ALIBABA CLOUD

Alibaba Cloud

High Availability Solution

Product Version:

Document Version: 20210322

Alibaba Cloud

Legal disclaimer

Alibaba Cloud reminds you to carefully read and fully understand the terms and conditions of this legal disclaimer before you read or use this document. If you have read or used this document, it shall be deemed as your total acceptance of this legal disclaimer.

- You shall download and obtain this document from the Alibaba Cloud website or other Alibaba Cloud-authorized channels, and use this document for your own legal business activities only. The content of this document is considered confidential information of Alibaba Cloud. You shall strictly abide by the confidentiality obligations. No part of this document shall be disclosed or provided to any third party for use without the prior written consent of Alibaba Cloud.
- No part of this document shall be excerpted, translated, reproduced, transmitted, or disseminated by any organization, company or individual in any form or by any means without the prior written consent of Alibaba Cloud.
- The content of this document may be changed because of product version upgrade, adjustment, or other reasons. Alibaba Cloud reserves the right to modify the content of this document without notice and an updated version of this document will be released through Alibaba Cloud-authorized channels from time to time. You should pay attention to the version changes of this document as they occur and download and obtain the most up-to-date version of this document from Alibaba Cloud-authorized channels.
- This document serves only as a reference guide for your use of Alibaba Cloud products and services. Alibaba Cloud provides this document based on the "status quo", "being defective", and "existing functions" of its products and services. Alibaba Cloud makes every effort to provide relevant operational guidance based on existing technologies. However, Alibaba Cloud hereby makes a clear statement that it in no way guarantees the accuracy, integrity, applicability, and reliability of the content of this document, either explicitly or implicitly. Alibaba Cloud shall not take legal responsibility for any errors or lost profits incurred by any organization,

company, or individual arising from download, use, or trust in this document. Alibaba Cloud shall not, under any circumstances, take responsibility for any indirect, consequential, punitive, contingent, special, or punitive damages, including lost profits arising from the use or trust in this document (even if Alibaba Cloud has been notified of the possibility of such a loss).

- By law, all the contents in Alibaba Cloud documents, including but not limited to pictures, architecture design, page layout, and text description, are intellectual property of Alibaba Cloud and/or its affiliates. This intellectual property includes, but is not limited to, trademark rights, patent rights, copyrights, and trade secrets. No part of this document shall be used, modified, reproduced, publicly transmitted, changed, disseminated, distributed, or published without the prior written consent of Alibaba Cloud and/or its affiliates. The names owned by Alibaba Cloud shall not be used, published, or reproduced for marketing, advertising, promotion, or other purposes without the prior written consent of Alibaba Cloud. The names owned by Alibaba Cloud include, but are not limited to, "Alibaba Cloud", "Aliyun", "HiChina", and other brands of Alibaba Cloud and/or its affiliates, which appear separately or in combination, as well as the auxiliary signs and patterns of the preceding brands, or anything similar to the company names, trade names, trademarks, product or service names, domain names, patterns, logos, marks, signs, or special descriptions that third parties identify as Alibaba Cloud and/or its affiliates.
- Please directly contact Alibaba Cloud for any errors of this document.

Document conventions

Style	Description	Example
Danger	A danger notice indicates a situation that will cause major system changes, faults, physical injuries, and other adverse results.	Danger: Resetting will result in the loss of user configuration data.
Warning	A warning notice indicates a situation that may cause major system changes, faults, physical injuries, and other adverse results.	Warning: Restarting will cause business interruption. About 10 minutes are required to restart an instance.
Notice	A caution notice indicates warning information, supplementary instructions, and other content that the user must understand.	Notice: If the weight is set to 0, the server no longer receives new requests.
Note	A note indicates supplemental instructions, best practices, tips, and other content.	Note: You can use Ctrl + A to select all files.
>	Closing angle brackets are used to indicate a multi-level menu cascade.	Click Settings> Network> Set network type.
Bold	Bold formatting is used for buttons , menus, page names, and other UI elements.	Click **OK**.
`Courier font`	Courier font is used for commands	Run the `cd /d C:/window` command to enter the Windows system folder.
Italic	Italic formatting is used for parameters and variables.	`bae log list --instanceid` *Instance_ID*
[] or [a|b]	This format is used for an optional value, where only one item can be selected.	`ipconfig [-all|-t]`
{} or {a|b}	This format is used for a required value, where only one item can be selected.	`switch {active|stand}`

Table of Contents

1.Disaster recovery design for core Alibaba Cloud products

1.1. SLB high availability design

This topic describes the high-availability architecture of Server Load Balancer (SLB) in terms of different system designs and product configurations to meet different business needs. You can also use SLB together with Alibaba Cloud DNS to achieve cross-region disaster recovery.

High availability of SLB

SLB instances are deployed in clusters to synchronize sessions and protect backend servers from single points of failure (SPOFs). This improves redundancy and ensures service stability. Layer 4 SLB uses Linux Virtual Server (LVS) and Keepalived software to balance loads, whereas Layer 7 SLB uses Tengine. Tengine, a web server project launched by Taobao, is based on NGINX and adds advanced features that are dedicated for high-traffic websites.

Requests from the Internet are forwarded to LVS clusters based on equal-cost multi-path (ECMP) routing. In an LVS cluster, each machine uses multicast packets to synchronize sessions with the other machines. This way, sessions are synchronized among all machines in the LVS cluster. LVS clusters also perform health checks on Tengine clusters. To ensure the availability of Layer 7 SLB, unhealthy devices are removed from Tengine clusters.

Best practice:

You can use session synchronization to prevent persistent connections from being affected by server failures within a cluster. However, for short-lived connections, server failures in the cluster may affect user access. Such incidents also happen when a connection does not trigger the session synchronization rule because the three-way handshake is not completed. To prevent session interruptions caused by server failures within the cluster, you can add a retry mechanism to the service logic. This reduces the impact of server failures on user access.

The high-availability solution with one SLB instance

Apsara Stack allows you to deploy SLB instances across multiple zones in different regions. You can deploy an SLB instance in primary/secondary zone mode. This ensures the high availability of the SLB instance. If the primary zone fails or becomes unavailable, a failover is triggered to redirect requests to servers in the secondary zone in about 30 seconds. After the primary zone recovers, traffic is automatically switched back to servers in the primary zone.

Note Zone-disaster recovery is implemented between the primary and secondary zones. SLB implements failovers only when the entire SLB cluster within the primary zone is unavailable due to factors such as power outage and optical cable failures. A failover is not triggered when a single instance in the primary zone fails.

Best practice:

1. We recommend that you create SLB instances in regions that support primary/secondary zone deployment for zone-disaster recovery.

2. You can determine the primary and secondary zones for an SLB instance based on the distribution of Elastic Compute Service (ECS) instances. Select the zone where most ECS instances are deployed as the primary zone to minimize access latency.

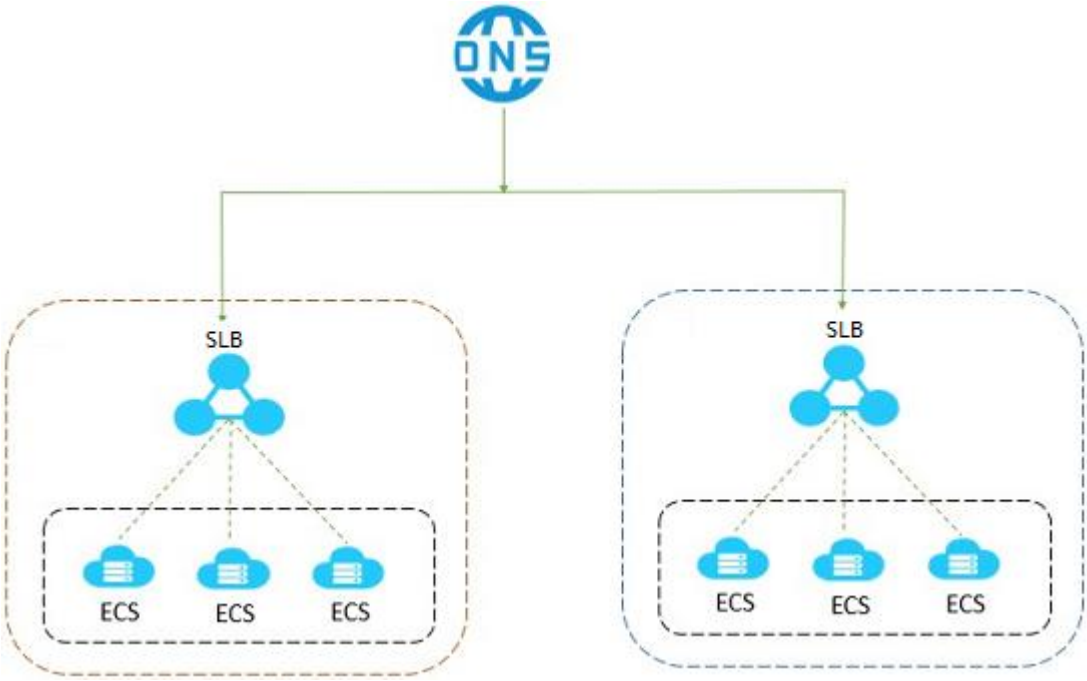
   However, we recommend that you do not deploy all ECS instances in the primary zone. You must deploy a small number of ECS instances in the secondary zone. This way, requests can be redirected to backend servers in the secondary zone when the primary zone becomes unavailable.

The high-availability solution with multiple SLB instances

If you require extremely high availability, the high-availability solution with one SLB instance may fail to suit your needs. If an SLB instance becomes unavailable due to network attacks or invalid configurations, failovers between the primary and secondary zones are not triggered. To avoid such incidents, you can create multiple SLB instances. Then, use Apsara Stack DNS to schedule requests, or use a global load balancing solution to achieve cross-region backup and disaster recovery.

Best practice:

You can deploy SLB instances and ECS instances in multiple zones within a region or across multiple regions, and schedule requests by using Apsara Stack DNS.



High availability of backend ECS instances

SLB checks the service availability of backend ECS instances by performing health checks. Health checks improve the overall availability of frontend services and help reduce the impact of service availability when backend servers are abnormal.

When SLB discovers that an instance is unhealthy, it distributes requests to other healthy ECS instances, and only resumes distributing requests to the instance when it has restored to a healthy status. For more information, see Health check overview.

Best practice:

You must enable and correctly configure the health check function. For more information, see Configure health checks.

1.2. ECS disaster recovery design

Disaster recovery solutions help guarantee the running stability and data security of your IT system. Specifically, the solutions incorporate data backup and disaster recovery of systems and applications. Alibaba Cloud ECS allows you to use snapshots and images for data backup.

Disaster recovery methods

- **Snapshot rollback**

  If exceptions occur in your system and you must roll back a disk to a previous state, you can roll back the disk so long as it has a corresponding snapshot created.

  Note

  - Rolling back a disk is an irreversible action. After disk rollback is completed, data cannot be restored. Exercise caution when performing this action.
  - After a disk is rolled back, data will be irretrievably lost from the creation time of the snapshot to the current time.

- **Image recovery**

  You can create custom images from snapshots to include the operating system and data environment in the image. The custom images can then be used to create multiple instances with the same operating system and data environment. For the configuration of snapshots and images, see Snapshots and Images.

  Note Custom images cannot be used across regions.

Technical metrics

RTO and RPO are related to the amount of data, typically on an hourly basis.

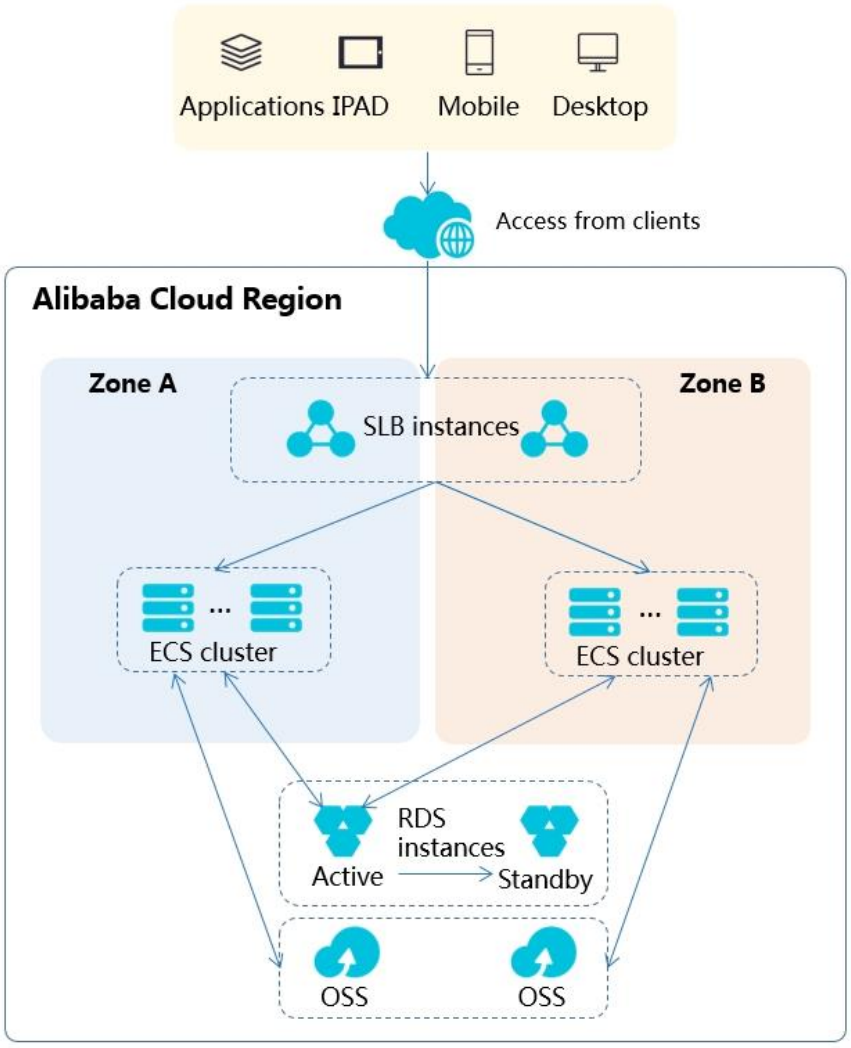
Scenarios

- Backup and recovery

  Alibaba Cloud ECS allows you to back up system disks and data disks with snapshots and images. If incorrect data is stored on a disk due to application errors or hackers' malicious access through application vulnerabilities, you can use the snapshot service to restore the disk to a desired state. In addition, Alibaba Cloud ECS allows you to reinitialize disks with images or create ECS instances from custom images.

- Disaster recovery

  Alibaba Cloud ECS supports the implementation of disaster recovery architecture. For example, you can buy and use an SLB instance at the frontend of an application, and deploy at least two ECS instances at the backend of the same application. Alternatively, you can use Auto Scaling provided by Alibaba Cloud to perform auto scaling by defining how to use ECS resources. This way, even if one of the ECS instances fails or is overloaded, disaster recovery can be implemented to ensure business continuity. The following figure provides an example in which ECS instances are deployed in data centers in two zones within the same region. All communications are implemented in the Alibaba Cloud Gigabit internal network to ensure fast response and reduce Internet traffic costs.



- SLB: SLB instances are used for load balancing between the two zones. Traffic is distributed to two or more data centers where ECS instance clusters are deployed.
- ECS cluster: ECS instances deployed in the two data centers are equivalent. The failure of a single instance does not affect data layer applications and the ECS control function. If a failure occurs, the system automatically performs hot migration so that other ECS instances can continue to provide services. This can prevent service interruptions caused by a single point of failure or hot migration failures. If hot migration fails, you will receive a notification about the failures based on system events so that you can deploy new nodes in a timely manner.

- Data layer: OSS is deployed at the region level. ECS nodes in data centers in different zones can access objects in OSS. For database applications, multi-zone ApsaraDB for RDS service is used. Primary nodes can perform read and write operations across zones without conflicting with application-layer traffic. In addition, secondary nodes can perform read operations across zones to prevent inability of ECS instances to read data in case of failures of the primary nodes.

1.3. OSS high-availability design

OSS ensures availability from system design, product configuration, and other aspects.

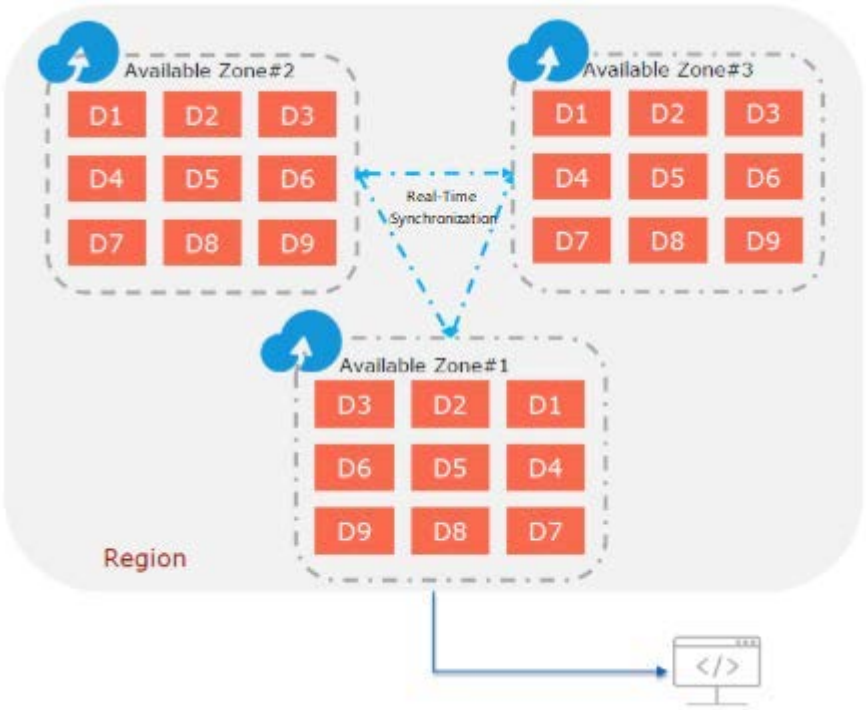
Data durability

The following figure shows the data durability- and availability-related metrics of OSS.

	Standard	IA	Archive
Designed durability	99.9999999999%	99.9999999999%	99.999999999%
Designed availability	99.995%	99.995%	99.99% (For restored objects)
Available Zone (AZ)	3	3	3

Zone-disaster recovery

OSS provides zone-redundant storage to achieve zone-disaster recovery. In the zone-disaster recovery mode, objects are stored as replicas across three zones within a region. OSS regularly checks the integrity of the stored data. Business data can be processed even if data in an entire zone is destroyed. Data can be synchronized and copied across three zones in real time. Business can be failed over from a faulty zone to a normal zone, which you are not aware of.



Features of OSS zone-disaster recovery are as follows:

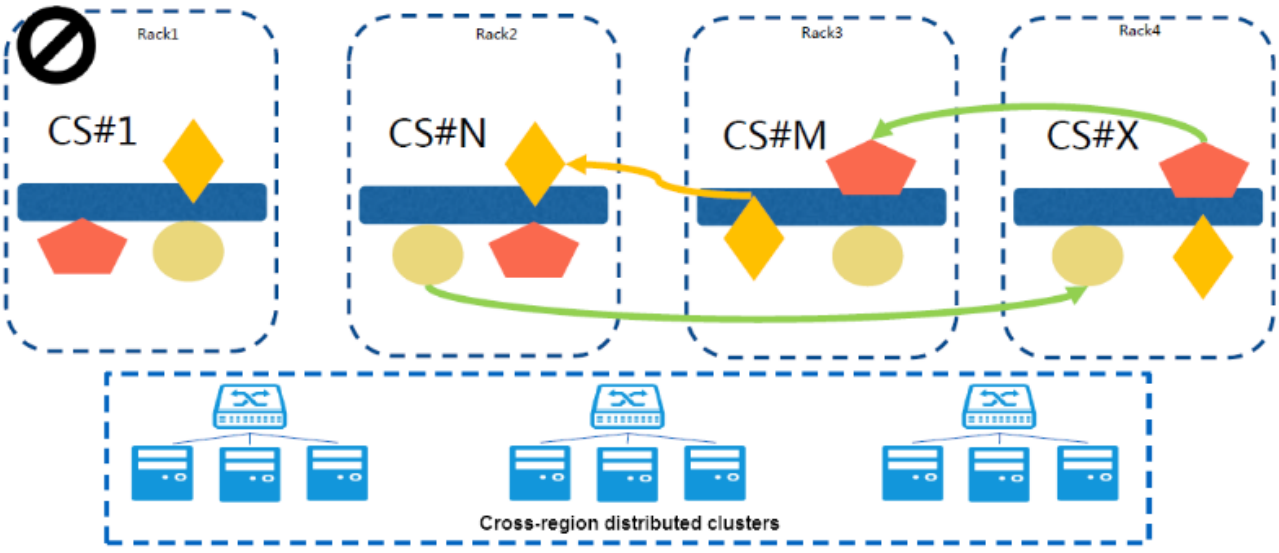
- Data center-level disaster recovery capabilities: Data reliability reaches 99.9999999999%. When a data center becomes unavailable due to hardware faults or disasters, OSS still maintains high consistency. This capability ensures that business is not affected and no data is lost. You are not aware of the failover process. This feature can meet requirements of key business systems that require zero Recovery Point Objective (RPO) and Recovery Time Objective (RTO).
- Higher SLA: OSS zone-redundant storage provides an SLA of 99.995%, which are five times higher than the SLA of the standard storage of data in a single zone.
- One-click activation: OSS-based zone-redundant storage allows you to build cloud-based zone-disaster recovery capabilities with a single click. You can enable zone-redundant storage when you create a bucket. OSS uses a multi-replica mechanism to automatically

store user data in three zones that are several kilometers away from each other within the same region.

OSS zone-redundant storage is available in China (Beijing), China (Shanghai), China (Hangzhou), and China (Shenzhen) on the China site (aliyun.com). This feature will be available in more regions in the future.

High reliability and stability

The three replicas are highly consistent and distributed across different zones. Data is automatically replicated when a fault occurs, as shown in the following figure.



Remote disaster recovery

The remote disaster recovery solution is mainly applicable to the following scenarios:

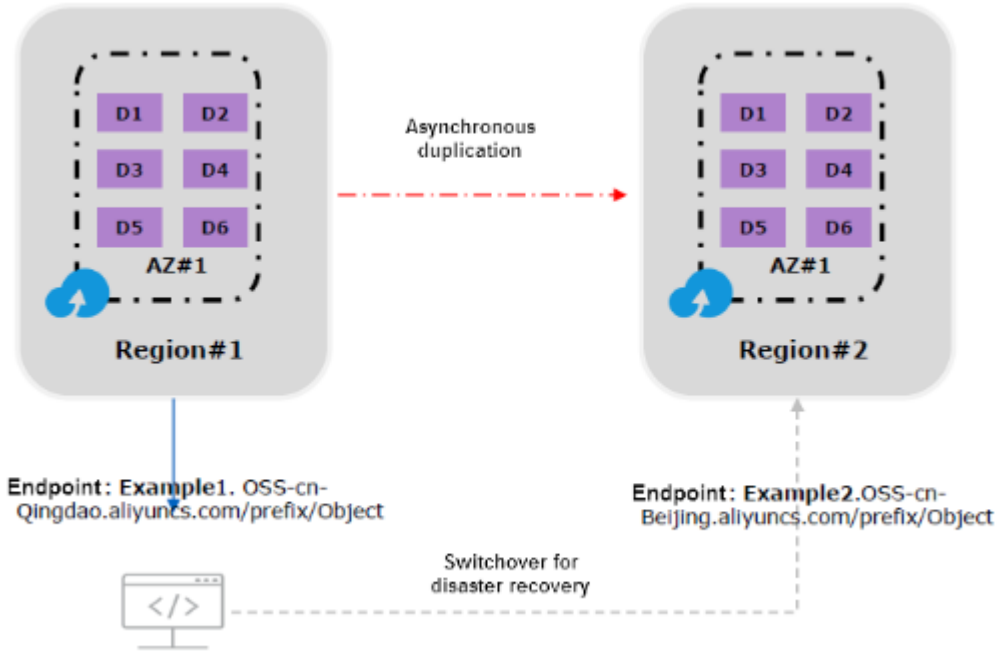
- Compliance requirements: According to some industrial compliance regulations, a replica of the data must be stored at a site that is a certain distance from the original site where the data is stored.

- Remote backup and disaster recovery: A replica of the data must be stored at a remote site in case of severe natural disasters, such as earthquakes and tsunamis.
- Data replication: For business reasons, you may need to migrate data from one OSS IDC to another.

OSS supports cross-region replication for remote disaster recovery. It provides the following features:

- In asynchronous replication mode, data latency is related to the amount of transmitted data and the transmission speed. Generally, the latency ranges from several minutes to several hours.
- You can view the progress of historical data synchronization tasks in the console.
- You can add, modify, or delete a synchronization task.
- Filter rules: You can set the configuration to synchronize all objects in the source bucket or synchronize only objects with a specified prefix.

For more information, see Cross-region replication.



1.4. ApsaraDB for RDS disaster recovery design

Backup and recovery

- RDS supports automatic and manual backups. You can set the automatic backup frequency or manually create backups at any time. For more information, see Backup and recovery.
- RDS supports data recovery by time or backup set. You can restore data of any point in time within the log retention period to a new instance, verify the data, and then transfer the data to the original instance. For more information, see Backup and recovery.

Local disaster recovery

Series	Description

Basic edition	• Data backup is stored on an OSS instance or a distributed cloud disk. Multi-replica redundancy is used to ensure that no data is lost. (This is applicable to all ApsaraDB for RDS instances.) • Only one node is available. No node is provided for hot backup. When a fault occurs, it takes a long time to restore service. **This edition is applicable to scenarios with less demanding availability requirements.**
High-availability edition	This solution adopts a dual-host hot architecture consisting of a primary node and a secondary node. It is applicable to more than 80% of scenarios. When the primary node fails, the traffic is switched to the secondary node within seconds. The failover process is transparent to the application. When the secondary node fails, ApsaraDB for RDS automatically creates a new secondary node to ensure the high availability of the service. • **Single-zone instances**: The primary and secondary nodes are in the same zone. The primary and secondary nodes are deployed on different physical servers. Redundant racks, air HVAC systems, circuits, and networks are available in the zone. In this way, the high availability of the service is ensured.

	• **Multi-zone instances** (local dual-IDC or local disaster recovery instances): The primary and secondary nodes are deployed in different zones of the same region. Cross-zone disaster recovery is supported without any additional charge. **Note** Single-zone instances can be converted to multi-zone instances and vice versa. For more information, see Migrate zones.
Cluster edition	The cluster edition provides a primary-secondary high-availability architecture and seven read-only nodes, allowing you to scale out the read capability of the cluster. The data of the secondary node and all read-only nodes is synchronized from the primary node. The cluster edition has the same high availability as the high-availability edition. The read-only nodes can be deployed in zones other than those of the primary and secondary nodes. **Note** • The cluster edition is only available for ApsaraDB RDS for SQL Server 2017. For more information, see Cluster edition. • For more information about the read-only Apsara RDS for MySQL instances, see Read-only instances.

	• For more information about the read-only nodes of the POLARDB cluster, see POLARDB clusters.
Enterprise edition	• This edition consists of one primary node and two secondary nodes, ensures strong data consistency through the synchronization of multiple replicas, and provides financial-level data reliability. It is applicable to the core production databases of large enterprises in all industries. • The three nodes in the finance edition are always deployed in three different zones of the same region. **Note** • Currently, enterprise edition is only available for MySQL 5.6. For more information, see Enterprise Edition. • Enterprise edition is only available for ApsaraDB for RDS instances deployed in the China (Beijing), China (Hangzhou), China (Shanghai), and China (Shenzhen) regions.

Remote disaster recovery

- You can synchronize data from a database in an on-premises data center or a user-created database on an ECS instance to an ApsaraDB for RDS instance in any region in real time.

Even if the data center is damaged, a data backup is always stored on the ApsaraDB for RDS instance. For more information about the operation, see Create real-time synchronization jobs.

1.5. ApsaraDB for Redis disaster recovery design

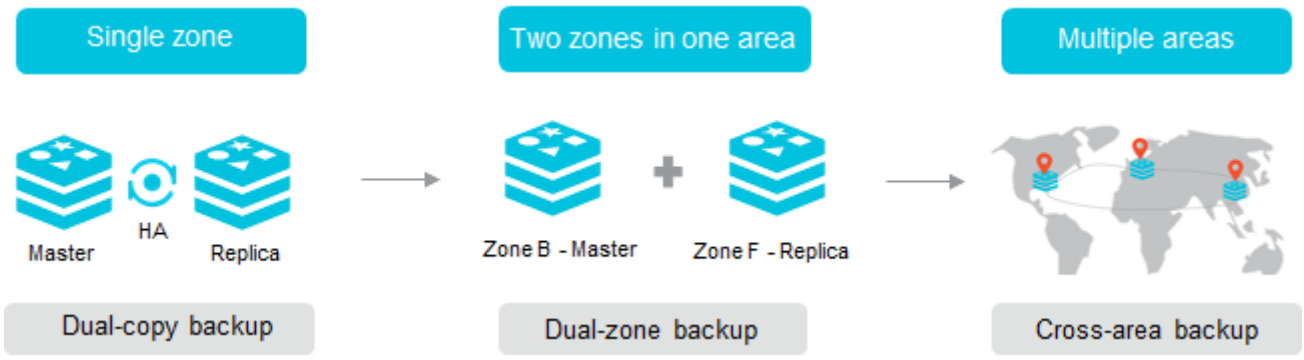
Data is a core element of many businesses, and as data storage systems, databases bear a critical responsibility. ApsaraDB for Redis is a high-performance key-value database that is often used to store large volumes of important service data. This topic describes the disaster recovery mechanism used by ApsaraDB for Redis in detail.

Roadmap of ApsaraDB for Redis for disaster recovery

A variety of problems may occur during data management, such as software bugs, device malfunctions, or power failures at data centers. A disaster recovery solution guarantees data consistency and service availability. ApsaraDB for Redis provides optimized disaster recovery solutions to achieve high availability in different scenarios.

The following figure shows the roadmap of ApsaraDB for Redis for disaster recovery.

Disaster recovery solutions



Three solutions are available in ApsaraDB for Redis to meet different requirements. The following sections describe these solutions in detail.

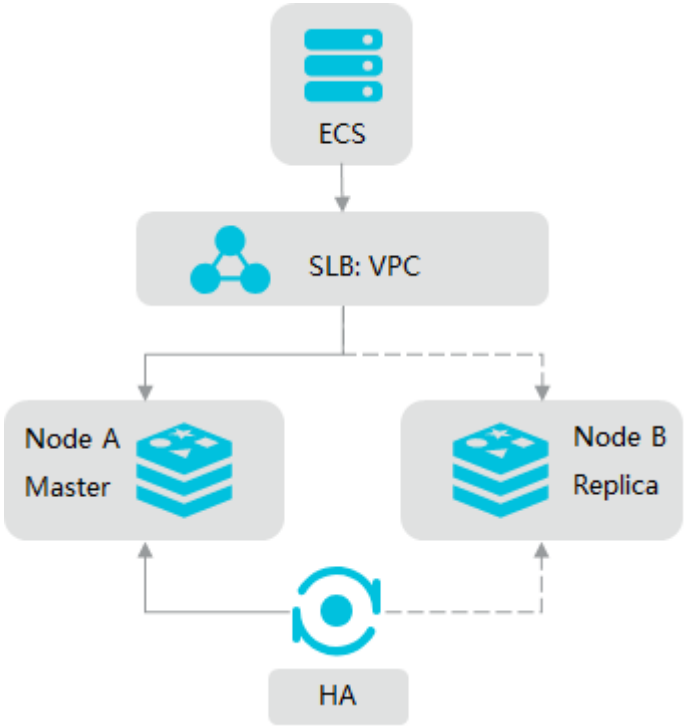
Single-zone high availability

All ApsaraDB for Redis instances support a single-zone HA architecture. The HA system runs on an independent platform to guarantee high availability across zones. Compared with self-managed Redis databases, ApsaraDB for Redis enables more stable database management.

Standard master-replica instances

A standard master-replica instance runs in a master-replica architecture. If the HA system detects a failure on the master node, the system switches the workloads from the master node to the replica node and the replica node takes over the role of the master node. The original master node works as the replica node after recovery. By default, data persistence is enabled for the instance. The system automatically creates backup files on the instance. You can use the backup files to roll back or clone the instance. This mechanism avoids data loss caused by user errors and enables reliable disaster recovery.

High-availability solution of a standard master-replica instance



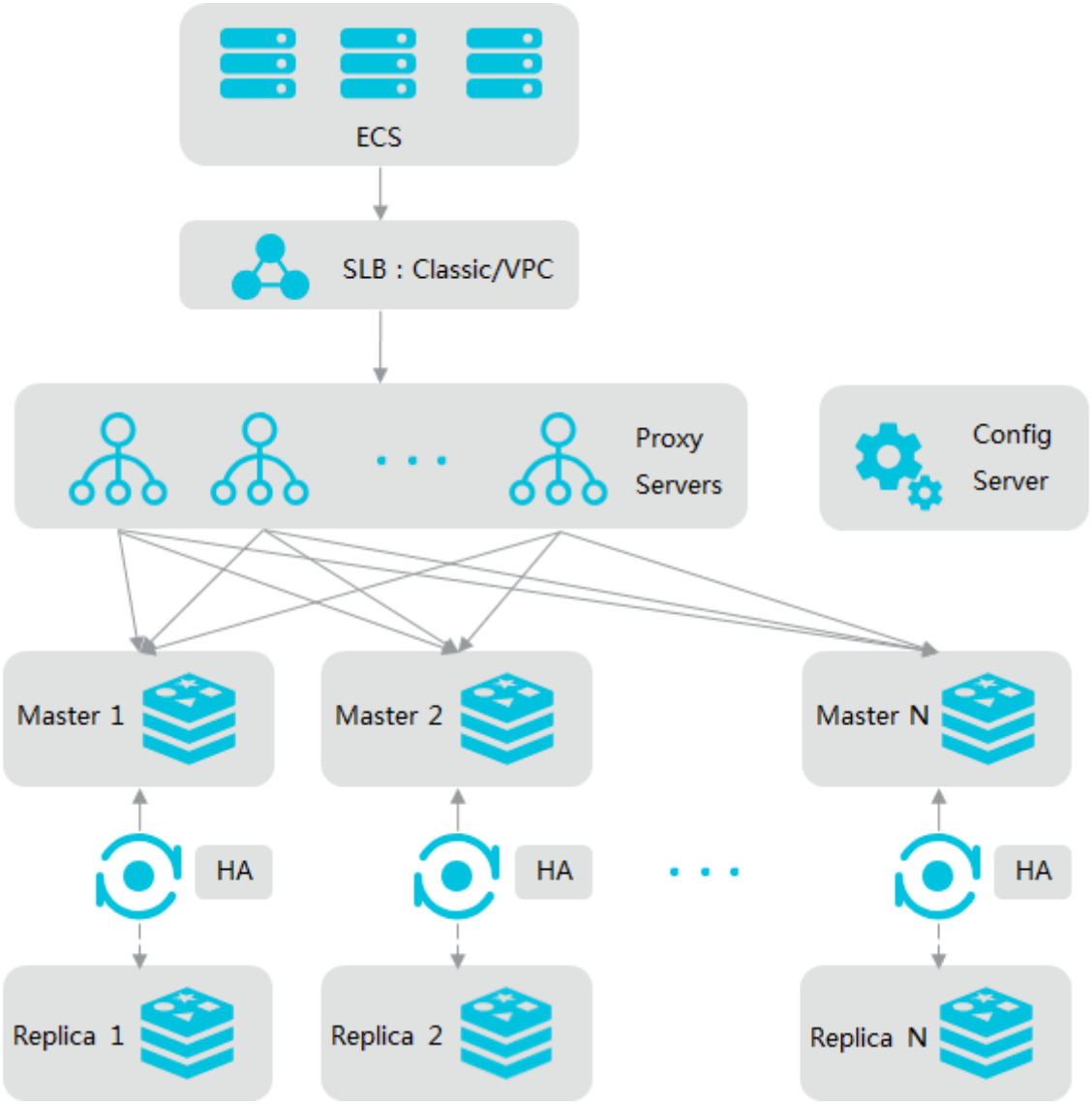
Master-replica cluster instances

A master-replica cluster instance consists of a configuration server, multiple proxy servers, and multiple data shards.

- The configuration server is a cluster management tool that provides global routing and configuration information. This server uses a cluster architecture with three replica nodes and follows the Raft protocol.
- A proxy server runs in a standalone architecture. A cluster contains multiple proxy servers. The cluster automatically balances loads and performs failovers among these proxy servers.
- A data shard runs in a master-replica high-availability architecture. Similar to a standard master-replica instance, if the master node fails, the HA system performs a failover to

ensure high availability, and updates the information on the proxy servers and configuration server.

High-availability solution of a master-replica cluster



Zone-disaster recovery

Standard instances and cluster instances support zone-disaster recovery across two data centers. If your workloads are deployed in a single region and require disaster recovery, you can select the zones that support zone-disaster recovery when you create an ApsaraDB for Redis

instance. For example, you can select **China (Hangzhou) Zone (B+F)** or **China (Hangzhou) (G+H)** from the Zone drop-down list in the console.

Create a zone-disaster recovery instance

When you create a multi-zone instance, the master node and replica node are deployed in different zones and provided with the same specifications. The master node synchronizes data to the replica node through a dedicated channel.

If a power failure or a network error occurs on the master node, the replica node takes over the role of the master node. The system calls an API operation on the configuration server to update routing information for proxy servers. The underlying network performs a failover based on the precision of the routing information available in a backbone network. The master node provides more specific CIDR blocks than the replica node. In normal conditions, the system transmits requests to the master node through specific CIDR blocks. If the master node fails, the master node does not upload routing information to the backbone network. The backbone network only provides less specific CIDR blocks of the replica node. The system routes requests to the replica node based on the available routing information.

ApsaraDB for Redis provides an optimized Redis synchronization mechanism. Similar to global transaction identifiers (GTIDs) of MySQL, ApsaraDB for Redis uses global operation identifiers

(OpIDs) to indicate synchronization offsets and runs lock-free threads in the background to search OpIDs. The system asynchronously synchronizes append-only file (AOF) binary logs (binlogs) from the master node to the replica node. You can throttle synchronization to ensure service performance.

2.Public cloud-based remote disaster recovery

When an enterprise has deployed its services on Alibaba Cloud and has high requirements for remote disaster recovery, we recommend Alibaba Cloud's public cloud-based remote disaster recovery solution. The enterprise can use DNS, SLB, and other Alibaba Cloud products to create a multi-zone architecture for remote disaster recovery.

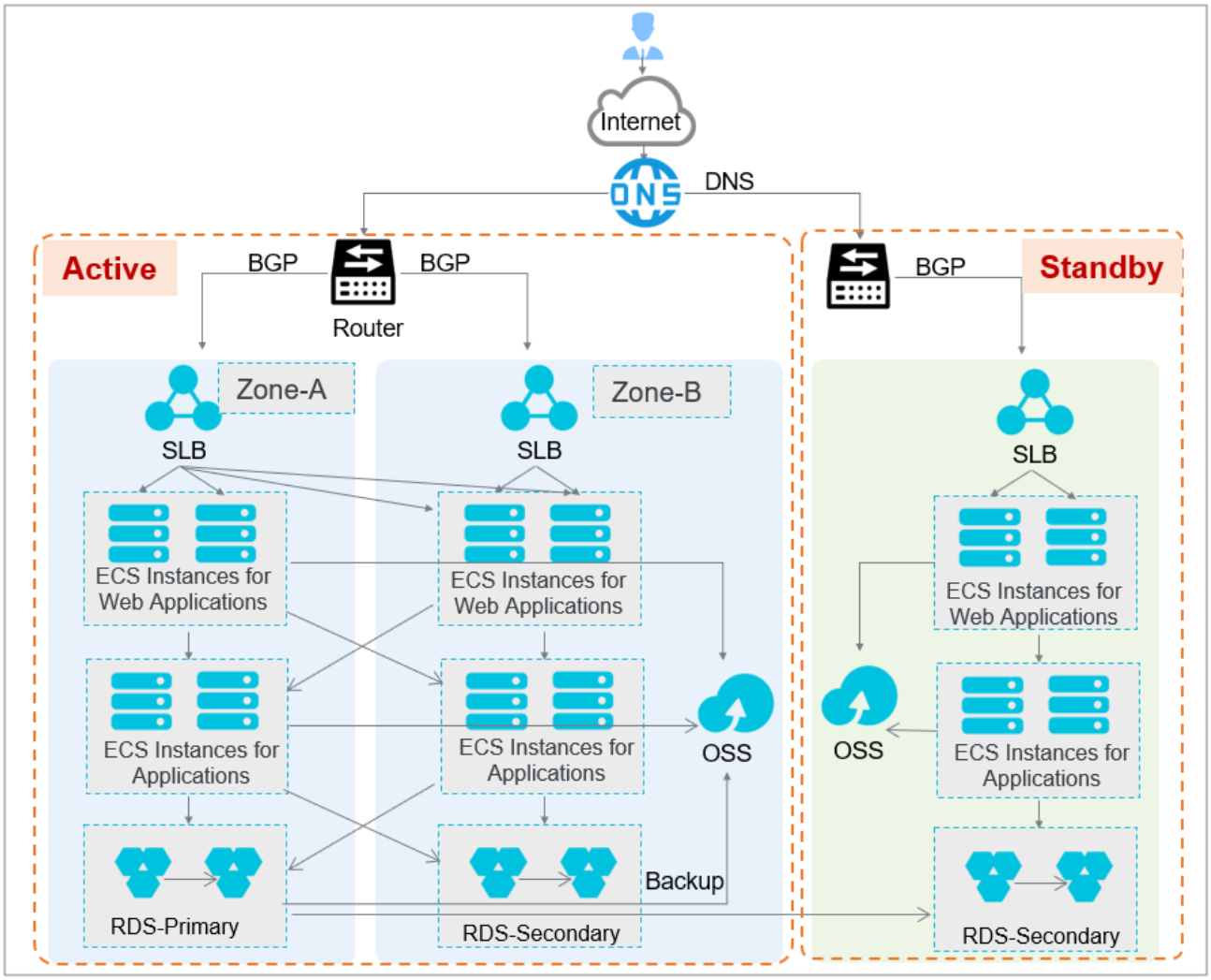
Scenarios

The public cloud-based remote disaster recovery solution is applicable to the following scenarios:

- Public cloud: An enterprise has deployed its services on Alibaba Cloud and wants to create a remote disaster recovery architecture involving multiple Alibaba Cloud regions.
- Application-level disaster recovery: An enterprise wants to back up all its applications for disaster recovery, instead of a single database or storage system.
- Cloud-based remote disaster recovery: This solution applies to public clouds that may experience a fault in a certain region. For example:
  - The whole region is unavailable due to a natural disaster, such as an earthquake.
  - The whole region is unavailable for an extended period due to infrastructure failures.

Recommended architecture

For large enterprises that require both local and remote disaster recovery solutions to ensure service security, service availability, and data reliability, we recommend the following remote disaster recovery solution.



Architecture description:

- A complete backup of the original application architecture is created and stored in each region and zone.
- Alibaba Cloud Express Connect is used for private network communication between different regions. This ensures real-time synchronization and minimizes transmission latency between databases in different regions.
- When a fault occurs, the front-end DNS instance is used to implement service failover within seconds, ensuring that the service can be restored in a timely manner.

- This architecture can resolve faults of a single IDC or faults caused by disasters such as earthquakes.

Advantages of the architecture

- Alibaba Cloud DNS supports intelligent resolution and facilitates traffic distribution or service failover for disaster recovery.
- This architecture supports communication between VPCs through Express Connect and allows you to publish and deploy applications and modify application configurations in a unified manner.
- This architecture supports data replication between OSS instances in different regions.
- This architecture supports data synchronization between different regions through DTS.

3.Hybrid Disaster Recovery

Alibaba Cloud's HDR solution is based on the HDR system, a core Alibaba Cloud product. We recommend that you design the HDR solution and create the required environment by referring to the following architecture.

Applicable scenarios

The HDR solution is applicable to the following scenarios:

- Hybrid cloud: An enterprise wants to use a cloud platform in a hybrid cloud architecture as its remote disaster recovery center for local data.
- Application-level disaster recovery: An enterprise wants to back up all its applications for disaster recovery, instead of a single database or storage system.
- Hour-level RTO/RPO: An enterprise wants an RTO/RPO in hours for application disaster recovery.

  Note For more information about RTO/RPO, see Appendix: Basic concepts of disaster recovery.

Recommended HDR architecture

Architecture description:

- Your on-premises IDC is connected to an off-premises IDC through a leased line or VPN to build a hybrid cloud architecture.
- When an HDR system or BCDR gateway is deployed on the on-premises IDC:
  - The HDR system can deliver an RTO/RPO of hours.

  - The BCDR gateway can deliver an RTO/RPO of minutes.

    The HDR system or BCDR gateway can silently back up data from the physical machines or VMs of the on-premises IDC to the HDR system based on the backup policies.

- The HDR system or BCDR gateway can upload the backups to the disaster recovery database based on relevant policies. Additionally, the HDR system or BCDR gateway can restore backups from the disaster recovery database to the on-premises database.
- The HDR console can monitor the status of backup tasks and the consumption of cloud resources. You can deploy a BCDR gateway in the HDR console as needed.
- The BCDR gateway can restore backups of the server data to off-premises ECS instances. When the on-premises gateway is restored, data in the ECS instances can be restored to the on-premises servers.

You can select different HDR system models based on your disaster recovery needs:

Model	Applicable disaster recovery scenarios
HDR 1000	• The 1U model can be used for Remote Office/Branch Office (ROBO) and disaster recovery scenarios with small or medium loads.

	• A single HDR 1000 supports a maximum of 20 nodes and provides a data capacity of 20 TB. • The backup and restoration throughput is 0.5 TB/hour. • It provides double power supplies, a RAID card supporting power failure protection, and a high-performance hard disk.
HDR 2000	• The 2U model can be used for scenarios in which a large amount of data needs to be backed up. • A single HDR 2000 supports a maximum of 100 nodes and provides a maximum data capacity of 50 TB. • The backup and restoration throughput is above 1.5 TB/hour. • It provides double power supplies, a RAID card supporting power failure

	protection, and a high-performance hard disk.

The following describes the disaster recovery capabilities of the BCDR gateway:

Advantages of the HDR architecture

- **Support for abundant platforms**

  The HDR architecture supports the backup, disaster recovery, and migration of mainstream operating systems, file systems, and applications of all types in physical machines or VMs.

- **Integration of backup and disaster recovery**

  The HDR architecture supports local backup, cloud backup, and cloud recovery, allowing it to satisfy various disaster recovery requirements.

- **Security and stability**
  - The HDR architecture supports AES256-based data encryption, ensuring data security.
  - The HDR system is proven to be stable and reliable according to Alibaba Cloud's strict standards. The data availability of the cloud disaster recovery database is 99.99999999999999%.
  - The HDR architecture is based on selecting the optimal server from hundreds of thousands of servers purchased from Alibaba Group each year. Its software and hardware are reinforced and optimized.

- **In-depth storage and WAN optimization**

Global data deduplication and compression is performed locally, which greatly improves the storage efficiency. Both full and incremental backup modes provide highly efficient cloud migration. Additionally, cloud migration can be performed on a public network.

4.Hybrid Backup Recovery

When enterprises want to back up their data and applications on Alibaba Cloud through a hybrid cloud, we recommend the Hybrid Backup Recovery (HBR) solution. You can use Hybrid Backup, a core Alibaba Cloud product, to build an architecture as follows.

Backup is the basis for disaster recovery. When a Hybrid Disaster Recovery (HDR) system is used, you can refer to the architecture of the Hybrid Disaster Recovery for backup purposes. Additionally, Alibaba Cloud HBR allows enterprises to only back up their data to a cloud in a more cost-effective manner. When HBR is used, we recommend that you create a disaster recovery environment by referring to the following architecture:

Applicable scenarios

The HBR solution is applicable to the following disaster recovery scenarios:

- Hybrid cloud: An enterprise wants to use a cloud platform in a hybrid cloud architecture as its remote backup center for local data.
- Application-level disaster recovery: An enterprise wants to back up all its applications, instead of a single database or storage system.
- Hour-level RTO/RPO: An enterprise wants an RTO/RPO in hours for application disaster recovery.

  Note

Recommended HBR architecture

Architecture description:

- The on-premises IDC is connected to Alibaba Cloud through a leased line or VPN to build a hybrid cloud architecture.
- The HBR client is installed on a on-premises physical machine or virtual machine (VM) whose data will be backed up. The region where the off-premises backup database is located needs to be configured in the client. The data of the on-premises IDC is backed up to Alibaba Cloud through a leased line or VPN.
- If the on-premises IDC fails, enterprises can restore the data backups from the off-premises backup database to the on-premises IDC.

The following table lists data that can be backed up through HBR:

Advantages of HBR

- **Deduplication and compression**

  Based on the deduplication and compression technologies developed by Alibaba Cloud, HBR effectively reduces the I/O transmission and cloud backup storage workloads, improving backup speeds and reducing costs.

- **Security**

  Three data replicas are saved in HBR, effectively ensuring the security of data backups. Additionally, Alibaba Cloud provides an end-to-end data encryption and verification mechanism and a RAM system that assigns different permission levels to operators, reducing the number of backup access paths.

- **Unlimited scaling**

  The HBR repository can be scaled infinitely to store massive volumes of data.

- **Full backup**

  HBR supports the backup of directories and files in a physical or virtual environment and the backup of VMware VMs.

- **Fully-managed**

  Instead of building a backup system in the on-premises or off-premises IDC, you only need to host the backup data on the HBR repository. The HBR solution means you do not have to worry about hardware presets, configurations, cluster scaling, and security.

5.Hybrid cloud-based database disaster recovery solution

5.1. Applicable disaster recovery scenarios

Enterprises that have high requirements for database disaster recovery can use the hybrid cloud-based database disaster recovery solution. Enterprises can back up their databases to a hybrid cloud to form a remote database disaster recovery architecture. In this architecture, when the on-premises database fails, enterprises can restore the database from its backups on the cloud.

This solution can be used in the following scenarios:

- Hybrid cloud: An enterprise wants to use a cloud platform in a hybrid cloud architecture as its remote disaster recovery center.
- Data-level disaster recovery: An enterprise wants to back up its databases for disaster recovery.
- RTO/RPO: Alibaba Cloud's hybrid cloud-based cold database disaster recovery solution can deliver an RPO in seconds. The hot database disaster recovery solution supports two-way real-time synchronization between databases.

5.2. Hybrid cloud-based database backup——using DBS

For enterprises that want to back up their databases on Alibaba Cloud through a hybrid cloud, we recommend the hybrid cloud-based database backup solution. You can use Data

Transmission (DTS), Database Backup (DBS), and other core Alibaba Cloud products to build an architecture as follows.

You can use DTS and DBS to synchronize on-premises databases of all types to off-premises databases, or restore on-premises databases from the backups stored on the cloud.

Database synchronization tool	Disaster recovery scenario	Backup database	RPO
DTS	Recommended for the hot database architecture	Database	Two-way real-time synchronization between off-premises and on-premises databases
DBS	Recommended for the Recommended cold database architecture	OSS	In seconds

Recommended hot database architecture

Architecture description:

- Your on-premises IDC is connected to an off-premises IDC through a leased line or VPN to build a hybrid cloud architecture.

- When DTS is used:
  - During data synchronization:

Data sources	Description
From ApsaraDB RDS for MySQL to ApsaraDB RDS for MySQL	The solution allows you to reduce the synchronization delay to the second level when the distance between databases is longer than 1000 km.
Between ApsaraDB RDS for MySQL and ApsaraDB RDS for MySQL	Two-way data synchronization is supported.
From ApsaraDB RDS for MySQL to MaxCompute	When data is synchronized from ApsaraDB for RDS to MaxCompute, full data tables and incremental data tables are created. You can retrieve the full amount of data at any time by merging the full data tables and incremental data tables.
From ApsaraDB RDS for MySQL to DataHub	You can synchronize data from ApsaraDB RDS for MySQL to DataHub, and then synchronize data from DataHub to StreamCompute.

From ApsaraDB RDS for MySQL to AnalyticDB	Data synchronization from ApsaraDB RDS for MySQL to AnalyticDB can be used for real-time data warehouse analytics, such as real-time reporting and screen monitoring.
From MySQL to DRDS	Incremental data synchronization is used to synchronize data from MySQL to DRDS.
From MongoDB to MongoDB	An example is the geo-disaster recovery project for AutoNavi Software Co., Ltd.

- During data migration:

Data sources	Support for schema migration	Support for full data migration	Support for incremental data migration
From MySQL to MySQL (RDS and user-created databases)	Yes	Yes	Yes (DML and some DDL statements)
From MySQL to DRDS, PetaData, or OceanBase	Yes	Yes	Yes (DML statements)

From MySQL to Oracle	Yes	Yes	Yes (DML statements)
From Oracle to MySQL (RDS and user-created databases)	Yes	Yes	Yes (DML and some DDL statements)
From Oracle to DRDS	Yes (manually enabled)	Yes	Yes (DML statements)
From Oracle to ApsaraDB RDS for PPAS	Yes	Yes	Yes (DML statements)
From Oracle to AnalyticDB	Yes	Yes	Yes (DML statements)
From Oracle to OceanBase	Yes	Yes	Yes
From SQL Server to SQL Server	Yes	Yes	Yes (DML statements)

From PostgreSQL to PostgreSQL	Yes	Yes	Yes (DML statements for primary key tables)
From MongoDB to MongoDB	Yes	Yes	Yes
From Redis to Redis	Yes	Yes	Yes
From IBM Db2 to MySQL	Yes	Yes	Yes (DML and some DDL statements)

Recommended cold database architecture

Architecture description:

- Key conditions:
  - You have deployed two databases on your on-premises environment: production and recovery databases. The production database is used to store production data, and the recovery database is used to restore data if a fault occurs.
  - Storage services are available in two Alibaba Cloud regions, for example, China (Shenzhen) and China (Qingdao). The storage service can be Object Storage Service (OSS) or Network Attached Storage (NAS).

  - DBS is available to perform real-time hot backups to copy data from on-premises databases to the cloud.

- Backup of on-premises production data to the cloud: (You can use either of the following methods to back up on-premises production data to the cloud.)
  - You can deploy another storage system to back up production data to the on-premises IDC. Then, you can copy the backup data from the on-premises IDC to the cloud.
  - You can use DBS to perform hot backups to copy production data from your on-premises production database to the storage systems of the two Alibaba Cloud regions.
- Data recovery:
  - You can restore data from your on-premises storage system to the on-premises recovery database in the following scenario. A fault occurs on your on-premises production database, but the on-premises storage system runs properly.
  - You can use DBS to restore data from your backups in the cloud to your on-premises recovery database in the IDC in either of the following scenarios. One of the scenarios is that a fault occurs on your on-premises production database and storage system. The other scenario is that the on-premises storage system is not deployed.
- Features:
  - Advantages: high consistency and quick recovery.
  - Disadvantages: varying RTO. The RTO varies depending on the database size, because a new database instance is created for data backup.

- Scenarios: This mature backup solution is applicable to most relational databases.

Advantages of the architecture

- When DTS is used:
  - This architecture supports real-time synchronization between databases of different types and can be used for remote disaster recovery and remote multi-active scenarios.
  - The architecture supports two-way synchronization between ApsaraDB RDS for MySQL instances, helping you build a remote multi-active architecture.
  - After the synchronization link is established, you can dynamically modify the synchronization objects and the synchronization speed.
- When DBS is used:
  - The architecture provides strong protection for data stored in multiple environments, such as enterprise data centers, hybrid clouds, and public clouds provided by third-party cloud vendors.
  - DBS provides an overall data backup and recovery solution, which allows for real-time incremental backup and data recovery within seconds.

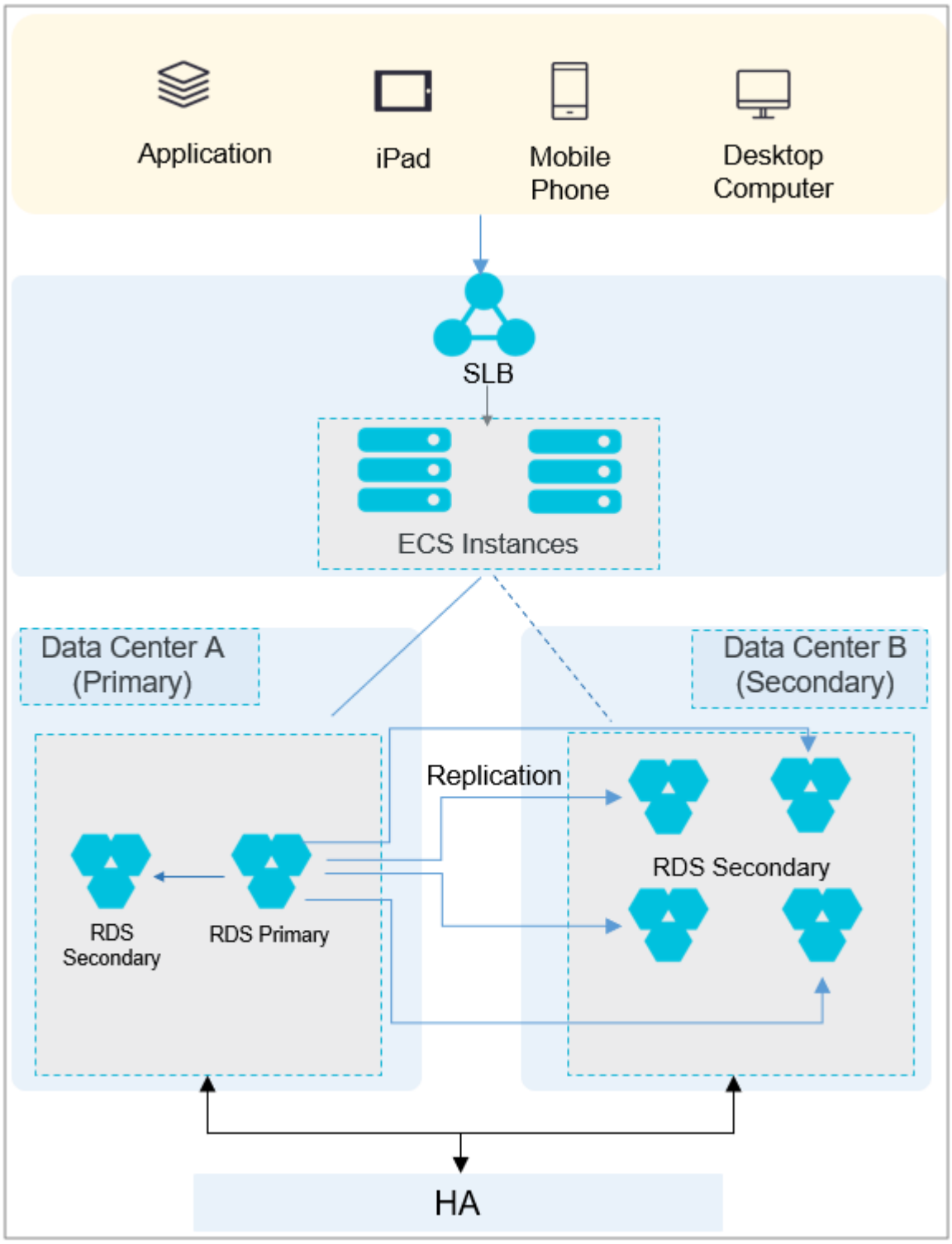
5.3. Dual zone disaster recovery and backup solution

When you need a database disaster recovery solution based on cloud services deployed in the same city, you can purchase an Alibaba Cloud DTS instance for data migration and real-time synchronization. The disaster recovery solution supports three modes: Replication and high availability, active/standby (A-S), and active-active (A-A).

- Replication and high availability: One database is deployed in each of the two data centers. You can copy data from the database of the primary data center to that of the secondary data center. If a fault occurs on the database of the primary data center, business switches to the database of the secondary data center.
- Active/standby mode: Both data centers are deployed with the same system. The secondary data center is used only to back up data. If a fault occurs on the primary data center, business switches to the secondary data center.
- Active/active mode: Both data centers are deployed with the same system, each with independent workloads and hosted services. Both data centers reserve some of their respective resources for data backup. If a fault occurs on either of the data centers, business switches to the other data center and the reserved resources are used. We recommend that you use the active/standby mode, if you have sufficient resources and have demanding requirements for zone-disaster recovery. If your resources are limited, we recommend that you use the active/active mode.

Replication and high availability

Relational Database Service (RDS) is used as an example to introduce the architecture of the replication and high availability mode.



Architecture:

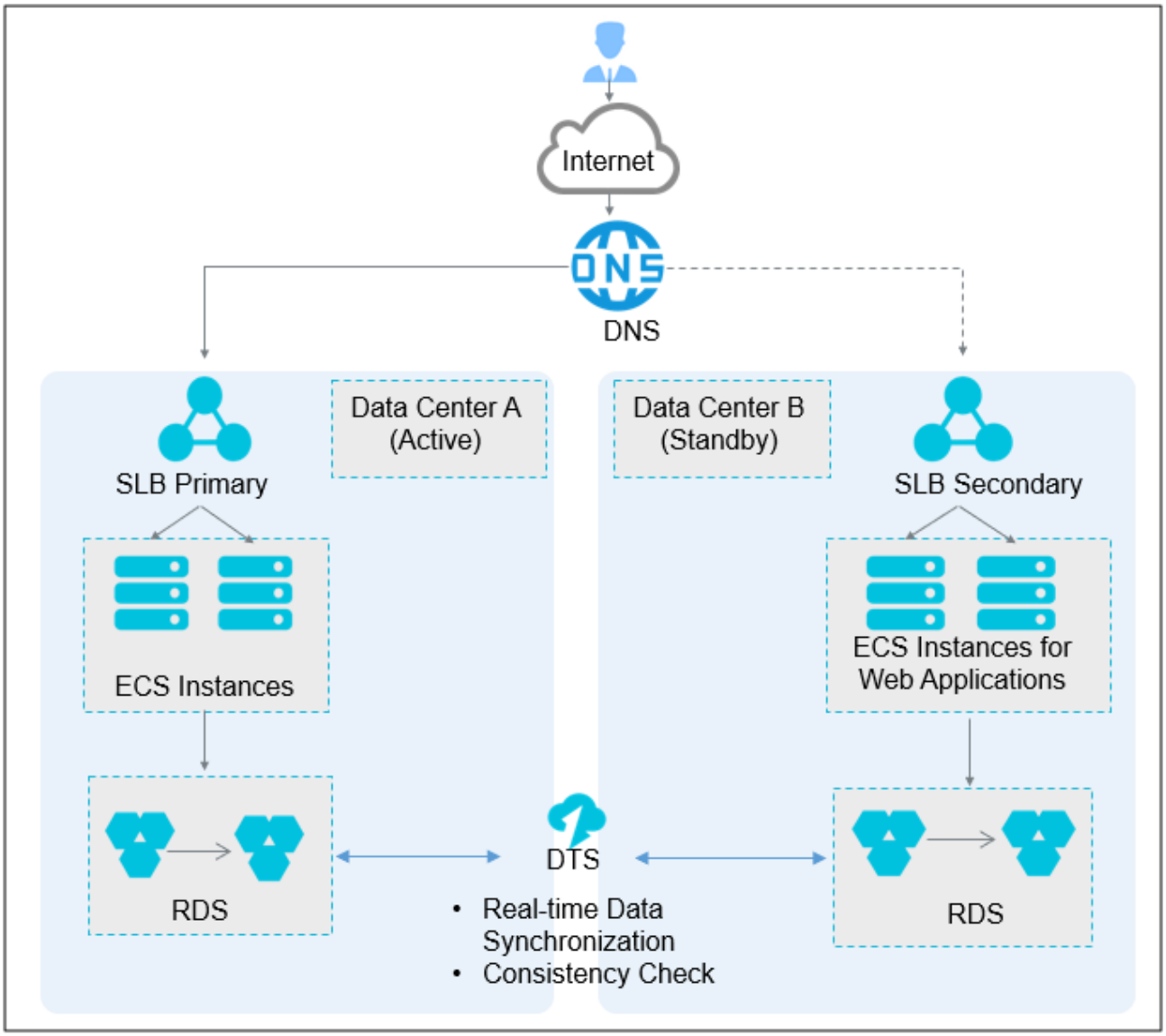
- Key conditions:
  - RDS clusters are deployed in data centers A and B, respectively.

  - SLB, ECS, and HA systems can be connected to applications across data centers. You only need to deploy one set of SLB, ECS, and HA systems in data centers A and B. You can manage and control the databases in the two data centers by using the HA system.

- Data backup and recovery:
  - This mode copies the data from the databases in Data Center A to the databases in Data Center B.
  - If a fault occurs on the databases in Data Center A, the HA system directs traffic to Data Center B. The resources of Data Center A remain unavailable if the fault persists. After Data Center A is recovered, you can specify it as a secondary zone.
- Features:
  - Advantages: lightweight switchover and low costs.
  - Disadvantages: risks of data inconsistency. A small amount of data may be inconsistent after the switchover, such as the loss of a transaction.

Active/standby mode

An example is used to introduce the architecture of the active/standby mode. This example assumes that you have deployed services in two Alibaba Cloud zones within the same region.



Architecture:

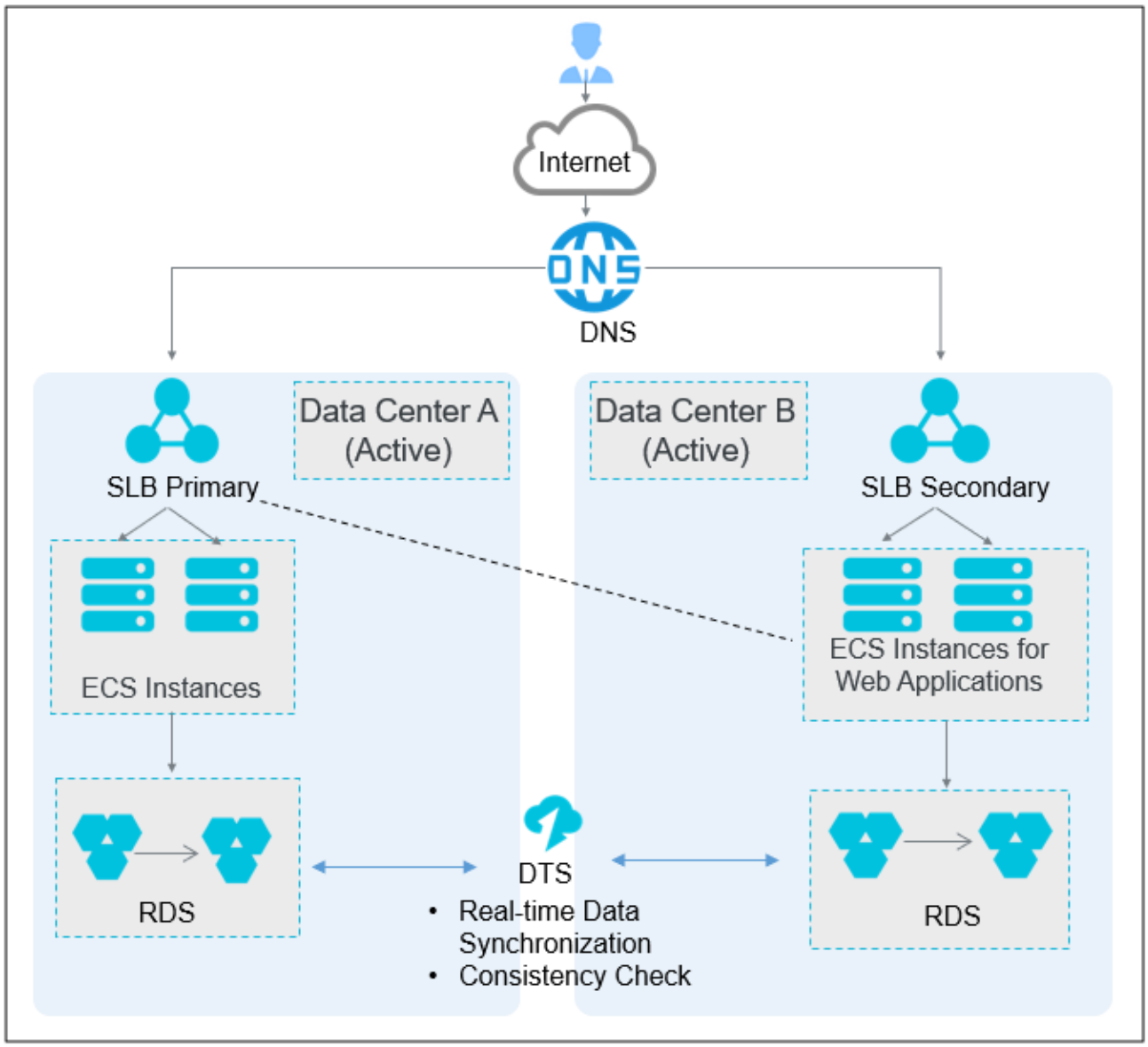
- Key conditions:
  - Data Center A has the same RDS clusters as Data Center B.
  - DTS is available to synchronize data between RDS instances to ensure data consistency.
- Data backup and recovery:
  - DTS must be stable and run properly to ensure real-time data synchronization. If a fault occurs, you must ensure the stability and accuracy of DTS data synchronization.

- If a fault occurs on Data Center A or the databases in Data Center A, traffic is directed to Data Center B. The resources of Data Center A remain unavailable if the fault persists. After Data Center A is recovered, you can specify it as a secondary zone.
- If a fault occurs on the applications, traffic is directed to Data Center B. A check must be performed to ensure the consistency of data segments between data centers A and B. After the check is completed, the databases in Data Center B becomes primary databases, and the databases in Data Center A becomes secondary databases. In this scenario, data is synchronized from Data Center B to Data Center A.
- You must configure DTS transmission links to synchronize data between RDS instances.

- Features:
  - Advantages: ultra-high performance, automatic switchover, and fewer manual operations.
  - Disadvantages: 50% resource utilization rate.

Active/active mode

An example is used to introduce the architecture of the active/active mode. This example assumes that you have purchased two simplified IT systems in an Alibaba Cloud region.



Architecture:

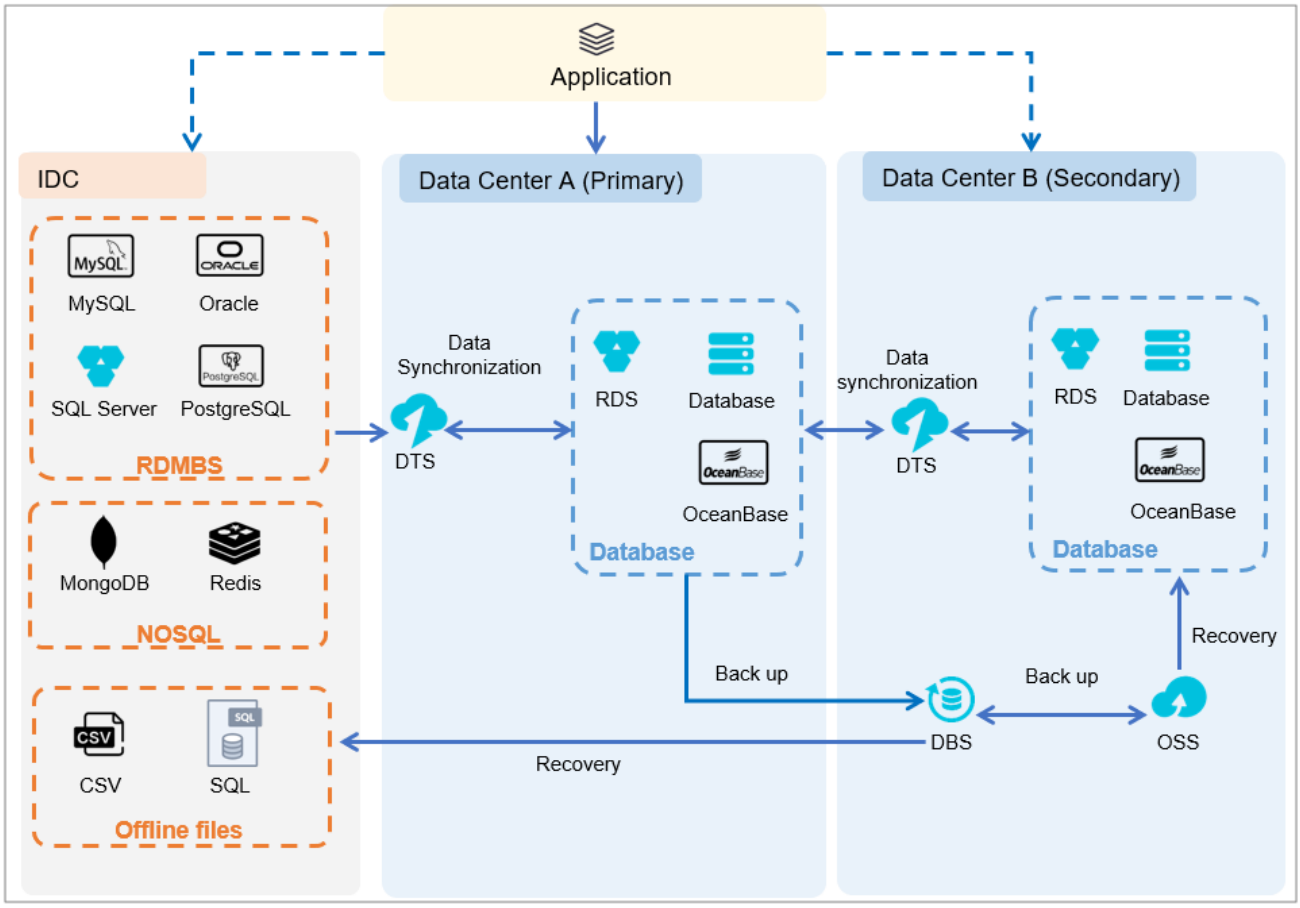
- Key conditions:
  - Two simplified IT systems are deployed in two zones (data centers A and B) of the same Alibaba Cloud region. Both data centers are primary zones.
  - You can deploy Alibaba Cloud DNS, SLB, and ECS to meet your domain management, traffic distribution, and cloud computing needs, respectively. You can deploy Alibaba Cloud RDS to meet your database needs.

  - DTS is available to synchronize incremental data in real time and migrate data between databases.

- Data backup and recovery:
  - If Data Center A works properly, you can use DTS to synchronize production data from the databases in Data Center A to the databases in Data Center B. You can also use DTS to check the consistency of data between the databases.
  - If a fault occurs on Data Center A, Alibaba Cloud DNS resolves traffic to Data Center B, and the production data is stored in the databases in Data Center B.
- Features:
  - Advantages: achieves a higher resource utilization rate than the active/standby mode does.
  - Disadvantages: cross-zone latency and manual switchover. If a fault occurs on Data Center A, a latency occurs when users access applications from Data Center B.

5.4. Hybrid cloud backup and recovery solution

Recommended architecture



Architecture description:

- This solution is based on a hybrid cloud consisting of an on-premises data center, a primary Alibaba Cloud zone, and a secondary Alibaba Cloud zone. The on-premises data center and cloud zones communicate through a leased line or VPN.
- "Stateless" applications are deployed in the on-premises data center and two zones for real-time synchronization and backup between on-premises and off-premises databases.
- If a failure occurs in any of the three databases (the on-premises data center or either of the two zones), traffic can be switched to one of the other two to ensure service continuity. After the fault is resolved, the service traffic can be switched back to the preferred service center.

6.Hybrid cloud-based multi-active solution

6.1. Scenarios

Faced with rapid business development, enterprises tend to build their disaster recovery architectures while considering service load balancing, O&M, expenditure, security, and other factors. The hybrid cloud-based multi-active disaster recovery solution provides a multi-active disaster recovery architecture based on a hybrid cloud consisting of an on-premises data center (IDC) and a cloud platform. This solution meets enterprises' requirements for service continuity.

Why use a hybrid cloud?

Based on a hybrid cloud architecture, the hybrid cloud-based multi-active solution can relieve enterprises of concerns about service scalability, O&M, costs, and security when they build their disaster recovery systems.

- Service scalability:

  Enterprises must quickly respond to changing service demands to stay in the game. Public cloud platforms support auto scaling and are able to respond to frequent business activities. For example, during the Double 11 Shopping Festival and other activities that produce traffic spikes, public cloud platforms should be able to tolerate these traffic peaks. By integrating the capabilities of traditional IDCs and cloud platforms, the hybrid cloud solution can meet the needs of enterprises for service expansion on their system architectures.

- O&M:

To resize an application in a traditional IDC, you have to apply to purchase a server, install the server and operating system, and deploy the application. The whole process is complex and time-consuming. Additionally, a series of problems may occur during resizing. For example, system faults may occur due to different server environments. To resize an application on a hybrid cloud, you only need to focus on O&M for the cloud-based application and do not need to consider O&M for the cloud platform infrastructure. This greatly reduces the difficulty of resizing and the complexity of O&M.

- Cost control:

  In a hybrid cloud solution, cloud resources are used to reduce the volume of idle resources present in traditional disaster recovery solutions, as well as the overall hardware and software O&M costs. You can wholly migrate your system to a public cloud without any additional investment, reducing the costs of system transformation and the migration time.

- Security control:

  In terms of security, the hybrid cloud solution allows you to migrate computing and cache nodes used for Internet access and front-end applications to Alibaba Cloud. While using Alibaba Cloud's mature and comprehensive security protection solutions, you can store core group data in your on-premises IDC to ensure its security.

Choose a local active-active or remote multi-active solution as needed

In the hybrid cloud-based multi-active solution, you can deploy two or more active IDCs in regions or zones in the same city or different cities based on your location and disaster recovery

policies. In the hybrid cloud-based remote active-active solution, only one remote cloud node is deployed. In the hybrid cloud-based remote multi-active solution, the services are provided in multiple regions to support disaster recovery. Therefore, users all over the country can access the service through the nearest region. When the service is interrupted in a region, the service traffic is immediately switched to unaffected regions.

Compared to the traditional local disaster recovery or active-active architecture, the hybrid cloud-based multi-active architecture solves the following problems:

- It is difficult to determine whether the traffic is switched to the disaster recovery center.
- When the disaster recovery center does not provide services, its resources are idle, resulting in high costs.
- The IDC that provides services is still deployed in a single region. When the service volume reaches a certain level, performance may slow due to the limited resources of a single region.
- In a traditional architecture, the two active IDCs are close to each other (usually within 50 km). Natural disasters or widespread power or network failures may result in both IDCs becoming unavailable.

The hybrid cloud-based multi-active architecture not only provides the functions of traditional remote disaster recovery solutions, but also improves the overall resource utilization and meets the remote deployment needs caused by business growth. Based on its years of experience with multi-active architecture during the Double 11 Shopping Festival, Alibaba Cloud provides you with a stable and cost-effective hybrid cloud-based multi-active disaster recovery solution.

6.2. Architecture

The following architecture is recommended for a multi-active disaster recovery solution on a hybrid cloud.

Key points of the architecture

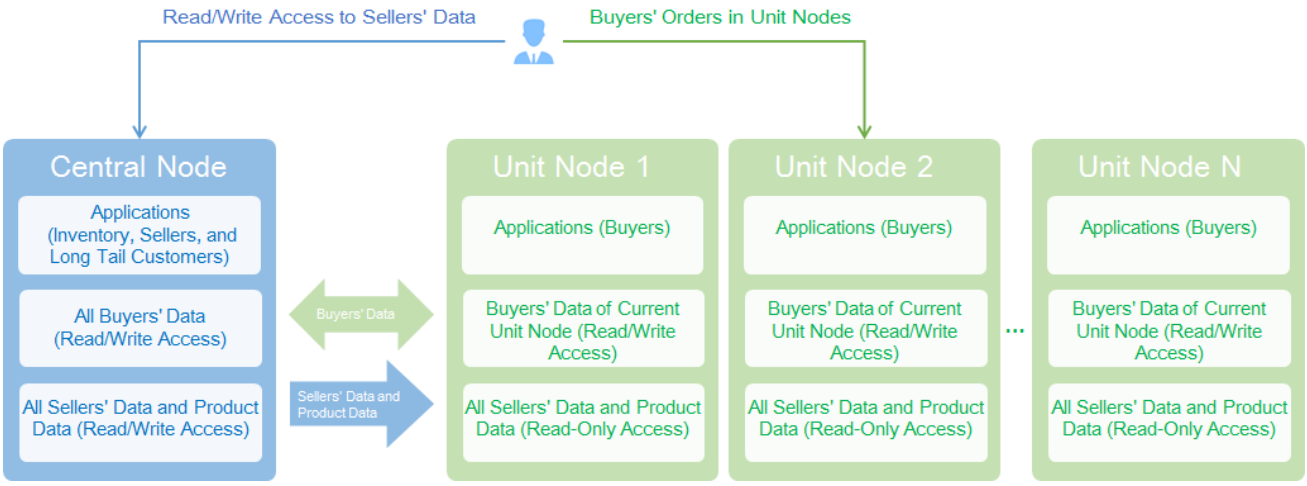
In a hybrid cloud-based multi-active architecture, remote redundant resources are used to ensure that services run properly even in extreme circumstances. When building an architecture running on a hybrid cloud that integrates your on-premises IDC and off-premises IDC, pay attention to the following points:

- **Unit node isolation**:

  Before deploying an active geo-redundancy solution, you must solve the latency issues that are often associated with geographic distribution. Latency issues may result in data inconsistency and inaccuracy if users submit requests to modify the same row of records in the databases of unit nodes in different regions. Additionally, a long latency may occur if one operation requires multiple requests for data in a unit node and the node has to call a chain of services with interdependencies. Therefore, you have to spend lots of time handling latency issues. Unit node isolation is at the core of an active geo-redundancy solution. To be more specific, each unit node has independent read/write access permissions, and multiple unit nodes cannot modify the same row of records in different regions. To isolate unit nodes, you must divide them into different categories based on a certain dimension.

- **Properly categorizing data into unit nodes**:

You need to analyze your business before determining how to manage read and write access permissions of each unit node. For example, placing orders is the most important process for e-commerce business. To reduce restructuring charges and improve user experience, the optimal choice is to categorize data into unit nodes based on the user ID.

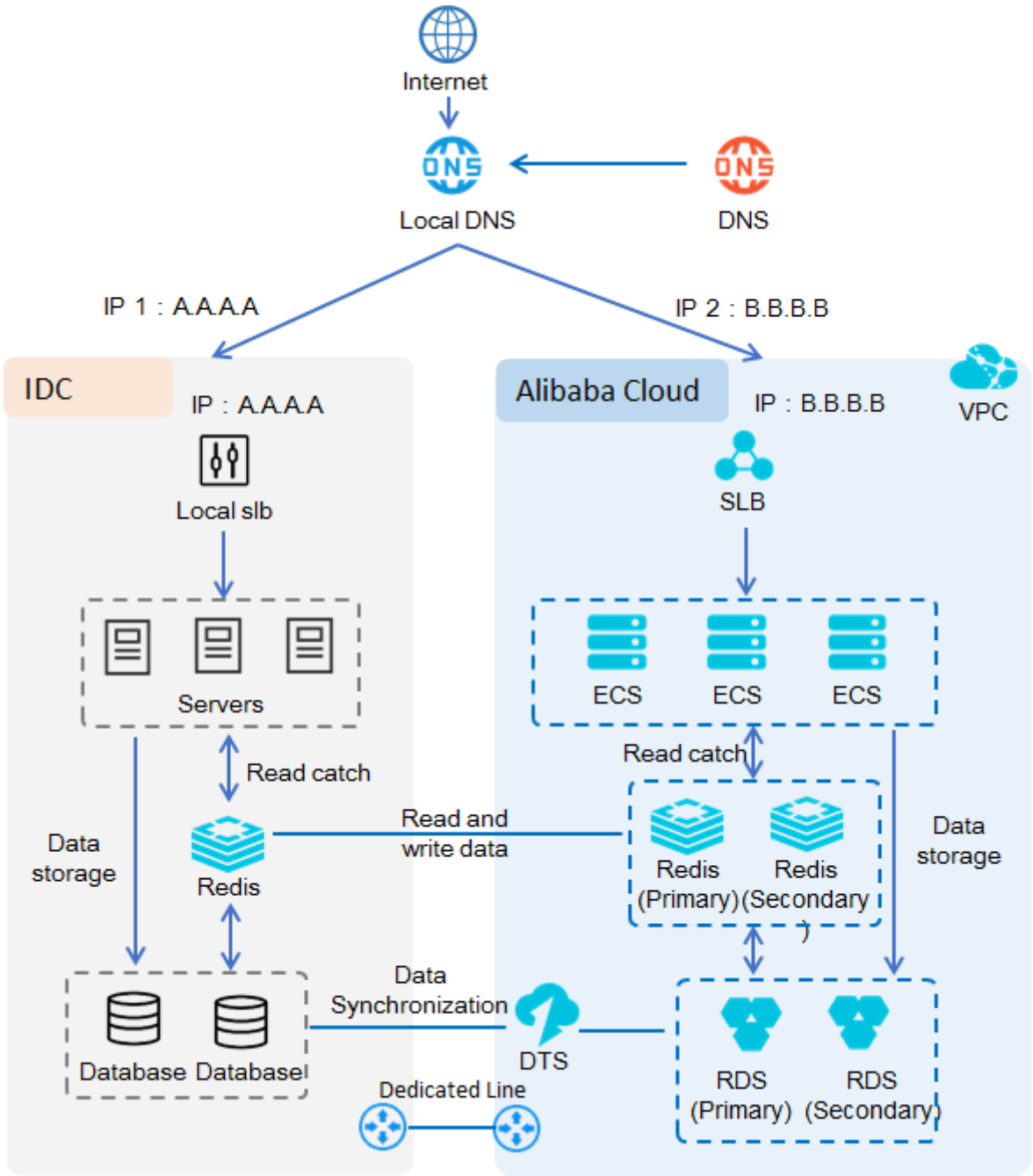


In this case, you can perform read/write operations on buyers' orders in the corresponding unit node, and data is not read or written across unit nodes. However, other data that is not related to buyers' information may be distributed across unit nodes. For example, sellers' operations of modifying product data may involve multiple unit nodes. If necessary, you can use read/write split to ensure the eventual consistency of buyers' and sellers' data. If eventual consistency cannot meet your needs, you must ensure that data can be read or written across unit nodes.

To provide the optimal service experience, the hybrid cloud-based multi-active architecture must be further optimized based on the service scenario and application implementation method. For assistance with architecture design, you can contact our Professional Services. By taking a simple IT system as an example, the following section describes how to build a hybrid

cloud-based active-active architecture in which instances are deployed in more than two Alibaba Cloud zones.

Recommended architecture



Architecture description:

- We recommend that you activate Express Connect, build a hybrid cloud through a leased line, and deploy your application system in your on-premises IDC and off-premises IDC in exactly the same manner. In this way, you can deploy a hybrid cloud-based active-active solution that consists of on-premises and off-premises IDCs.
- Both IDCs provide services to achieve load balancing.
    - **Access side**:

      Use intelligent DNS to distribute traffic to both IDCs and make your application stateless. Deploy your application system in both IDCs in exactly the same manner to overcome the carrier's regional restrictions and split traffic by region.
    - **Application deployment**:

      Deploy your application system in the on-premises and off-premises IDCs in exactly the same manner. In each IDC, mount the application cluster to the SLB instance in the IDC, which distributes the traffic to a node in the application cluster.
    - **Cache side**:

      We recommend that you use ApsaraDB for Redis to read the application cache. Alibaba Cloud ApsaraDB for Redis is compatible with the open source Redis protocol. When instances deployed in on-premises and off-premises IDCs are both ApsaraDB for Redis instances, two-way read-write synchronization is supported. The Conflict-free Replicated Data Type mechanism is used to detect and remove data conflicts and ensure data consistency. When open source Redis is used in the on-premises IDC, the ApsaraDB for

Redis instance can receive the one-way read-write synchronization information from the Redis instance in the on-premises IDC.

Note When open source Redis is used in the on-premises IDC, the on-premises Redis instance may be incompatible with the enhanced cache processing capability of the ApsaraDB for Redis instance. The ApsaraDB for Redis instance can read data from, but cannot write data to, the on-premises Redis instance. We recommend that you deploy an ApsaraDB for Redis instance in your on-premises IDC.

- **Data side**:

  Application data is stored in off-premises and on-premises databases, and data is synchronized between the databases through DTS to ensure mutual data consistency.

Advantages of the architecture

- **Multiple IDCs**: Alibaba Cloud deploys multiple IDCs around the world. You can purchase Alibaba Cloud products and deploy them in the nearest or most appropriate region.
- **Stability**: Each region and each product are stable. After multiple rounds of iteration, SLB, ECS, ApsaraDB for Redis, ApsaraDB for RDS, and other key Alibaba Cloud products now provide excellent disaster recovery capabilities. Fine-grained disaster recovery control can be achieved through additional functional product modules.
- **Scalability**: You can scale your existing services out or in, or up or down, or purchase additional services based on your needs.

7.Operation example: Cross-zone high availability solution on a public cloud

For an enterprise, whether or not their services are on the cloud, service stability and continuity have always been crucial. To reduce the impact of uncontrollable factors on normal service operations, you must improve the availability and disaster recovery capabilities of your products. Although your products may already be highly available, you cannot ignore the important task of improving service availability and disaster recovery capabilities.

To improve service availability and disaster recovery capabilities, many users take advantage of these cloud products: Elastic Compute Service (ECS), Server Load Balancer (SLB), ApsaraDB for RDS, and Object Storage Service (OSS).

Zone

Zones are physical areas in the same region that have independent power grids and networks. The network latency is lower for ECS instances in the same zone.

Intranet communication is available across different zones in the same region, and fault isolation is supported between zones. The choice to deploy ECS instances in the same zone is a tradeoff that depends on factors such as network performance and disaster recovery requirements.

- If your applications require high disaster recovery capabilities, we recommend that you deploy your ECS instances in different zones of the same region.

- If your applications require low network latency between instances, we recommend that you deploy your ECS instances in the same zone.

In the Region List, you can view the number of zones in each region. Alternatively, you can use the Region List API in OpenAPI Explorer to view the list of all zones.

Product introduction

ECS

ECS is a basic cloud computing service provided by Alibaba Cloud. An ECS instance is a virtual computing environment that incorporates a CPU, memory, operating system, disks, bandwidth, and other basic server components. It is the operating entity presented to each user.

You can create ECS instances at any time according to your business needs, without having to purchase hardware in advance. As your service grows, you can resize the disks and increase the bandwidth of your ECS instances. When you no longer need an ECS instance, you can release it to reduce costs.

ECS instances themselves do not have high availability and disaster recovery capabilities. Instead, these capabilities are implemented through architecture construction.

SLB

SLB is a traffic distribution control service that distributes traffic to multiple backend ECS instances based on the routing algorithms. SLB extends application service capabilities and enhances application availability.

SLB sets a virtual service address to virtualize ECS instances into an application service pool with high performance and high availability. Then, it distributes requests from clients to ECS instances in the ECS instance pool based on the routing algorithms.

The following features allow SLB to improve the availability and disaster recovery capabilities of ECS instances:

- SLB is deployed in clusters. Each cluster has a certain number of backend ECS instances to eliminate single point of failure (SPOF). This means that SLB is not affected if one or several backend ECS instances fail.

  The Layer-4 SLB (LVS) service, Layer-7 SLB (Tengine) service, control system, and other key components in the SLB system are all deployed in clusters to improve their scalability and availability.

- Currently, most SLB instances are multi-zone instances, with primary and secondary instances located in the IDCs of different zones in the same city. When the IDC in which the primary instance is located experiences faults, services can quickly fail over to a secondary instance, supporting disaster recovery and the high availability of services. Click here for more information on the distribution of multiple zones in each region.

ApsaraDB for RDS

ApsaraDB for RDS is a stable, reliable, and scalable online database service. Based on the distributed file system and high-performance storage of Alibaba Cloud, ApsaraDB for RDS

supports MySQL, SQL Server, PostgreSQL, and PPAS (Postgres Plus Advanced Server, a database highly compatible with Oracle) engines. It provides a complete set of solutions for disaster recovery, backup, monitoring, migration, and other functions, allowing you to focus on services rather than database O&M.

- For more information about the basic edition of ApsaraDB for RDS, click here.
- In the dual-host high-availability version of ApsaraDB for RDS, primary and secondary instances can be deployed in the same zone. When the primary instance experiences a fault, it fails over to a secondary instance, providing high availability and disaster recovery capabilities.
- In multi-zone ApsaraDB for RDS, primary and secondary instances are deployed in different zones.
- You can use Data Transmission Service (DTS) to synchronize and migrate data between ApsaraDB for RDS instances.

OSS

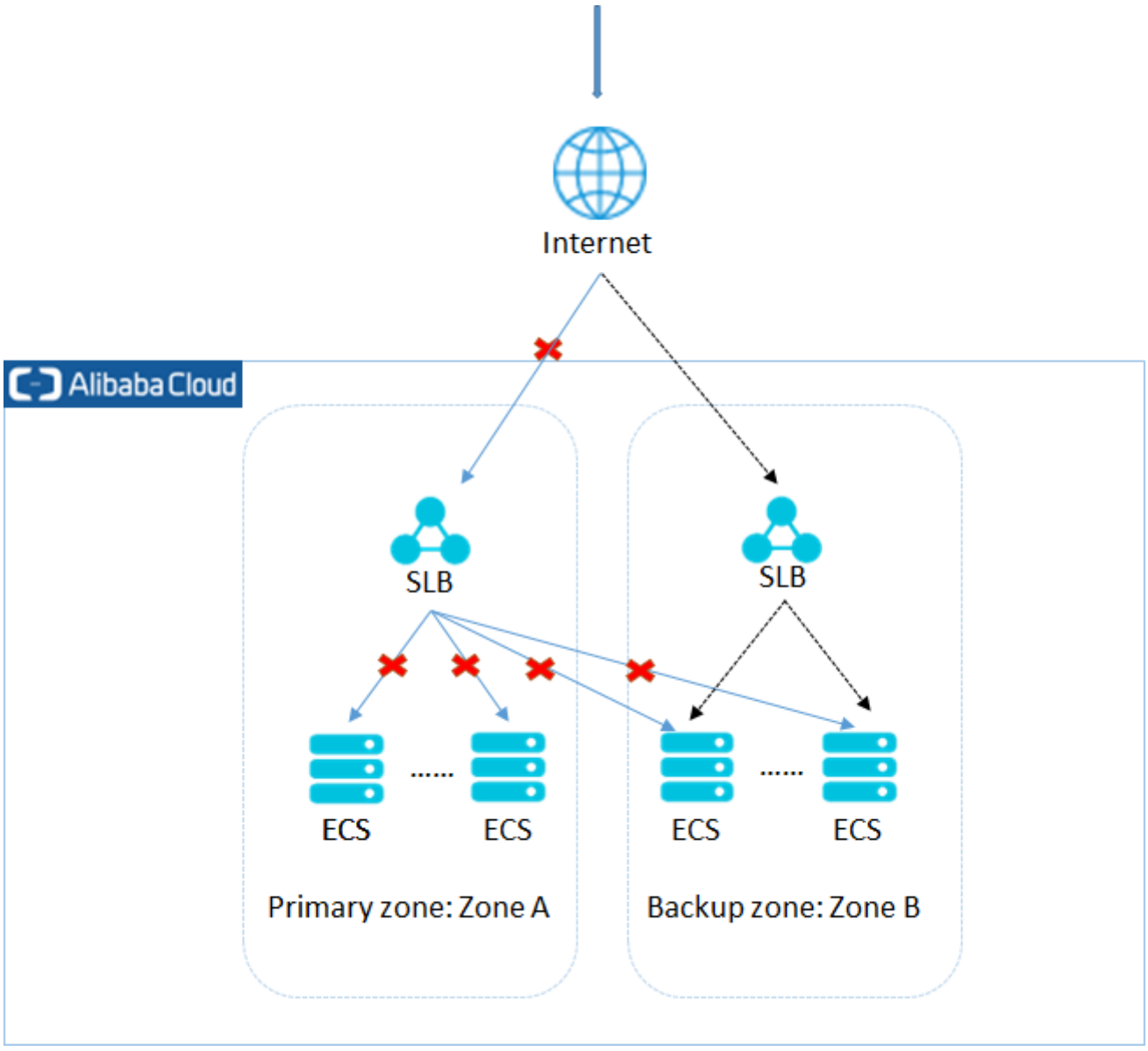
OSS is a massive, secure, cost-effective, and highly reliable cloud storage service provided by Alibaba Cloud. You can upload and download data for any application, anytime, anywhere by calling APIs. In addition, you can perform simple data management operations in the web console. OSS can store any type of file and is therefore suitable for various websites, development enterprises, and developers. Your OSS instance is only billed for the capacity that you actually use, allowing you to focus on your core services.

Files are chunked for storage. By default, three replicas of each chunk are saved on chunkserver nodes in different racks. In the Apsara Distributed File System cluster, up to one master and two chunkserver nodes can fail without affecting services, while multiple KVServers and WS nodes can fail.

The following describes the architecture and construction process for services with high availability and disaster recovery capabilities in detail.

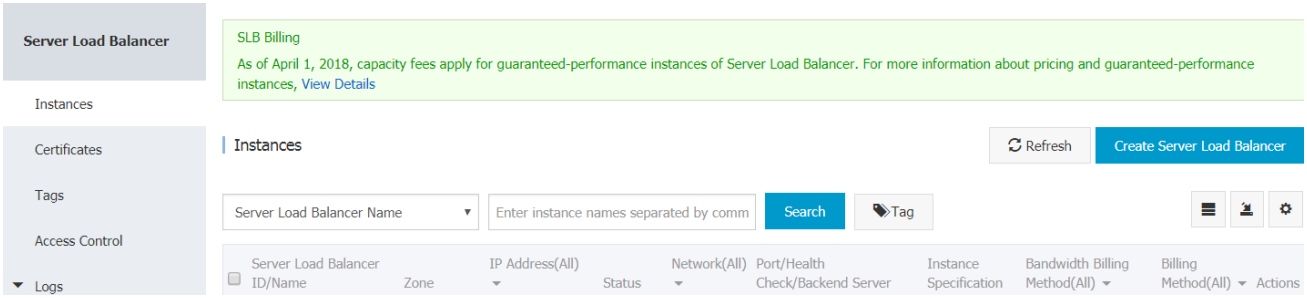
Multi-zone SLB instances + ECS instances in different zones

In the following figure, ECS instances are bound to different zones under an SLB instance. This way, when Zone A works normally, user access traffic follows the path of the blue solid line shown in the figure. When a fault occurs in Zone A, user access traffic is distributed to the path of the black dotted line. This prevents a fault in a single zone from causing service unavailability, and reduces latency by selecting zones between different products.

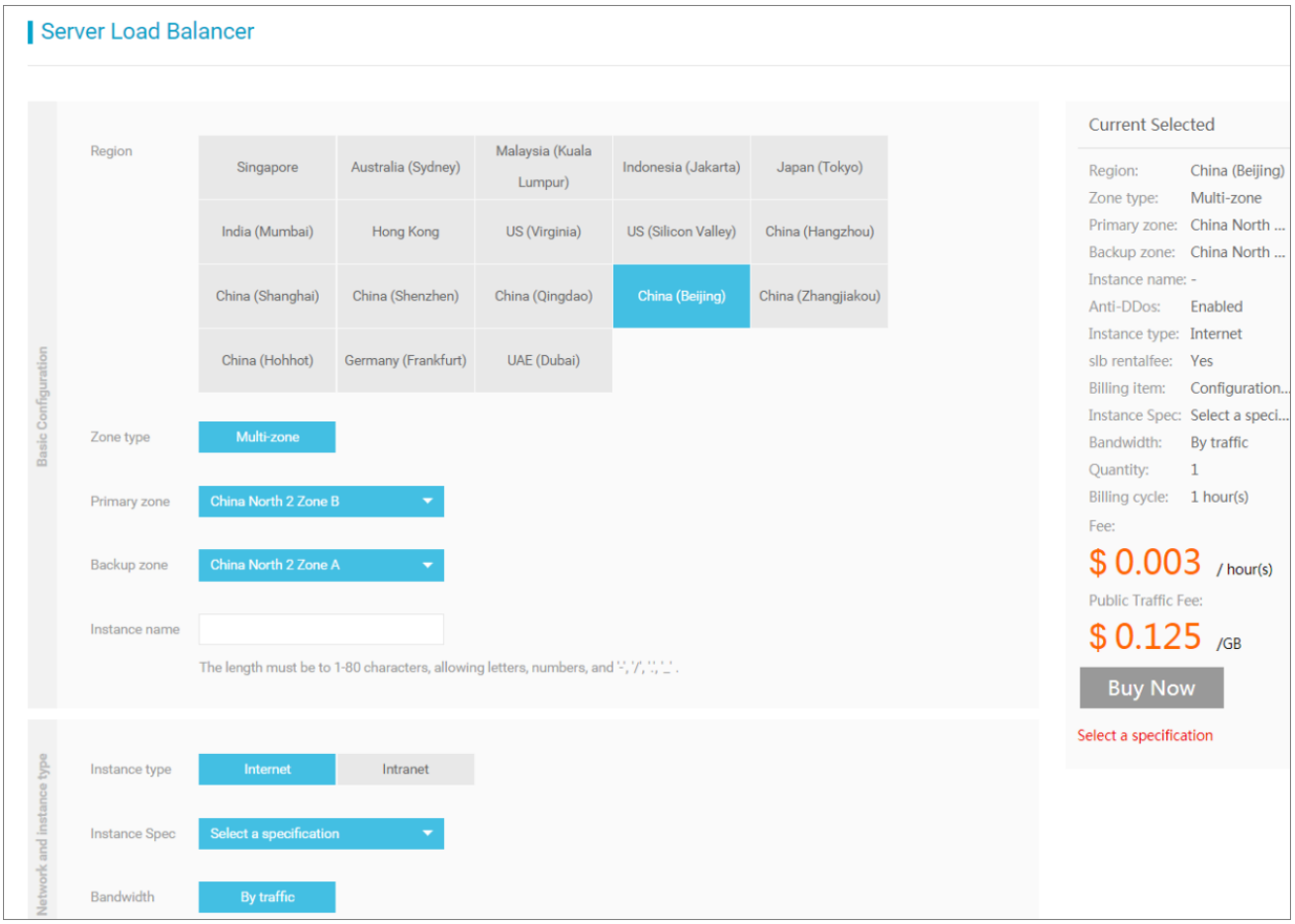


Perform the following steps to construct this architecture:

1. Log on to the Alibaba Cloud console and click **Server Load Balancer**. On the page that appears, click **Create Server Load Balancer**.

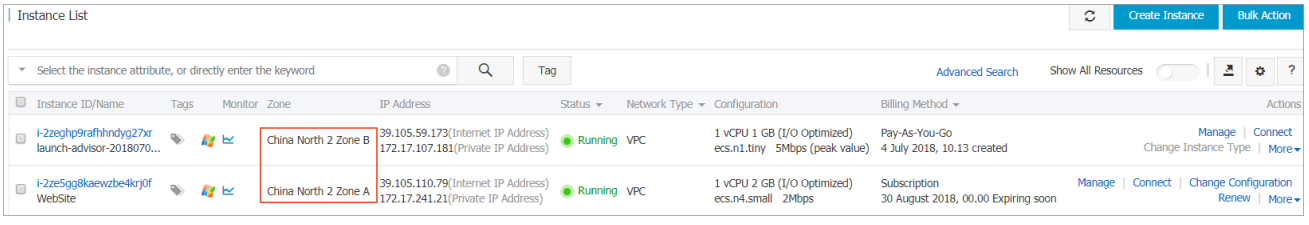


Here, we use the China (Beijing) region as an example and purchase a multi-zone instance, with primary zone B and secondary zone A.



2. Create ECS instances in both the primary and secondary zones of the SLB instance.

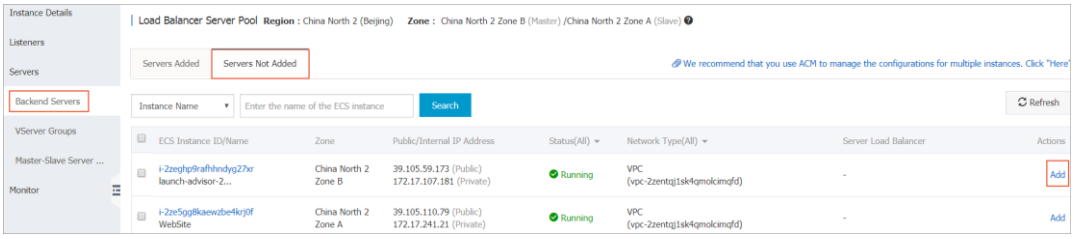
   Create a test instance in zone A and zone B of the China (Beijing) region. In this example, we use the default security group and VPC network with a 1-core 2-GB memory CentOS 7.2 configuration.



3. Create listeners and add backend servers (ECS instances).

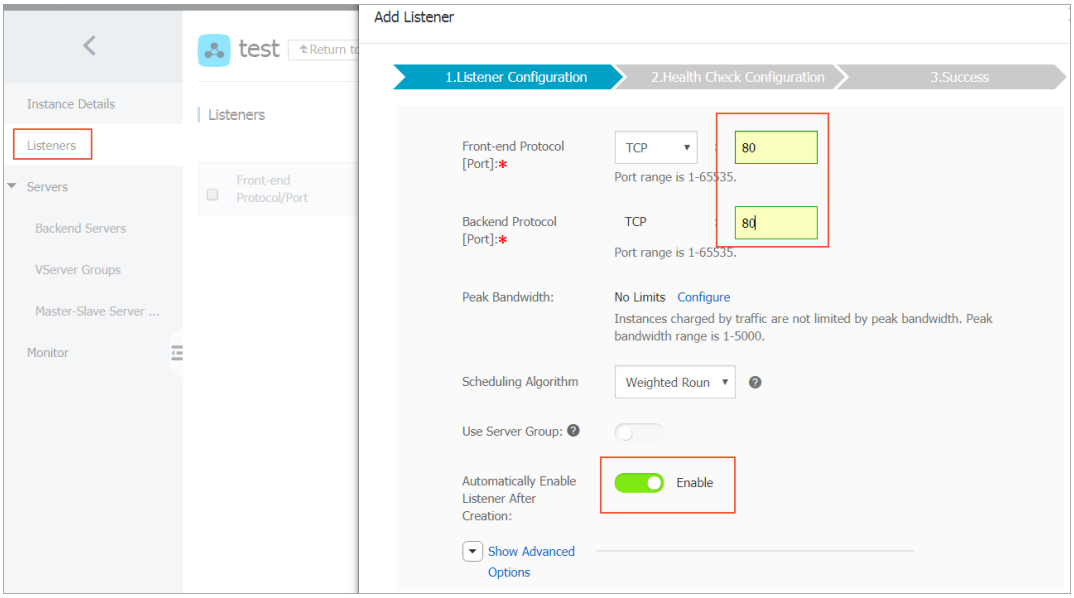
i. In the SLB console, locate the instance you created, and click **Manage**.

ii. Click **Backend Server** and select Excluded Servers. Then, find your instance and click **Add**.



iii. After completing the process, you can view your ECS instances and their weights on the Included Servers page.

iv. Click the Listener tab on the left. On the tab page that appears, click **Add Listener**. Set listener attributes as needed. In this example, we use the Layer-4 TCP mode, set the listener port to port 80, set the backend forwarding port to port 80, and use the default weighted round robin method. We also enable session persistence and use the default 1,000-second time-out period.



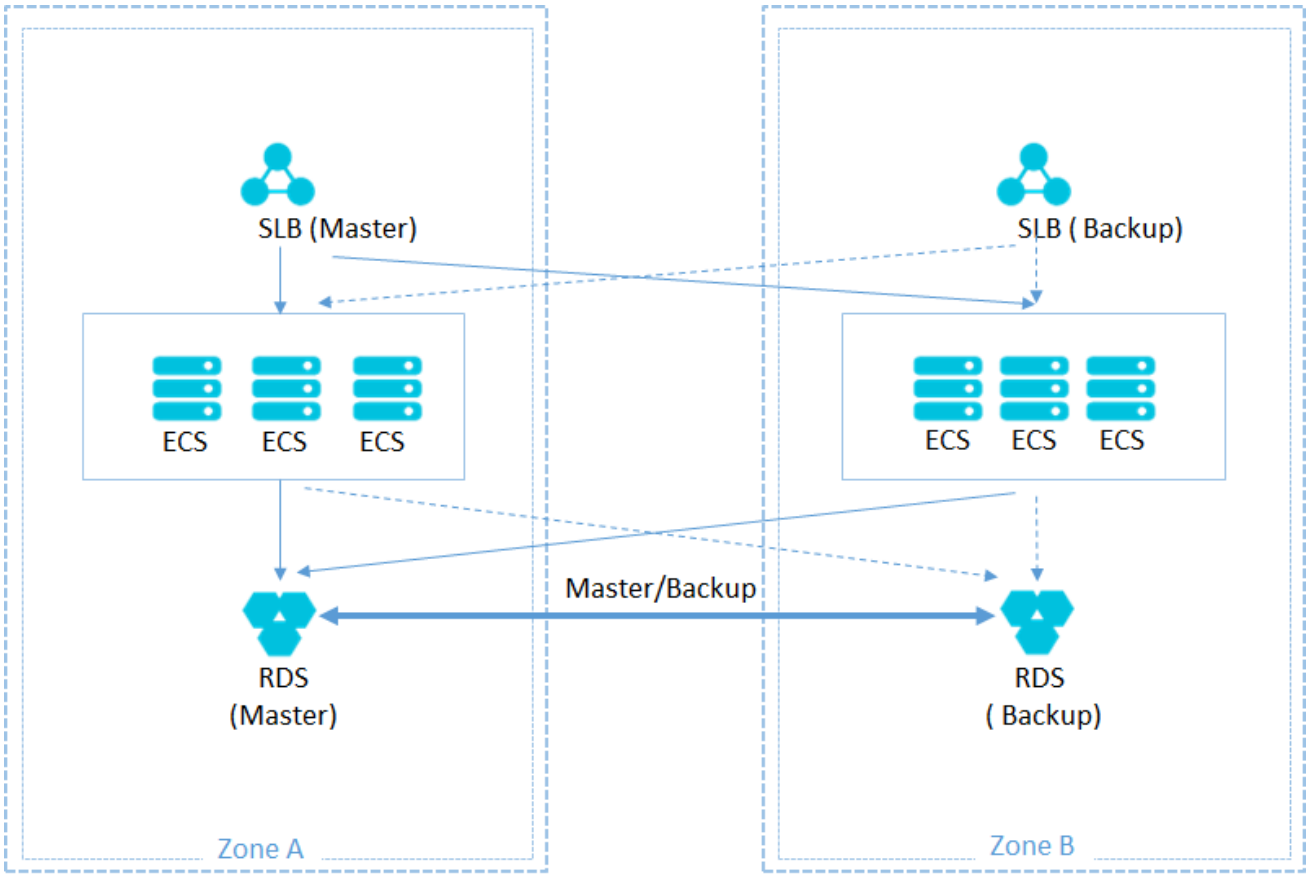
v. Set the health check mode to TCP and the backend check port to 80.

vi. After completing these steps, you can view the added listener and its status on the Listener tab page.

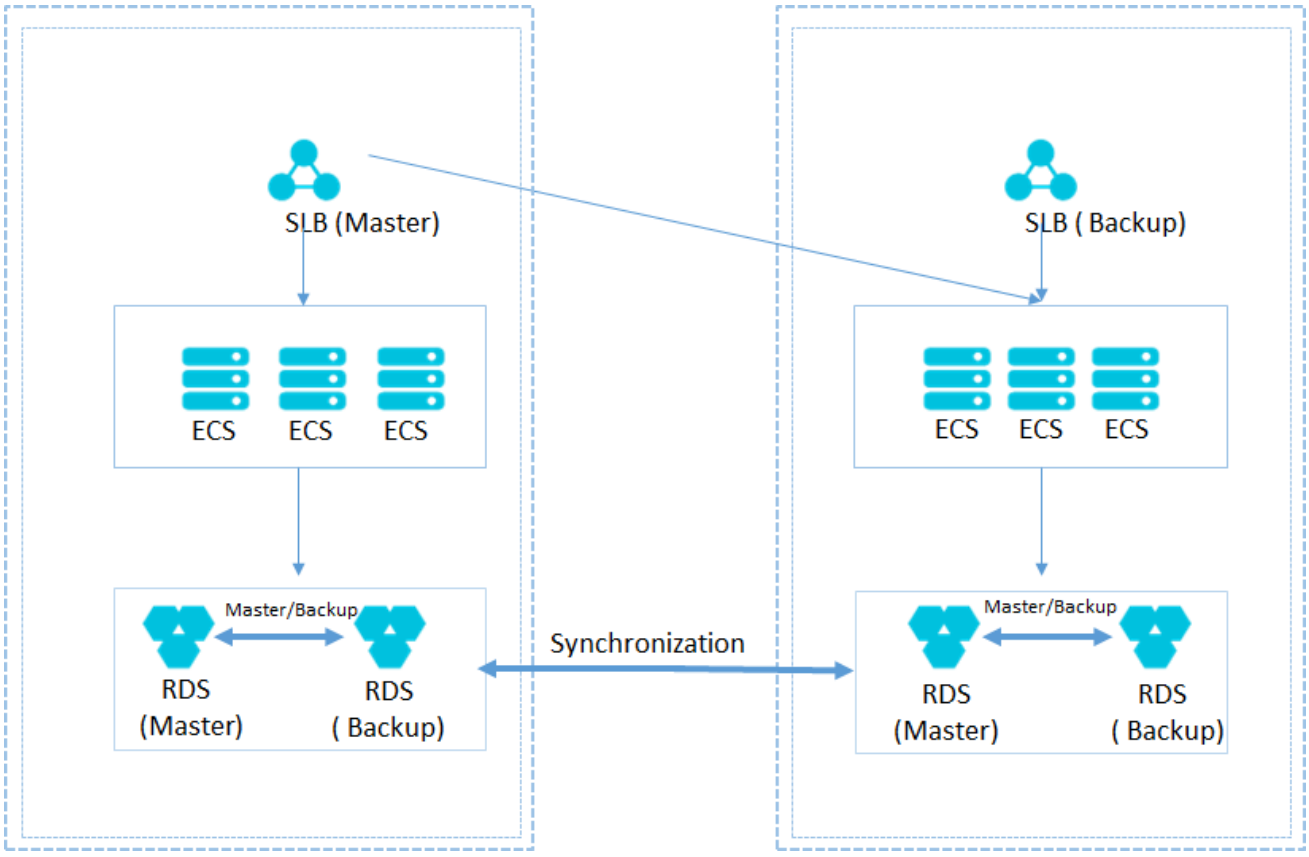
Note You only need to deploy the relevant service on the ECS instances and listen to port 80. Then, resolve the domain name to the public IP address of the SLB instance, so the SLB instance can forward requests to backend ECS instances and provide service.

Multi-zone SLB instances + ECS instances in different zones + highly available ApsaraDB for RDS instances

The following figure shows the multi-zone ApsaraDB for RDS architecture.

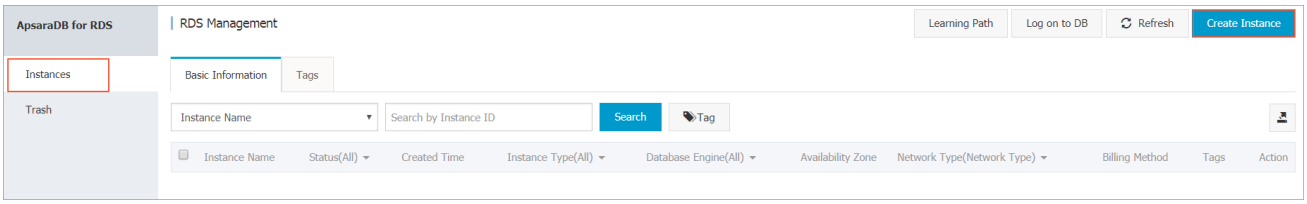


In regions where multi-zone ApsaraDB for RDS is not supported, you can create an ApsaraDB for RDS instance in each zone, with the secondary zone used as the backup database. This database is synchronized with the ApsaraDB for RDS instance in the primary zone.



Perform the following steps to construct the multi-zone RDS architecture:

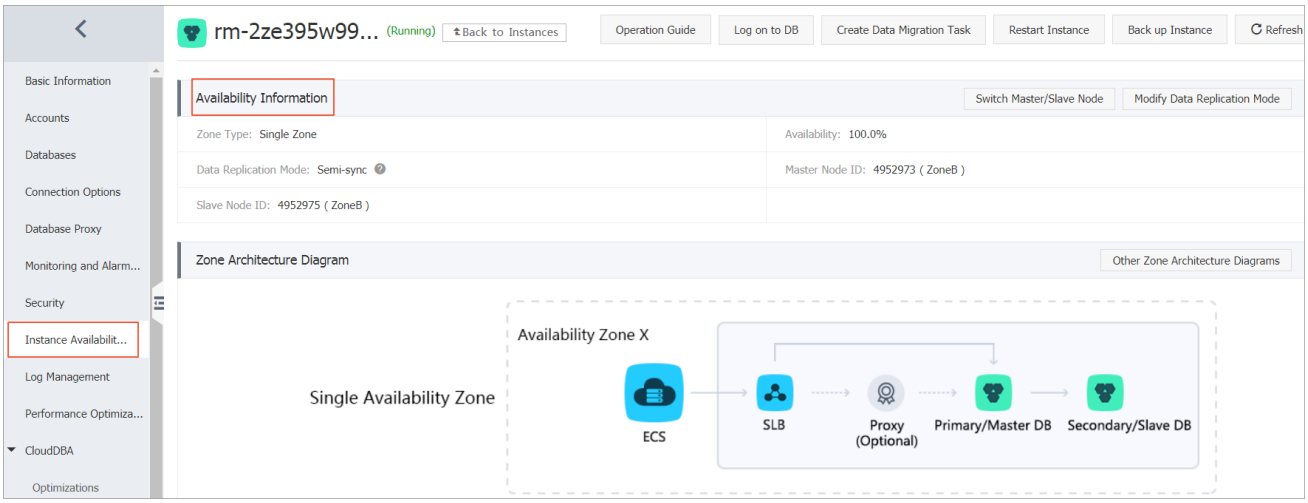
1. After deploying a multi-zone SLB instance and multiple ECS instances in different zones, purchase ApsaraDB for RDS instances.



2. Select a region that supports multi-zone ApsaraDB for RDS, as shown in the following figure.

3. After purchasing ApsaraDB for RDS instances, you can view them in the console.

   In addition, you can view the high availability information of ApsaraDB for RDS instances and switch between primary and secondary instances in the console, as shown in the following figure.



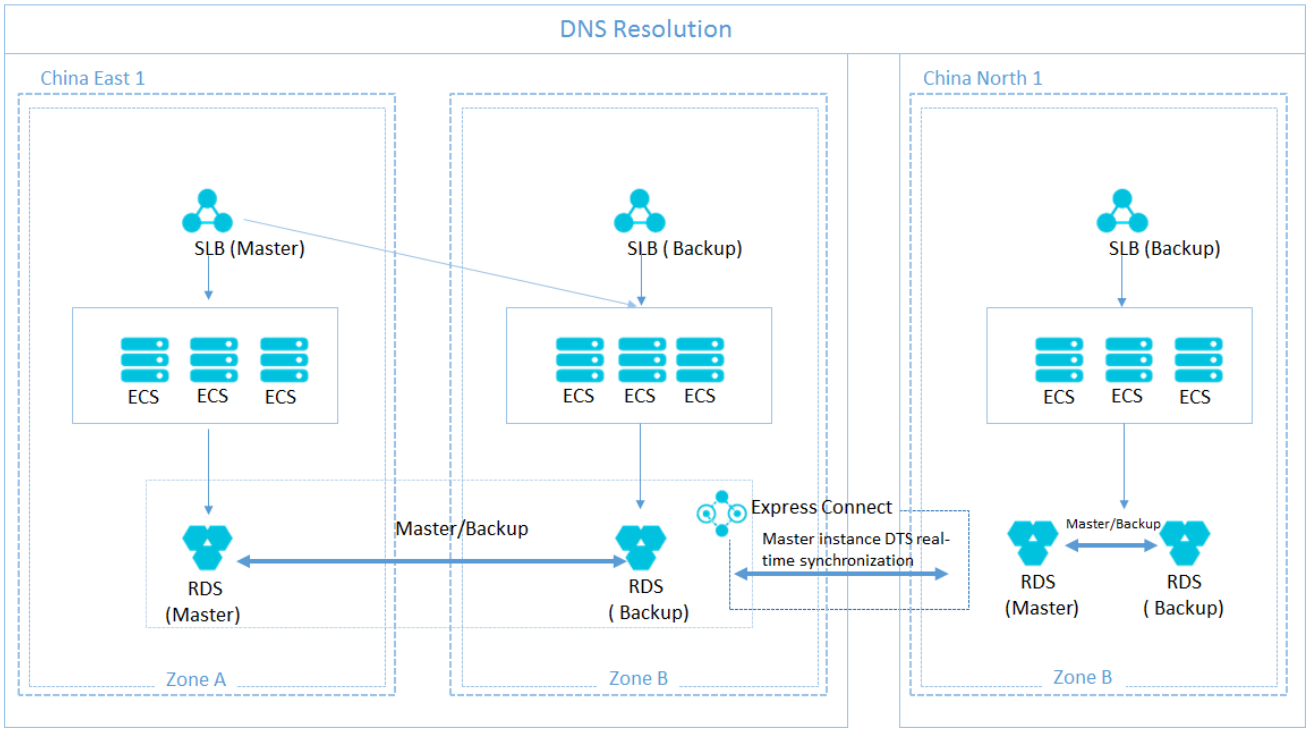
The following describes the example in which an ApsaraDB for RDS instance is deployed in each zone.

1. Purchase dual-host highly available ApsaraDB for RDS instances in Zone A and Zone B, respectively.

2. Create a DTS synchronization task.

High availability - remote disaster recovery

When multiple zones are available in the same city and an environment is deployed in a remote region as well, the resulting architecture greatly increases the service availability and achieves remote disaster recovery.

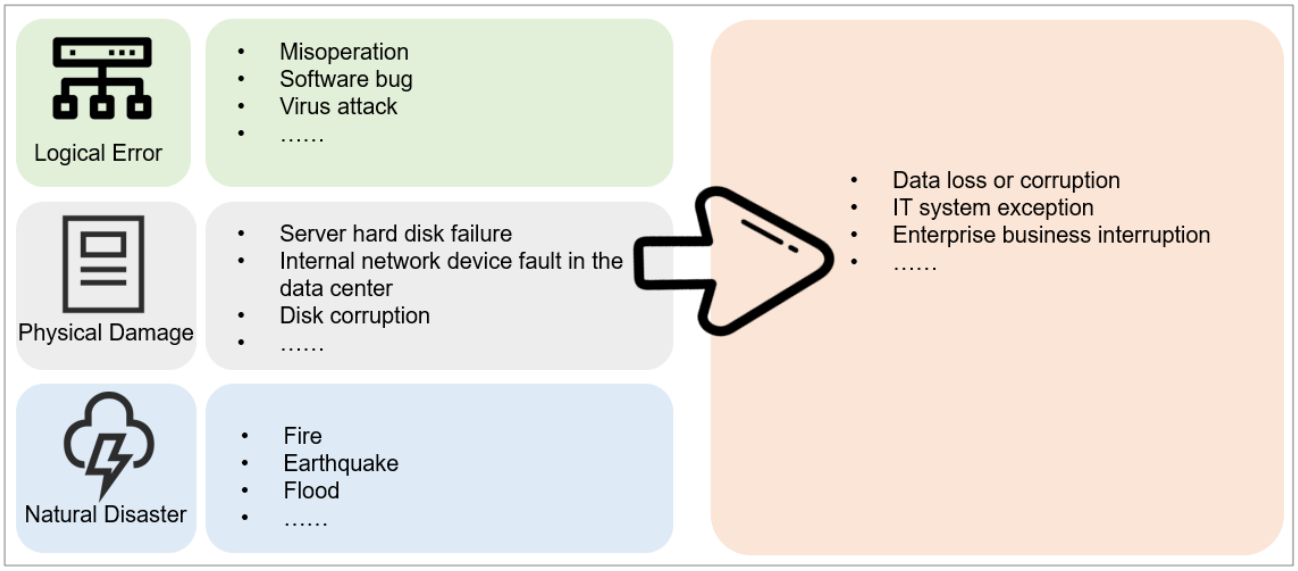


Note Configure the DNS resolution to specify the ultimate service access region and use DTS for data synchronization between ApsaraDB for RDS instances.

8.Appendix: Trends and basic concepts in disaster recovery

8.1. Industry trends and challenges

User data and system data are the core and most important assets of enterprises in all industries. Stable service operation and normal IT system functionality are the most important development demands of enterprises. However, these demands often cannot be satisfied due to unpredictable natural or man-made disasters, including:



Therefore, it is very important to guarantee the service stability, proper operation of IT system functions, and data security for enterprises. In this context, solutions that support both system and application data backup and disaster recovery are emerging and developing rapidly.

Note A disaster recovery solution implements both disaster recovery and backup.

- Backup refers to the creation of one or more replicas of the important data generated by application systems or critical original data.

- Disaster recovery refers to the deployment of two or more IT systems with the same functions in two separate locations in the same city or in different cities. These systems mutually monitor each other's health and support failover and failback. In the event that a system stops working due to an accident (a natural or man-made disaster), all application system services fail over to another system so that they can be provided without interruption.

Shortcomings of traditional disaster recovery solutions

Traditionally, enterprises built their own disaster recovery centers based on their individual needs. However, such solutions are resource-intensive and face many challenges.

Traditional disaster recovery solution	Feature	Challenge
Use of application-level replication software	Support for application-level disaster recovery	• A remote IDC for disaster recovery must be built and maintained at a high cost. • Relevant software and hardware must be deployed, which are hard to maintain. • A large number of operations and tests are required for failback.

Use of Continuous Data Protection (CDP) technology	• Support for application-level disaster recovery • Satisfactory RPO and RTO	• A remote IDC for disaster recovery must be built and maintained at a high cost. • Expensive CDP devices are required. • Relevant software and hardware must be deployed, which are hard to maintain. • Data cached in applications cannot be replicated. • A large number of operations and tests are required for failback.
Use of storage devices to replicate data. Support for application-level disaster recovery	• Effective support for application-level disaster recovery • Guaranteed RPO and RTO	• A remote IDC for disaster recovery must be built and maintained at a high cost. • Expensive active-active storage devices and supporting network devices are required.

		• Relevant software and hardware must be deployed, which are hard to maintain. • The application awareness capability is limited and many scripts must be executed manually.

However, as their data scales increase rapidly and data value grows exponentially, enterprises have higher requirements for service continuity. Considering their huge investments, long construction times, and high maintenance costs, traditional disaster recovery solutions cannot meet the future development needs of IT systems.

Advantages of cloud disaster recovery solutions

In recent years, the steady and rapid development of cloud computing has given rise to cloud disaster recovery solutions. Cloud disaster recovery solutions boast lower costs and faster system recovery. The following table compares traditional disaster recovery solutions with cloud disaster recovery solutions.

Consideration	Traditional disaster recovery solution	Cloud disaster recovery solution

Implementation method	The solutions are mostly based on physical devices. Relevant physical resources must be deployed at multiple sites. The disaster recovery sites and their sizes vary greatly with the system reliability requirements.	The disaster recovery sites are deployed on a hybrid cloud or a public cloud.
Construction duration	It takes several months to deploy the solutions.	It takes several days to deploy the solutions.
Investment	A large number of servers, storage devices, and physical network devices are required. The investment is huge.	The initial investment is small. The solutions support auto scaling as actual service demands increase in the future.
O&M costs	The devices must be maintained by a large number of professionals. The O&M costs are high.	Maintenance by O&M professionals is not required.

Cloud disaster recovery solutions are more cost-effective, efficient, and scalable. They will be the mainstream disaster recovery solutions of the future.

8.2. Basic disaster recovery concepts

Disaster recovery solutions ensure the high availability of enterprise data and services by combining disaster recovery with data backup. Disaster recovery protection levels are classified by risk and RTO/RPO.

Key technical indicators

Take the following key technical indicators for disaster recovery into account when designing your disaster recovery solution:

- **Recovery time objective (RTO)**: The period of time within which IT systems and services must be restored after an outage. RTO indicates the timeliness of service recovery, that is, the maximum recovery time for IT systems that enterprises can tolerate. A smaller RTO indicates a higher disaster recovery capability, but requires a higher enterprise investment.
- **Recovery point objective (RPO)**: The point in time to which data is restored by the disaster recovery system after an outage. RPO indicates the amount of data loss, that is, the maximum amount of data loss that enterprises can tolerate. A smaller RPO indicates less data loss and less harm to the enterprise.

Disaster recovery protection levels

According to the National Standard of the People's Republic of China GB/T 20988-2007 Information Security Technology - Disaster Recovery Specifications for Information Systems, the protection levels are determined as follows:

Protection level	Data backup	Measure	Preventable risk	RTO	RPO
Level 1: Basic support	• All the data is backed up once a week. • The backup media are stored offsite.	-	The service data is damaged.	Two days or more	One to seven days
Level 2: Secondary site support	• All the data is backed up once a day. • The backup	Backups are called after an outage.	• The service data is damaged. • The service	24 hours or more	One to seven days

	media are stored offsite. • The data is regularly synchronized in batches several times each day.		processing site is not available.		
Level 3: Electronic transmission and partial device support	• All the data is backed up once a day. • The backup media are	Backups are provided for some data processing devices.	• The service data is damaged. • The service processing site is	12 hours or more	Several hours to one day

	stored offsite. • The data is regularly synchronized in batches several times each day.		not available. • Some devices or networks fail.		
Level 4: Electronic transmission and full device support	• All the data is backed up once a day. • The backup media are stored offsite.	Backups are provided for all devices in the available status (cold site).	• The service data is damaged. • The service processing site is not	Several hours to two days	Several hours to one day

	• The data is regularly synchronized in batches several times each day.		availabl e. • All the backup devices or network s fail.		
Level 5: Real-time data transmissi on and full device support	• All the data is backed up once a day. • The backup media are stored offsite. • The data is replicated	Backups are provided for all devices in the ready or running status (warm site).	• The service data is damage d. • The service processi ng site is not availabl e.	Severa l minute s to two days	0 to 30 minute s

	in real time.		• All the backup devices or networks fail.		
Level 6: Zero data loss and remote cluster support	• All the data is backed up once a day. • The backup media are stored offsite. • The data is synchronized and backed up	• The disaster recovery site and the production site have the same processing capability and are compati	• The service data is damaged. • The service processing site is not available. • All the backup	Several minutes	0

	in real time to ensure zero data loss.	ble with each other. • Software clusters are used to implement seamless failover and failback. • Real-time monitoring and automatic failover to the remote	devices or networks fail.		

		cluster system are supported (hot-active).			

Key factors for disaster recovery

The core aim of a disaster recovery solution is to help enterprises balance the needs of RTO and RPO and find the optimal technologies and means of implementation. From an economic perspective, the solution helps enterprises optimize their Total Cost of Ownership (TCO) and Return on Investment (ROI).