Virtual Private Cloud

Best Practices

Best Practices

VPC (Virtual Private Cloud) is a network type that Alibaba Cloud recommends users to use. VPCs are logically isolated from other VPCs in Alibaba Cloud. More and more users use VPCs because of the following features:

Isolated network environment

Based on tunneling technology, VPCs isolate the data link layer and provide an independent, isolated, and safe network for each user. Different VPCs are completely isolated from each other. Resources within a VPC can communicate with each other over the intranet, but cannot directly communicate with resources in other VPCs unless you have configured a public IP or bound an EIP to the VPC resources.

Controllable network configurations

You have full control over your own VPC network. You can select its IP address range, create subnets, and configure route tables.

For more details, refer to VPC overview.

You can start planning and designing your VPC by answering the following questions:

Q1: How many VPCs are need?

Q2: How many VSwitches are needed?

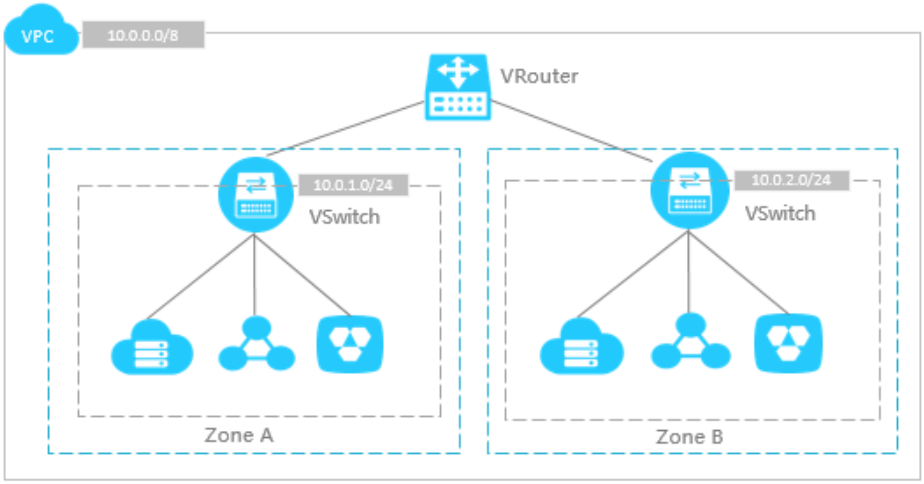
Q3: How to select CIDR blocks?

Q4: How to select CIDR blocks when planning to connect a VPC to on-premises data centers or other VPCs?

Question 1: How many VPCs are needed?

One VPC

If you do not have requirements to deploy your systems in multiple regions and do not need to isolate the systems, we recommend that you create only one VPC. Currently, a VPC can run up to 10,000 cloud product instances, which should be large enough to meet your requirements.

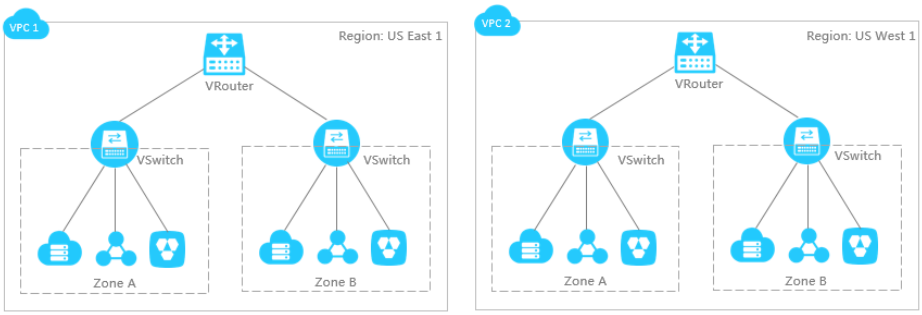


Multiple VPCs

We recommend creating multiple VPCs if you have the following requirements:

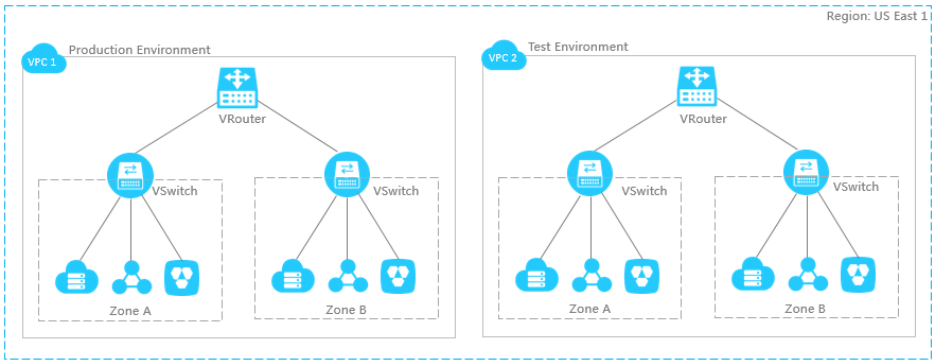
Deploy cloud product resources in different regions

VPCs are a region-specific resource that cannot be deployed across regions. If you need to deploy different systems in different regions, you have to create multiple VPCs. Then, use Express Connect to connect VPCs.



Isolate different systems

If you need to isolate your systems, such as a production environment and test environment, you have to create multiple VPCs. Different VPCs are completely isolated.



Question 2: How many VSwitches are needed?

In general, we recommend creating at least two VSwitches for each VPC, and deploying these two VSwitches in two different zones. This protects your applications from failure of a single location.

Network latency between different zones in the same region is very small. You need to verify the network latency in your real business system. The network latency might be larger than expected due to complicated systems calls or cross-zone calls. We recommend optimizing and adjusting the system to find a balance between high availability and low latency.

> Zones are physical areas with independent power supplies and networks in a region. Zones within the same region are interconnected over the intranet.

Additionally, the number of VSwitches is related to the system size and system design. If the front-end system requires mutual access to the Internet, consider deploying different front-end systems under different VSwitches for better disaster tolerance. Then, deploy the backend system under other VSwitches.

Question 3: How to select CIDR blocks?

When creating VPCs and VSwitches, you have to specify the IP address range for the VPC and the VSwitch in the form of a Classless Inter-Domain Routing (CIDR) block.

VPC CIDR block

Apply the following rules when specifying CIDR blocks for VPCs. Note that the CIDR block cannot be modified after a VPC has been created.

> Use the CIDR blocks provided in the table or their subsets as the IP address range for your VPC.

CIDR block	Number of available IP	Notes

	addresses	
192.168.0.0/16	65,532	Exclude system reserved addresses
172.16.0.0/12	1,048,572	Exclude system reserved addresses
10.0.0.0/8	16,777,212	Exclude system reserved addresses

> If you have special requirements on the IP address range, open a ticket or contact your customer manager.

If you have multiple VPCs, or you want to build a hybrid cloud composed of one or more VPCs and on-premises data centers, we recommend using the subset of these standard CIDR blocks as the IP address range for your VPC and ensure that the netmask is no larger than /16.

If you have only one VPC and it does not need to communicate with an on-premises data center, you are free to use any IP address range as listed in the previous table.

If you plan to connect cloud product instances in a classic network with a VPC, we recommend that you do not use the CIDR block 10.0.0.0/8, which is also used as the CIDR block by the classic network.

VSwitch CIDR block

When specifying CIDR blocks for VSwitches, apply the following rules. Note that the CIDR block cannot be modified after a VSwitch is created.

The allowed block size for a VSwitch is between a /16 netmask and /29 netmask, which can provide 8 to 65,536 IP addresses. The reason for this limitation is that the /16 netmask is larger enough to use with 65,532 IP addresses, while the /29 netmask is too small.

The CIDR block of a VSwitch can be the same as the CIDR block of the VPC that it belongs to, or a subset of the CIDR block of the VPC. For example, if the CIDR block of a VPC is 192.168.0.0/16, then the CIDR block can be 192.168.0.0/16, or any CIDR block from 192.168.0.0/17 to 192.168.0.0/29.

> Note: If the CIDR block of the VSwitch is the same as that of the VPC, you can just create one VSwitch.

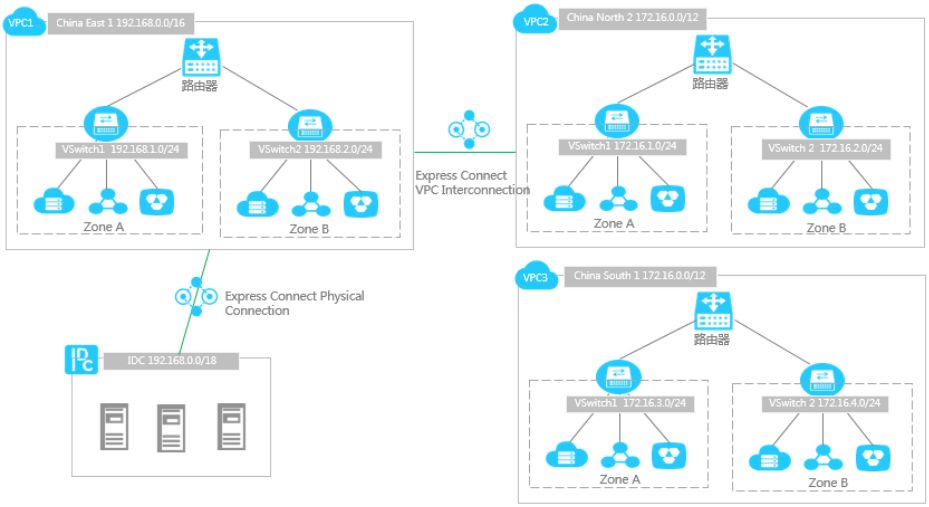
The first and last three IP addresses are reserved by the system. For example, for the CIDR block 192.168.1.0/24, the IP addresses 192.168.1.0, 192.168.1.253, 192.168.1.254, and 192.168.1.255 are reserved by the system.

Alibaba Cloud is planning on providing the ClassicLink feature, which allows ECS instances in the classic network to communicate with other cloud product instances in a VPC. If the CIDR block of the linked VPC is 10.0.0.0/8, the CIDR block of the VSwitch that requires communication to the classic ECS instance must be 10.111.0.0/16.

Additionally, when specifying the CIDR block for the VSwitch, you need to consider the number of cloud product instances running in the VSwitch.

Question 4: How to select CIDR blocks when planning to connect a VPC to on-premises data centers or other VPCs?

As shown in the following figure, assume that you have three VPCs: VPC1 in China East 1, VPC2 in China North 2, and VPC3 in China South 1. VPC1 and VPC2 can communicate with each other over the intranet through Express Connect. VPC3 has no requirements to communicate with other VPCs, but might need to communicate with VPC2 in the future. In addition, you have a self-built data center, which is connected to VPC1 by applying for a leased line.



When you need to connect a VPC with other VPCs or on-premises data centers, ensure that the CIDR block for each VPC is not the same. Therefore, in the previous example, the CIDR blocks of VPC1 and VPC2 are different, while the CIDR block of VPC3 is the same as VPC2. This is because VPC3 does not need to communicate with VPC2. However, the VSwitches in these two VPCs use completely different CIDR blocks to meet future demands for communication with each other.

VPC interconnection requires that the CIDR blocks cannot conflict with each other. Therefore, the CIDR block is the CIDR block of the VSwitch in each VPC. In other words, the CIDR blocks of VPCs can be the same. But for best practices, we recommend using different CIDR blocks for VPCs.

When specifying CIDR blocks for VPCs requiring communication with other VPCs or on-premises data centers, apply the following rules.

Try to use different CIDR blocks for different VPCs. You can use the subsets of the standard CIDR blocks to increase the number of VPCs.

If the CIDR blocks of the VPCs are the same, then use different CIDR blocks for VSwitches in the VPCs.

If neither of the first or second rule is met, ensure that the CIDR blocks of VSwitches requiring communication are different.

In the case where the CIDR blocks for both the on-premises data center and the VPC have been determined and cannot be modified, but they still need to communicate with each other, you can use the leased line gateway to solve this problem. Alibaba Cloud is planning to provide this function.

Currently, Alibaba Cloud does not provide a dedicated access control for VPC. The access control in a VPC depends on the cloud products themselves. This section introduces the access control features of ECS, ApsaraDB for RDS and Server Load Balancer.

ECS: security groups

A security group is a virtual firewall that provides stateful packet inspection (SPI). Security groups are used to set network access control for one or more ECS instances. As an important method of security isolation, security groups are used to divide security domains on the cloud.

Security rules

A security group is a logical group that groups instances in the same region, with the same security requirements, and mutual trust. Each instance belongs to at least one security group, which must be specified at the time of creation. Instances in the same security group can communicate through the network, but instances in different security groups cannot communicate through the intranet by default. However, mutual access can be authorized between two security groups.

A VPC ECS instance can only join the security group in this VPC. You can authorize or cancel security group rules at any time. Your changes to the security group rule will be automatically applied to ECS instances associated with the security group.

You can control the inbound and outbound network traffic through the VPC ECS instances by configuring corresponding inbound and outbound rules:

Inbound: Accept or deny the access of the inbound traffic from a specified IP or CIDR block to a VPC ECS instance.

Outbound: Accept or deny the access of the outbound traffic from a specified IP or CIDR block to a VPC ECS instance.

If the rules conflict, the rule with the higher priority takes effect. When their priorities are the same, the Deny rule takes effect.

Configuration policies

When using a security group as a whitelist, follow the principle of minimum authorization. For example, the instance used as the jump server in O&M generally has high permissions for intranet access. The instance is exposed to the Internet and allows SSH login. Such instances have higher risks, we recommend managing them separately. You can design security group rules of the jump server instance as follows:

Deny all accesses to this instance from all addresses (0.0.0.0/0).

Allow access to this instance from the IP address of the specific O&M staff through SSH (TCP Port 22).

ECS instances within the same security group are interconnected over the intranet. We recommend placing the ECS instances with different business requirements and access controls into different security groups:

Place the instances that offering Internet services and intranet services in the different security groups.

Apply different security groups to different applications.

Apply different security groups to different deployment environment.

Use the following methods to allow intranet access for the ECS instances in a VPC network:

Configure mutual authorization between security groups.

Configure a security group rule to authorize the inbound and outbound access from a specific CIDR block.

When the number of security groups is small or the network plan and design are not strict, the first method is relatively simple. When the number of security groups increases with the complexity growth in business deployment, authorization on the specific IP address range can effectively simplify the security group configuration and cut down management costs with reasonable planning.

For more information about security groups, refer to Security groups.

ApsaraDB for RDS: whitelist

Powered by the whitelist, you can customize the IP addresses allowed to access RDS. Access from unspecified IP addresses will be denied. When you use RDS products in a VPC, you need to add the IP address of the ECS instance that requires the access to the RDS products to the whitelist of the RDS.

For more configuration on ApsaraDB for RDS whitelist, refer to Set a whitelist.

Server Load Balancer: whitelist

Powered by whitelist, you can customize the IP addresses allowed to access the Server Load Balancer listener. It applies to scenarios where the application only allows access from some specific IP addresses. Server Load Balancer is a traffic distribution control service that distributes access traffic to multiple backend ECS servers based on scheduling algorithms and forwarding rules. Server Load Balancer service access is usually open to Internet or intranet users. When the service is only open to specified users, or only allows intranet access, the whitelist feature can effectively control access to the service. To configure a whitelist, you need to add the user’ s IP address requiring access to the Server Load Balancer service in the whitelist of the Server Load Balancer listener.

For more information, refer to Set whitelist access control.

Currently, Alibaba Cloud provides two Internet-facing products, Server Load Balancer (SLB) and Elastic Public IP address (EIP).

These products are applicable to different scenarios. By using these products, the VPC resources can communicate with the Internet and provide external services.

Product	Function	Scenario keywords	Pricing
SLB	Supports port mapping of the	DNAT+Layer-4 load balancing/Layer-7	Pay by bandwidth

	inbound traffic (DNAT). A Server Load Balancer distributes Internet traffic to backend servers based on configured forwarding rules. You can dynamically increase or decrease the number of the backend servers according to business traffic changes. A Server Load Balancer provides the health check function, and automatically blocks unhealthy servers to ensure a high availability of the system. A Server Load Balancer is applicable in a scenario with high availability demands and uses multiple ECS servers to provide business services.	load balancing Health check/high availability	
EIP	An EIP can dynamically bind to a VPC ECS instance. It provides an independent public IP address and bandwidth for each ECS instance to use.	SNAT+DNAT One EIP can bind to only one ECS instance.	Pay by bandwidth

Suggestions:

If you have more than one ECS instance and need to distribute traffic loads based on ports, choose Server Load Balancer.

If you only need a very small number of ECS instances to provide external services, choose EIP. You can directly bind the EIP to the VPC ECS instance without other configurations.

Typical scenarios and solutions

The following shows how to choose an Internet-facing product based on different scenarios. Assume all ECS instances use the VPC network.

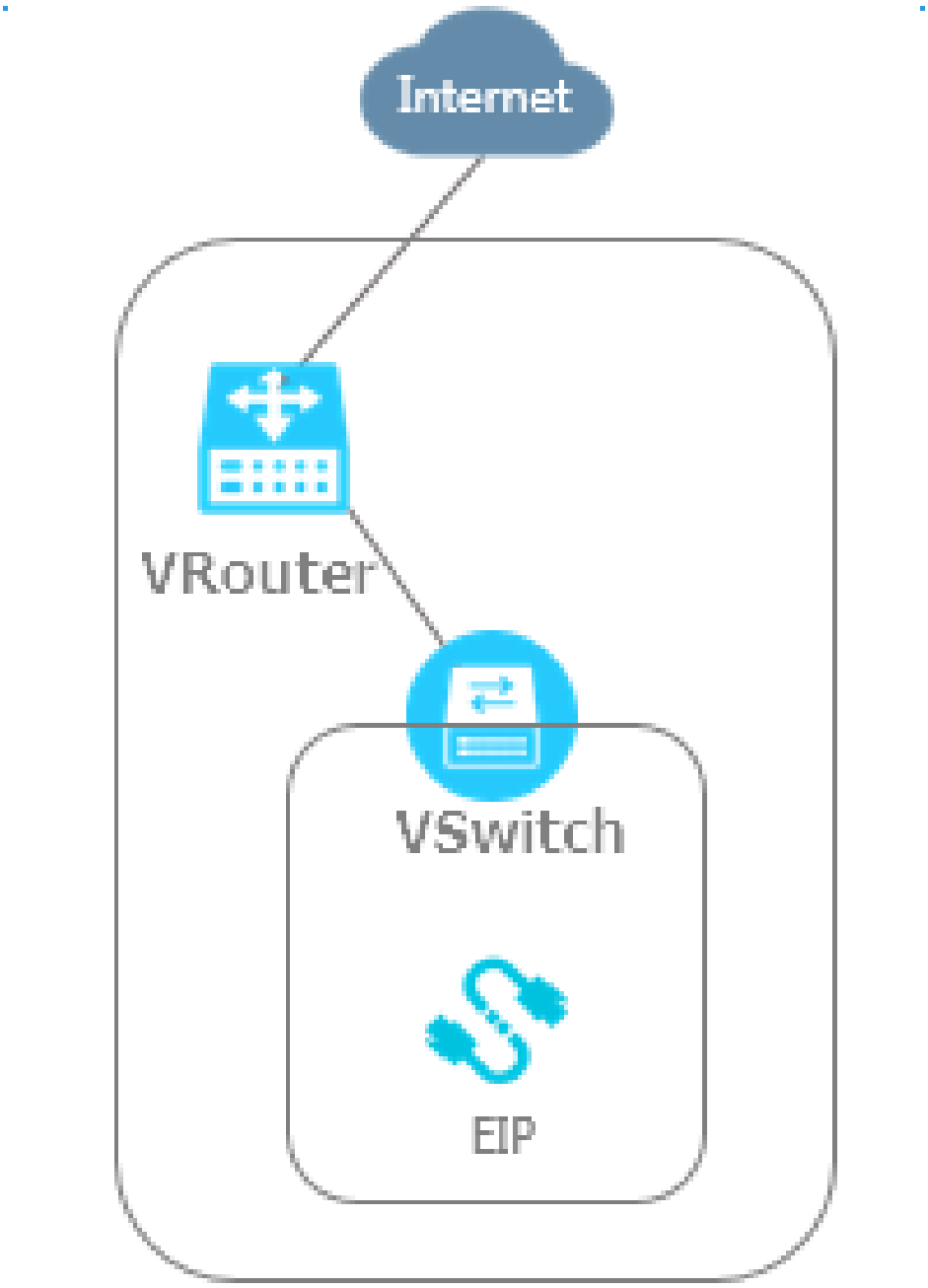
Provide external services

Access the Internet

Provide external services

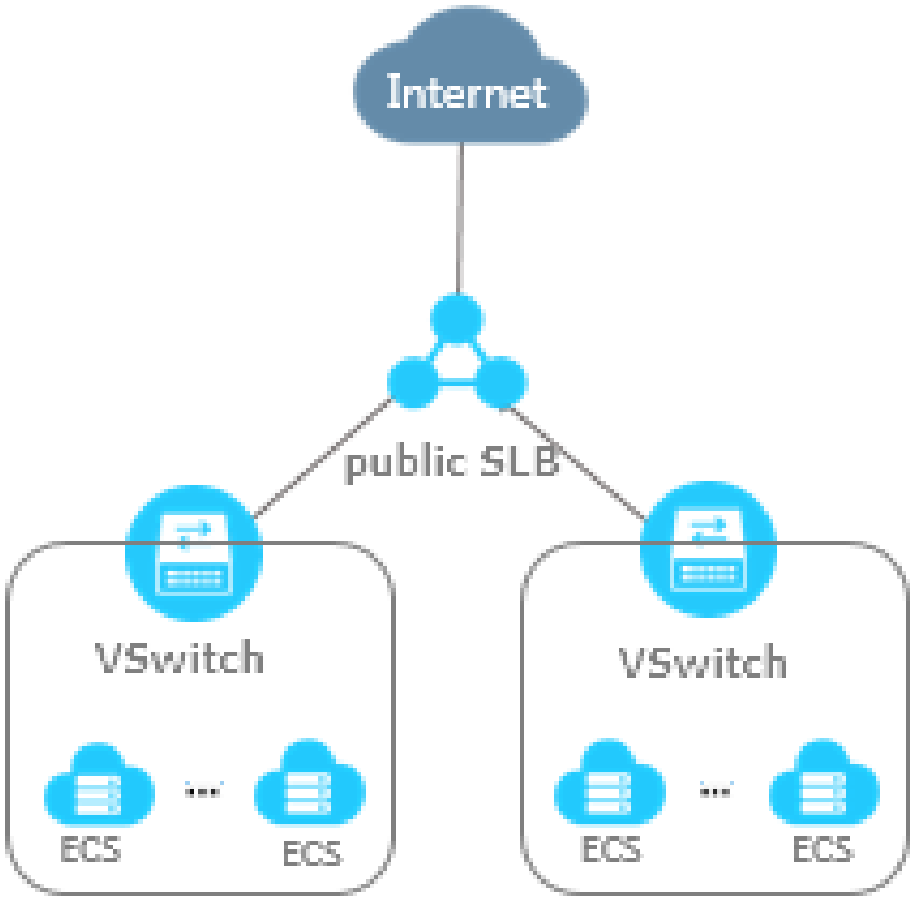
Provide external services with a single ECS instance

When you only have a single application and the access traffic is small, you can deploy the application, the database, and the files on one ECS instance. Then, bind an EIP to the ECS instance to provide external services.



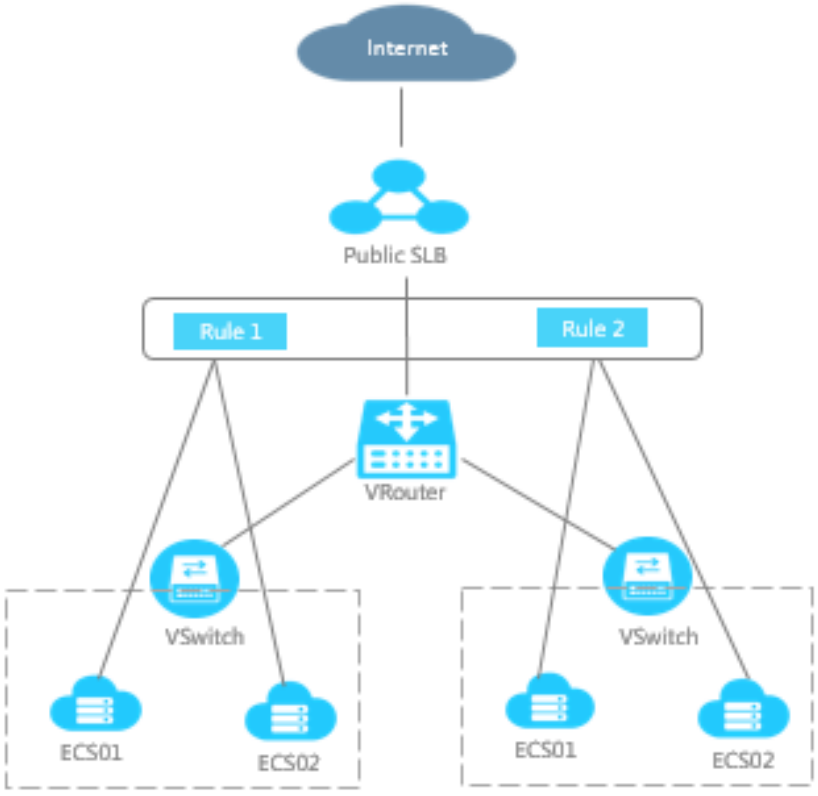
Provide a public Layer-4 load balancing service

When the ECS instance used to provide external services is unable to support the high access traffic, you need to use more ECS instances. In this situation, comparing with the more advanced Layer-7 listener, the simple Layer-4 (TCP) listener can meet your requirements. You can create a Layer-4 (TCP/UDP) listener and add multiple backend ECS instances to build the entire business architecture.



Provide a public Layer-7 load balancing service

In addition to the basic traffic distribution requirement, if you want to distribute the traffic from different business applications to different servers, choose the Layer-7 Server Load Balancer service. You can create a Layer-7 listener with multiple forwarding rules configured.



Access the Internet

Bind EIPs

If an ECS instance requires the Internet access but does not have a public IP, you can bind an EIP to this ECS. You can unbind the EIP whenever you do not need the Internet access.

How to choose to use VPC?

A VPC is an isolated private network. By default, VPCs cannot communicate with each other over the intranet. ECS instances in a VPC cannot access the Internet or be accessed from Internet, and a VPC cannot access a classic network through the intranet. But most of Alibaba Cloud products provide capabilities of both Internet access and intranet access, more than 95% of the cloud products support VPCs.

Note: The network type must be the same for the cloud product instances requiring intranet communication. For example, if a VPC ECS instance needs to access a Server Load Balancer instance and an RDS instance through the intranet, the network type of the Server Load Balancer

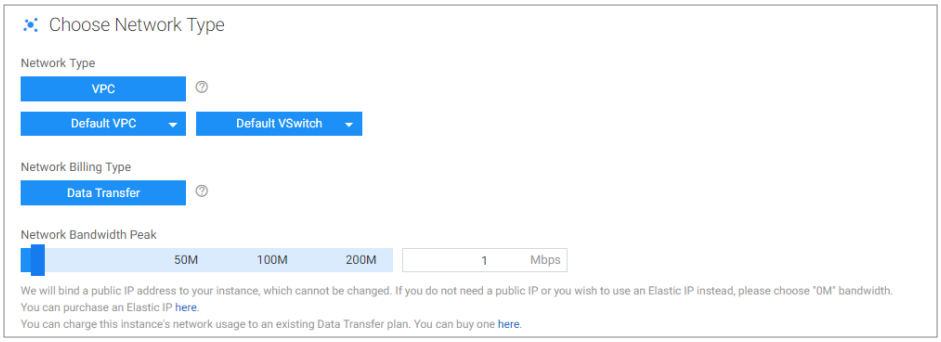
and RDS must also be VPC and belong to a same VPC. Otherwise, they cannot communicate with each other.

The ways to use VPC are different for different cloud products.

Select VPC on the purchase page

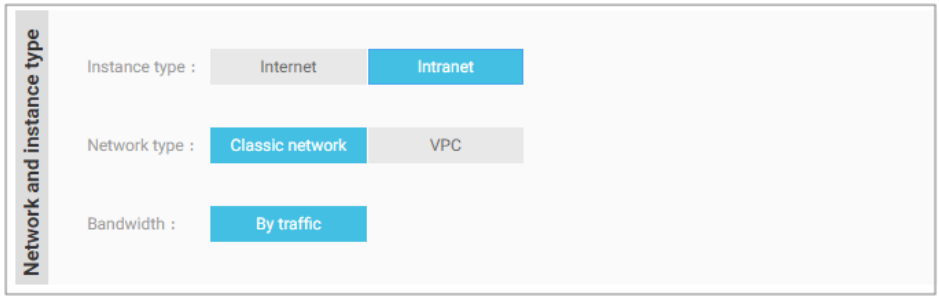
This approach mainly applies to instance-type cloud products, such as ECS, RDS, and Server Load Balancer. The purchase pages of such cloud products offer an option to select the network type.

The following figure shows the network type option on the ECS purchase page.



The following figure shows the network type option on the RDS purchase page.

The following figure shows the network type option on the Server Load Balancer purchase page.



Select the VPC access on the console

This approach applies to the cloud products such as OTS, Container Service, E-MapReduce and NAS.

You can set a VPC access domain for OTS instances on the OTS console. For Container Service and E-MapReduce, you can select a VPC when creating a cluster on the console. The Network Attached Storage product provides the VPC mount point.

Check the VPC access endpoint through documentation

For cloud products, such as Log Service and OSS, refer to the following documents:

- Regions and endpoints
- Service endpoint

How to switch the network type?

Some instance-type cloud products support the switching of network types. For example, you can switch the network type of an RDS instance on the console.

ECS products will also be able to switch from the classic network to VPC.

Server Load Balancer does not support switching from the classic network to VPC. As an alternative, you can purchase a VPC instance.

Note that some cloud products, such as CDN and Situation Awareness, do not need VPC.

Migrate to VPC

Note: In this document, a classic ECS refers to an ECS instance created the classic network, while a VPC ECS refers to an ECS instance created in VPC.

Virtual Private Cloud (VPC) is a private network logically isolated from other virtual networks. Alibaba Cloud VPC allows you to build an isolated network environment with full control over your private network. With these benefits, VPC has become a preferred networking choice for cloud users.

Alibaba Cloud provides two solutions to migrate from classic network to VPC. You can use them separately, or in combination, to meet different migration scenarios.

- Hybrid access and hybrid addition
- VPC ClassicLink

Hybrid access and hybrid addition

The hybrid access and hybrid addition solution is a seamless migration solution. Firstly, you need to create the required cloud product instances (such as ECS) in the VPC, and then use this solution to smoothly migrate your system to VPC. After all the systems have been migrated to VPC, you can release resources in classic network to complete the whole migration process. For details, see Example of hybrid acess and addition.

Hybrid access

Hybrid access is that a database can be accessed by both a classic ECS and a VPC ECS. With hybrid access, you can migrate your ApsaraDB from the classic network to VPC without affecting the continuity of your service. For details, refer to Hybrid access of ApsaraDB.

Hybrid addition

Hybrid addition is that you can add both the classic ECS instances and VPC ECS instances to a Server Load Balancer (SLB) instance as the backend servers for handling the distributed the front-end requests. The VServer group also supports hybrid addition.

> Note: If a VPC SLB instance adds both the classic ECS instances and VPC ECS instances as the backend servers, for a layer-4 listener, you cannot obtain the real IP of the client on the classic ECS instance, but can obtain it on the VPC ECS instance. For the layer-7 listener, no impact on obtaining the real IP of the client.

Note the following when using the hybrid access and hybrid addition solution:

This solution meets most of the system migration requirements except for the system that has communication requirement between the classic ECS and VPC ECS. For such system, you can complete the migration in conjunction with the VPC ClassicLink solution.

This solution can only be used to migrate from classic network to VPC.

VPC ClassicLink

VPC provides you with the ClassicLink function to link a classic ECS to a VPC. Therefore, the classic

ECS can access the VPC resources. For details, see ClassicLink overview.

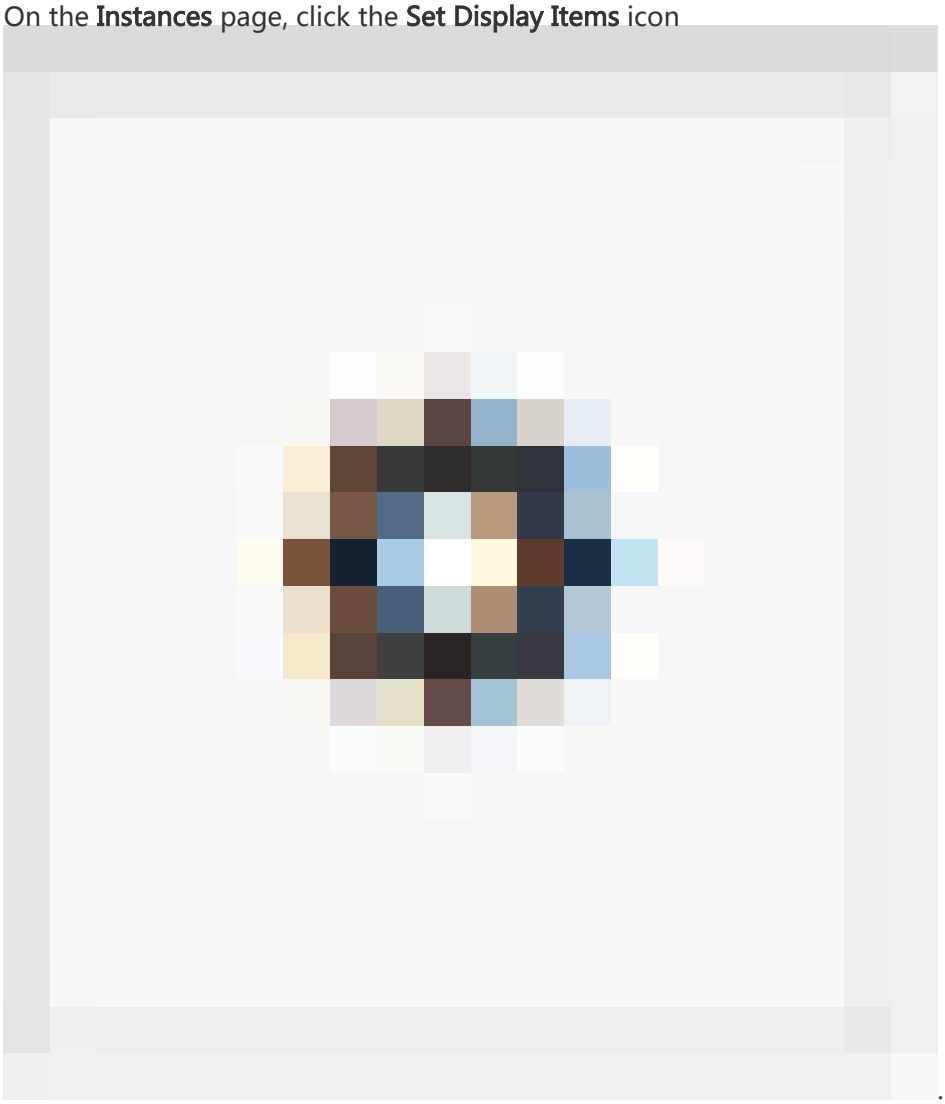
ClassicLink

View link status

Log on to the ECS console.

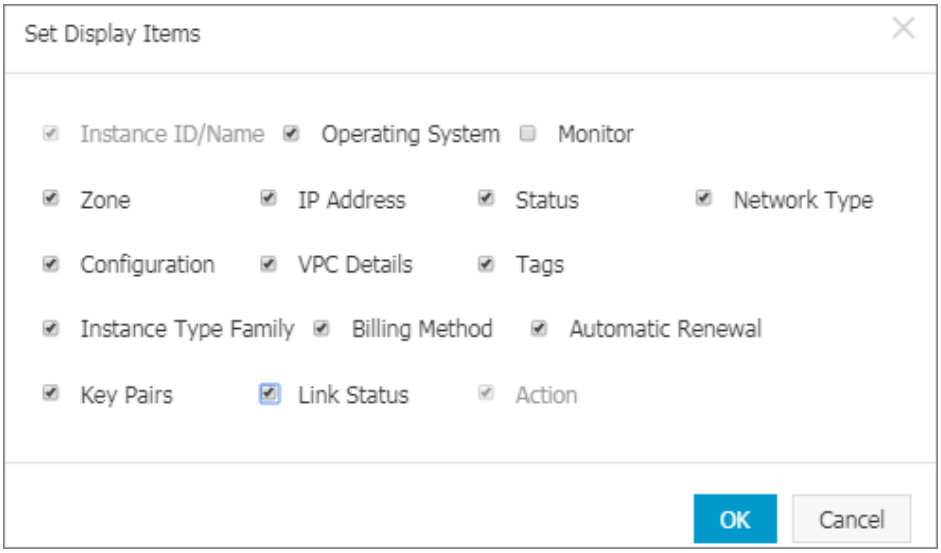
In the left-side navigation pane, click **Instances**.

On the **Instances** page, click the **Set Display Items** icon

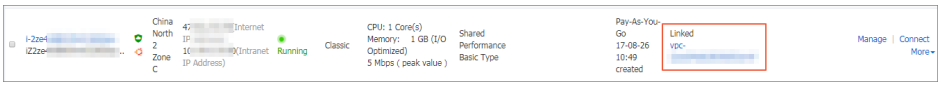


.

Select **Link Status** in the displayed dialog and click **OK**.



View the link status.



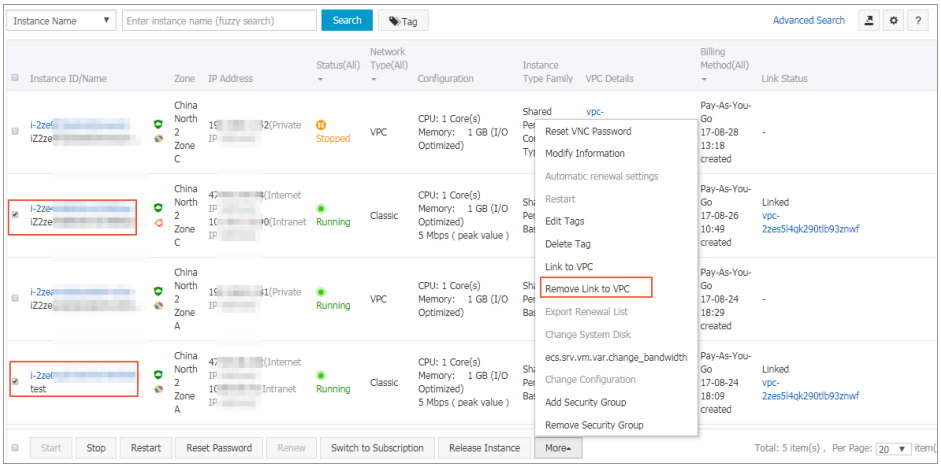
Remove the VPC link

Log on to the ECS console.

In the left-side navigation pane, click **Instances**.

Find the target ECS, click **More** > **Remove Link to VPC**.

You can remove the VPC link for multiple classic ECS instances at one time. Select the target ECS instances, and then click **More** > **Remove Link to VPC** at the lower corner of the instance list.



Click **OK**.

After removing the link, the link status changes to **Not Linked**.

This document provides an example of hybrid access and hybrid addition.

Before you begin

Before migrating to VPC, ensure that:

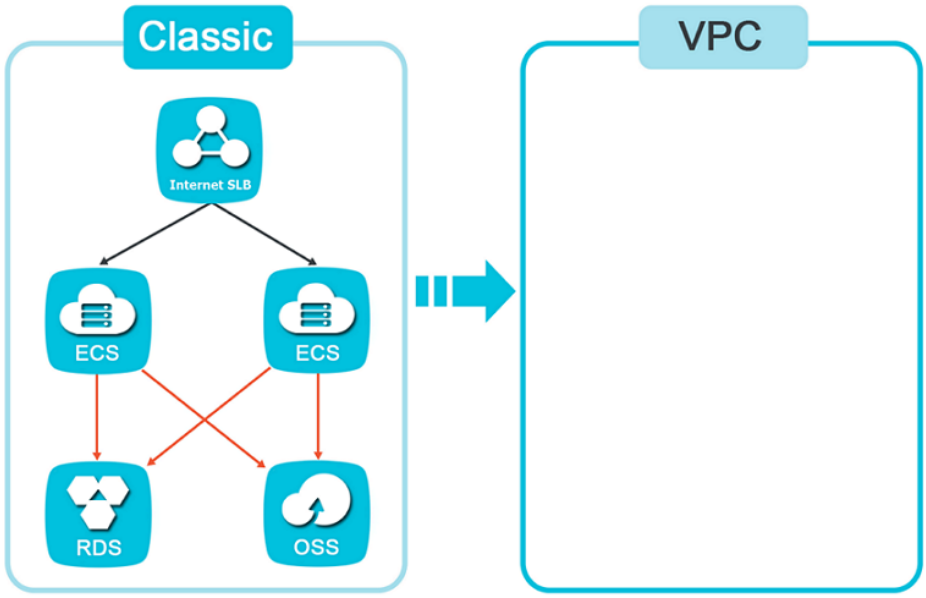
You are familiar with the VPC and the related products. VPC and the classic network are very different. Apart from the network isolation, VPC enables you to control your private network by using other related products together.

The migration system in this example is quite simple. Many systems are more complex than the example. Before the migration, you need to carefully evaluate the migration system and sort out dependencies of the system.

Example systems

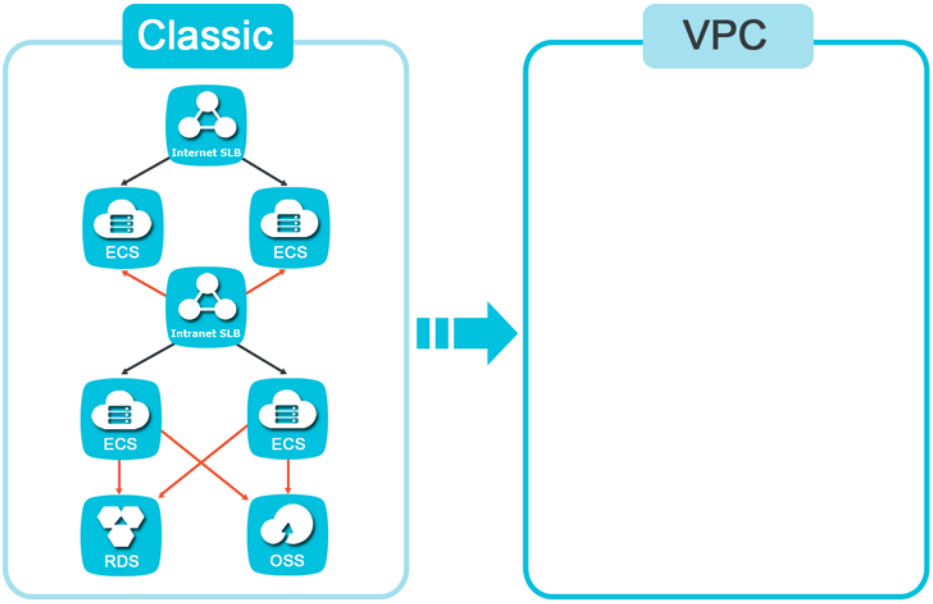
System 1

As shown in the following figure, the system to be migrated consists of SLB, ECS, RDS and OSS. The Internet SLB instances adds two ECS instances as the backend servers and the applications deployed on these two ECS instances have to access RDS and OSS。



System 2

The system 2 is relatively complex. As shown in the following figure, the Internet SLB instance adds two ECS instances (ECS 1 and ECS 2) as the backend servers. And these two ECS instances have to access the service of an Intranet SLB instance. Similarly, the Intranet SLB instance also adds two ECS instances (ECS 3 and ECS 4), on which the applications have to access RDS and OSS.

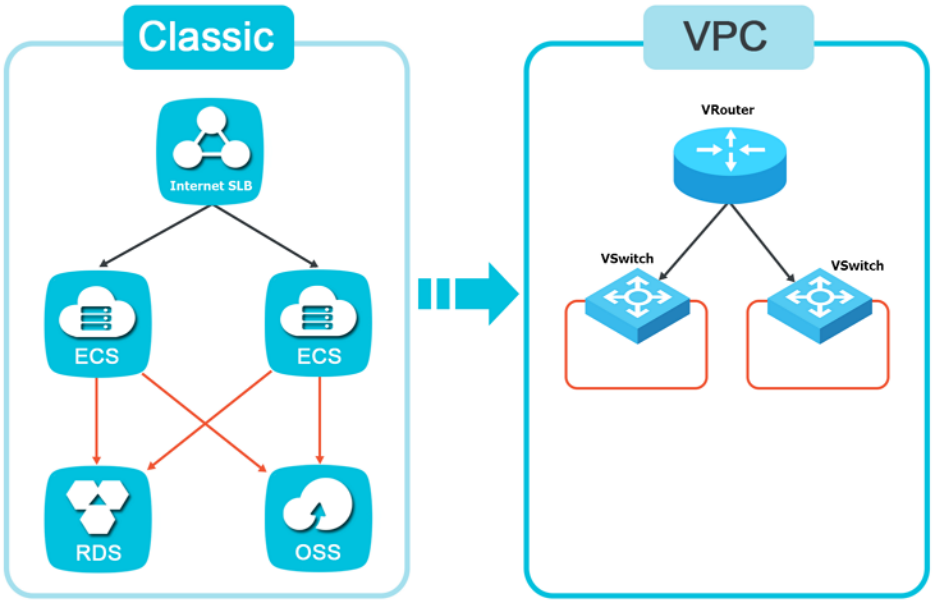


Migrate system 1 to VPC

Prepare the VPC environment.

Firstly, you have to create a VPC and VSwitch to which the system is migrated.

For details, see Build VPC.

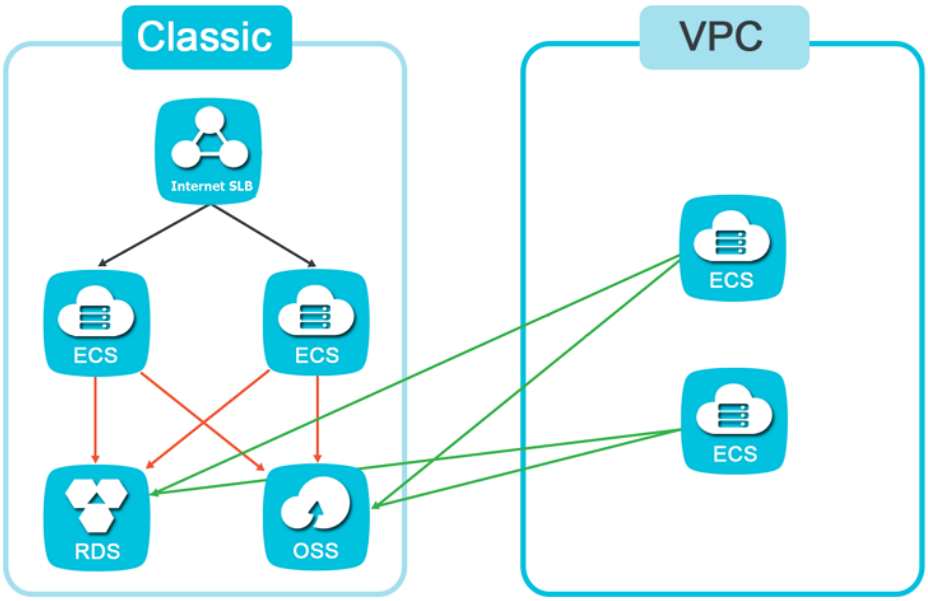


Obtain the VPC access endpoint of OSS and RDS.

> You can migrate the RDS instance to the VPC either by the API or the console and reserve the classic network endpoint at the same time. For details, see Hybrid access of ApsaraDB for RDS.
>
> After the migration, the classic network endpoint will not be changed. Therefore, the service is not interrupted during the migration. Additionally, the classic network endpoint will be released once the reservation time is reached.
>
> OSS itself has two endpoints, no need to migrate. To obtain the VPC endpoint, see Regions and endpoints.

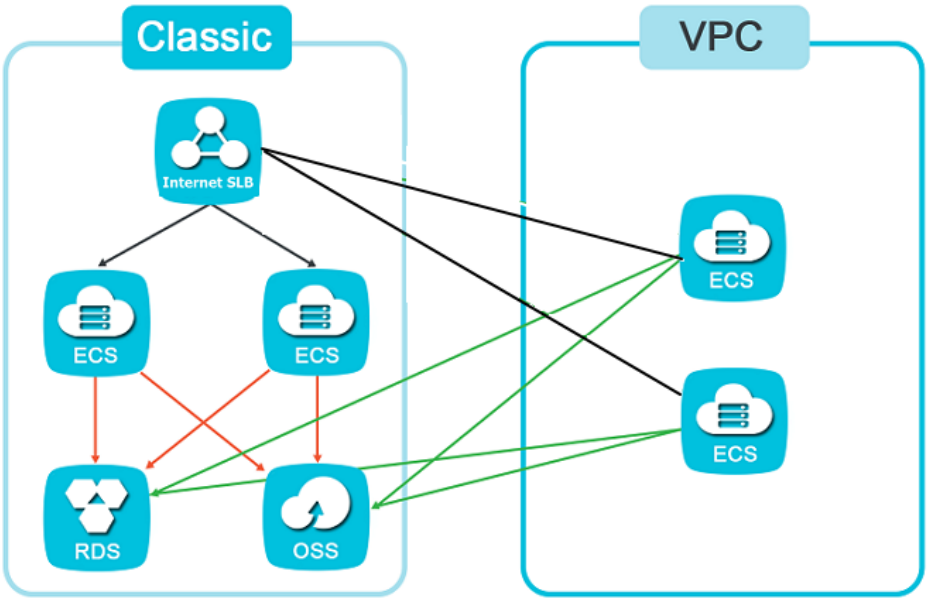
Create two ECS instances in the VPC and configure the ECS.

As shown in the following figure, set the access endpoint for the applications on the ECS to the VPC endpoint of the RDS and OSS. After configuring, test if the application in the VPC ECS can access the RDS and OSS.



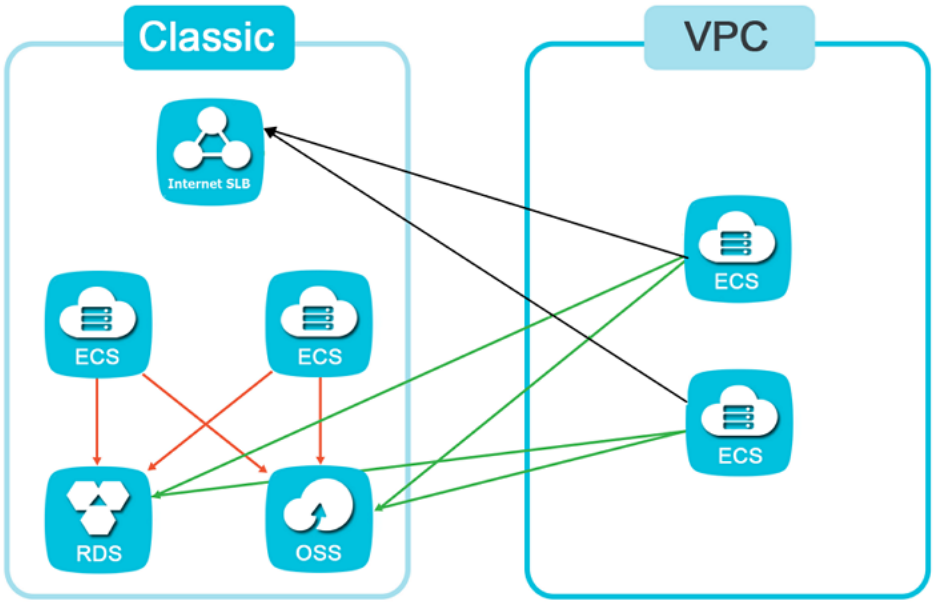
Add the VPC ECS instances to the Internet SLB instance as the backend servers.

Check the health check status of the VPC ECS instances. You can set a smaller weight for the VPC ECS instances to avoid the service interruption due to other exceptions which are not captured by the health check.



Remove the classic ECS instances from the Internet SLB instance.

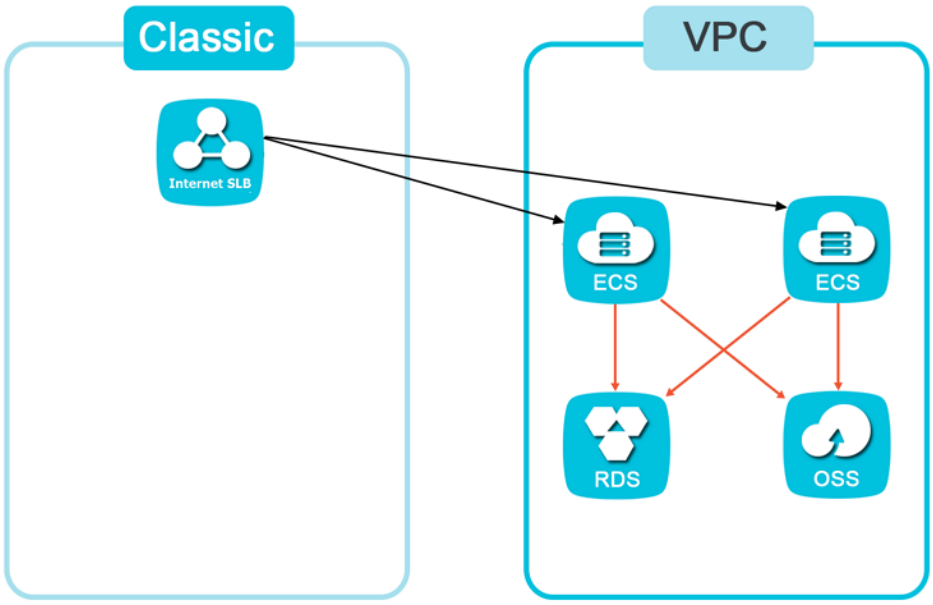
You can set the weight values of the classic ECS instances to 0 and then remove the classic ECS instances when no more traffic is distributed to them.



Release the classic link ECS instances.

When the system has run smoothly for a while, release the classic ECS instances. Because the SLB instance can add VPC ECS instances as the backend server, no need to migrate SLB. The whole migration process is completed for now.

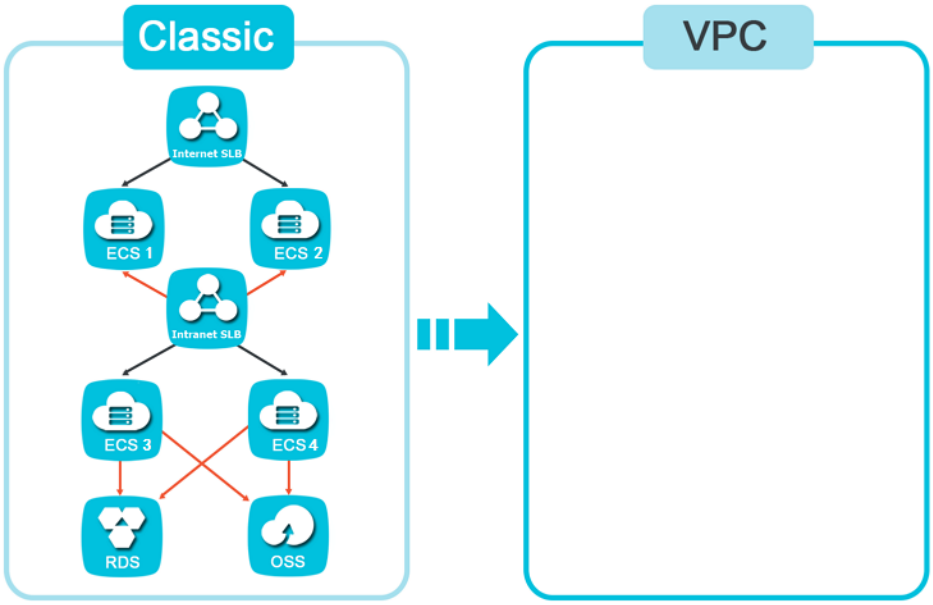
> **Note**: The classic endpoint of the RDS instance will be automatically deleted when the reservation time is reached.



Migrate system 2 to VPC

When migrating a relative complex system as shown in the following figure, if use the same

procedure as migrating the system 1, the VPC ECS instances cannot access the classic intranet SLB instance because the SLB does not support hybrid access.



The steps for migrating the system 2 is as follows:

Create two ECS instances in the VPC to migrate the applications deployed in the ECS 3 and ECS 4 in the classic network.

Configure the newly created VPC ECS instances using the VPC endpoints of RDS and OSS.

Create an intranet SLB instance in the VPC used to replace the intranet SLB in the classic network.

Add the VPC ECS instances created in step 1 to the intranet SLB instance as the backend server.

Create another two ECS instances in the VPC to migrate the applications deployed in the ECS 1 and ECS 2 in the classic network.

Configure these two ECS instances created in step 5, replacing the SLB IP address to the IP address of the intranet SLB instance in the VPC.

The following steps are similar to the migration of the system 1.