ApsaraDB for RDS

Product Introduction

Product Introduction

Product overview

ApsaraDB Relational Database Service (RDS) is a stable, reliable, and auto-scaling online database service. Based on the Apsara distributed system and high-performance storage, the RDS supports MySQL, SQL Server, and PostgreSQL engines. It provides a complete set of solutions for disaster recovery, backup, recovery, monitoring, migration, and other functions to eliminate troubles of database operation and maintenance.

ApsaraDB for MySQL

MySQL is the world's most popular open source database. As an important part of LAMP, a combination of open source software (Linux + Apache + MySQL + Perl/PHP/Python), MySQL is widely used in various applications.

BBS software system Discuz and blogging platform Wordpress that swept the Internet in the Web 2.0 era are both built on MySQL-based underlying architecture. In the Web 3.0 era, leading Internet companies such as Alibaba, Facebook and Google have all built a mature and large cluster of databases based on the more flexible MySQL.

Based on the MySQL source code branch of Alibaba, ApsaraDB for MySQL has withstood highly concurrent and massive data traffic during many November 11 shopping festivals with proven excellent performance and throughput. Besides, ApsaraDB for MySQL provides a range of advanced functions such as optimized read/write splitting, data compression, and intelligent optimization.

The current RDS for MySQL supports versions 5.5 and 5.6.

ApsaraDB for SQL Server

As one of the first commercial databases and also an important part of the Windows platform (IIS + .NET + SQL Server), the SQL Server supports a wide range of enterprise applications. Management Studio supplied with the SQL Server comes built-in with a rich set of graphic tools and script editors. You can get a quick start with various database operations through visual interfaces.

Powered by a high-availability architecture and anytime data recovery capabilities, ApsaraDB for SQL Server provides strong support for various enterprise applications and also covers Microsoft's license fee without additional cost to you.

The current RDS for SQL Server supports versions 2008 R2 and 2012.

ApsaraDB for PostgreSQL

PostgreSQL is the world's most advanced open source database. As the forerunner of academic relational database management system, PostgreSQL sets itself part with full compliance with SQL specifications and robust support for a diverse range of data formats (JSON, IP and geometric data, which are not supported by most commercial databases).

In addition to perfect support for such features as transactions, subqueries, Multi-Version Concurrency Control (MVCC) and data integrity check, ApsaraDB for PostgreSQL integrates a series of important functions including high availability and backup recovery that help mitigate your operation and maintenance burden.

The current RDS for PostgreSQL supports versions 9.4.

ApsaraDB for PPAS

Postgres Plus Advanced Server (PPAS) is a stable, secure, and scalable enterprise-level relational database. Based on PostgreSQL, the world's most advanced open source database, PPAS delivers enhanced performance, application solutions and compatibility, and provides the ability to run Oracle applications directly. You can run various enterprise applications on the PPAS stably and more cost-effectively.Armed with functions such as account management, resource monitoring, backup recovery and security control, ApsaraDB for PPAS will continue to be updated and improved.

The current RDS for PPAS supports versions 9.3.

System Architecture

Data link service

ApsaraDB provides full data link services, including DNS, SLB and Proxy. Since the RDS uses the source-originated DB Engine and database operations are highly similar, there is basically no learning cost.

DNS

The DNS module supports dynamic resolution from domain name to IP address, so as to avoid the

impact caused by IP address change of the ApsaraDB instance. After the domain name is configured in the connection pool, the ApsaraDB instance still can be accessed even if the corresponding IP address changes.

For example, the domain name of an ApsaraDB instance is test.rds.aliyun.com, and the IP address corresponding to this domain name is 10.10.10.1. If test.rds.aliyun.com or 10.10.10.1 is configured in the connection pool of a program, the ApsaraDB instance can be accessed.

If zone migration or version upgrade is performed for this ApsaraDB instance, the IP address may change to 10.10.10.2. If the domain name configured in the connection pool of this program is test.rds.aliyun.com, the ApsaraDB instance can be accessed. If, however, the IP address configured in the connection pool of this program is 10.10.10.1, the ApsaraDB instance can not be accessed.

SLB

The SLB module provides instance IP addresses (including the Intranet IP address and Internet IP address) so as to screen the impact caused by change of the physical server.

For example, the Intranet IP address of an RDS instance is 10.1.1.1, and the corresponding Proxy or DB Engine runs on 192.168.0.1. In the normal case, the SLB module will redirect the traffic that flows to 10.1.1.1 to 192.168.0.1. When 192.168.0.1 fails, 192.168.0.2 in the hot standby status takes over the job of 192.168.0.1. In this case, the SLB module will redirect the traffic that flows to 10.1.1.1 to 192.168.0.2, and the RDS instance still provides services normally.

Proxy

The Proxy module provides data routing, traffic detection, session holding and other functions.

- Data routing function: supports distributed complex query aggregation under the big data scenario and corresponding capacity management.
- Traffic detection function: reduces the SQL injection risk and supports SQL log backtracking when necessary.
- Session holding function: solves the problem of database connection interruption if a fault occurs.

DB engine

The RDS fully supports mainstream database protocols, as detailed in the table below:

Database Type	Version
MySQL	5.1(Deprecated), 5.5, 5.6
SQL Server	2008 R2, 2012
PostgreSQL	9.4
PPAS	9.3, highly compatible with Oracle

High-availability service

The high-availability service consists of the Detection, Repair, Notice and other modules. They guarantee availability of data link services and also process internal exceptions of databases.

In addition, the RDS improves its high-availability service by migrating to a region that supports multiple zones and by adopting proper high-availability policies.

Detection

The Detection module detects whether the master node and slave node of the DB Engine provide services normally. The HA node can easily obtain the health status of the master node through the heartbeat information at an interval of 8 to 10 seconds. With reference to the health status of the slave node and the heartbeat information of other HA nodes, the Detection module can eliminate the misjudgment risk induced by exceptions such as network jitter, and complete the switching operation within 30 seconds.

Repair

The Repair module maintains the replication relationship between the master node and the slave node of the DB Engine and also repairs the error that occurs to the master node or the slave node in daily operations.

For example:

        - Automatic repair of abnormal disconnection of the master and slave replication
        - Automatic repair of the table level damage of the master node and the slave node
        - On-site saving and automatic repair for crash of the master node and the slave node

Notice

The Notice module informs the SLB or Proxy of the status change of the master node and the slave node and ensures that you can access the correct node.

For example, when the Detection module finds an exception with the master node, it notifies the Repair module to repair it. If the Repair module fails in its attempts to repair the master node, it notifies the Notice module to perform traffic switching. The Notice module forwards the switching request to the SLB or Proxy, and then all of your traffic is directed to the slave node. Meanwhile, the Repair module recreates a new slave node on another physical server and synchronizes the change to the Detection module. The Detection module starts to recheck the health status of the instance.

Multi-zone

Multi-zone means a physical area that is formed by combining multiple single zones in the same region. Compared with a single-zone RDS instance, a multi-zone RDS instance can withstand disasters at higher levels.

For example, a single-zone RDS instance can withstand faults at the server and rack levels, while a multi-zone RDS instance can bear faults at the equipment room level.

At present, no extra charge is required for multi-zone RDS instances. Users in the region where the multi-zone has been enabled can directly purchase multi-zone RDS instances or convert a single-zone RDS instance into a multi-zone RDS instance through inter-zone migration.

Note: Because network delay exists to some extent between multiple zones, the response time to a single update may be longer than that of a single-zone instance when a multi-zone RDS instance adopts the semi-synchronous data replication solution. In this case, it is preferred to increase the overall throughput by enhancing the concurrency.

High-availability policy

The high-availability policy is formed by using different combinations between service priorities and data replication modes according to characteristics of your business.

There are two service priorities:

- RTO (Recovery Time Objective) priority: The database should restore services as soon as possible, that is, the maximum available time is reached. The RTO priority policy should be used for users with a higher requirement for online duration of databases.
- RPO (Recovery Point Objective) priority: The database should guarantee data reliability as much as possible, that is, the minimum data loss is achieved. The RPO priority policy should be used for users with a higher requirement for data consistency.

There are three data replication modes:

- Asynchronous replication (Async): The application initiates an update (including the adding, deletion and modification operations) request, the Master node responds to the application immediately after completing the corresponding operation, and the Master node replicates data to the Slave node asynchronously. Therefore, in the asynchronous replication mode, unavailability of the Slave node does not affect the operation on the primary database, but there is a slight probability for unavailability of the Slave node to cause data inconsistency.
- Forced synchronous replication (Sync): The application initiates an update (including the adding, deletion and modification operations) request, the Master node replicates data to the Slave node after completing the corresponding operation, the Slave node returns the success message to the Master node after receiving the data, and the Master node responds to the application after receiving the feedback from the Slave node. Since the Master node replicates data to the Slave node synchronously, unavailability of the Slave node will affect

the operation on the Master node, but unavailability of the Master node will not cause data inconsistency.

- Semi-synchronous replication (Semi-Sync): Normally data is replicated in the forced synchronous replication mode. When an exception occurs (unavailability of the Slave node or a network exception between the two nodes) when the Master node replicates data to the Slave node, the Master node will suspend response to the application, till the replication mode times out and degrades to asynchronous replication. If the application is allowed to update data now, unavailability of the Master node will cause data inconsistency. When data replication between the two nodes resumes normal (the Slave node or network connection is recovered), asynchronous replication will be changed to forced synchronous replication. The time of restoring to forced synchronous replication depends on the implementation mode of semi-synchronous replication. ApsaraDB for MySQL5.5 is different from ApsaraDB for MySQL5.6 in this regard.

You can select different combination modes of service priorities and data replication modes to improve availability according to their business characteristics.

Cloud Data Engine	Service Priority	Data Replication Mode	Combination Characteristics
MySQL 5.1	RPO	Async	If the Master node fails, switching will take place after the Slave node applies all the Relay Logs. If the Slave mode fails, application operations on the Master node is not affected. The data on the Master node will be synchronized after the Slave node recovers.
MySQL 5.5	RPO	Async	If the Master node fails, switching will take place after the Slave node applies all the Relay Logs. If the Slave mode fails, application operations on the Master node is not affected.The data on the Master node will be synchronized after the Slave node recovers.
MySQL 5.5	RTO	Semi-Sync	If the Master node fails and data replication does not degrade, the RDS

			will immediately trigger the switching operation to direct the traffic to the Slave node because data consistency has been guaranteed. If the Slave node fails, application operations on the Master node will time out, and then the data replication mode degrades to asynchronous replication; after the Slave node recovers and the data on the Master node is synchronized completely, the data replication mode is restored to forced synchronization. When the data on the two nodes is inconsistent and the data replication mode degrades to asynchronous replication, if the Master node fails, switching will take place after the Slave node applies all the Relay Logs.
MySQL 5.6	RPO	ASync	If the Master node fails, switching will take place after the Slave node applies all the Relay Logs. If the Slave mode fails, application operations on the Master node is not affected.The data on the Master node will be synchronized after the Slave node recovers.
MySQL 5.6	RTO	Semi-Sync	If the Master node fails and data replication does not degrade, the RDS will immediately trigger the switching operation to direct

			the traffic to the Slave node because data consistency has been guaranteed. If the Slave node fails, application operations on the Master node will time out, and then the data replication mode degrades to asynchronous replication; after the Slave node recovers and the data on the Master node is synchronized completely, the data replication mode is restored to forced synchronization. When the data on the two nodes is inconsistent and the data replication mode degrades to asynchronous replication, if the Master node fails, switching will take place after the Slave node applies all the Relay Logs.
MySQL 5.6	RPO	Semi-Sync	If the Master node fails and data replication does not degrade, the RDS will immediately trigger the switching operation to direct the traffic to the Slave node because data consistency has been guaranteed. If the Slave node fails, application operations on the Master node will time out, and then the data replication mode degrades to asynchronous replication; when the Slave node obtains the Master node information (the Slave node or network fault

			recovers) again, the data replication mode is restored to forced synchronization. When the data on the two nodes is inconsistent and the data difference on the Slave node cannot be supplemented completely, if the Master node fails, you can obtain the time point of the Slave node through the API and decide the switching time and the data supplementing method.
SQL Server 2008 R2	X	X	Adjustment not supported at present
SQL Server 2012	X	X	Adjustment not supported at present
PostgreSQL	X	X	Adjustment not supported at present

Backup recovery service

The backup service supports offline backup, dumping and recovery of data.

Backup

The Backup module compresses and uploads the data and logs on the master node and the slave node. The RDS uploads the backup data to the OSS by default. In a specific scenario, the backup files can also be dumped to a cheaper and more persistent OAS. When the slave node operates normally, backup is always initiated on the slave node to avoid the impact on the service provided by the master node; if the slave node is unavailable or damaged, the Backup module will create backup through the master node.

Recovery

The Recovery module recovers the backup file on the OSS to the target node.

- Master node roll-back function: After a customer initiates a wrong operation on data, the roll-back function can be used to recover data by time point.
- Slave node fixing function: Automatically creates a new slave node to reduce risks when an irreparable fault occurs to the slave node.
- Read-only instance creation function: Creates a read-only instance by using backup.

Storage

The Storage module is responsible for uploading, dumping and downloading backup files. Currently all the backup data is uploaded to the OSS for storage, and the customer can obtain a temporary link to download the data as needed. In a specific scenario, the Storage module supports dumping the backup files on the OSS to the OAS to implement cheaper and more persistent offline storage.

Monitoring service

ApsaraDB provides multidimensional monitoring services at the physical layer, network layer and application layer to ensure business availability.

Service

The Service module tracks statuses of the service level. It monitors whether the SLB, OSS, OAS, SLS and other cloud products on which the RDS depends are normal, including the function and response time. In addition, the Service module also determines whether internal services of the RDS operate normally through logs.

Network

The Network module tracks statuses at the network layer, including monitoring on connectivity between the ECS and RDS, connectivity between RDS physical servers, and packet loss rates of the router and the switch.

OS

The OS module tracks statuses at the hardware and OS kernel layer, including:

- Hardware overhaul: Constantly checks the operation status of such devices as CPU, memory, main board and storage, pre-judges whether a fault will occur, and automatically submits a repair report in advance.
- OS kernel monitoring: Tracks all the database calling operations and analyzes the causes of slowness or error in calling according to kernel statuses.

Instance

The Instance module collects the information at the RDS instance level, including:

- Available information of instances
- Capacity and performance indicators of instances
- SQL execution records of instances

Scheduling service

The scheduling service consists of the Resource module and Version module. It mainly implements resource allocation and instance version management.

Resource

The Resource module mainly allocates and integrates the RDS underlying resources, which means instance enabling and migration to you. For example, when you creates an instance through the RDS console or the API, the Resource module will calculate the most suitable physical server to carry traffic. The Resource module also allocates and integrates underlying resources required for inter-zone migration of the RDS instance. After instance creation, deletion and migration for a long time, the Resource module will calculate the resource fragmentation degree in the zone and regularly initiate resource integration to improve the service carrying capacity of the zone.

Version

The Version module is responsible for version upgrade of the RDS instance. For example:

- MySQL major version upgrade: MySQL 5.1 upgraded to MySQL 5.5, MySQL 5.5 upgraded to MySQL 5.6, etc.
- MySQL minor version upgrade: fixing the bug for the MySQL source code.

Migration service

The migration service helps you to migrate data from a local database to ApsaraDB, or to migrate an ApsaraDB instance to another instance. ApsaraDB provides a Data Transfer Service (DTS) tool to facilitate quick database migration for you.

DTS

The DTS is cloud data transfer service, which can quickly migrate an instance in a local database or the RDS to another RDS instance. Currently, the DTS supports three types of databases: MySQL, SQL Server and PostgreSQL.

The DTS also provides three migration modes: structure migration, full migration and incremental migration:

> Structure migration: The DTS will migrate the structure definition of migration objects to the target instance. At present, the objects that support structure migration include tables, views, triggers, stored procedures, and stored functions.

> Full migration: The DTS will migrate all of the existing data of the source database migration object to the target instance.

> **Notice:** To ensure data consistency, non-transaction tables without a primary key will be locked during the full migration process. The locked tables cannot be written, and the locked duration depends on the data volume of these tables. The locks will be released only after these non-transaction tables without a primary key are migrated.

> Incremental migration: The DTS will synchronize data changes made in the migration process to the target instance.

> **Notice:** If a DDL operation is performed during the migration, the structure changes will not be synchronized to the target instance.

Product Strengths

Cheap and easy-to-use

Out-of-the-box way

You can customize the RDS specification through the official website of Alibaba Cloud or the API. After the order is sent, the RDS generates the target instance in real time. The RDS can work with the ECS to reduce the application response time and save the public traffic fee.

On-demand upgrade

In the initial stage of business, you can purchase the RDS instance of small specification to cope with the business pressure. Along with changes in the database pressure and data storage capacity, you can flexibly adjust the instance specification, and the RDS will not interrupt the data link service in the upgrade period.

Transparent and compatible

The using method of the RDS is the same as that of the native database engine. You can get started easily without secondary learning. In addition, the RDS is compatible with your current programs and tools. General data import and export tools can be used to migrate data to the RDS, and the manpower cost during the migration is very low.

Ease of management

Alibaba Cloud is responsible for routine maintenance and management of the RDS, including but not limited to such work as hardware/software fault processing and database patch update, so as to ensure normal operation of the RDS. You can also independently perform database adding, deletion, restart, backup, recovery and other management operations through the Alibaba Cloud console.

High performance

Parameter optimization

Alibaba Cloud has gathered top database experts in China, and the parameters of all the RDS instances have been obtained from many years experience in production and optimization. In the lifecycle of the RDS instance, the DBA optimizes it continuously to ensure that the RDS keeps running based on the best practice.

SQL optimization recommendations

In view of features of your application scenario, the RDS will lock low-efficiency SQL statements and put forward optimization recommendations so that you can optimize the business code.

High-end hardware investment

All servers used by the RDS have been evaluated by multiple parties to ensure that they top the like products in performance and stability.

High security

Anti-DDoS attack

When using an Internet connection to access RDS instances, users may experience DDoS attacks. When considering that a user's instance is under DDoS attack, the RDS security system first enables the flow cleaning function. If the attack reaches the black hole threshold or the cleaning operation fails, black hole processing will be triggered.

The following shows how flow cleaning and black hole processing work and when these activities are triggered:

- Flow cleaning: This applies only to inbound traffic from the Internet. During this process, the RDS instance can be normally accessed. The system automatically triggers and terminates flow cleaning. Flow cleaning is triggered if a single ApsaraDB instance meets any of the following conditions:
  - PPS (Package Per Second) reaches 30,000;
  - BPS (Bits Per Second) reaches 180 Mbps;
  - The number of concurrent connections created per second reaches 10,000;
  - The number of concurrent active connections reaches 10,000;
  - The number of concurrent non-active connections reaches 10,000.
- Black hole processing: This only applies to inbound traffic from the Internet. During this process, RDS instances cannot be accessed from the Internet. Generally, this means their applications will be unavailable. Black hole processing provides a way to ensure the availability of the overall RDS service. Black hole triggering and termination conditions are as follows:
  - BPS (Bits Per Second) reaches 2 Gbps;
  - Flow cleaning is ineffective;
  - The black hole will be removed automatically after 2.5 hours.

Note: We suggest users access RDS instances via intranet to avoid the risk of DDoS attack.

Access control policy

You can define the IP addresses allowed to access the RDS. Access of an IP address other than the

specified ones will be denied.

Each account can only view and operate its own database.

System security

The RDS is under protection of multiple layers of firewall and can effectively resist various malicious attacks and ensure data security.

Direct login to the RDS server is not allowed. Only the port required by the specific database service is open.

The RDS server cannot actively initiate an external connection, but can only accept accesses passively.

Professional security team

Alibaba Group's security department provides security technology support for the RDS.

High reliability

Hot standby

The RDS adopts a hot standby architecture. If a physical server fails, the service is switched over in seconds. The entire switching process is transparent to the application.

Multi-copy redundancy

The data on the RDS server is built on the RAID, and the data backup is stored on the OSS.

Data backup

The RDS provides an automatic backup mechanism. You can independently select a backup cycle or initiate temporary backup at any time according to your business characteristics.

Data recovery

The data can be recovered by the backup set and specified time point. In most scenarios, you can recover the data at any time point in 7 days to a temporary RDS instance. After the data is verified, the data can be migrated back to the primary RDS instance so as to implement data backtracking.

Typical Application

Diversified data storage

The RDS can work with storage products such as the RDS for Memcached, the RDS for Redis, and the RDS for OSS to implement diversified storage extension.
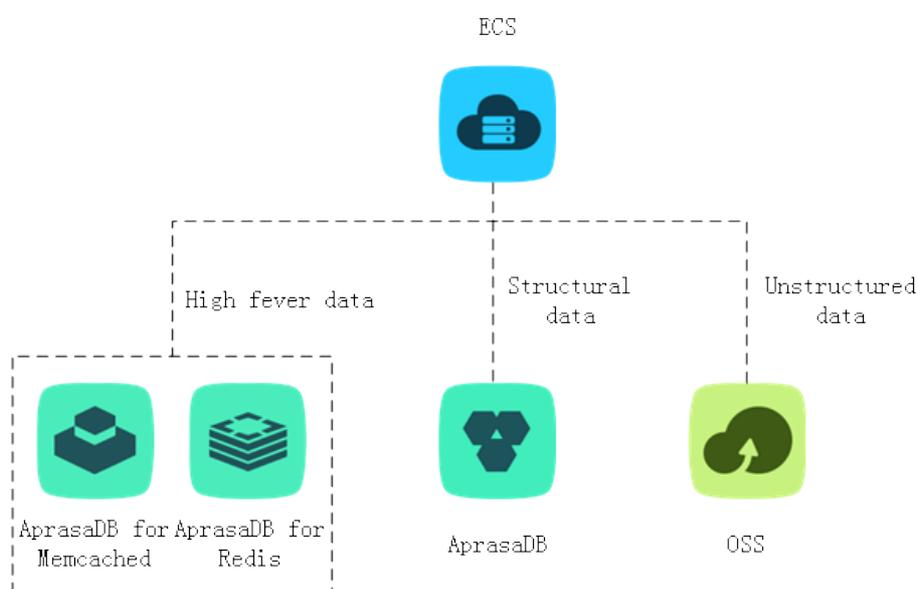
Cache data persistence

The RDS can work with the RDS for Memcached and the RDS for Redis to form a high-throughput and low-latency storage solution. Compared with the RDS, the RDS cache product has two features:

- High response speed: The request delay of the RDS for Memcached and the RDS for Redis is usually within several milliseconds
- The cache area can support a higher QPS (Requests Per Second) than the RDS

For cases related to cache data persistence, please refer to Cache Data Persistence.

Multi-structure data storage

The OSS is a cloud storage service provided by Alibaba Cloud, featuring massive capacity, security,
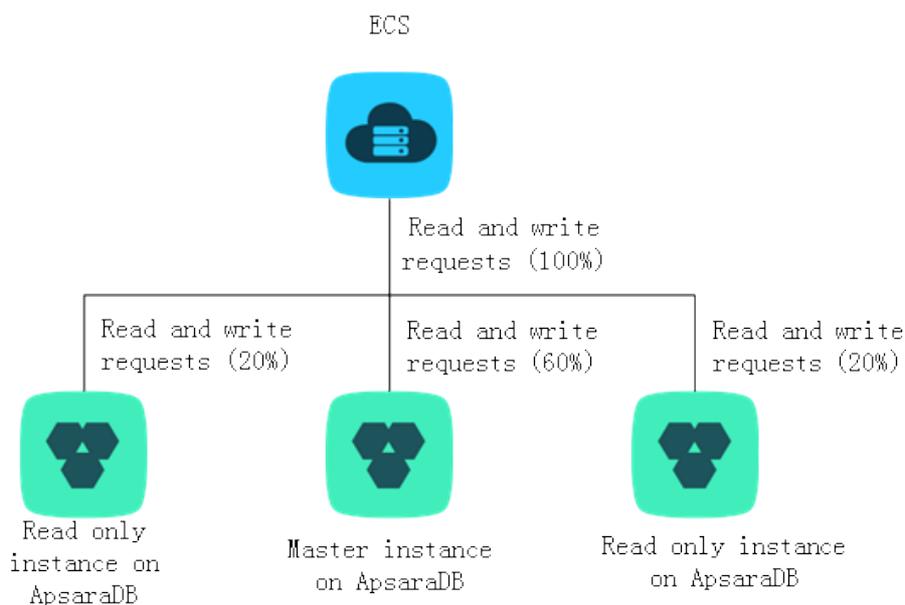
low cost, and high reliability. The RDS can work with the OSS to form multiple types of data storage solutions.

For example, when the business application is a forum and the RDS works with the OSS, resources such as registered users' images and post content images can be stored in the OSS to reduce the storage pressure of the RDS.

For cases related to the joint use of the RDS and the OSS, please refer to Multi-structure Data Storage .

Read/Write splitting

ApsaraDB for MySQL supports direct attaching of read-only instances to share the read pressure of primary instances. Each read-only instance has an independent connection string, and the read pressure can be automatically distributed at the application end.

ECS

Read and write
requests (100%)

Read and write
requests (20%)

Read and write
requests (60%)

Read and write
requests (20%)

Read only
instance on
ApsaraDB

Master instance
on ApsaraDB

Read only instance
on ApsaraDB

For the method for creating a read-only instance of the ApsaraDB for MySQL, please refer to Creating a Read-only Instance.

History

Date	Key events
2011-06	RDS for MySQL 5.1 was launched
2011-11	RDS for SQL Server 2008 R2 was launched

2012-07	RDS supported the high security access mode
2012-11	RDS supported e-commerce cloud (Jushita cloud service) and processed some orders of Taobao and Tmall on November 11
2012-12	RDS supported SQL detail collection
2013-02	RDS supported backup dumping to the OSS
2013-05	RDS for MySQL 5.5 was launched
2013-07	RDS supported recovery of the specified time point
2013-08	RDS supported AntCloud (Jubaopen cloud service)
2013-11	RDS processed 50% of orders of Taobao and Tmall on November 11
2014-02	DMS data management service was launched
2014-07	RDS for MySQL 5.6 was launched
2014-10	RDS for MySQL read-only instance was launched
2014-10	RDS opened the Open API and supported the Pay-As-You-Go instance
2014-11	RDS processed 98% of orders of Taobao and Tmall on November 11
2014-12	RDS supported the multi-zone disaster recovery instance
2015-02	RDS manager was launched
2015-04	RDS supported the Virtual Private Cloud (VPC)
2015-06	RDS for PostgreSQL 9.4 was launched
2015-07	RDS for MySQL supported TokuDB, with 5 times of the data compression capability
2015-07	RDS for PPAS 9.3 was launched
2015-09	The DTS was launched
2015-12	The RDS for MySQL permission was open
2016-01	RDS for MySQL remote disaster recovery instance was launched
2016-03	RDS supports 730-day backup and monitoring at an interval of 60 seconds.
2016-03	RDS received the new version ISO 20000 certification
2016-04	RDS received ISO 22301 certification
2016-04	RDS passed Service Organization Control

	(SOC) independent audit
2016-05	RDS passed CNAS national cloud computing standards test
2016-05	RDS released new instance type
2016-05	RDS supports SSL\TDE security encryption
2016-06	RDS received the Payment Card Industry Data Security Standard (PCI - DSS) certification
2016-06	RDS received the Singapore national standard MTCS tier 3 certification