot an Unhealthy Backend Server?](#)
- If the troubleshooting fails, click **Troubleshoot Again** or perform further checks as guided by [How Do I Troubleshoot an Unhealthy Backend Server?](#)

13.3 Other Issues

You can also use ELB self-service troubleshooting to find the answers to the following issues:

- [ELB Billing](#)
- [Differences Between Dedicated and Shared Load Balancers](#)

ELB Billing

You can learn more about ELB billing as described in [Table 13-3](#).

Table 13-3 ELB billing

Scenario	Reference
Billing rules	<ul style="list-style-type: none">• Billing (Dedicated Load Balancers)• Billing (Shared Load Balancers)
Billing modes	Changing the Billing Mode or Bandwidth Billing Option
Specifications	Changing the Specifications of a Dedicated Load Balancer

Differences Between Dedicated and Shared Load Balancers

Learn more about the advantages of each type of load balancer as described in [Table 13-4](#).

Table 13-4 Differences

Scenario	Reference
Feature comparison	Differences Between Dedicated and Shared Load Balancers
Creating a backend server group	<ul style="list-style-type: none">• Creating a Backend Server Group (Dedicated Load Balancers)• Creating a Backend Server Group (Shared Load Balancers)
Adding a backend server	<ul style="list-style-type: none">• Overview• Overview

14 Monitoring

14.1 Monitoring Metrics

Overview

This section describes the namespace, the metrics that can be monitored by Cloud Eye, and dimensions of these metrics. You can view the metrics reported by ELB and the generated alarms on the Cloud Eye console. For details, see [Viewing Metrics](#).

Namespace

SYS.ELB

Metrics

Table 14-1 Metrics supported by ELB

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
m1_cps	Concurrent Connections	<p>Load balancing at Layer 4: total number of TCP and UDP connections from the monitored object to backend servers</p> <p>Load balancing at Layer 7: total number of TCP connections from the clients to the monitored object</p> <p>Unit: N/A</p>	≥ 0	<ul style="list-style-type: none"> • Dedicated load balancer • Shared load balancer • Dedicated load balancer - listener • Shared load balancer - listener 	1 minute
m2_act_conn	Active Connections	<p>Number of TCP and UDP connections in the ESTABLISHED state between the monitored object and backend servers</p> <p>You can run the following command to view the connections (both Windows and Linux servers): netstat -an</p> <p>Unit: N/A</p>	≥ 0	<ul style="list-style-type: none"> • Shared load balancer - listener 	

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
m3_inact_conn	Inactive Connections	Number of TCP connections between the monitored object and backend servers except those in the ESTABLISHED state You can run the following command to view the connections (both Windows and Linux servers): netstat -an Unit: N/A	≥ 0		
m4_ncps	New Connections	Number of connections established between clients and the monitored object per second Unit: Count/s	≥ 0/ second		
m5_in_pps	Incoming Packets	Number of packets received by the monitored object per second Unit: Packet/s	≥ 0/ second		
m6_out_pps	Outgoing Packets	Number of packets sent from the monitored object per second Unit: Packet/s	≥ 0/ second		
m7_in_Bps	Inbound Rate	Traffic used for accessing the monitored object from the Internet per second Unit: byte/s	≥ 0 bytes/s		

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
m8_out_Bps	Outbound Rate	Traffic used by the monitored object to access the Internet per second Unit: byte/s	≥ 0 bytes/s		
m9_abnormal_servers	Unhealthy Servers	Number of unhealthy backend servers associated with the monitored object Unit: N/A	≥ 0	<ul style="list-style-type: none"> Dedicated load balancer Shared load balancer 	1 minute
ma_normal_servers	Healthy Servers	Number of healthy backend servers associated with the monitored object Unit: N/A	≥ 0		
m1e_server_rps	Reset Packets from Backend Servers	(TCP listener metrics) Number of reset packets forwarded by the monitored object from backend servers to clients Unit: Packet/s	≥ 0 /second	<ul style="list-style-type: none"> Shared load balancer Shared load balancer - listener 	1 minute
m21_client_rps	Reset Packets from Clients	(TCP listener metrics) Number of reset packets forwarded by the monitored object from clients to backend servers Unit: Packet/s	≥ 0 /second		
m1f_lvs_rps	Reset Packets from Load Balancers	(TCP listener metrics) Number of reset packets generated by the monitored object per second Unit: Packet/s	≥ 0 /second		

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
m22_in_bandwidth	Inbound Bandwidth	Bandwidth used for accessing the monitored object from the Internet Unit: bit/s	≥ 0 bit/s	<ul style="list-style-type: none"> Shared load balancer Shared load balancer - listener 	1 minute
m23_out_bandwidth	Outbound Bandwidth	Bandwidth used by the monitored object to access the Internet Unit: bit/s	≥ 0 bit/s		
mb_l7_queries	Layer-7 Query Rate	Number of requests the monitored object receives per second Unit: Query/s	≥ 0 query/s	<ul style="list-style-type: none"> Dedicated load balancer Shared load balancer Dedicated load balancer - listener Shared load balancer - listener 	1 minute
md_l7_http_3xx	Layer-7 3xx Status Codes	Number of 3xx status codes returned by the monitored object Unit: Count/s	≥ 0 /second	<ul style="list-style-type: none"> Dedicated load balancer Shared load balancer Dedicated load balancer - listener Shared load balancer - listener 	1 minute

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
mc_l7_http_2xx	Layer-7 2xx Status Codes	Number of 2xx status codes returned by the monitored object Unit: Count/s	≥ 0/second	<ul style="list-style-type: none"> • Dedicated load balancer • Shared load balancer • Dedicated load balancer - listener • Shared load balancer - listener 	1 minute
me_l7_http_4xx	Layer-7 4xx Status Codes	Number of 4xx status codes returned by the monitored object Unit: Count/s	≥ 0/second		
mf_l7_http_5xx	Layer-7 5xx Status Codes	Number of 5xx status codes returned by the monitored object Unit: Count/s	≥ 0/second		
m10_l7_http_other_status	Layer-7 Other Status Codes	Number of status codes returned by the monitored object except 2xx, 3xx, 4xx, and 5xx status codes Unit: Count/s	≥ 0/second		
m11_l7_http_404	Layer-7 404 Not Found	Number of 404 Not Found status codes returned by the monitored object Unit: Count/s	≥ 0/second		
m12_l7_http_499	Layer-7 499 Client Closed Request	Number of 499 Client Closed Request status codes returned by the monitored object Unit: Count/s	≥ 0/second		
m13_l7_http_502	Layer-7 502 Bad Gateway	Number of 502 Bad Gateway status codes returned by the monitored object Unit: Count/s	≥ 0/second		

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
m14_l7_rt	Average Layer-7 Response Time	<p>Average response time of the monitored object</p> <p>The response time starts when the monitored object receives requests from the clients and ends when it returns all responses to the clients.</p> <p>Unit: ms</p> <p>NOTE The average response time it takes to establish a WebSocket connection may be very high. This metric cannot be used as a reference.</p>	≥ 0 ms		
m15_l7_upstream_4xx	4xx Status Codes Backend	<p>Number of 4xx status codes returned by the monitored object</p> <p>Unit: Count/s</p>	≥ 0/second	<ul style="list-style-type: none"> • Dedicated load balancer • Shared load balancer • Dedicated load balancer - listener • Shared load balancer - listener 	1 minute
m16_l7_upstream_5xx	5xx Status Codes Backend	<p>Number of 5xx status codes returned by the monitored object</p> <p>Unit: Count/s</p>	≥ 0/second		

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
m17_l7_upstream_rt	Average Server Response Time	<p>Average response time of backend servers</p> <p>The response time starts when the monitored object routes the requests to the backend server and ends when the monitored object receives a response from the backend server.</p> <p>Unit: ms</p> <p>NOTE The average response time it takes to establish a WebSocket connection may be very high. This metric cannot be used as a reference.</p>	≥ 0 ms		
m1a_l7_upstream_rt_max	Maximum Server Response Time	<p>Maximum response time of backend servers (This metric is available only when the frontend protocol is HTTP or HTTPS.)</p> <p>The response time starts when the monitored object routes the requests to the backend server and ends when the monitored object receives a response from the backend server.</p> <p>Unit: ms</p>	≥ 0 ms	<ul style="list-style-type: none"> • Dedicated load balancer • Shared load balancer • Dedicated load balancer - listener • Shared load balancer - listener 	1 minute

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
m1b_l7_upstream_rt_min	Minimum Server Response Time	<p>Minimum response time of backend servers (This metric is available only when the frontend protocol is HTTP or HTTPS.)</p> <p>The response time starts when the monitored object routes the requests to the backend server and ends when the monitored object receives a response from the backend server.</p> <p>Unit: ms</p>	≥ 0 ms		
m1c_l7_rt_max	Maximum Layer-7 Response Time	<p>Maximum response time of the monitored object (This metric is available only when the frontend protocol is HTTP or HTTPS.)</p> <p>The response time starts when the monitored object receives requests from the clients and ends when it returns all responses to the clients.</p> <p>Unit: ms</p>	≥ 0 ms	<ul style="list-style-type: none"> • Dedicated load balancer • Shared load balancer • Dedicated load balancer - listener • Shared load balancer - listener 	1 minute

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
m1d_l7_rt_min	Minimum Layer-7 Response Time	<p>Minimum response time of the monitored object (This metric is available only when the frontend protocol is HTTP or HTTPS.)</p> <p>The response time starts when the monitored object receives requests from the clients and ends when it returns all responses to the clients.</p> <p>Unit: ms</p>	≥ 0 ms		
l7_con_usage	Layer-7 Concurrent Connection Usage	<p>Ratio of HTTP and HTTPS connections established between the monitored object and backend servers per second, to the maximum number of concurrent connections allowed per second</p> <p>Unit: percent (%)</p>	≥ 0%	Dedicated load balancer	1 minute

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
l7_in_bps_usage	Layer-7 Inbound Bandwidth Usage	<p>Ratio of the bandwidth that the monitored object uses to return response to clients over HTTP and HTTPS, to the maximum outbound bandwidth allowed</p> <p>Unit: percent (%)</p> <p>CAUTION If the inbound bandwidth usage reaches 100%, the load balancer performance has reached the upper limit. If the inbound bandwidth keeps higher than the bandwidth that the load balancer can provide, the service availability cannot be guaranteed.</p>	≥ 0%		
l7_out_bps_usage	Layer-7 Outbound Bandwidth Usage	<p>Ratio of the bandwidth that the monitored object uses to return response to clients over HTTP and HTTPS, to the maximum outbound bandwidth allowed</p> <p>Unit: percent (%)</p> <p>CAUTION If the outbound bandwidth usage reaches 100%, the load balancer performance has reached the upper limit. If the outbound bandwidth keeps higher than the bandwidth that the load balancer can provide, the service availability cannot be guaranteed.</p>	≥ 0%		

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
l7_ncps_usage	Layer-7 New Connection Usage	Ratio of HTTP and HTTPS connections established between clients and the monitored object per second, to the maximum number of new connections allowed per second Unit: percent (%)	≥ 0%		
l7_qps_usage	Layer 7 QPS Usage	Ratio of HTTP and HTTPS queries per second on the monitored object, to the maximum number of queries allowed per second Unit: percent (%)	≥ 0%		
m18_l7_upstream_2xx	2xx Status Codes_Backend	Number of 2xx status codes returned by the monitored object (This metric is available only when the frontend protocol is HTTP or HTTPS.) Unit: Count/s	≥ 0/second	<ul style="list-style-type: none"> Dedicated load balancer - backend server group Shared load balancer - backend server group 	1 minute
m19_l7_upstream_3xx	3xx Status Codes_Backend	Number of 3xx status codes returned by the monitored object (This metric is available only when the frontend protocol is HTTP or HTTPS.) Unit: Count/s	≥ 0/second	<ul style="list-style-type: none"> Shared load balancer - backend server group 	

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
m25_l7_resp_Bps	Backend Server Response Bandwidth	The bandwidth that the monitored object uses to return response to clients Unit: bit/s NOTE When HTTP/2 is enabled for a listener, this metric cannot be used as a reference.	≥ 0 bit/s		
m24_l7_req_Bps	Backend Server Request Bandwidth	The bandwidth that the monitored object uses to receive requests from clients Unit: bit/s NOTE When HTTP/2 is enabled for a listener, this metric cannot be used as a reference.	≥ 0 bit/s		
l4_con_usage	Layer-4 Concurrent Connection Usage	Ratio of TCP and UDP connections established between the monitored object and backend servers per second, to the maximum number of concurrent connections allowed per second Unit: percent (%)	$\geq 0\%$	Dedicated load balancer	1 minute

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
l4_in_bps_usage	Layer-4 Inbound Bandwidth Usage	<p>Ratio of the bandwidth that the monitored object uses to receive requests from clients over TCP and UDP, to the maximum inbound bandwidth allowed</p> <p>Unit: percent (%)</p> <p>CAUTION If the inbound bandwidth usage reaches 100%, the load balancer performance has reached the upper limit. If the inbound bandwidth keeps higher than the bandwidth that the load balancer can provide, the service availability cannot be guaranteed.</p>	≥ 0%		
l4_out_bps_usage	Layer-4 Outbound Bandwidth Usage	<p>Ratio of the bandwidth that the monitored object uses to return response to clients over TCP and UDP, to the maximum outbound bandwidth allowed</p> <p>Unit: percent (%)</p> <p>CAUTION If the outbound bandwidth usage reaches 100%, the load balancer performance has reached the upper limit. If the outbound bandwidth keeps higher than the bandwidth that the load balancer can provide, the service availability cannot be guaranteed.</p>	≥ 0%		

Metric ID	Name	Description	Value	Monitored Object	Monitoring Period (Raw Data)
l4_ncps_usage	Layer-4 New Connection Usage	Ratio of TCP and UDP connections established between clients and the monitored object per second, to the maximum number of new connections allowed per second Unit: percent (%)	≥ 0%		



Dimensions

Key	Value
lbaas_instance_id	<ul style="list-style-type: none">ID of a dedicated load balancerID of a shared load balancer
lbaas_listener_id	<ul style="list-style-type: none">ID of a listener added to a dedicated load balancerID of a listener added to a shared load balancer
lbaas_pool_id	ID of the backend server group

14.2 Setting an Alarm Rule

You can add, modify, and delete alarm rules. For details, see the [Cloud Eye User Guide](#).

14.2.1 Creating an Alarm Rule

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Click  in the upper left corner and choose **Management & Governance > Cloud Eye**.
4. In the navigation pane on the left, choose **Alarm Management > Alarm Rules**.

5. On the displayed **Alarm Rules** page, click **Create Alarm Rule**.
[Table 14-2](#) describes how to create an alarm rule.

Table 14-2 Configuring parameters



Parameter	Setting
Resource Type	Select Elastic Load Balance .
Dimension	Select from the following options: <ul style="list-style-type: none">• Elastic Load Balancers• Listeners• Backend Server Group NOTE For a shared load balancer, Listeners cannot be selected as a dimension.
Other Parameters	Set this parameter as required.

Once the alarm rule is created and the notification function has been enabled, the system automatically sends you a notification when an alarm is generated.

 **NOTE**

For more information about alarm rules of load balancers and listeners, see the [Cloud Eye User Guide](#).

14.2.2 Modifying an Alarm Rule

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Click  in the upper left corner and choose **Management & Governance > Cloud Eye**.
4. In the navigation pane on the left, choose **Alarm Management > Alarm Rules**.
5. On the **Alarm Rules** page, locate the alarm rule and click **Modify** in the **Operation** column.
 - a. On the **Modify Alarm Rule** page, modify the parameters.
 - b. Set other parameters as required and then click **Modify**.

Once the alarm rule is set and you have enabled the notification function, the system automatically sends you a notification when an alarm is generated.

 **NOTE**

For more information about alarm rules of load balancers and listeners, see the [Cloud Eye User Guide](#).

14.3 Viewing Metrics

Scenarios

Cloud Eye provided by the public cloud platform monitors the running statuses of load balancers.

You can view the metrics of each load balancer on the ELB console or the Cloud Eye console.

The transmission of monitoring data takes a while, so the status of each load balancer displayed on the Cloud Eye dashboard is not its real-time status. For a newly created load balancer or a newly added listener, you need to wait for about 5 minutes to 10 minutes before you can view its metrics.

Prerequisites



- The load balancer is running properly.
If backend servers are stopped, faulty, or deleted, no monitoring data is displayed.

NOTE

Cloud Eye stops monitoring a load balancer and removes it from the monitored object list if its backend servers have been deleted or are in stopped or faulty state for over 24 hours. However, the configured alarm rules will not be automatically deleted.



- You have interconnected ELB with Cloud Eye and configured an alarm rule for the load balancer on the Cloud Eye console.
Without alarm rules, there is no monitoring data. For details, see [Setting an Alarm Rule](#).
- If an IAM user wants to view the ELB monitoring data on the Cloud Eye console, the IAM user must be granted the **ELB Administrator** permission. Otherwise, the IAM user cannot view all monitoring data.

Viewing Monitoring Metrics on the ELB Console

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Hover on  in the upper left corner to display **Service List** and choose **Networking > Elastic Load Balance**.
4. Locate the load balancer and click its name.
5. View the metrics of each load balancer and listener.
 - a. Load balancer: Click **Monitoring** tab and select **Load balancer** for **Dimension**.
 - b. Listener (two ways):
 - i. Click **Monitoring** tab, select **Load listener** for **Dimension**, locate the target listener, and view the monitoring metrics.

- ii. Click the name of the target listener, switch to the **Monitoring** tab, and view the monitoring metrics.

Viewing Monitoring Metrics on the Cloud Eye Console

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.
3. Click  in the upper left corner and choose **Management & Governance > Cloud Eye**.
4. In the navigation pane on the left, choose **Cloud Service Monitoring > Elastic Load Balance**.
5. On the **Cloud Service Monitoring** page, click the name of the load balancer. Alternatively, locate the load balancer and click **View Metric** in the **Operation** column.
6. Select the time period during which you want to view metrics. You can select a system-defined time period (for example, last 1 hour) or specify a time period.
7. Click **Select Metric** in the upper right corner and select the metrics to be viewed.

NOTE

For more details, see the [Cloud Eye User Guide](#).

15 Auditing

15.1 Key Operations Recorded by CTS

You can use CTS to record operations on ELB for query, auditing, and backtracking.

[Table 15-1](#) lists the operations recorded by CTS.

Table 15-1 ELB operations recorded by CTS

Action	Resource Type	Trace
Configuring access logs	logtank	createLogtank
Deleting access logs	logtank	deleteLogtank
Creating a certificate	certificate	createCertificate
Modifying a certificate	certificate	updateCertificate
Deleting a certificate	certificate	deleteCertificate
Creating a health check	healthmonitor	createHealthMonitor
Modifying a health check	healthmonitor	updateHealthMonitor
Deleting a health check	healthmonitor	deleteHealthMonitor
Adding a forwarding policy	l7policy	createL7policy
Modifying a forwarding policy	l7policy	updateL7policy
Deleting a forwarding policy	l7policy	deleteL7policy
Adding a forwarding rule	l7rule	createL7rule


Action	Resource Type	Trace
Modifying a forwarding rule	l7rule	updateL7rule
Deleting a forwarding rule	l7rule	deleteL7rule
Adding a listener	listener	createListener
Modifying a listener	listener	updateListener
Deleting a listener	listener	deleteListener
Creating a load balancer	loadbalancer	createLoadbalancer
Modifying a load balancer	loadbalancer	updateLoadbalancer
Deleting a load balancer	loadbalancer	deleteLoadbalancer
Adding a backend server	member	createMember
Modifying a backend server	member	updateMember
Removing a backend server	member	batchUpdateMember
Creating a backend server group	pool	createPool
Modifying a backend server group	pool	updatePool
Deleting a backend server group	pool	deletePool

15.2 Viewing Traces

Scenarios

CTS records the operations performed on ELB and allows you to view the operation records of the last seven days on the CTS console. To query these records, perform the following operations.

Procedure

1. Log in to the management console.
2. In the upper left corner of the page, click  and select the desired region and project.


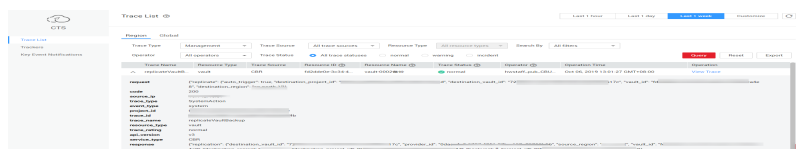
- Under **Management & Governance**, click **Cloud Trace Service**.
- In the navigation pane on the left, choose **Trace List**.
- Specify the filters used for querying traces. The following filters are available:
 - **Trace Type, Trace Source, Resource Type, and Search By**
Select a filter from the drop-down list.
If you select **Trace name** for **Search By**, you need to select a specific trace name.
If you select **Resource ID** for **Search By**, select or enter a specific resource ID.
If you select **Resource name** for **Search By**, select or enter a specific resource name.
 - **Operator**: Select a specific operator (at the user level rather than the tenant level).
 - **Trace Status**: Available options include **All trace statuses, Normal, Warning, and Incident**. You can only select one of them.
 - **Time range**: You can query traces generated at any time range of the last seven days.
- Click  on the left of the required trace to expand its details.

Figure 15-1 Expanding trace details

- Click **View Trace** in the **Operation** column to view trace details.

Figure 15-2 View Trace

```
"context": {
  "code": "204",
  "source_ip": "10.45.152.59",
  "trace_type": "ApiCall",
  "event_type": "system",
  "project_id": "0503dda89700fed2f78c00909158a4d",
  "trace_id": "116a2aff-deb8-11e9-95f5-d5c0b02a9b97",
  "trace_name": "deleteMember",
  "resource_type": "member",
  "trace_rating": "normal",
  "api_version": "v2.0",
  "service_type": "ELB",
  "response": "{\"member\": {\"project_id\": \"0503dda89700fed2f78c00909158a4d\", \"name\": \"9646e73b-338c-4d27-a17c-219be532812c\"}}, \"resource_id\": \"9646e73b-338c-4d27-a17c-219be532812c\"}",
  "tracker_name": "system",
  "time": "1569321775225",
  "resource_name": "9646e73b-338c-4d27-a17c-219be532812c",
  "record_time": "1569321775903",
  "user": {
    "domain": {
      "name": "0503dda89700fed2f78c00909158a4d",
      "id": "0503dda87800fed0f75c0096d70a960"
    }
  }
},
```

For details about key fields in the trace, see the [Cloud Trace Service User Guide](#).

Example Traces

- Creating a load balancer

```
request {"loadbalancer":{"name":"elb-test-
zcy","description":"","tenant_id":"05041fffa40025702f6dc009cc6f8f33","vip_subnet_id":"ed04fd93-
e74b-4794-b63e-e72baa02a2da","admin_state_up":true}}
code 201
source_ip 124.71.93.36
trace_type ConsoleAction
event_type system
project_id 05041fffa40025702f6dc009cc6f8f33
trace_id b39b21a1-8d49-11ec-b548-2be046112888
trace_name createLoadbalancer
resource_type loadbalancer
trace_rating normal
api_version v2.0
service_type ELB
response {"loadbalancer":{"description":"","provisioning_status":"ACTIVE","provider":"vlb",
"project_id":"05041fffa40025702f6dc009cc6f8f33","vip_address":"172.18.0.205","pools":[],
"operating_status":"ONLINE","name":"elb-test-zcy","created_at":"2022-02-14T03:53:39",
"listeners":[],"id":"7ebe23cd-1d46-4a49-b707-1441c7f0d0d1","vip_port_id":
"5b36ff96-3773-4736-83cf-38c54abedeea","updated_at":"2022-02-14T03:53:41","tags":[],
"admin_state_up":true,"vip_subnet_id":"ed04fd93-e74b-4794-b63e-e72baa02a2da","tenant_id":
"05041fffa40025702f6dc009cc6f8f33"}}
resource_id 7ebe23cd-1d46-4a49-b707-1441c7f0d0d1
tracker_name system
time 2022/02/14 11:53:42 GMT+08:00
resource_name elb-test-zcy
record_time 2022/02/14 11:53:42 GMT+08:00
request_id
user {"domain":{"name":"CBUInfo","id":"0503dda87802345ddafed096d70a960"},"name":"zcy",
"id":"09f106afd2345cdeff5c009c58f5b4a"}
```

- **Deleting a load balancer**

```
request
code 204
source_ip 124.71.93.36
trace_type ConsoleAction
event_type system
project_id 05041fffa40025702f6dc009cc6f8f33
trace_id 4f838bbf-8d4a-11ec-a1fe-1f93fdaf3bec
trace_name deleteLoadbalancer
resource_type loadbalancer
trace_rating normal
api_version v2.0
service_type ELB
response {"loadbalancer":{"listeners":[],"vip_port_id":"5b36ff96-3773-4736-83cf-38c54abedeea",
"tags":[],"tenant_id":"05041fffa40025702f6dc009cc6f8f33","admin_state_up":true,"id":
"7ebe23cd-1d46-4a49-b707-1441c7f0d0d1","operating_status":"ONLINE","description":"","pools":
[],"vip_subnet_id":"ed04fd93-e74b-4794-b63e-e72baa02a2da","project_id":
"05041fffa40025702f6dc009cc6f8f33","provisioning_status":"ACTIVE","name":"elb-test-zcy",
"created_at":"2022-02-14T03:53:39","vip_address":"172.18.0.205","updated_at":
"2022-02-14T03:53:41","provider":"vlb"}}
resource_id 7ebe23cd-1d46-4a49-b707-1441c7f0d0d1
tracker_name system
time 2022/02/14 11:58:03 GMT+08:00
resource_name elb-test-zcy
record_time 2022/02/14 11:58:03 GMT+08:00
request_id
user {"domain":{"name":"CBUInfo","id":"0503dda87802345ddafed096d70a960"},"name":"zcy","id":
"09f106afd2345cdeff5c009c58f5b4a"}
```

16 Permissions Management

16.1 Creating a User and Granting Permissions

Use [IAM](#) to implement fine-grained permissions control over your ELB resources. With IAM, you can:

- Create IAM users for employees based on your enterprise's organizational structure. Each IAM user will have their own security credentials for accessing ELB resources.
- Grant only the permissions required for users to perform a specific task.
- Entrust another Huawei Cloud account or cloud service to perform efficient O&M on your ELB resources.

Skip this section if your Huawei Cloud account does not need individual IAM users.

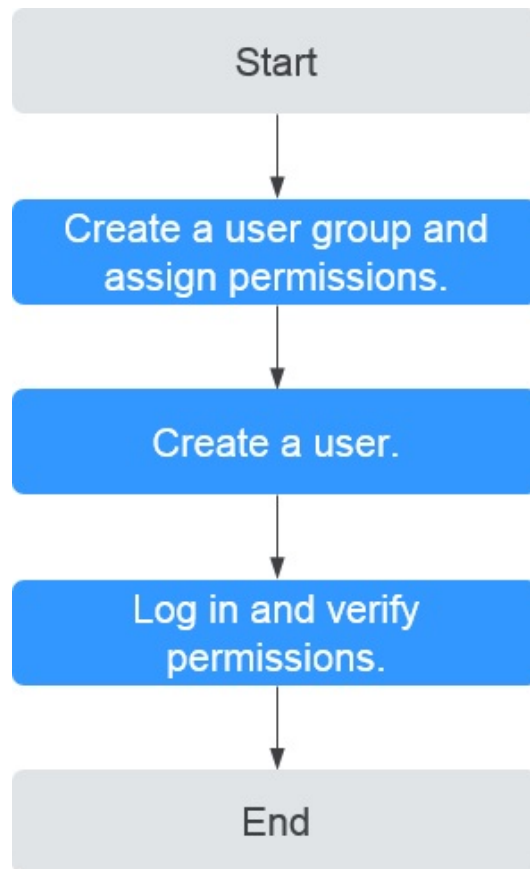
This following describes the procedure for granting permissions.

Prerequisites

You have learned about ELB policies and can select the appropriate policies based on service requirements. Learn about [permissions](#) supported by ELB. For the permissions of other services, see [System Permissions](#).

Process Flow

Figure 16-1 Process for granting ELB permissions



1. **Create a user group and assign permissions.**
Create a user group on the IAM console and assign the **ELB ReadOnlyAccess** policy to the group.
2. **Create a user and add it to a user group.**
Create a user on the IAM console and add the user to the group created in 1.
3. **Log in** and verify permissions.
Log in to the ELB console by using the created user, and verify that the user only has read permissions for ELB.
 - Choose **Service List** > **Elastic Load Balance**. Then click **Buy Elastic Load Balancer** on the ELB console. If you cannot create a load balancer, the **ELB ReadOnlyAccess** policy has taken effect.
 - Choose any other service in **Service List**. If a message appears indicating that you have insufficient permissions to access the service, the **ELB ReadOnlyAccess** policy has already taken effect.

16.2 Creating a Custom Policy

Custom policies can be created as a supplement to the system policies of ELB. For the actions supported for custom policies, see "Permissions Policies and Supported Actions" in the [Elastic Load Balance API Reference](#).

You can create custom policies in either of the following ways:

- Visual editor: Select cloud services, actions, resources, and request conditions. This does not require knowledge of policy syntax.
- JSON: Edit JSON policies from scratch or based on an existing policy.

For details, see [Creating a Custom Policy](#). The following section contains examples of common ELB custom policies.

Example Custom Policies

- Example 1: Allowing users to update a load balancer

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "elb:loadbalancers:put"
      ]
    }
  ]
}
```

- Example 2: Denying load balancer deletion

A deny policy must be used in conjunction with other policies to take effect. If the permissions assigned to a user contain both Allow and Deny actions, the Deny actions take precedence over the Allow actions.

If you grant the system policy **ELB FullAccess** to a user but do not want the user to have the permission to delete load balancers defined in the policy, you can create a custom policy that rejects the deletion of load balancers and grant the **ELB FullAccess** and deny policies to the user, so that the user can perform all operations on ELB except deleting load balancers. The following is an example deny policy:

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Deny",
      "Action": [
        "elb:loadbalancers:delete"
      ]
    }
  ]
}
```

- Example 3: Defining permissions for multiple services in a policy

A custom policy can contain the actions of multiple services that are of the global or project-level type. The following is an example policy containing actions of multiple services:

```
{
  "Version": "1.1",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "elb:loadbalancers:get",
        "elb:loadbalancers:list",
        "elb:loadbalancers:delete",
        "ecs:cloudServers:delete"
      ]
    }
  ]
}
```

```
}  
]  
}
```

17 Quotas

What Is Quota?

Quotas can limit the number or amount of resources available to users, such as the maximum number of ECSs or EVS disks that can be created.

If the existing resource quota cannot meet your service requirements, you can apply for a higher quota.

How Do I View My Quotas?


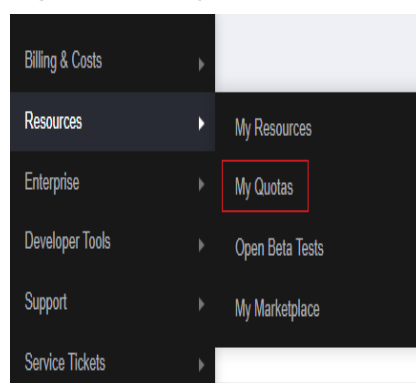
1. Log in to the management console.
2. Click  in the upper left corner and select the desired region and project.
3. In the upper right corner of the page, choose **Resources > My Quotas**.
The **Service Quota** page is displayed.

Figure 17-1 My Quotas



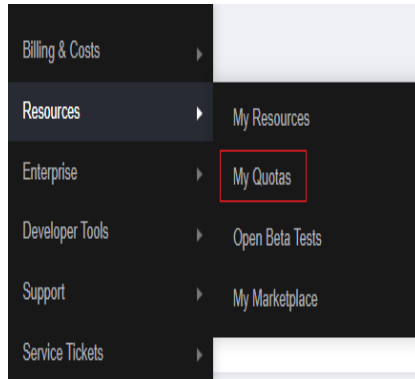
4. View the used and total quota of each type of resources on the displayed page.
If a quota cannot meet service requirements, apply for a higher quota.

How Do I Apply for a Higher Quota?

1. Log in to the management console.

- In the upper right corner of the page, choose **Resources > My Quotas**. The **Service Quota** page is displayed.

Figure 17-2 My Quotas



- Click **Increase Quota** in the upper right corner of the page.

Figure 17-3 Increasing quota

The screenshot shows the 'Service Quota' page with a table listing various services and their quotas. A red 'Increase Quota' button is visible in the top right corner. The table has four columns: Service, Resource Type, Used Quota, and Total Quota.

Service	Resource Type	Used Quota	Total Quota
Auto Scaling	AS group	0	
	AS configuration	0	
Image Management Service	Image	0	
Cloud Container Engine	Cluster	0	
FunctionGraph	Function	0	
	Code storage(MB)	0	
Elastic Volume Service	Disk	3	
	Disk capacity(OB)	120	
	Snapshots	4	
Storage Disaster Recovery Service	Protection group	0	
	Replication pair	0	
Cloud Server Backup Service	Backup Capacity(OB)	0	
	Backup	0	
Scalable File Service	File system	0	
	File system capacity(OB)	0	
CDN	Domain name	0	
	File URL refreshing	0	
	Directory URL refreshing	0	
	URL prefetching	0	

- On the **Create Service Ticket** page, configure parameters as required. In the **Problem Description** area, fill in the content and reason for adjustment.
- After all necessary parameters are configured, select **I have read and agree to the Ticket Service Protocol and Privacy Statement** and click **Submit**.

18 Appendix

18.1 Configuring the TOA Module

Scenarios

ELB provides customized strategies for managing service access. Before these strategies can be customized, the clients' IP addresses contained in the requests are required. To obtain the IP addresses, you can install the TCP Option Address (TOA) kernel module on backend servers.

This section provides detailed operations for you to compile the module in the OS if you use TCP to distribute incoming traffic.

The operations for Linux OSs with kernel version of 2.6.32 are different from those for Linux OSs with kernel version of 3.0 or later.

NOTE

- TOA does not support listeners using the UDP protocol.
- The module can work properly in the following OSs and the methods for installing other kernel versions are similar:
 - CentOS 6.8 (kernel version 2.6.32)
 - SUSE 11 SP3 (kernel version 3.0.76)
 - CentOS 7 and CentOS 7.2 (kernel version 3.10.0)
 - Ubuntu 16.04.3 (kernel version 4.4.0)
 - Ubuntu 18.04 (kernel version 4.15.0)
 - Ubuntu 20.04 (Kernel version 5.4.0)
 - OpenSUSE 42.2 (kernel version 4.4.36)
 - Debian 8.2.0 (kernel version 3.16.0)

Prerequisites

- The development environment for compiling the module must be the same as that of the current kernel. For example, if the kernel version is kernel-3.10.0-693.11.1.el7, the kernel development package version must be kernel-devel-3.10.0-693.11.1.el7.

- Servers can access OS repositories.
- Users other than **root** must have sudo permissions.

Procedure

- In the following operations, the Linux kernel version is 3.0 or later.
1. Prepare the compilation environment.

NOTE

- During the installation, download the required module development package from the Internet if it cannot be found in the source.
- If the kernel development package (kernel-devel) cannot be obtained, contact the image provider.

The following are operations for compiling the module in different Linux OSs. Perform appropriate operations.

– CentOS

- i. Run the following command to install the GCC:

```
sudo yum install gcc
```

- ii. Run the following command to install the make tool:

```
sudo yum install make
```

- iii. Run the following command to install the module development package (the package header and module library must have the same version as the kernel):

```
sudo yum install kernel-devel-`uname -r`
```

NOTE

- During the installation, download the required module development package from the following address if it cannot be found in the source:
https://mirror.netcologne.de/oracle-linux-repos/ol7_latest/getPackage/
For example, to install 3.10.0-693.11.1.el7.x86_64, run the following command:

```
rpm -ivh kernel-devel-3.10.0-693.11.1.el7.x86_64.rpm
```
- If the kernel development package (kernel-devel) cannot be obtained, contact the image provider.

– Ubuntu and Debian

- i. Run the following command to install the GCC:

```
sudo apt-get install gcc
```

- ii. Run the following command to install the make tool:

```
sudo apt-get install make
```




- iii. Run the following command to install the module development package (the package header and module library must have the same version as the kernel):

```
sudo apt-get install linux-headers-`uname -r`
```

– SUSE

- i. Run the following command to install the GCC:

```
sudo zypper install gcc
```

- ii. Run the following command to install the make tool:
sudo zypper install make
 - iii. Run the following command to install the module development package (the package header and module library must have the same version as the kernel):
sudo zypper install kernel-default-devel
 2. Compile the module.
 - a. Use the git tool and run the following command to download the module source code:
git clone https://github.com/Huawei/TCP_option_address.git
 **NOTE**
If the git tool is not installed, download the module source code from the following link:
https://github.com/Huawei/TCP_option_address
 - b. Run the following commands to enter the source code directory and compile the module:
cd src
make
If no warning or error code is prompted, the compilation was successful. Verify that the **toa.ko** file was generated in the current directory.
 **NOTE**
If error message "config_retpoline=y but not supported by the compiler, Compiler update recommended" is displayed, the GCC version is too old. Upgrade the GCC to a later version.
 3. Load the module.
 - a. Run the following command to load the module:
sudo insmod toa.ko
 - b. Run the following command to check the module loading and to view the kernel output information:
dmesg | grep TOA
If **TOA: toa loaded** is displayed in the command output, the module has been loaded.
 **NOTE**
After compiling the CoreOS module in the container, copy it to the host system and then load it. The container for compiling the module shares the **/lib/modules** directory with the host system, so you can copy the module in the container to this directory, allowing the host system to use it.
 4. Set the script to enable it to automatically load the module.
To make the module take effect when the system starts, add the command for loading the module to your startup script.
You can use either of the following methods to automatically load the module:
 - Add the command for loading the module to a customized startup script as required.

- Perform the following operations to configure a startup script:
 - i. Create the **toa.modules** file in the **/etc/sysconfig/modules/** directory. This file contains the module loading script.
The following is an example of the content in the **toa.modules** file.

```
#!/bin/sh
/sbin/modinfo -F filename /root/toa/toa.ko > /dev/null 2>&1
if [ $? -eq 0 ]; then
/sbin/insmod /root/toa/toa.ko
fi
```

/root/toa/toa.ko is the path of the module file. You need to replace it with their actual path.
 - ii. Run the following command to add execution permissions for the **toa.modules** startup script:

```
sudo chmod +x /etc/sysconfig/modules/toa.modules
```

NOTE

If the kernel is upgraded, the current module will no longer match. Compile the module again.

5. Install the module on multiple servers.

To load the module in the same OS, copy the **toa.ko** file to servers where the module is to be loaded and then perform the operations in [3](#).

After the module is successfully loaded, applications can obtain the real IP address contained in the request.

NOTE

The OS of the server must have the same version as the kernel.

6. Verify the module.

After the module is successfully installed, the source address can be directly obtained. The following provides an example for verification.

Run the following command to start a simple HTTP service on the backend server where Python is installed:

```
python -m SimpleHTTPServer port
```

The value of **port** must be the same as the port configured for the backend server, and the default value is **80**.

Access the IP address of the load balancer from a client. Access logs on the server are as follows:

```
192.168.0.90 - - [06/Aug/2020 14:24:21] "GET / HTTP/1.1" 200 -
```

NOTE

192.168.0.90 indicates the client's source IP address that is obtained by the backend server.

- In the following operations, the Linux kernel version is **2.6.32**.

NOTE

The TOA plug-in supports the OSs (CentOS 6.8 image) with a kernel of 2.6.32-xx. Perform the following steps to configure the module:

1. Obtain the kernel source code package **Linux-2.6.32-220.23.1.el6.x86_64.rs.src.tar.gz** containing the module from the following link:
http://kb.linuxvirtualserver.org/images/3/34/Linux-2.6.32-220.23.1.el6.x86_64.rs.src.tar.gz
2. Decompress the kernel source code package.
3. Modify compilation parameters.
 - a. Open the **linux-2.6.32-220.23.1.el6.x86_64.rs** folder.
 - b. Edit the **net/toa/toa.h** file.
Change the value of **#define TCPOPT_TOA200** to **#define TCPOPT_TOA254**.
 - c. On the shell page, run the following commands:
sed -i 's/CONFIG_IPV6=m/CONFIG_IPV6=y/g' .config
echo -e '\n# toa\nCONFIG_TOA=m' >> .config
After the configuration, the IPv6 module is compiled into the kernel. TOA is compiled into a separate module and can be independently started and stopped.
 - d. Edit **Makefile**.
You can add a description to the end of **EXTRAVERSION =**. This description will be displayed in **uname -r**, for example, **-toa**.

4. Run the following command to compile the software package:

```
make -j n
```

 **NOTE**

n indicates the number of vCPUs. For example, if there are four vCPUs, *n* must be set to 4.

5. Run the following command to install the module:

```
make modules_install
```

The following information is displayed.

Figure 18-1 Installing the module

```
INSTALL /lib/firmware/kaweth/trigger_code_fix.bin
INSTALL /lib/firmware/ti_3410.fw
INSTALL /lib/firmware/ti_5052.fw
INSTALL /lib/firmware/mts_cdma.fw
INSTALL /lib/firmware/mts_gsm.fw
INSTALL /lib/firmware/mts_edge.fw
INSTALL /lib/firmware/edgeport/boot.fw
INSTALL /lib/firmware/edgeport/boot2.fw
INSTALL /lib/firmware/edgeport/down.fw
INSTALL /lib/firmware/edgeport/down2.fw
INSTALL /lib/firmware/edgeport/down3.bin
INSTALL /lib/firmware/whiteheat_loader.fw
INSTALL /lib/firmware/whiteheat.fw
INSTALL /lib/firmware/keyspan_pda/keyspan_pda.fw
INSTALL /lib/firmware/keyspan_pda/xircom_pgs.fw
DEPMOD 2.6.32-toa
```

6. Run the following command to install the kernel:

```
make install
```

The following information is displayed.

Figure 18-2 Installing the kernel

```
INSTALL /lib/firmware/keyspan_pda/xircom_pgs.fw
DEPMOD 2.6.32-toa
[root@SZX1000167219 linux-2.6.32-220.23.1.el6.x86_64.rs]# make install
sh /root/humin/linux-2.6.32-220.23.1.el6.x86_64.rs/arch/x86/boot/install.sh 2.6.32-toa arch/x86/boot/bzImage \
    System.map "/boot"
ERROR: modinfo: could not find module xen_procfs
ERROR: modinfo: could not find module ipv6
ERROR: modinfo: could not find module xen_scifront
ERROR: modinfo: could not find module xen_hcall
ERROR: modinfo: could not find module xen_balloon
[root@SZX1000167219 linux-2.6.32-220.23.1.el6.x86_64.rs]#
```

7. Open the **/boot/grub/grub.conf** file and configure the kernel to start up when the system starts.
 - a. Change the default startup kernel from the first kernel to the zeroth kernel by changing **default=1** to **default=0**.
 - b. Add the **nohz=off** parameter to the end of the line containing the **vmlinuz-2.6.32-toa** kernel. If **nohz** is not disabled, the CPU0 utilization may be high and overload the kernel.

Figure 18-3 Configuration file

```
default=1
timeout=5
splashimage=(hd0,1)/boot/grub/splash.xpm.gz
hiddenmenu
title Red Hat Enterprise Linux Server (2.6.32-toa)
    root (hd0,1)
    kernel /boot/vmlinuz-2.6.32-toa ro root=UUID:
et nohz=off
    initrd /boot/initramfs-2.6.32-toa.img
```

- c. Save the modification and exit. Restart the OS.
During the restart, the system will load the **vmlinuz-2.6.32-toa** kernel.
8. After the restart, run the following command to load the module:
modprobe toa

Add the **modprobe toa** command to both the startup script and the system scheduled monitoring script.

Figure 18-4 Adding the **modprobe toa** command

```
[root@SZX1000167219 ~]# modprobe toa
[root@SZX1000167219 ~]# lsmod |grep toa
toa                4203  0
[root@SZX1000167219 ~]#
```

After the module is loaded, query the kernel information.

Figure 18-5 Querying the kernel

```
[root@SZX1000167219 ~]# uname -a
Linux SZX1000167219 2.6.32-toa #1 SMP Sat Oct 15 11:50:05 CST 2016 x86_64 x86_64 x86_64 GNU/Linux
```

9. Verify the module.
After the module is successfully installed, the source address can be directly obtained. The following provides an example for verification.

Run the following command to start a simple HTTP service on the backend server where Python is installed:

```
python -m SimpleHTTPServer port
```

The value of **port** must be the same as the port configured for the backend server, and the default value is **80**.

Access the IP address of the load balancer from a client. Access logs on the server are as follows:

```
192.168.0.90 - - [06/Aug/2020 14:24:21] "GET / HTTP/1.1" 200 -
```

 **NOTE**

192.168.0.90 indicates the client's source IP address that is obtained by the backend server.

19 Change History

Released On	Description
2024-03-15	This issue is the twenty-ninth official release. Added the descriptions about tag policies in Creating a Dedicated Load Balancer and Tag .
2024-01-05	This issue is the twenty-eighth official release. Added Self-service Troubleshooting .
2023-09-07	This issue is the twenty-seventh official release. Updated the following sections: <ul style="list-style-type: none">• Creating a Dedicated Load Balancer• Creating a Shared Load Balancer• Managing IP Address Groups
2023-07-14	This issue is the twenty-sixth official release. <ul style="list-style-type: none">• Updated Transfer Client IP Address (Dedicated Load Balancers).• Added Transfer Client IP Address (Shared Load Balancers).
2023-05-18	This issue is the twenty-fifth official release. Added the following sections: <ul style="list-style-type: none">• Backend Server Group• Backend Server (Dedicated Load Balancers)• Backend Server (Shared Load Balancers)
2023-02-03	This is the twenty-fourth official release. Added Enabling Guaranteed Performance for a Shared Load Balancer .

Released On	Description
2022-12-30	This issue is the twenty-third official release. Modified the following sections: Modifying the Bandwidth <i>Changing the Specifications of a Dedicated Load Balancer</i>
2022-09-07	This issue is the twenty-second official release. Modified the following section: Added restrictions on ping verification for load balancers in sections <i>Creating a Dedicated Load Balancer</i> , <i>Creating a Shared Load Balancer</i> , and Access Control .
2022-06-30	This is the twenty-first official release. Added the following section: <ul style="list-style-type: none">• Adding a UDP Listener, UDP listeners do not support fragmentation.• Adding a UDP Listener (with a QUIC Backend Server Group Associated), UDP listeners using QUIC protocol do not support fragmentation.
2022-05-30	This issue is the twentieth official release. Added the following section: Routing Traffic to Backend Servers in the Same VPC as the Load Balancer
2022-03-30	This issue is the nineteenth official release. Added the following section: Using Advanced Forwarding for Application Iteration
2022-03-18	This issue is the eighteenth official release. Deleted the FAQ "What Is the Maximum Size of Files that Can Be Transferred Using HTTP or HTTPS?"
2022-03-07	This issue is the seventeenth official release, which incorporates the following changes: Added Does ELB Support IPv6 Networks?
2022-02-14	This issue is the sixteenth official release, which incorporates the following changes: Updated Viewing Metrics . Added Example Log .

Released On	Description
2022-01-10	This issue is the fifteenth official release. Added the following sections: Routing Traffic Across Cloud Servers and On-Premises Servers Transfer Client IP Address.
2022-01-04	This issue is the fourteenth official release. Added the following sections: <ul style="list-style-type: none">• Forwarding Policy (Dedicated Load Balancers)• Advanced Forwarding (Dedicated Load Balancers)
2021-12-29	This issue is the thirteenth official release. Added the following sections: <ul style="list-style-type: none">• Changing the Billing Mode or Bandwidth Billing Option• Configuring Timeout Durations• HTTP/2• Introduction to Certificates• Querying Listeners by Certificate• Protection for Mission-Critical Operations
2021-12-14	This issue is the twelfth official release, which incorporates the following changes: Added What Functions Will Become Unavailable If a Dedicated Load Balancer Is Frozen?
2021-12-09	This issue is the eleventh official release, which incorporates the following changes: Added the diagram of timeout durations at layer 4.
2021-10-28	This issue is the tenth official release, which incorporates the following changes: Added Permissions Management .
2021-10-21	This issue is the ninth official release, which incorporates the following changes: <ul style="list-style-type: none">• Can Both the Listener and Backend Server Group Use HTTPS?• Do Shared Load Balancers Have Specifications?

Released On	Description
2021-09-02	<p>This issue is the eighth official release, which incorporates the following changes:</p> <ul style="list-style-type: none">• Optimized Differences Between Dedicated and Shared Load Balancers.• Added the following sections: Can Backend Servers Access the Ports of a Load Balancer? Can I Bind a Public IP Address Purchased from a Third-Party Cloud Provider to My Load Balancer? Do I Need to Configure Bandwidth for My Load Balancers? Can I Bind Multiple EIPs to a Load Balancer? Why Does a Dedicated Load Balancer Need Multiple IP Addresses? Why Are Requests from the Same IP Address Routed to Different Backend Servers When the Load Balancing Algorithm Is Source IP Hash? Can Backend Servers Access the Internet Using the EIP of the Associated Load Balancer? Why Can't I Select the Target Backend Server Group When Adding or Modifying a Listener? Why Must the Subnet Where the Load Balancer Resides Have at Least 16 Available IP Addresses When I Enable the IP as a Backend Function? Why Is a Forwarding Policy in the Faulty State? How Can I Add a Forwarding Policy to a Listener? What Are Status Codes for Normal Health Checks?
2021-07-16	<p>This issue is the seventh official release, which incorporates the following changes:</p> <p>Changed Management & Deployment to Management & Governance and Computing to Compute based on the latest console product catalog.</p>
2021-06-18	<p>This issue is the sixth official release, which incorporates the following changes:</p> <p>Deleted all descriptions and operations related to classic load balancers.</p>

Released On	Description
2021-02-28	This issue is the fifth official release, which incorporates the following changes: <ul style="list-style-type: none">• Optimized the meaning of Concurrent Connections in section "Monitoring Metrics".• Added section "Configuring Security Group Rules for Backend Servers (Dedicated Load Balancers)".• Added information about dedicated load balancers in FAQ "How Do I Troubleshoot an Unhealthy Backend Server?"
2020-05-30	This issue is the fourth official release, which incorporates the following changes: Changed the name of enhanced load balancers to shared load balancers.
2019-03-30	This issue is the third official release, which incorporates the following changes: Added the content related to enterprise project management.
2018-12-30	This issue is the second official release, which incorporates the following changes: Modified the content and changed some figures in the document based on the latest console.
2018-10-31	This issue is the first official release.