Alibaba Cloud

Alibaba Cloud - A Silver Lining to Your Cloud Application Architecture Design

Contents

01 Introduction

Cloud computing services are services around servers, applications, databases, and virtual machines delivered through a network or the Internet. They aim to simplify how organizations manage, maintain and run their applications, servers, data, and data storage in a cost-effective manner. This enables organizations to run their operations with agility and without having to worry about the security, management, and connectivity of their Information Technology (IT) environment and related applications.

Unlike traditional IT infrastructures that require on-premise IT teams to ensure months' of deployment and to run operations, cloud services are accessible irrespective of location. Cloud computing services let an organization concentrate on its core business and pay only for what it uses while enhancing the user experience by providing easy access from any location, at any given time, as required by the user.

This whitepaper introduces cloud application architecture design and the procedure of transforming applications to a cloud architecture. It presents the universal practices for various scenarios on the cloud platform and explores how different Alibaba Cloud product portfolios are tailor-made for these scenarios.

Please note that operational scenarios and complexity are not the same for each transformation, and this document outlines each scenario without in-depth technical discussion.

02 Cloud Service Framework

The Cloud service framework has multiple components including the application architecture. The design of the cloud architecture includes the front-end platform, back-end platform, connecting network, and a cloud-based delivery. A "front-end" application is the user interface that enables interaction between the user and the back-end applications. A "back-end" application or program acts as the back-end support to the user and the front-end application.

The back-end application interacts directly with the IT resource and performs the corresponding function(s). At times, there are intermediate programs that help with the communication between the front-end and back-end applications. The connecting network can be via the public Internet, or in cases where security is of primary concern, via virtual private connections established by the cloud service provider. The cloud-based delivery depends on the requirements of the user organization and the services of the cloud service provider.

03 Traditional Versus Cloud Architecture

Traditional architecture primarily works on two formats: dedicated and shared. With dedicated architecture, the organization owns the complete hardware or servers required for storage of data and applications. In a shared setup, organizations share the resources with other businesses and pay only for a set amount of space on the servers. For both traditional architecture types, an organization needs to provision additional hardware in advance and scale up instances manually to handle the expected increase in website traffic. This carries the uncertainty of under-provisioning or over-provisioning of hardware, which in turn leads to greater cost on servers and resources.

The following diagram depicts a commonly used traditional architecture:

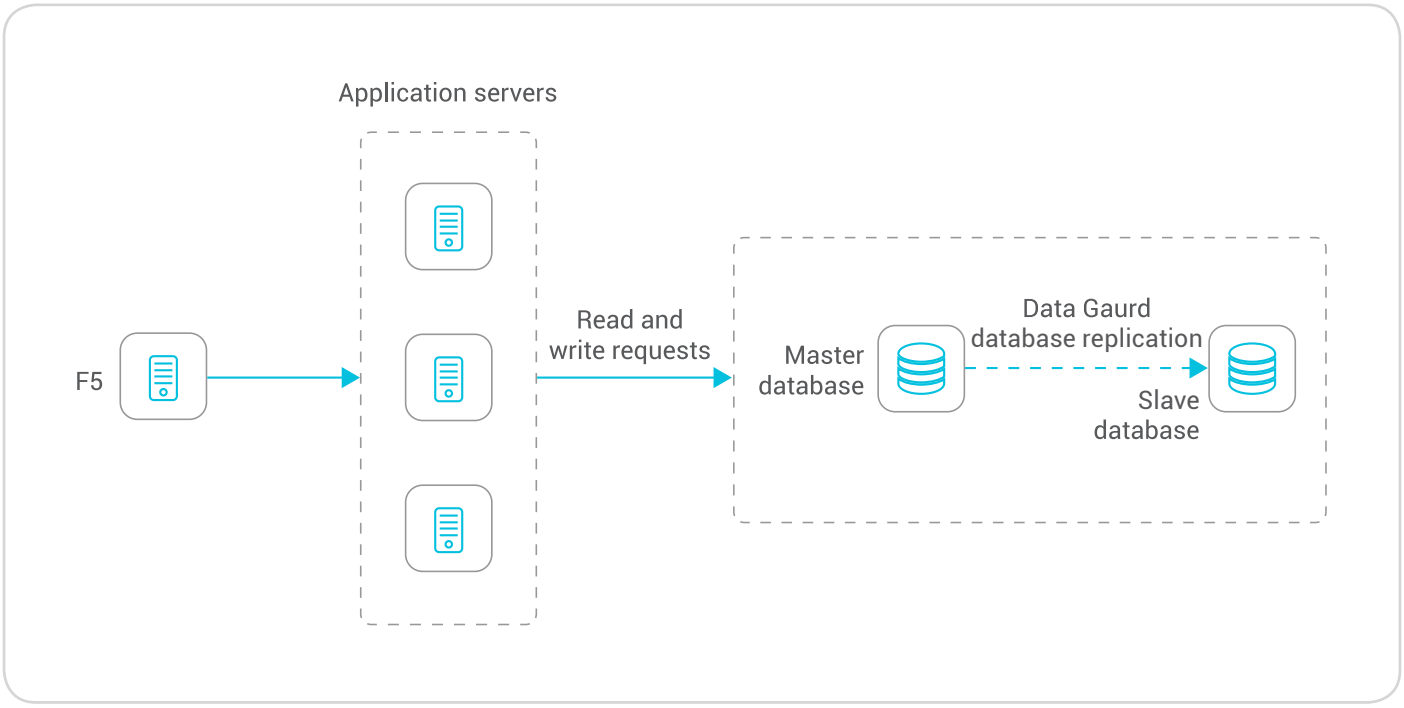


Diagram 1: Common Traditional Architecture

Components of a traditional architecture are briefly described below:

Firewall: Hardware or software solution to open standard port (80, 443)

- **Web Load Balancer:** Hardware or software solution to distribute traffic over web servers
- **Web Tier:** Servers handling Hypertext Transfer Protocol (HTTP) requests
- **App Load Balancer:** Hardware or software solution to distribute traffic over app servers
- **App Server Tier:** Servers handling app specific workloads

- **Database Tier:** Database servers to serve data storage needs
- **Backups:** Optional periodic backups stored on tapes

This architecture follows a standard three-tier web application model with the presentation, application, and persistence layers.

One can launch additional servers manually at any of the three layers to achieve scalability as per business requirements. Further, this architecture includes a load balancer that requires manual operation and maintenance for high availability during peak traffic. Installed within every layer of the system architecture are firewalls to block malicious traffic.

Traditional Architecture Challenges

With most applications experiencing unpredictable traffic, manual deployment and load balancing of traffic using hardware balancers can lead to either under-provisioning or over-provisioning of infrastructure. Under provisioning occurs when applications are unable to meet business requirements during peak traffic periods, which hampers customer satisfaction.

Over-provisioning occurs during off periods when there is little to no traffic due to the underutilization of hardware, which leads to inefficient use of capital. In addition, probable security breaches in shared hosting can result in performance issues. For an enterprise business the hardware works on geographic distribution over various locations. Though most of the companies manage this through multiple agencies to reduce costs, management of all the agencies present at distributed data centers from a remote location is both cumbersome and time-consuming.

In addition, the distributed setup incurs greater costs to manage IT staff and monitoring requires more network, storage, and power resources that can prove to be a challenge.

Cloud Architecture

According to recent trends, companies are making the pivotal cloud shift for enhanced application hosting experience. A cloud architecture is an extension of an organization's IT infrastructure and its business. It uses simple application program interfaces (APIs) to provide the user with accessibility to its data, application and related services via the Internet or a virtual private network.

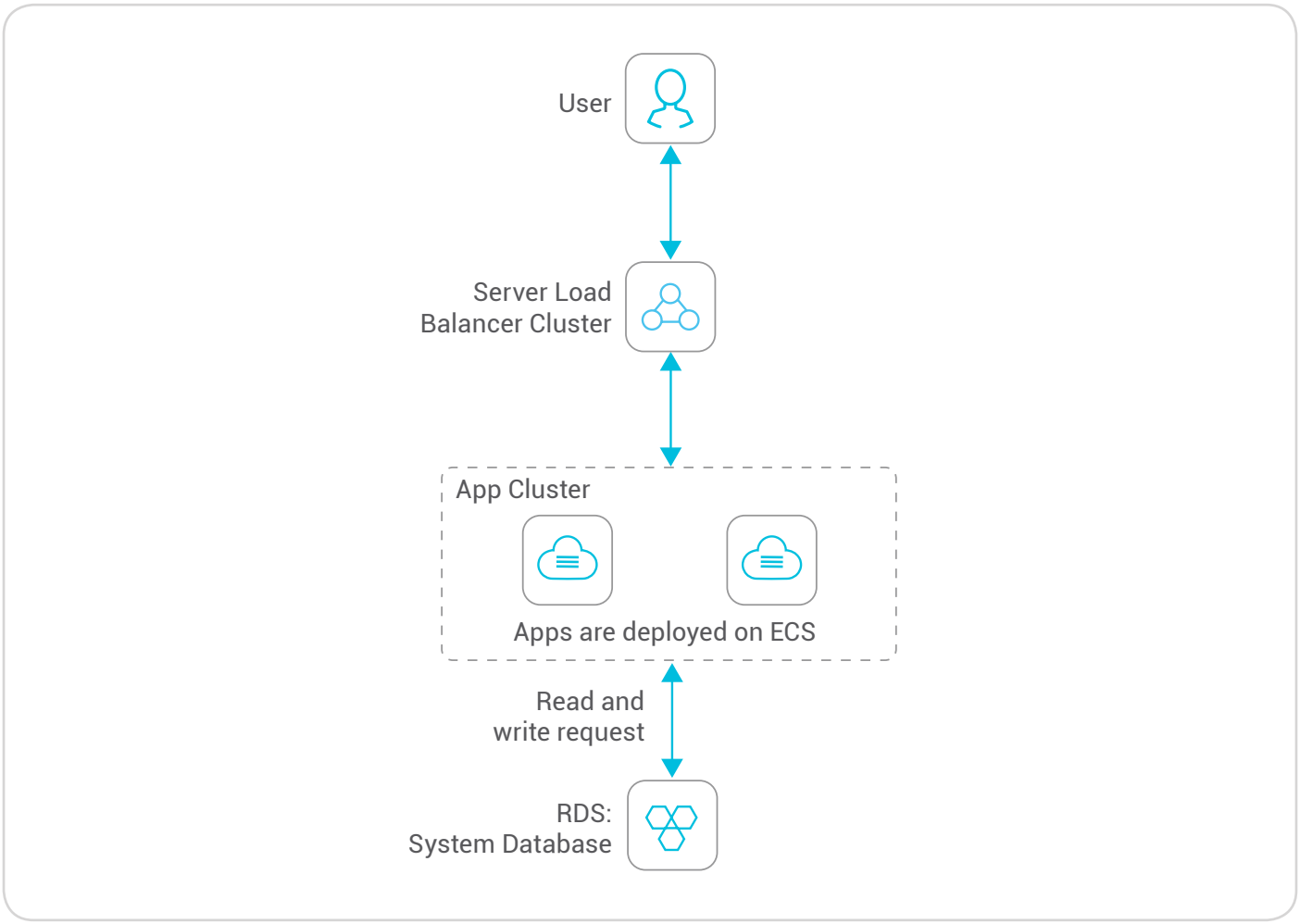


Diagram 2: Cloud Architecture

Components of a Cloud Architecture are:

Relational Database Service (RDS)

RDS is an on-demand database service that frees businesses from the administrative task of managing a database and leaves users with more time to focus on their core business. It protects against network attacks and intercepts Structured Query Language (SQL) injections, brute force attacks and other types of database attacks. It also provides automatic failover in case primary instances are not available while offering remote and local disaster recovery. Furthermore, it ensures reliability with automated backups, snapshots, and automatic host replacement.

Elastic Computer Service (ECS)

ECS cluster provides high-speed memory and the latest Intel CPUs to help achieve faster results with low latency to power your cloud applications. Scale capacity up or down based on your business and web application's real-time demands. Pay only for the resources you use and avoid the need to provision expensive IT infrastructure and hire large network teams.

Server Load Balancer

Manages sudden spikes in traffic, minimizes response time, and maintains the availability of your web application with a Server Load Balancer. This product monitors the health of servers and ensures high availability by automatically distributing application requests to servers with optimal performance in different availability zones.

A cloud-based environment allows your applications to utilize on-demand, virtual resources or servers and saves you from investing money on in-house server storage and application requirements as well as operational, maintenance and resources associated with it.

The cloud service provider manages your end-to-end IT infrastructure, including the global reach, continuity, and scalability of your business as per your business requirements. It takes complete responsibility of sudden spikes in demand while ensuring security for your data and applications. You can control the required resources via a user management portal while only paying for the resources you use including server space and bandwidth.

04 Alibaba Cloud Solution

Alibaba Cloud offers a comprehensive cloud solution with virtual load balancers, security services, and necessary virtual resources for hassle-free deployment on the cloud. This solution balances web traffic across a cluster of virtual servers to ensure zero downtime. It allows you to host, run, and maintain your applications on the cloud through a single global account in a secure and cost-effective manner.

The following diagram depicts how the traditional architecture functions when deployed using Alibaba Cloud products and services.

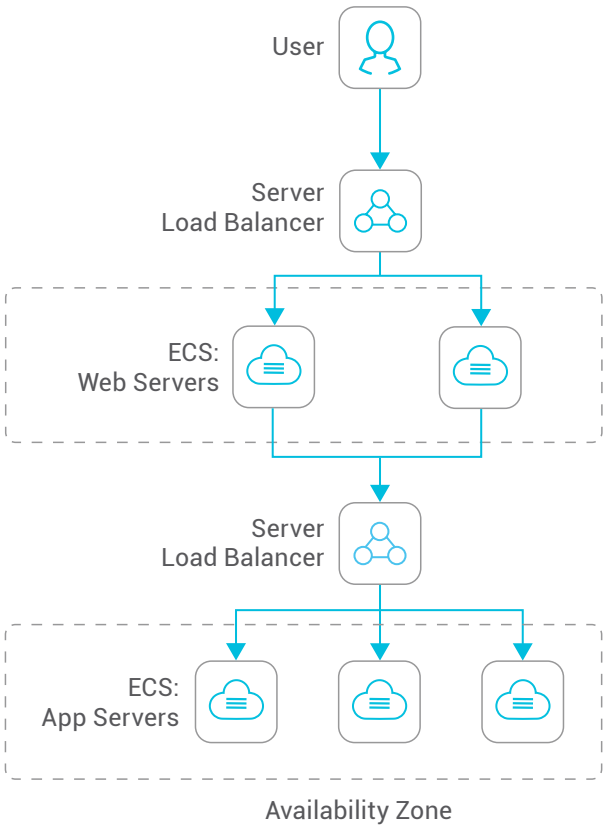


Diagram 3: Alibaba Cloud Architecture

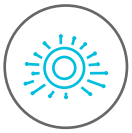
The following are advantages of cloud architecture over traditional architecture

Multiple Data Centers in a Region

Within each region, Alibaba Cloud has at least two data centers called Availability Zones (AZs). As Elastic Compute Service (ECS) in different AZs is logically and physically separated, Alibaba Cloud provides an easy-to-use model for deploying applications across AZs for higher availability and reliability. In the case of a hardware failure in one zone, the application will still be functional using the resources located in a different zone to prevent loss of service.

High Security for Web Applications

With more than 90 percent of applications being vulnerable to security attacks, application security is a major concern for organizations. These attacks have the capability to exploit applications and inherent servers, which increases the risk of financial loss for businesses. To protect applications from such malicious attacks, Alibaba Cloud provides a suite of network and application security services.



Anti-DDoS

Provides layer-4 to layer-7 (pro version) network defense capabilities, which includes protection against attack types like DDoS), SYN flood, and user datagram protocol (UDP) flood. It offers 500M of free Anti-DDoS protection for each ECS instance and 1000G+ Anti-DDoS protection with Alibaba Cloud advanced pro version.

Web Application Firewall (WAF)

Protects applications from common web exploits that cause application unavailability, compromised security, or excessive consumption of resources. WAF gives the users complete control of applications by letting them define customizable security rules to block malicious traffic from reaching the applications. The user can create custom rules that prevent common attacks, such as SQL injection attacks, cross-site scripting, and protection against Trojan virus injection.

Virtual Firewalls

When compared with traditional application hosting DMZ models, ECS offers additional security through virtual firewalls called security groups. Security Groups are similar to an inbound/outbound network firewall, for which protocols, ports, and source IP ranges can be specified and allowed to reach your resources.

The Alibaba Cloud solution provides a reliable, scalable and secure cloud-based architecture for organizations and users, alike. It comprises of various cloud products, applications and data centers, which perform different functions in making it an effective and recommended cloud solution for organizations.

05 Alibaba Cloud Products - Introduction and Overview

Cloud products comprise of various components that connect the application-related and storage-related products to facilitate the working of the entire architecture.

Alibaba Cloud offers a varied portfolio of products that include products that act as an alternative to the original on-premise physical servers, those that effectively balance the load, as per requirement and scalability. It provides for customized architecture functions based on products which are storage, database, network and security related while taking into account the latest developer and management tools for enterprises and startups as per their needs. Let us dive into Alibaba Cloud offerings.

5.1 Application-related Products

5.1.1 Elastic Compute Service (ECS)

Cloud servers provide virtual hosts (CPU/memory/ephemeral disk/network interface card) and you may use the ECS as a physical server.

Application Scenario: ECS acts as an alternative to the original physical servers or on premise servers.

5.1.2 Server Load Balancer

Server Load Balancer provides load-balancing services for multiple ECSs by distributing the traffic. It can expand the application system's service capabilities through traffic distribution.

Application Scenario: An alternative to traditional load balancing servers. It eliminates the single points of failures of application systems, improving the system's availability.

5.1.3 Virtual Private Cloud (VPC)

Virtual Private Cloud (VPC) establishes an isolated network environment based on the cloud environment.

Application Scenarios: Scenarios requiring high network security and network isolation.

5.1.4 Alibaba Cloud Security

Alibaba Cloud Security provides one-stop security services such as security vulnerability detection and Trojan detection, as well as host intrusion detection and anti-DDoS protection for ECS.

Application Scenario: Websites and ECS with security requirements on Alibaba Cloud.

5.2 Storage-related Products

5.2.1 Relational Database Service (RDS)

Relational Database Service (RDS) is compatible with MySQL and SQL Server database engines and provides sound relational database management.

Application Scenario: Applicable to OLTP businesses to provide real-time, high-concurrency, and strongly consistent online.

5.2.2 Distributed Relational Database Service (DRDS)

Distributed Relational Database Service (DRDS) is an online distributed database service compatible with MySQL databases, supporting horizontal splitting, smooth scaling, and read-write splitting.

Application Scenario: Applicable to online transaction systems in which standalone MySQL databases are not competent.

5.2.3 ApsaraDB for Memcache

ApsaraDB for Memcache is capable of high-speed caching of highly accessed data, in the key-value data structure and compatible with the Memcached protocol.

Application Scenario: Reducing the excess stress on the database and caching users' highly accessed data.

5.2.4 Analytic DB

Analytic DB performs multi-dimensional instant analysis and operational exploration for hundreds of billions of data records within milliseconds (Real-time OLAP).

Application Scenario: Provides massive, real-time, and high-concurrency online analysis.

5.2.5 MaxCompute

MaxCompute is an offline processing and analysis service targeting terabytes/petabytes of data with low real-time requirements (within seconds).

Application Scenario: Provides distributed processing and targeting of terabytes/petabytes of data with low real-time requirements.

5.2.6 Table Store

Table Store is a NoSQL database service that provides storage of and access to massive structured data sets.

Application Scenario: Quick processing of a large number of read or write requests to massive amounts of data (logs/ monitoring data).

5.2.7 Object Service (OSS)

Object Service (OSS) is a cloud storage service featuring massive capacity (petabytes of data), security, a low cost, and high reliability.

Application Scenario: An alternative to NAS file storage.

06 Cloud Application Practices

Cloud applications are a set of distributed applications and services or APIs integrated to create a composite application/service that is available to the customer organization. The effectiveness of the cloud application and the architecture, as a whole, is partially determined by how these applications interact with each other and the products.

Three main types of cloud application practices exist:

Transaction Application

Transaction applications have an implicit relationship with database transformation and migrating the database system to the cloud. The scheme of replacing traditional databases with RDS reduces the cost considerably. At the same time, for systems with larger amounts of data, Alibaba Cloud utilizes DRDS to achieve highly scalable system architectures while OSS is exclusively for solving problems in traditional storage systems. Highly available Server Load Balancers solve the load balancing challenges faced by applications, establishing a set of highly scalable, reliable, and high-performance distributed cloud applications.

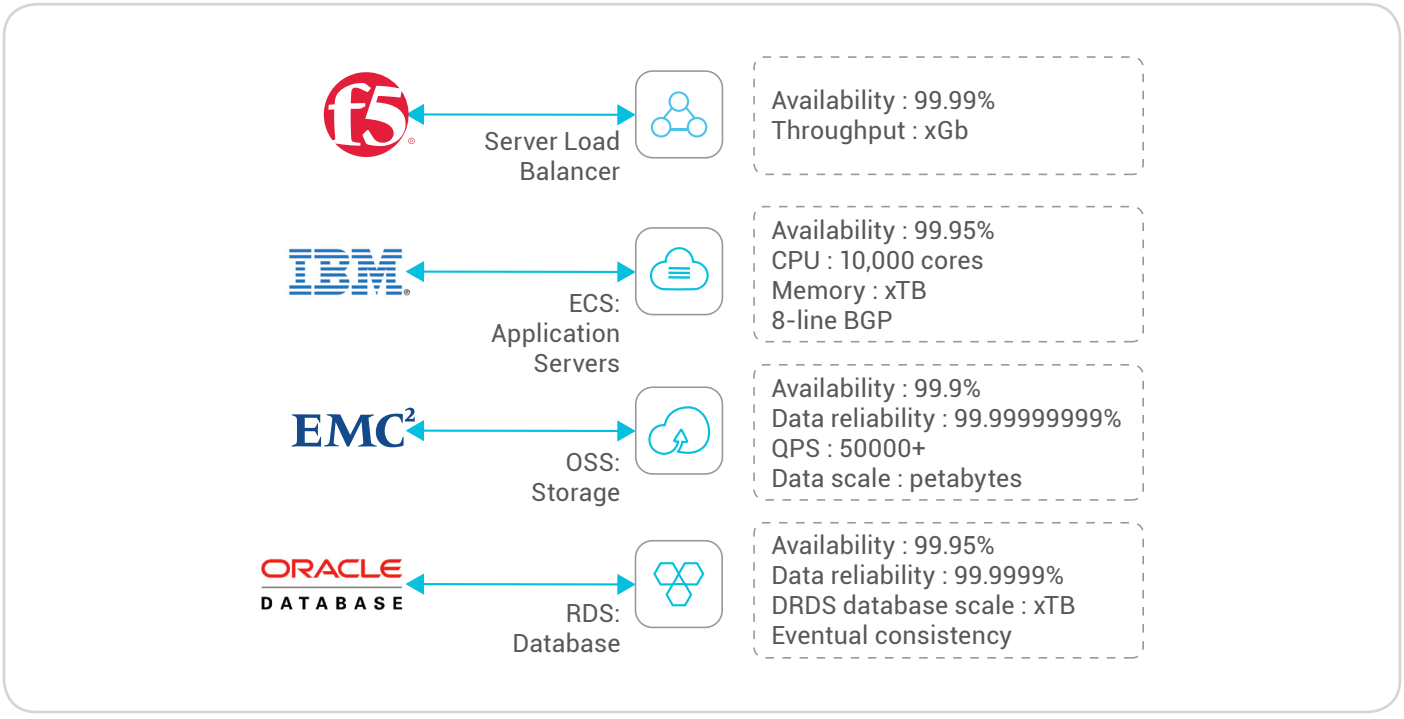


Diagram 4: Transaction Application

Content-based Application

Content-based applications, similar to transaction applications, feature a huge demand for bandwidth resources. The maintenance for self-built or leased resources in traditional industries is complex as it is highly demanding for skilled IT staff in enterprises. Most of these applications involve videos, web pages, among others, and have a higher volume of traffic.

Alibaba Cloud has a wealth of resources with a high degree to cater to the challenges faced by traditional IT Operations and Maintenance (O&M) personnel.

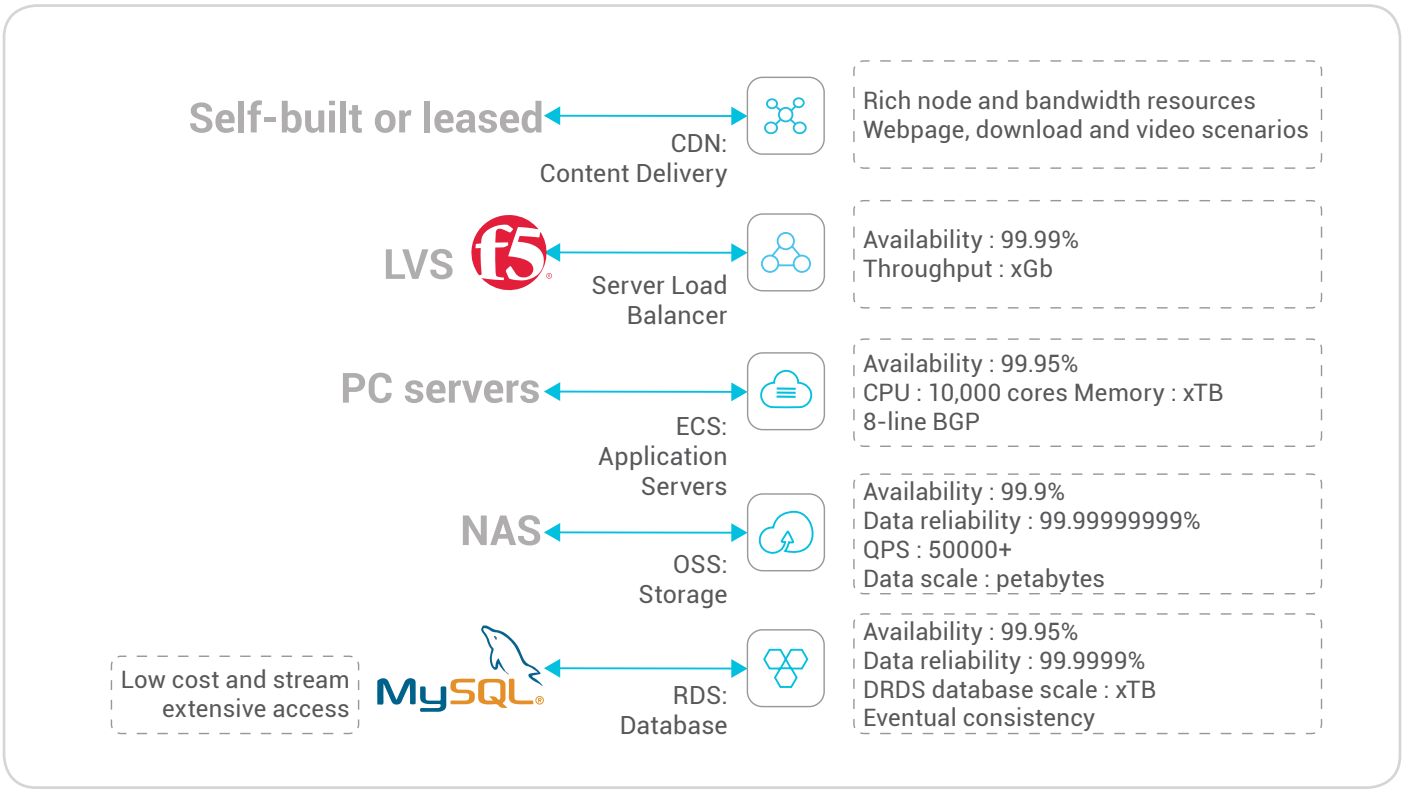


Diagram 5: Content-based Application

Analysis Application

Analysis application mainly serves the Business Intelligence (BI) department in traditional industries and has higher latency. The T + 1-time requirement is a common phenomenon (calculated on the next day).

Alibaba Cloud provides a portfolio of big data solutions that cater to the BI function applications in traditional industries, with many improvements. For example, the computing can reach almost real-time levels (CDP + ODPS), and the efficiency of queries can be rapid (within seconds) (ADS). More importantly, the Alibaba Cloud big data platform is an end-to-end data-sharing platform that provides centralized data exchange services, so that enterprise applications can easily leverage it for innovative data value applications and other benefits.

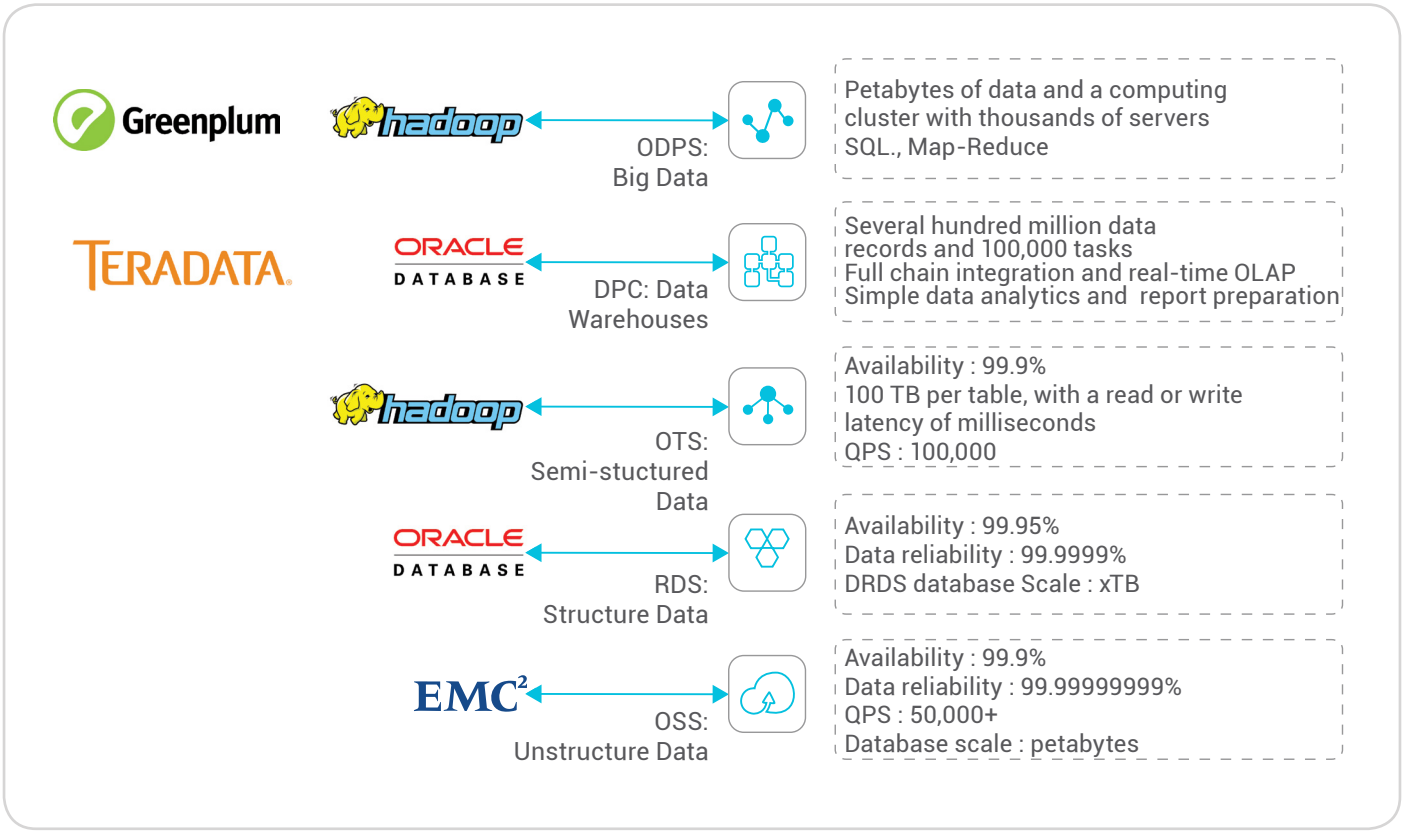


Diagram 6: Analysis Application

07 Cloud Service-based Application Architecture Design

The architecture design of a cloud service is based on the front-end and back-end platform, which comprises of cloud applications, servers, databases and the connecting network. The way in which these applications interact with each other affect the functioning and performance of the architecture.

With time and popularity, the cloud-based application architecture design has advanced and evolved like most widespread technologies. The basic three-layer architecture has given way to more complex and flexible cloud infrastructure capabilities that facilitate scaling, load balancing, security, and other varied storage and computing services. These more complex architectures require intense decision-making on matters concerning migration and its execution.

The new architecture of the application system on the cloud, with migration schemes determined system checks and risk assessment results, combined with the characteristics of the cloud platform. It allows efficient decision-making on whether the system needs direct migration to the cloud platform or whether a series of transformations (e.g., de-O) are necessary. In addition, whether the file system needs to migrate to the OSS, whether the data analysis system should be compatible with the cloud platform, and what the estimated transformation schedules are. Eventually, we can establish the cloud architecture design and transformation scheme more effectively.

08 Basic Steps for Transformation of Applications into the Cloud Architecture

To be part of the deployed cloud architecture and effectively perform its corresponding functions, each application requires transformation to the cloud. Discussed below are the basic steps that each application requires to go through while transforming into the cloud.

8.1 Server Load Balancer Transformation

The original system web/application servers adopt F5 or open-source LVS hardware equipment to achieve load balancing for web/application servers. When the system migrates to the cloud, the load-balancing scheme should transform to the cloud Server Load Balancer service as well. The on-cloud Server Load Balancer service can support Transmission Control Protocol (TCP), Hypertext Transfer Protocol (HTTP), and HTTP Secure (HTTPS) protocols to achieve traffic load balancing. At the same time, it supports automatic health checks for web/application servers, automatic blocking of servers in abnormal states, and automatic unblocking of servers after the server returns to a normal state, to resume services.

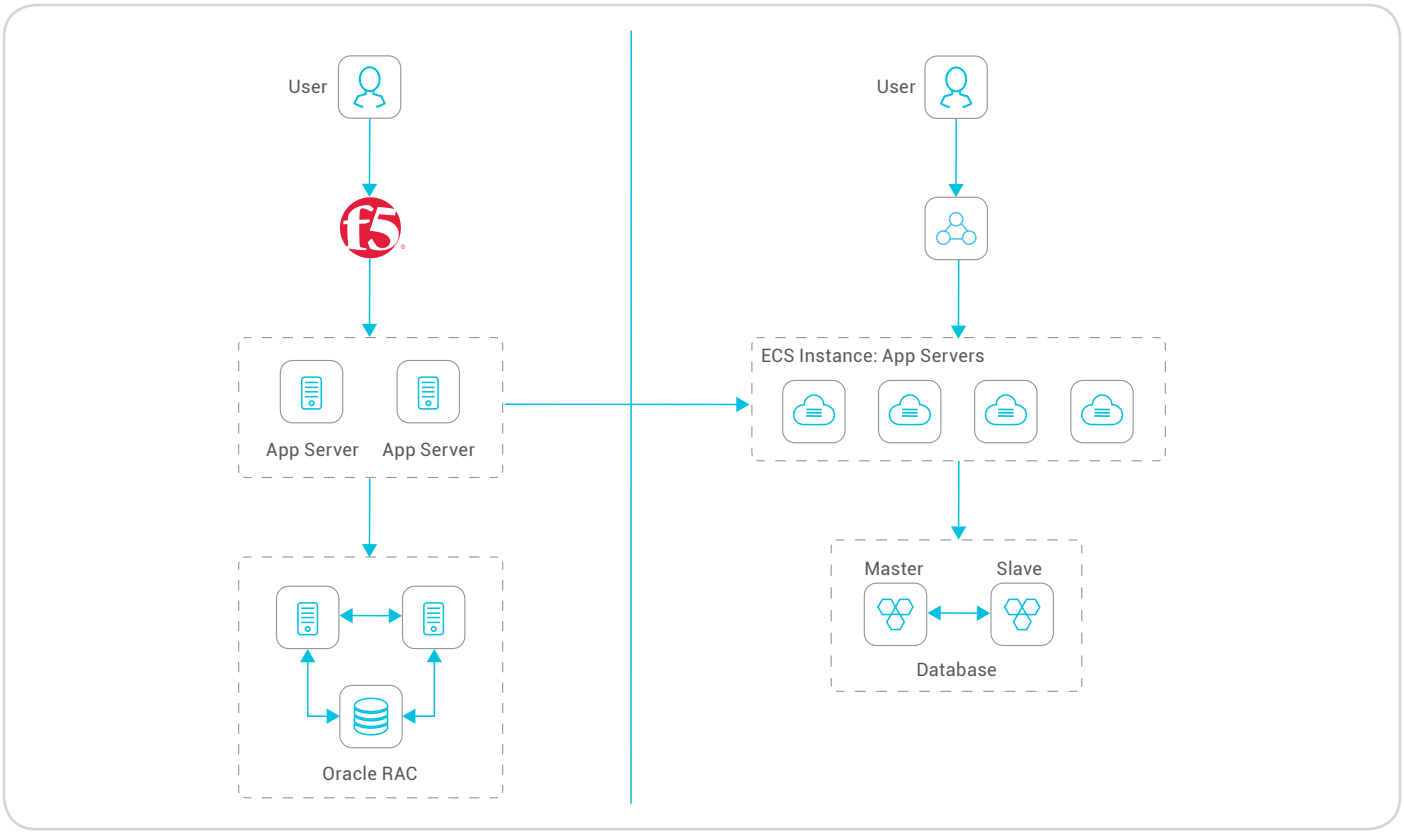


Diagram 7: Server Load Balancer Transformation

8.2 Web and Application Layer Transformation

We can deploy the original system web/application services deployed on minicomputers, PC servers, commercial or open-source virtualized servers directly on the cloud's ECS. Currently, ECS supports both Linux and Windows operating systems. To ensure high availability of web/application services, we recommend the deployment of the same type of services on at least two ECS servers, using the Server Load Balancer for load balancing and service fault tolerance.

Features

- The performance requirement for a single server is not high. The server is stateless and horizontally scalable.
- There are a rich variety of apps/services, dominated by the B/S model
- Different systems choose different types of servers, dominated by X86 PC servers.

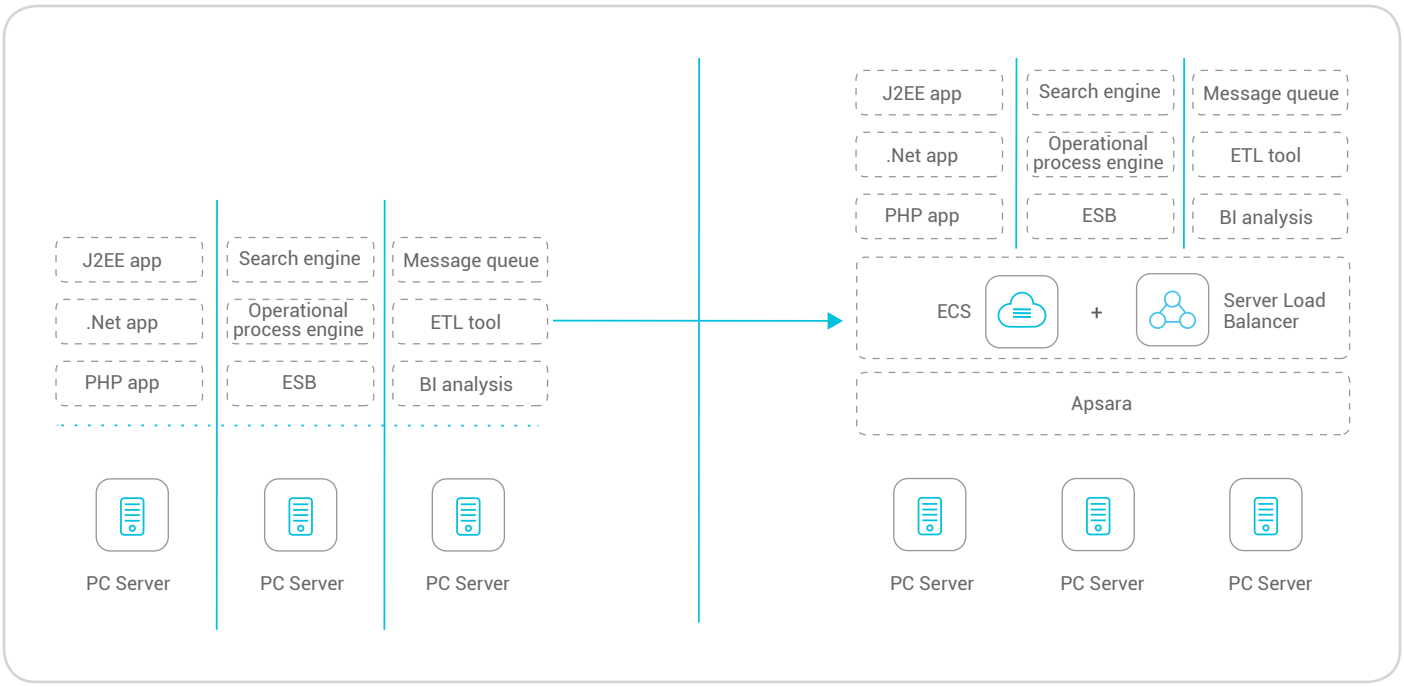


Diagram 8: Web and Application Layer Transformation

8.3 File Storage Transformation

Content management application systems involve the storage and management of a large number of file objects. Traditional solutions include:

- Ephemeral disk storage and regular data backups: This solution faces problems with capacity and performance and has many shortcomings when it comes to scalability and high availability.

- The adoption of IP-SAN and NAS: The use of this solution is for centralized data storage, but is a much more costly solution.
- Storing files in a database: This solution is expensive and has lower performance. Relational databases are not suitable for storing such data types.

Targeting the file object storage, the cloud platform provides open Object Storage Service (OSS) with higher availability, greater scalability, higher efficiency, and in a more cost-efficient manner, effectively solving the storage problems of application files, images, and videos. To use the OSS service, there is a need to transform the original application file system, primarily including:

- Plan buckets and the file directory structure in the OSS according to the storage structure of the application system files.
- Set the bucket access permissions (public-read-write/public-read/private). For applications with high-security requirements, you can configure file storage to be in the form of cipher text in the OSS.
- Scan the program code, find the code involving reading data from or writing data to storage from a file, and transform the code to be implemented through the OSS SDK interface.

  Note: For smaller files (<100M), read and write the file directly by calling the object provided by the SDK. For larger files (>100M), it is recommended to use the Multipart Upload interface, provided by the SDK, to upload the file in large blocks, in multiple threads, to improve the file upload efficiency.

The storage requirements are as follows:

- The enterprise follows a digitalization developmental path and the file storage requirements grow.
- Requirements of data backup and file sharing databases for file storage.
- Various types and huge scale.
- The database blog storage adopted is an anti-pattern.
- HA, performance and cost are key elements of consideration.

The specific system architecture is shown in the figure below.

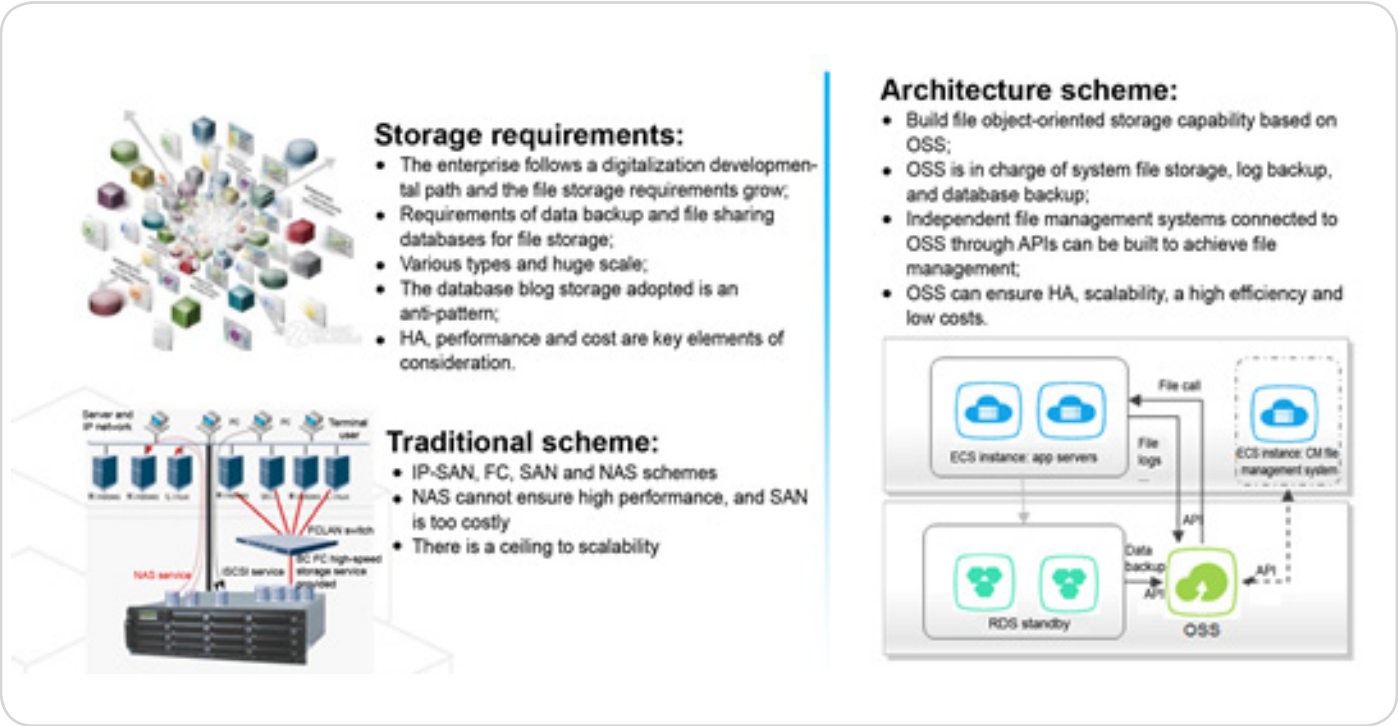


Diagram 9: File Storage Transformation

Architecture Scheme:

The build file object storage capability is based on OSS. Here, OSS is in charge of system file storage, log backup, and database backup. The independent file management systems connected to OSS through APIs can be built to achieve file management. This enables OSS to ensure HA, scalability, a high efficiency and low costs.

8.4 Relational Database Transformation

8.4.1 Architecture Transformation of OLTP-type Relational Databases

For systems with a Low Access Volume:

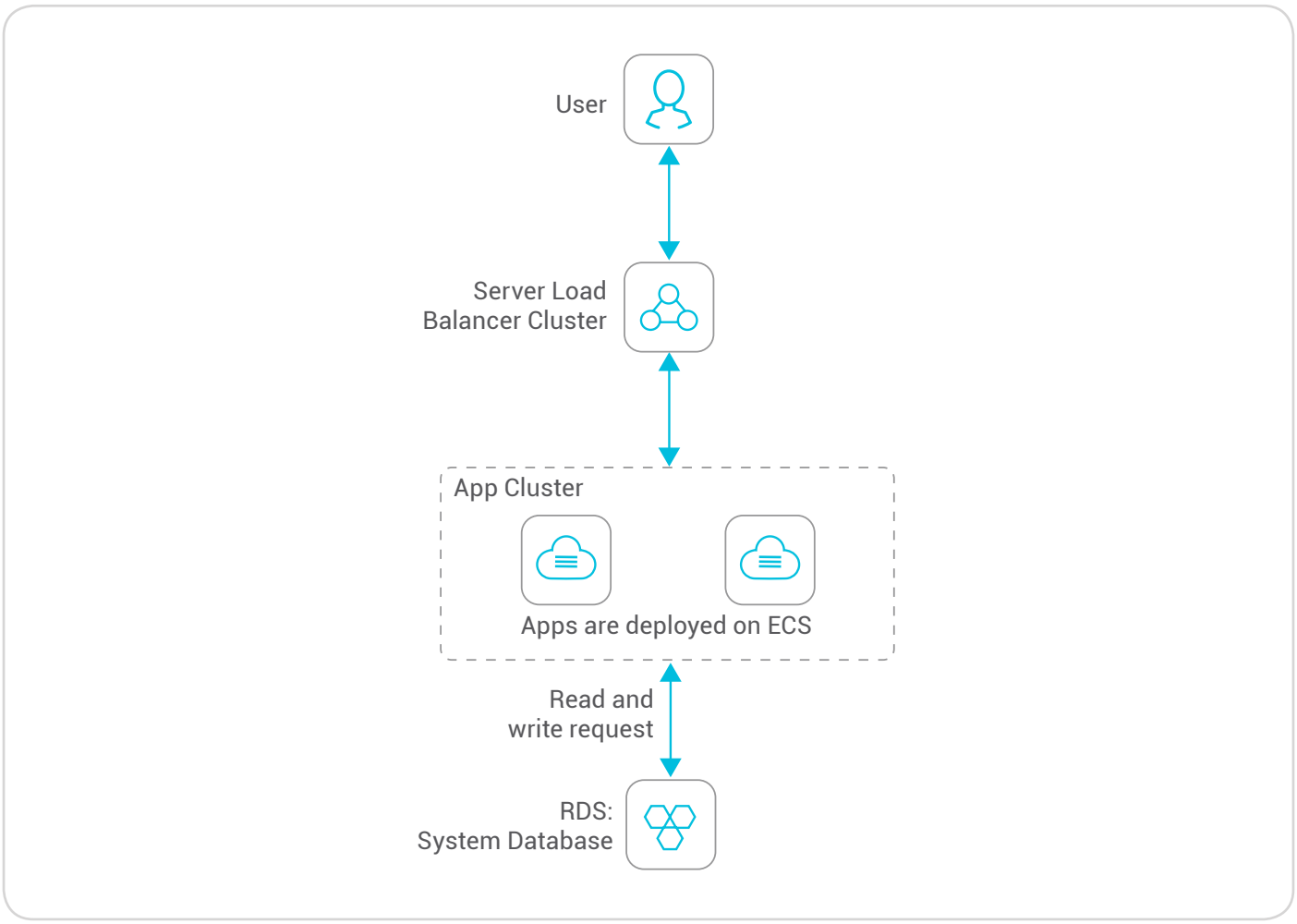


Diagram 10: OLTP-type Relational Databases for Systems with Low Access Volume

With ever-increasing business (database) access pressures, a single RDS is unable to meet the business read and write requests, especially for systems with excessive read pressures. In such cases, consider using cache systems (like OCS, similar to the Memcached service) to share the read pressure and provide quick access through caching highly read data, or consider transforming the database into a read-write application architecture. The RDS product supports read-write splitting. For data analysis scenarios, read-write splitting is an effective solution, as shown in the figure below.

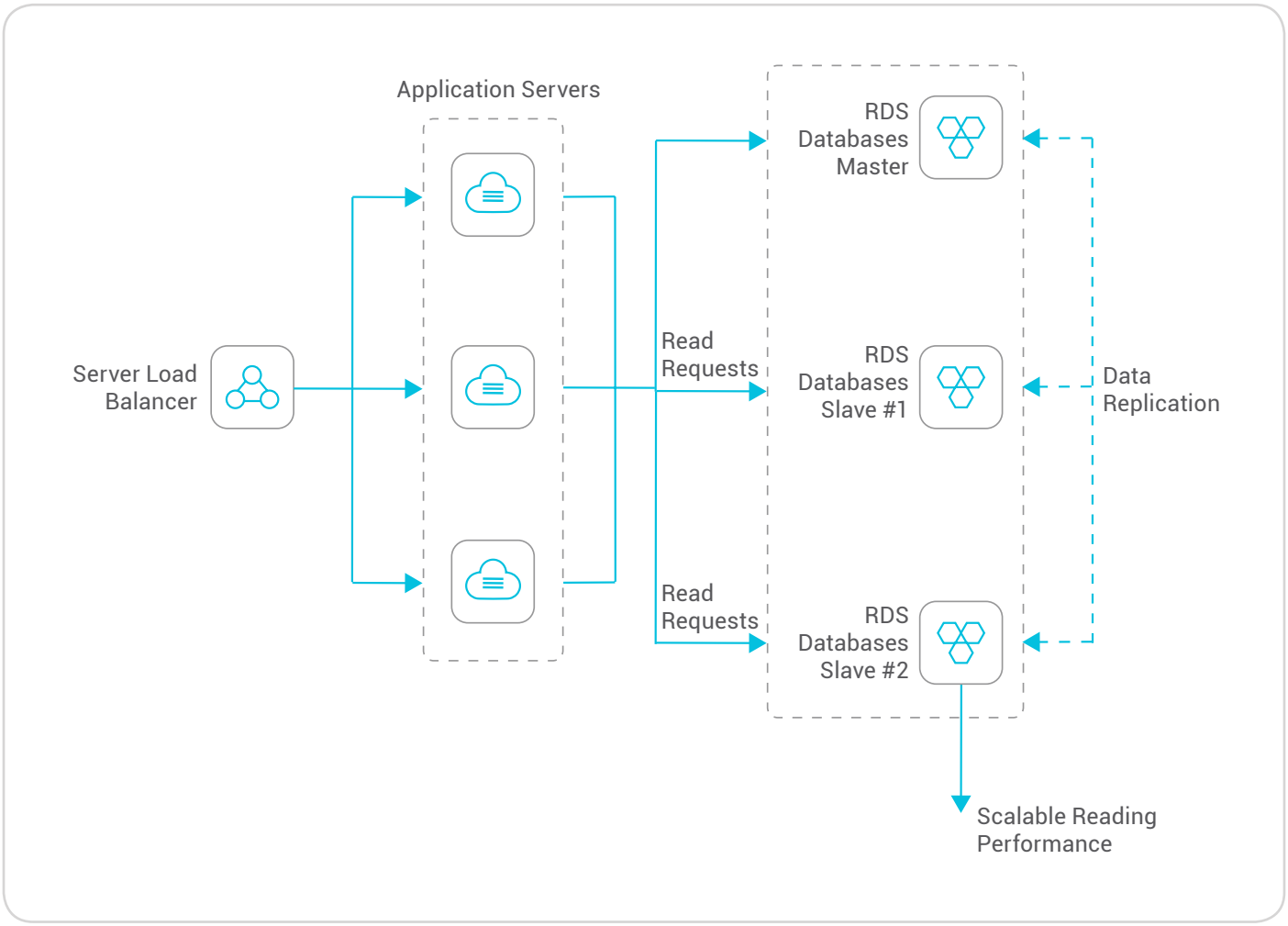


Diagram 11: OLTP-type Relational Databases for Systems with Low Access Volume - Read-write Splitting Scenario

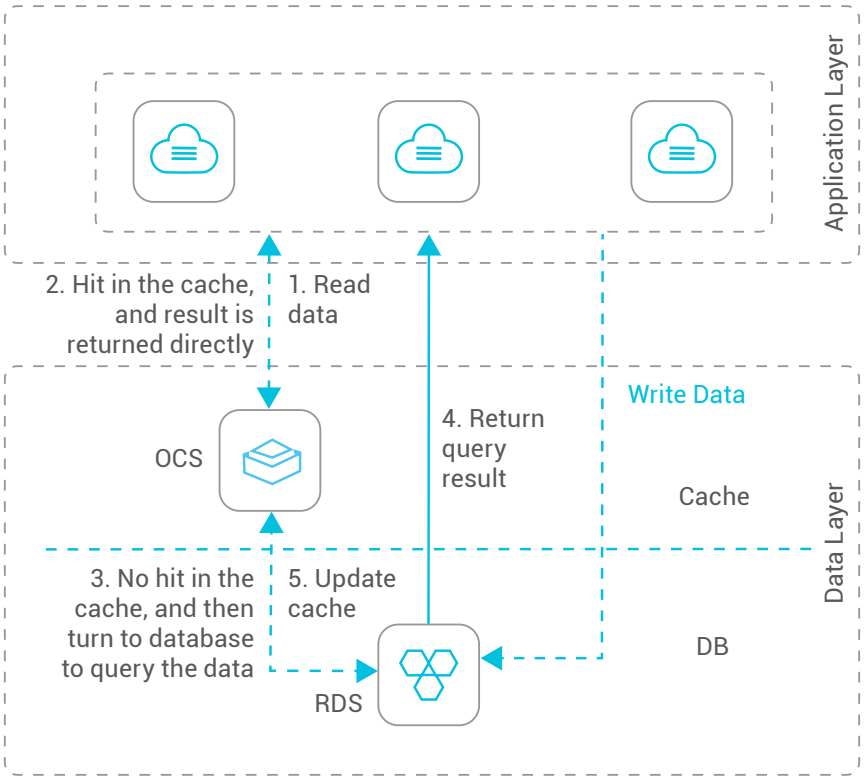


Diagram 12: OLTP-type Relational Databases for Systems with Low Access Volume - Scenario with Cache System Used

Considering that the RDS capacity should not exceed the 2TB limit, and considering the performance constraints of a single database instance when the single database instance cannot meet the storage space requirement, or the writing TPS draws closer to the ceiling capacity of the database, the database usually needs scale-up and scale-out splitting. Vertical splitting (splitting data into different databases, according to different businesses) is comparatively less difficult, while horizontal splitting (splitting a single business) is usually tougher to carry out. Database and table based splitting approaches, distributed transactions, cross-database joins, the global uniqueness of the ID, and the high global availability among other variables are worthy of attention.

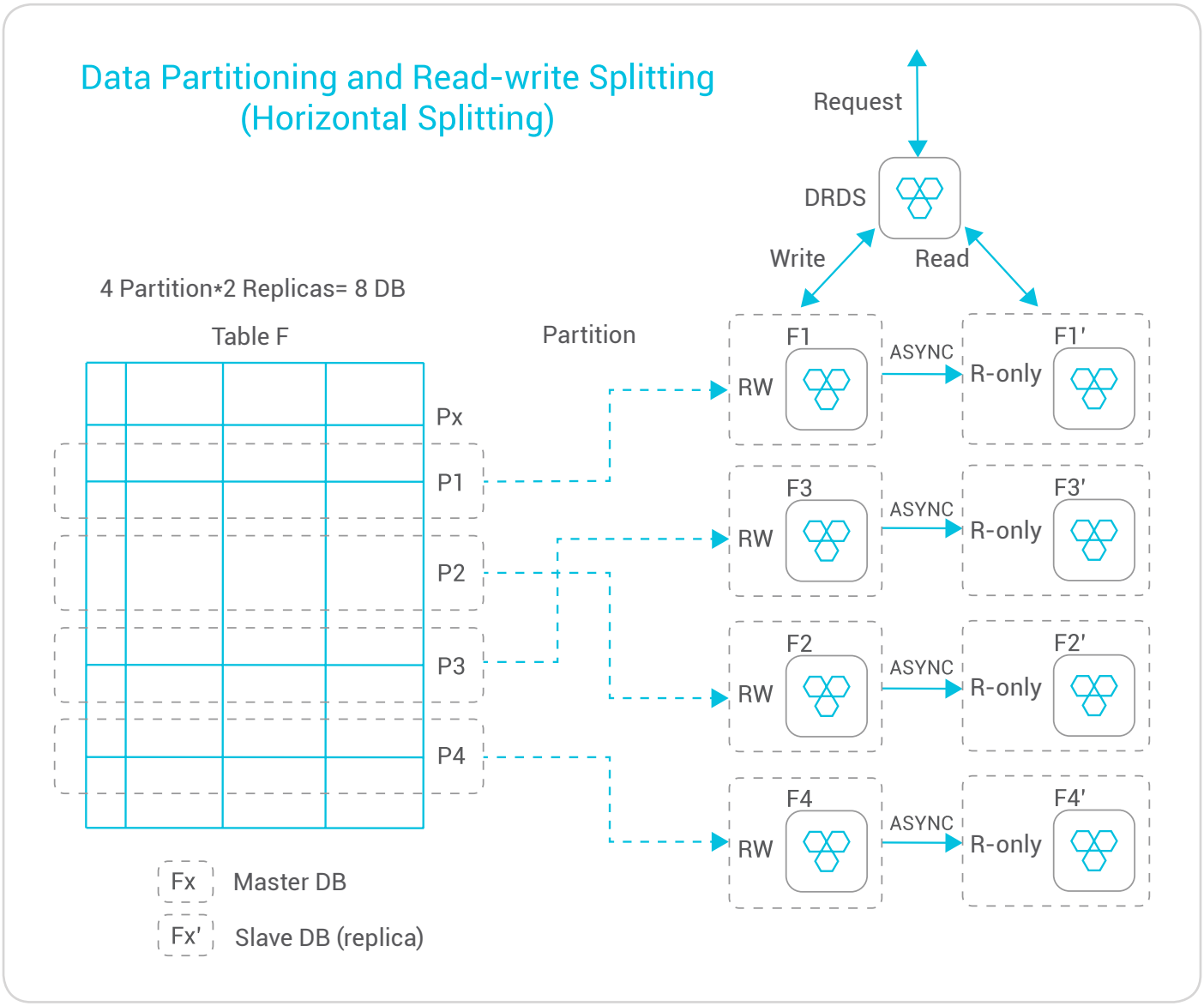


Diagram 13: OLTP-type Relational Databases for Systems with Low Access Volume - Data Partitioning and read-write Splitting Scenario

The cloud platform provides off-the-shelf distributed database services (DRDS) that allow transparent horizontal and vertical splitting of databases to the application. It significantly reduces the cost of splitting, and using it directly is an option worth considering.

8.4.2 Architecture Transformation of OLAP-type Relational Databases

The On-Line Analytical Processing (OLAP) system is the primary application for data warehouse systems designed to support complex analysis operations. It focuses on decision-making support for decision makers and enables complex query processing with a large volume of data quickly and flexibly, according to the analyst's requirements. The query

results are accessible to decision makers in an intuitive and easy to understand form so decision makers can accurately grasp the enterprise's business conditions, understand the customer's needs, and develop appropriate operation plans. The cloud platform offers various solutions for different sizes of OLAP-type applications.

Small-scale systems are OLAP systems, but they only perform real-time analysis for a specific type of a business's historical data. The data size is several hundred gigabytes, and the quantity of data analysis dimensions is a dozen or less. For such application systems, we use RDS or ODPS databases directly and OLAP analysis tools built on the databases. The large-scale real-time analysis is also a type of OLAP system designed for a data storage scale of the 100TB level, and the number of data records in a single table may reach into the hundreds of billions. To achieve real-time analysis of large-scale data sets users should utilize Alibaba Cloud Analytic DB.

- **Small-scale Analysis Systems:** Small-scale OLAP systems target analytical applications with a data size smaller than a few hundred gigabytes and the number of data records in a single table less than 10 million. For such applications, one may consider using RDS as the data analysis database, and deploying the OLAP analysis tools or applications on the RDS, to build the OLAP capabilities.
- **Large-scale Real-time Analysis System:** For large-scale real-time analysis, Alibaba Cloud provides Analytic DB. Analytic DB is an independently developed cloud computing service by Alibaba that targets real-time high-concurrency online analysis of massive data sets (real-time OLAP). It is capable of instant multidimensional analyses and business explorations for several hundreds of billions of data records within seconds. ADS boasts free calculations of massive data sets and speedy responses to massive requests, allowing you to explore data flexibly in an instant, and to find data values quickly. Ideally, you can embed Analytic DB into the business system to provide analytic services for end user clients.

  BI products such as BIEE, QlikView, and Cognos usages are required, but analysis of the on-cloud data (in place of Teradata, SAP Hana, GreenPlum, and other engines) is desired.

  a. Operations depend on massive data analyses, but the computing speed and response time of existing products cannot meet the system's operational requirements.

  b. Cloud internet companies wish to boost business growth from massive data analyses, particularly applications that hope to achieve expansion in user theme applications and geographical theme applications. Other scenarios that require substantial data (more than 100 million data records) processing in real-time (responses within three seconds) and high-availability solutions.

The scenario analysis is described below:

1. The target data queries are large in size
2. The amount of data increases rapidly, posing a high scalability requirement

3. The amount of data increases rapidly, posing a high scalability requirement, for example, the query to tens of billions of data records should be completed within seconds
4. Data model is required to be flexible
5. High availability should be ensured

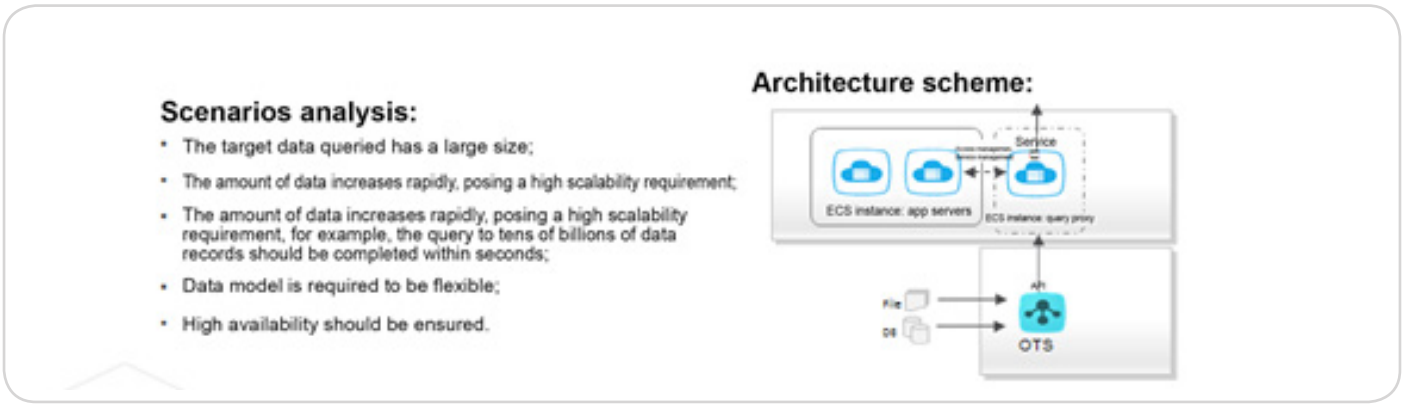


Diagram 14: OLAP-type Relational Databases - Large-scale Real-time Analysis System

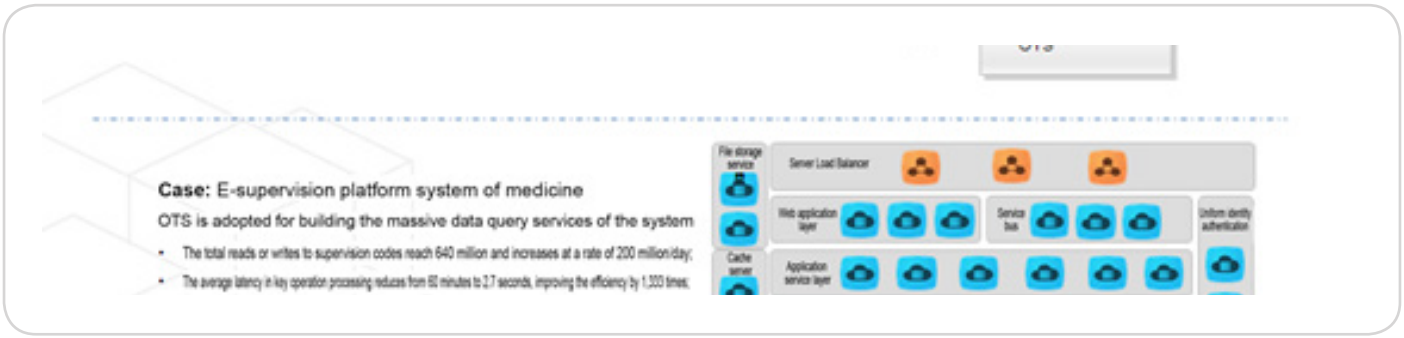


Diagram 15: OLAP-type Relational Databases - E-Supervision Platform System of Medicine Architecture

8.5 Service-oriented Transformation

The messaging middleware may be used to design the service-oriented architecture of the system, such as the service-oriented architecture of the Taobao series:

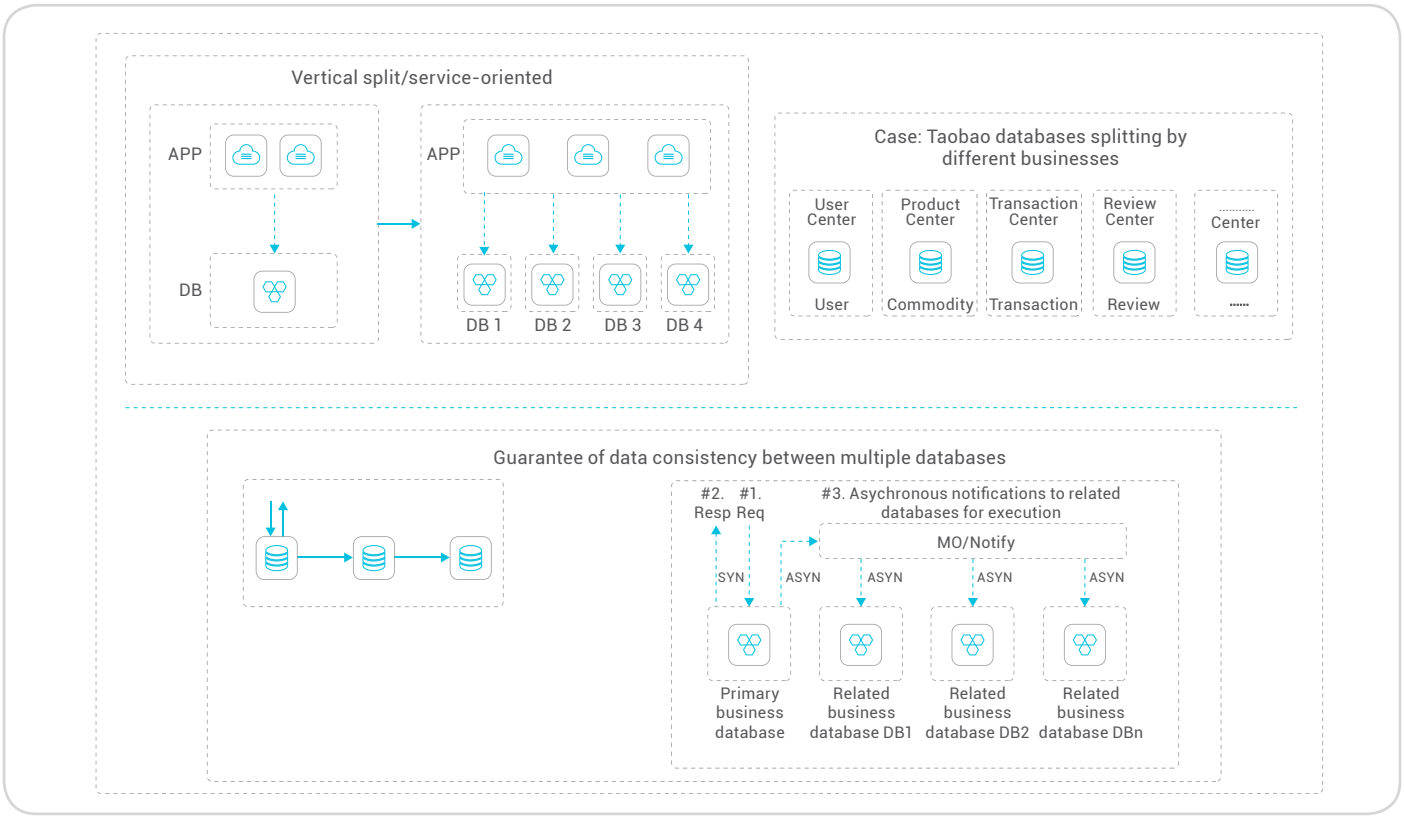


Diagram 16: Service-oriented Architecture of Taobao Series

Major scenarios of messaging middleware are briefly mentioned below:

- **Distributed transactions:** Provides distributed transaction support for service-oriented architectures (SOA) and ensures global data consistency.
- **Delay queue:** The messaging middleware serves as a reliable delay queue and a timer in a distributed environment.
- **Broadcasting and notifications:** Reliable in-cluster broadcasting notifications as well as to notify the existence of invalidated caches or other events.
- HSF-based communication service architecture.

Below is the Alibaba Taobao series basic architecture for your reference.

Diagram 17: Basic Architecture of Taobao

09 Common Scenarios with Recommended Solutions by Alibaba Cloud

This section focusses on the recommended Alibaba Cloud architectures for common scenarios.

9.1 Scenarios with Low Access Concurrency to the System

When the application has low access concurrency, with QPS/TPS less than 500, and a moderate data size, namely less than 1TB, but has complex features, such as cloud architecture designs for administrative examinations and approvals, office collaborations, OA, and other systems.

- Use the Server Load Balancer for system load balancing to achieve high availability.
- Deployment of applications, or middleware, occur on the ECS cluster (we recommend at least two ECS servers).
- Deploy the database service on RDS.

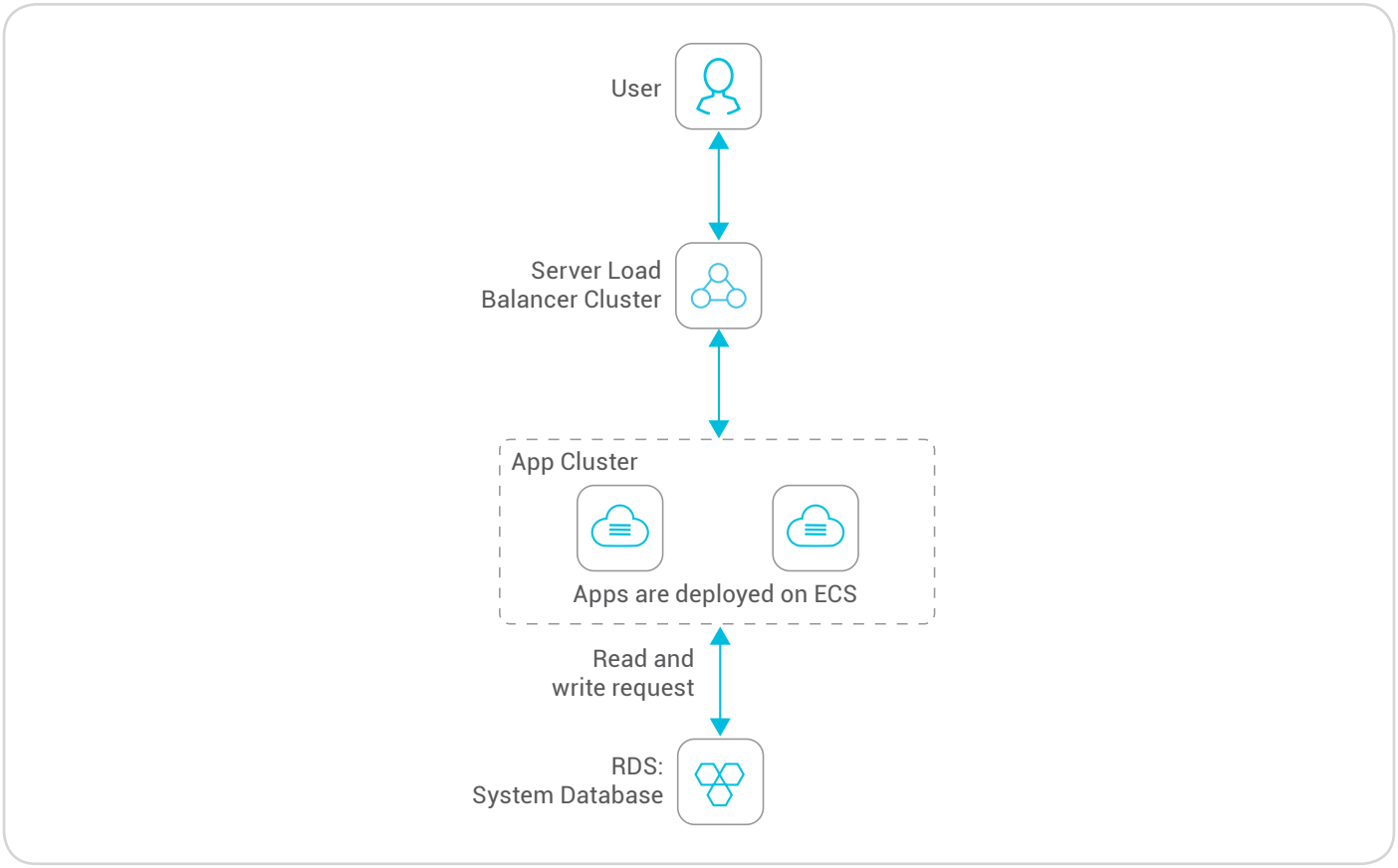


Diagram 18: Scenario with Low Access Concurrency to the System

9.2 Scenarios with Moderate Access Concurrency to the System

When the application has moderate access concurrency, with QPS/TPS between 500 and 2000, and a moderate data size of less than 1TB, and requires video and image storage, such as in health, tourism, and websites.

Below is the recommended system architecture diagram:

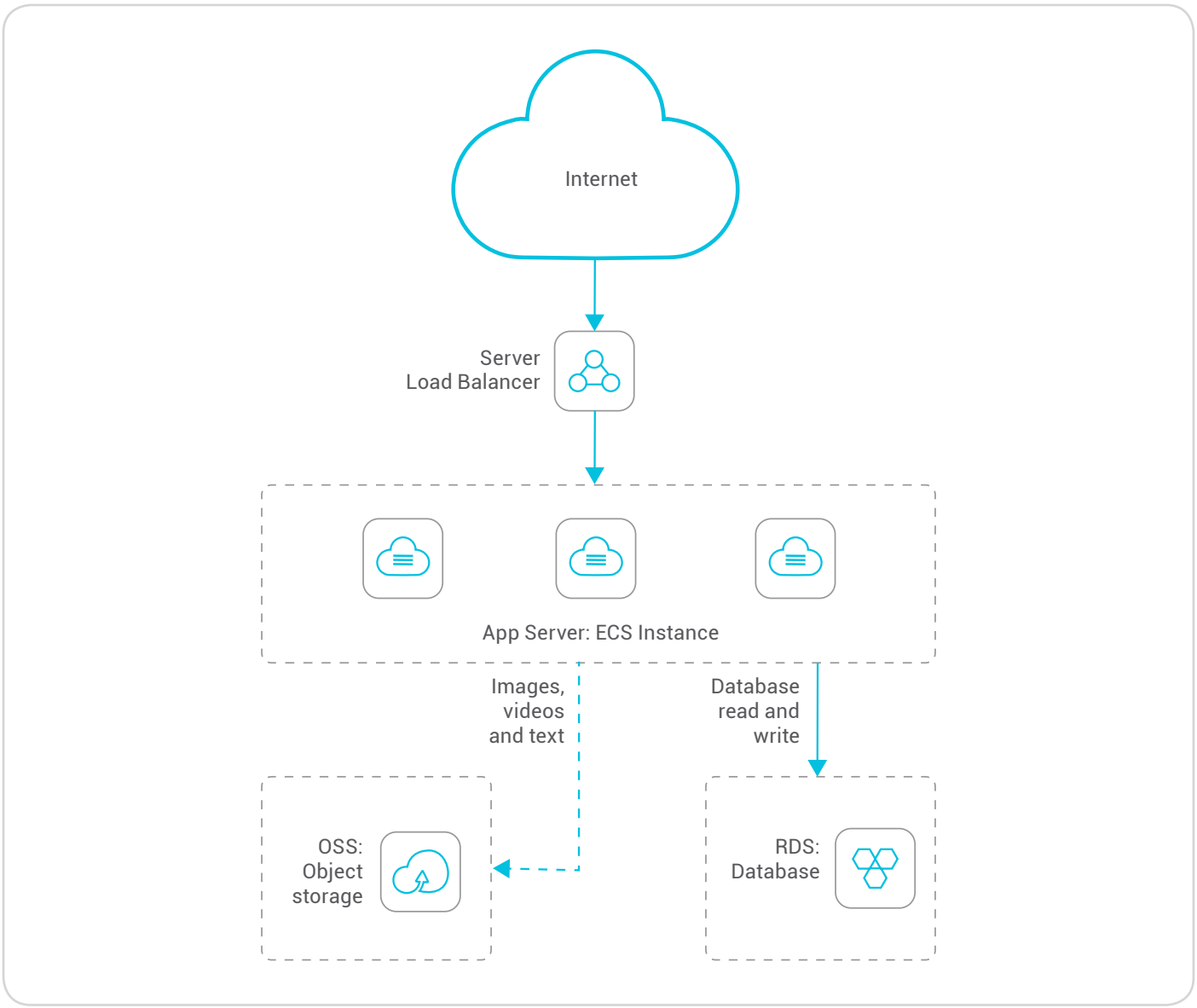


Diagram 19: Scenario with Moderate Access Concurrency to the System

- Use the Server Load Balancer service to achieve high system availability.
- Deploy applications, or middleware, on the ECS (we recommend at least two ECS servers).
- The deployment of the database service takes place on RDS.
- Use OSS to store texts, videos, and images.
- If pressure on the database is high, consider using RDS read-only instances to share the read pressure.

9.3 Scenarios with High Access Concurrency to the System

When the system has moderate access concurrency, with the QPS/TPS higher than 2000, and a database capacity higher than 1TB, and applications that require BI and data warehouses such as scenarios in transportation, environmental protection, and weather systems, the system architecture is as follows:

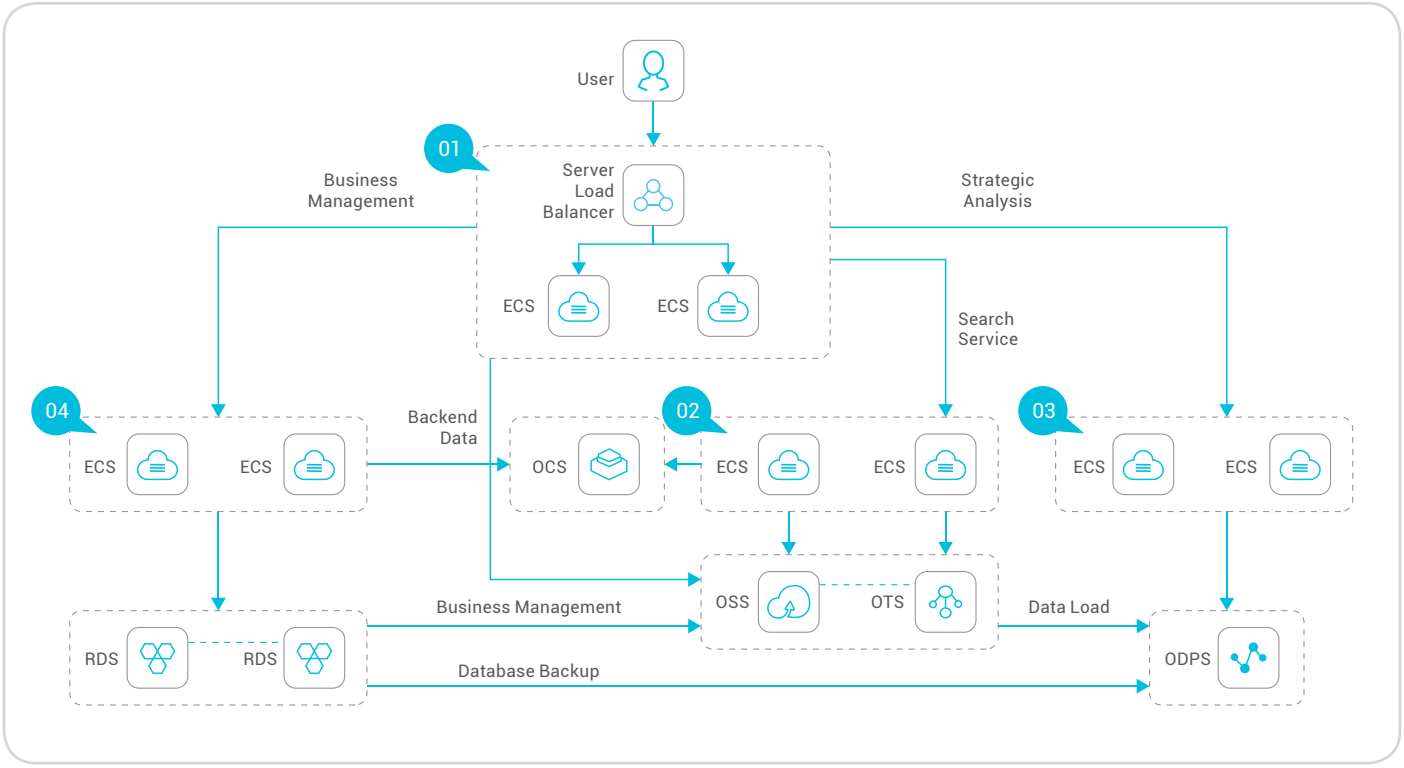


Diagram 20: Scenario with High Access Concurrency to the System

- **Load balancing layer:** Server Load Balancer provides traffic load balancing.
- **Web service layer:** Web server cluster on an ECS instance is flexible enough to allow expansion (we recommend at least two units).

- **Application service layer:** The application server cluster running on the ECS instances. The application servers are grouped by function and adopt Server Load Balancers for load balancing and are flexibly scalable (we recommend at least two units).
- **Data storage layer:** A highly available, scalable RDS/DRDS relational database service based on MySQL databases. OSS supports massive unstructured distributed storage, and OTS supports high-performance distributed queries to massive structured data sets. In addition, ADS supports real-time big data analysis and queries and ODPS supports data warehouses, as well as big data storage and analysis.
- **Data synchronization service:** CDP on cloud data pipelines open up the data products' channels on the cloud, achieving data exchanges between cloud data products.

9.4 Data Warehouse Scenarios

Common processing procedures of data warehouse businesses are as follows:

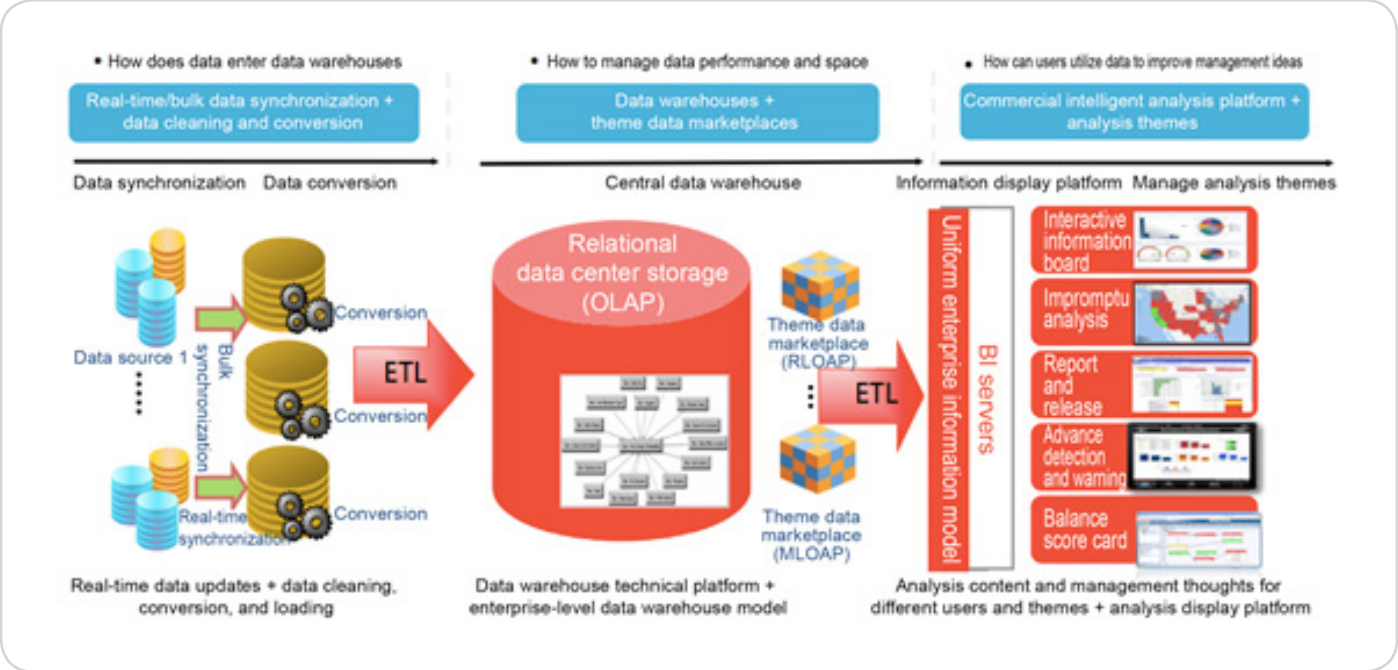


Diagram 21: Common Data Warehouse Business Process

The ETL is responsible for synchronizing data (such as relational databases, file storage systems, and other data storage systems) stored in the online business system to the central data warehouse. Data cleaning and conversion, dimensional modeling, offline computing, and data mining is achieved in the central data warehouse. Finally, synchronize the cleaned, converted, and computed data to the online database system on the data presentation layer

using the ETL tool to fulfill the report presentation, instant query, data analysis, mining, and estimation functions. Through secure access control of the application system, the system incorporates user access permissions. The implementation of all the above data warehouse system functions and processes occur on the Alibaba Cloud platform.

The figure below shows the architecture and implementation of data warehouse businesses based on the Alibaba Cloud platform.

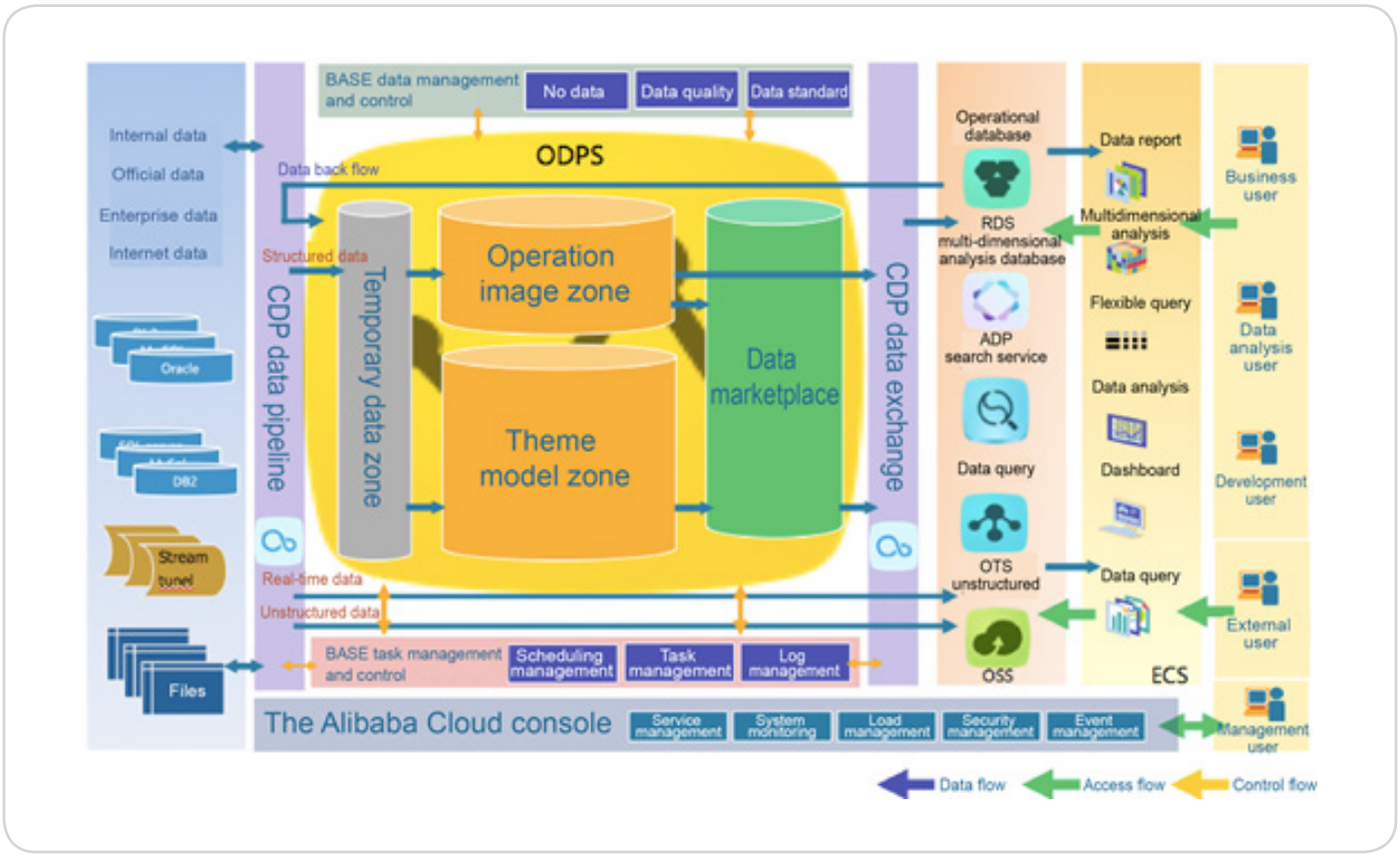


Diagram 22: Alibaba Cloud Architecture - Data Warehouse Businesses

10 Conclusion

Unlike traditional architecture, the cloud ensures high availability of services through its distributed architecture. Through its dynamic horizontal scalability, it meets the constantly increasing demands of operations while integrating a series of O&M services including backups, monitoring, HA, and auditing. The cloud platform is accessible in the form of immediately available services, and organizations can purchase the product for immediate use without going through a series of complicated underlying O&M work. This allows organizations to focus on their core business and research and development (R&D).

Through exploring various architecture scenarios, we have outlined the Alibaba Cloud computing platform and its cloud product portfolio. Alibaba Cloud provides an integrated fault management platform and capacity management through the use of technical policies and best practices. We have explored how to transform specific applications to the cloud, following which the cloud service-based architecture applications ensure higher scalability, lower costs, higher performance, as well as the higher availability of the overall system for an organization.

Alibaba Cloud