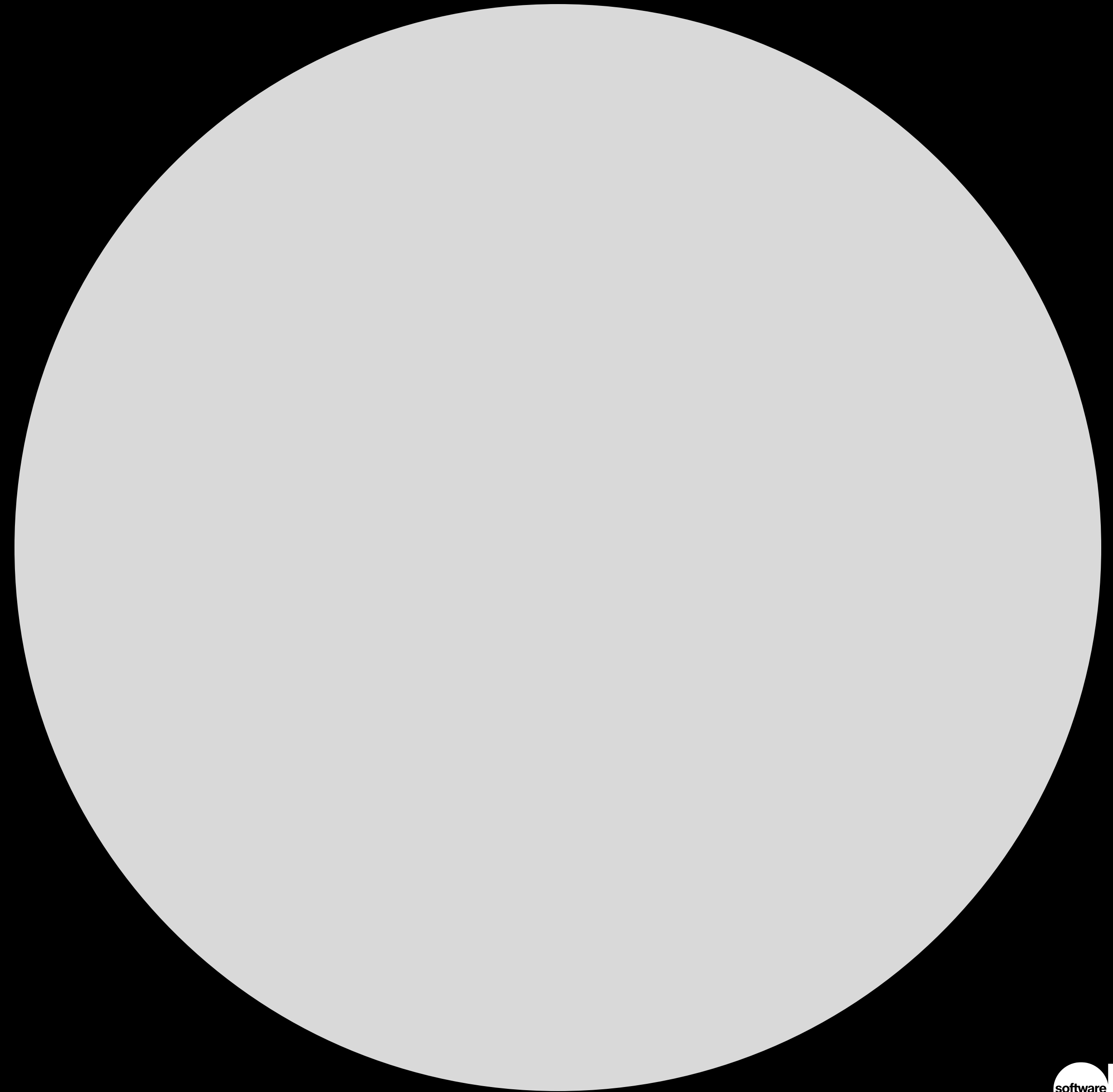




Alibaba Managed Services

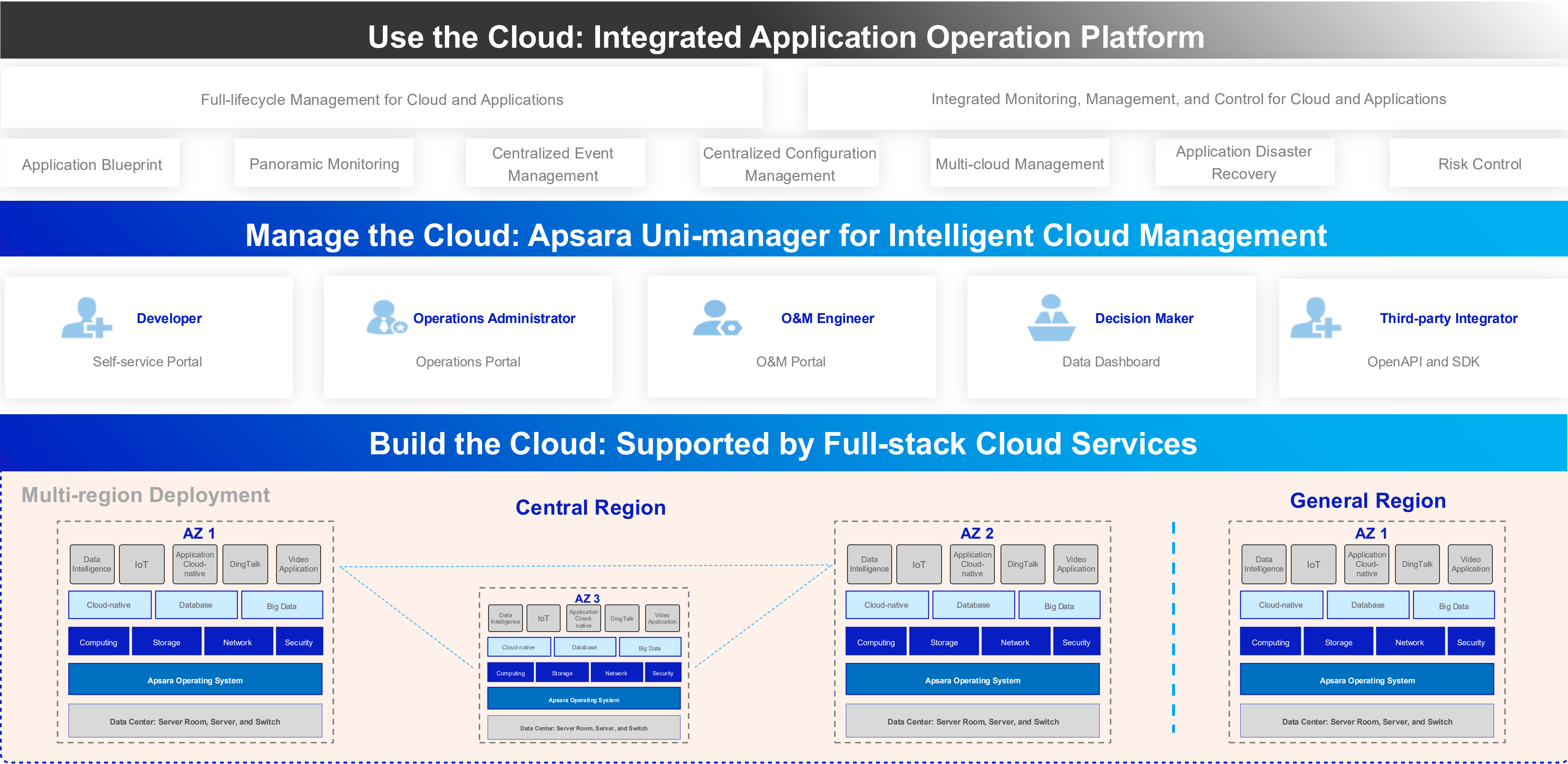
G-Cloud 15

Aspara Stack: Custom Private Cloud



Apsara Stack: Integrate Building, Management, and Use of Cloud and Help Public Sectors and Enterprises for Digital Intelligence Innovation

ALIBABA CLOUD INTELLIGENCE GROUP

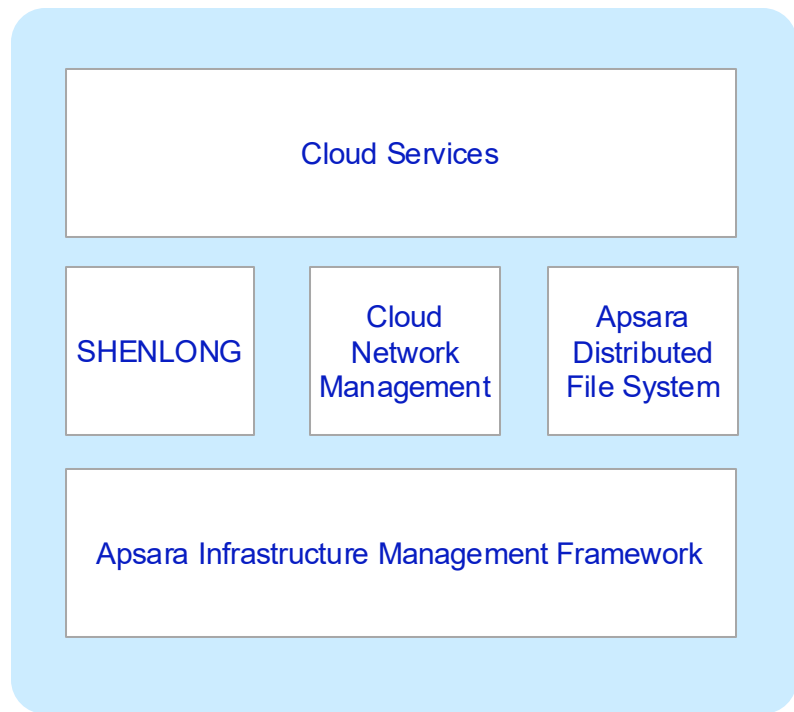


Three Core Competitive Advantages to Build the Capabilities of Apsara Stack

ALIBABA CLOUD INTELLIGENCE GROUP

Unified architecture: cloud technologies proven to be mature by Alibaba Cloud business

Apsara Stack can commercialize both Alibaba Cloud public cloud services and Alibaba Cloud Apsara Stack services and synchronously publish their features based on architecture and platform evolution. It strengthens the capability of unified scheduling and autonomy across data centers. It also provides cloud-edge integrated management, cross-region computing and scheduling, disaster recovery and backup with quick switchover and recovery, and native Apsara Stack in conjunction with Alibaba Cloud public cloud.



Apsara Stack

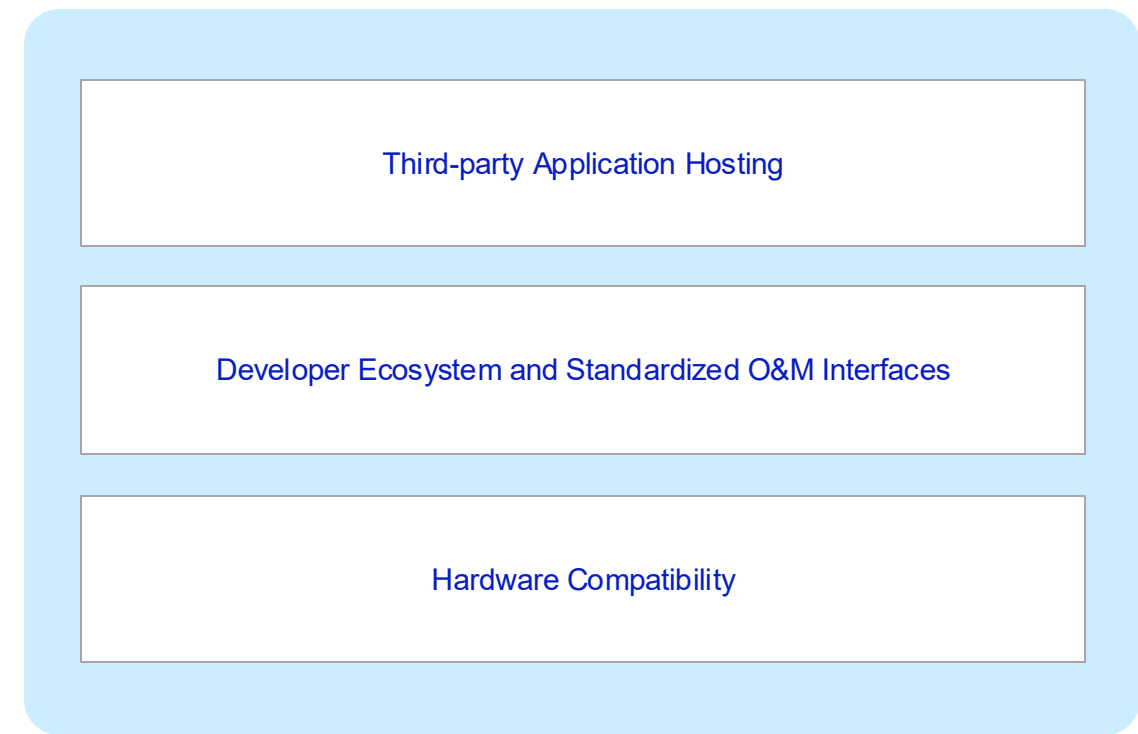
Best practices: develop services based on platform operations and O&M capabilities

Platform for accumulating the experience of Alibaba Cloud and customer operations and O&M cloud platforms, helping customers provide the automated O&M capabilities and digital management and operations capabilities of cloud platforms for business based on the 1-10-30 rule.



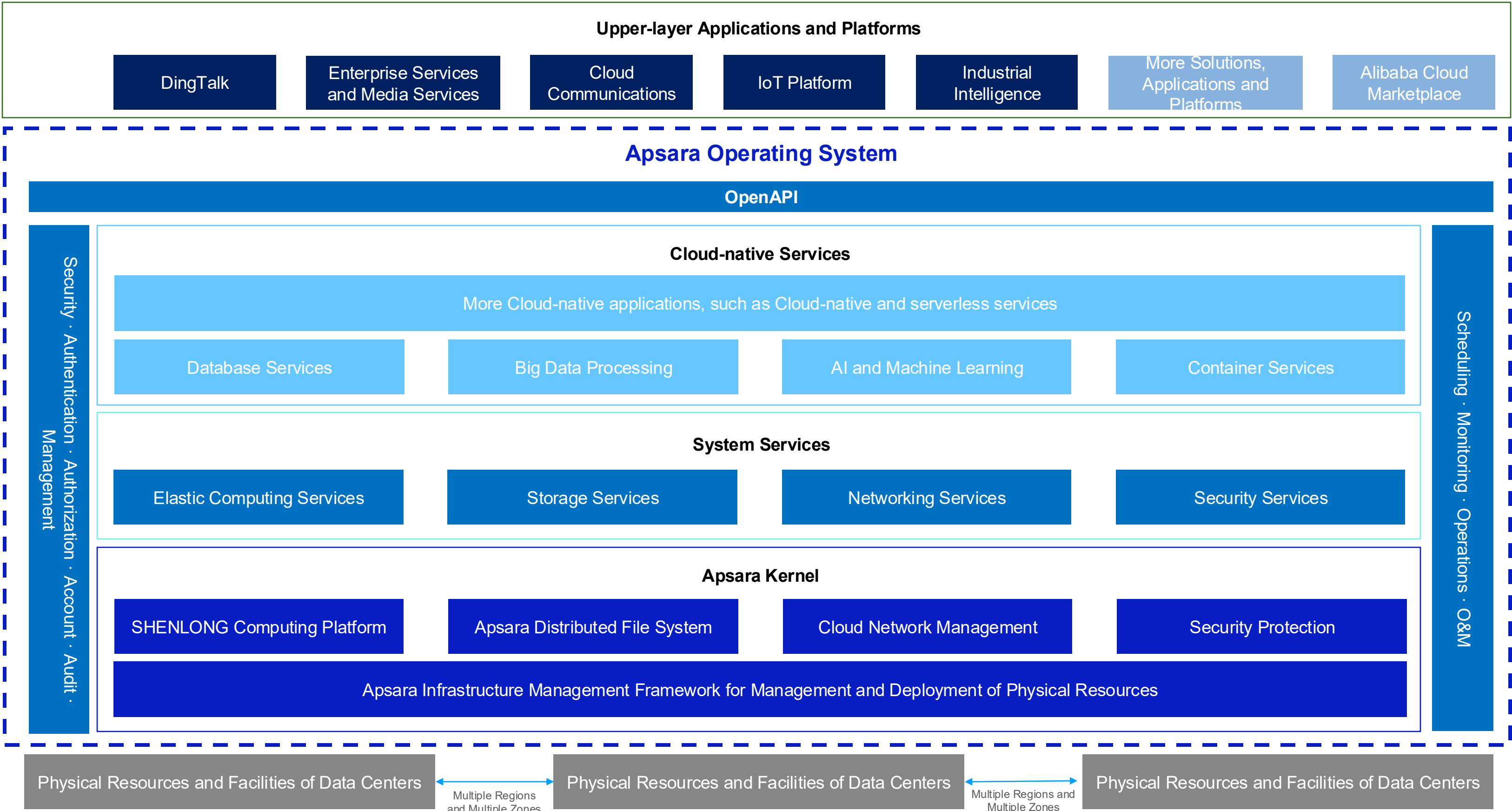
Be integrated: open and compatible enterprise ecosystem

Build an open and compatible Apsara Stack service ecosystem by building cloud platforms and services that offer openness and compatibility. This ensures that all cloud platforms and services are fully compatible with the developer ecosystem, and support more than 80% of mainstream hardware manufacturers and more than 1,000 third-party cloud services of mainstream enterprises.



Apsara Operating System

Apsara operating system is **the core basic software that powers applications on the Internet**. It manages resources in data centers at the underlying layer and provides public services and APIs for the upper-layer services.

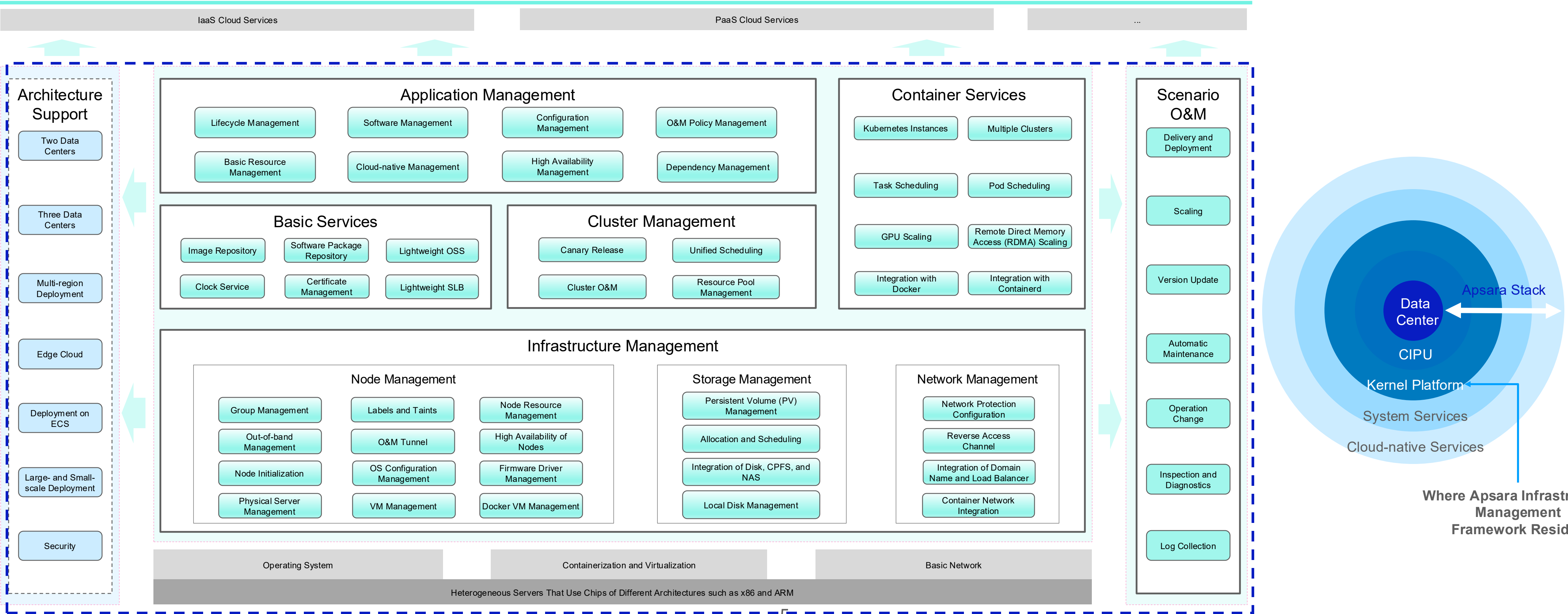


What Is Apsara Infrastructure Management Framework?

ALIBABA CLOUD INTELLIGENCE GROUP

Apsara Infrastructure Management Framework

Apsara Infrastructure Management Framework is a system that is **fully developed by Alibaba Cloud** for **automatic management of cloud computing hardware and software**. It is a **desired-state-oriented** platform that support automatic O&M on clusters and services. As the kernel of the Apsara operating system, Apsara Infrastructure Management Framework provides upper-layer services of the platforms in different deployment modes with capabilities such as **desired state planning, deployment configuration, canary release, resource scheduling, and automatic O&M**. These capabilities simplify and streamline the building of a digital economy base.

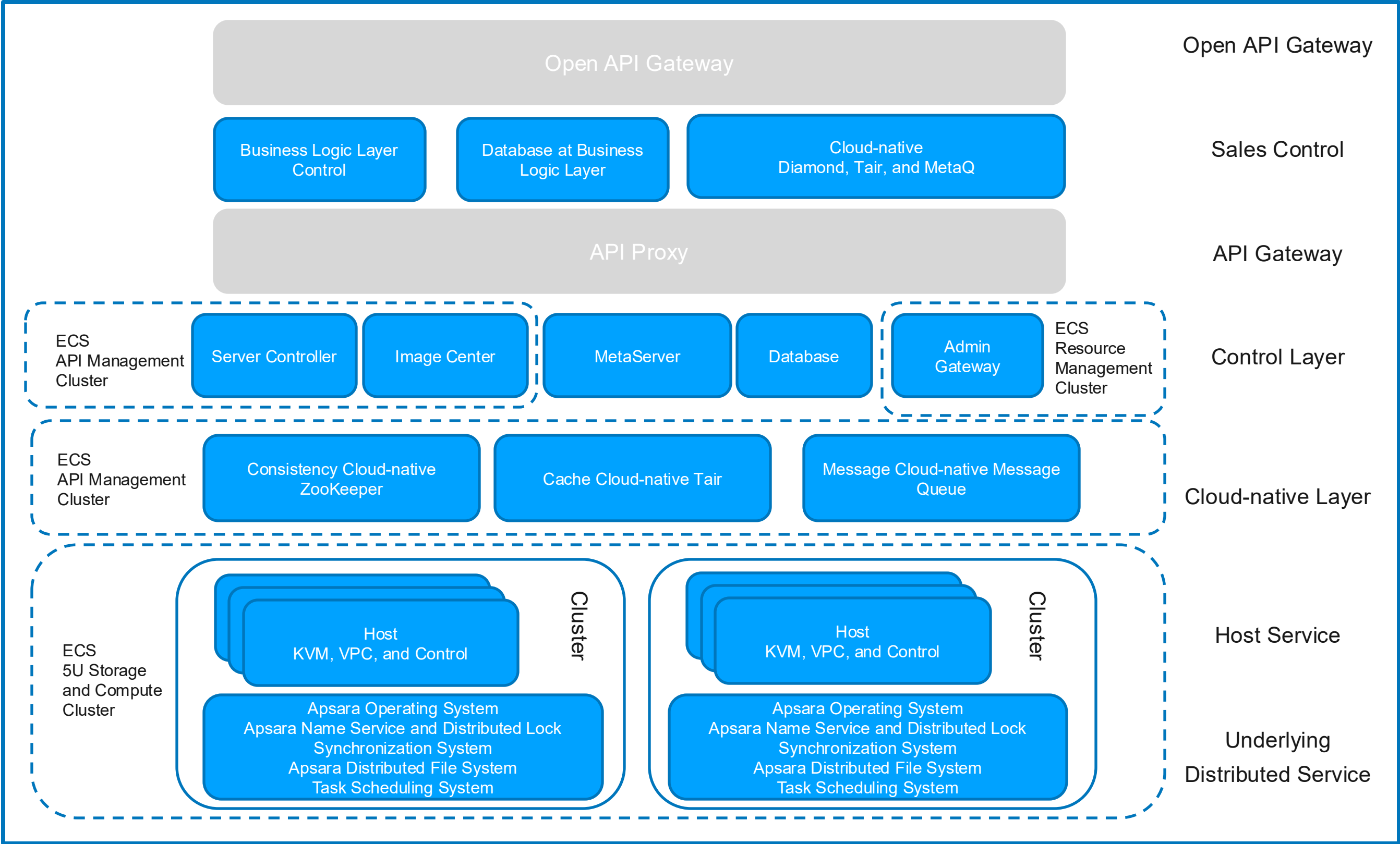


Elastic Compute Service (ECS)

Application Scenarios

- In most cases, ECS is used as the runtime environment for applications and is a required service for Apsara Stack.
- ECS is used together with other cloud computing services.

Architecture



Overview

- ECS is a secure, reliable, elastic, and scalable cloud computing service. It has features such as fast creation, ease of use, excellent performance, high stability and reliability, and reduced IT costs.

Benefits

- **Large-scale deployment:** A single zone supports more than 5,000 ECS instances.
- **Hot migration:** ECS supports hot migration of VMs. A single VM can be hot migrated with a network interruption of less than 200 microseconds.
- **VM configuration change:** You can upgrade or downgrade the configurations of a VM offline. You can upgrade the configurations, including CPU and memory, of a VM online.
- **Crash-consistent snapshots:** You can create snapshots for a VM. You can also create crash-consistent snapshots for the system disk and multiple data disks of the VM.
- **High resource reuse rate:** The number of vCPUs of a shared cluster can be increased by up to three times.
- **Multi-level security protection:** ECS supports isolated VPCs and security groups. The modified configurations of security groups immediately take effect to block established connections. You can specify the priorities of security group rules. ECS supports multi-level password protection for access to ECS instances by using VNC and allows you to change VNC login passwords.
- **Multi-chip architecture in one cloud:** ECS supports Intel, Hygon, Kunpeng and Phytium CPUs. ECS also supports hybrid deployment that uses the chips of different architectures in one cloud, the chips that are created by Chinese vendors, and IT application innovation.

Object Storage Service (CDS-OSS)



Application Scenarios

- Storage of a large amount of unstructured data such as images, short videos, audio, and documents.
- Backup storage and data migration storage.



Overview

- CDS-OSS is a secure, cost-effective, and highly reliable storage service that can store a large amount of unstructured data. CDS-OSS uses RESTful API operations to provide services.

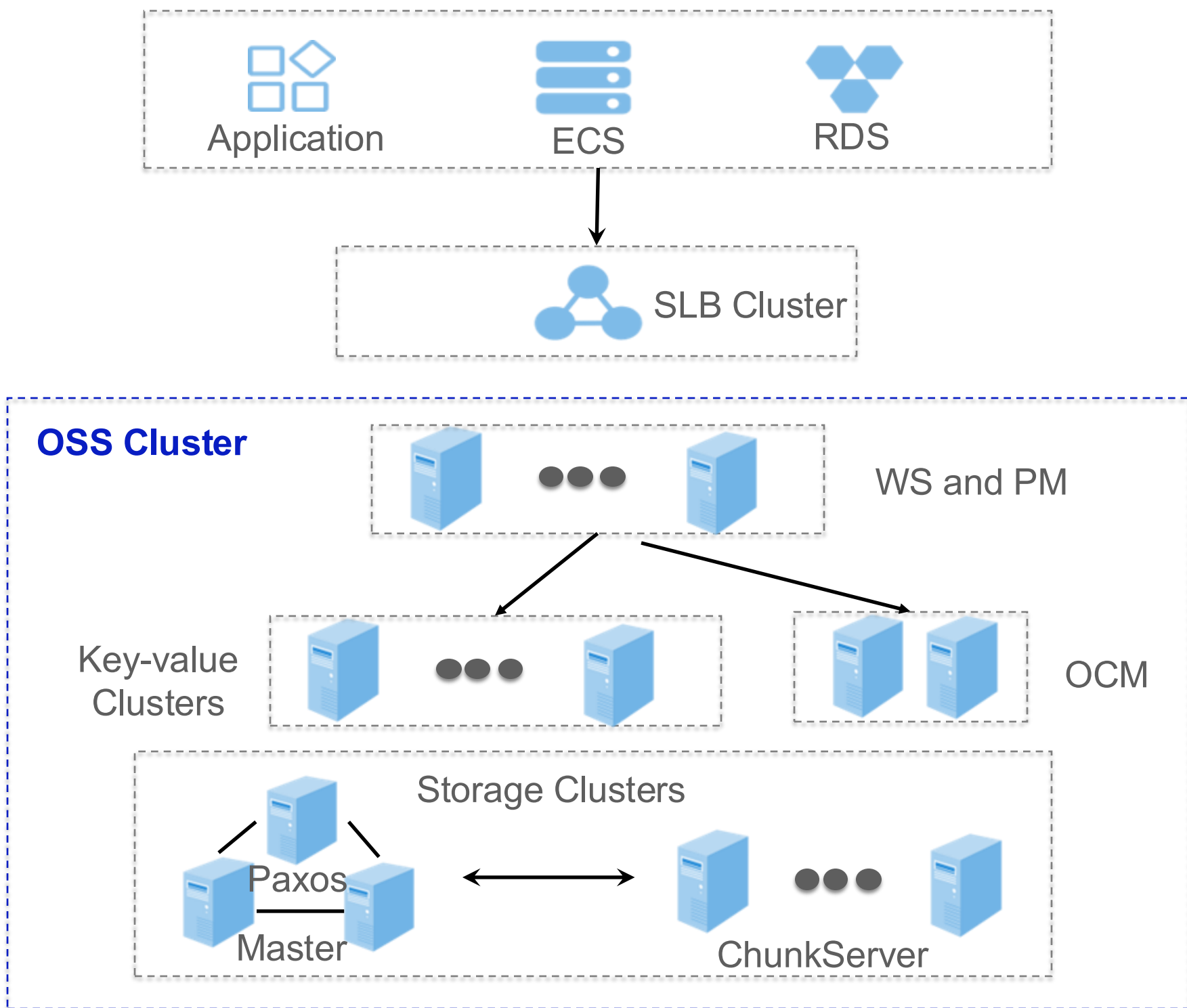


Benefits

- **Large scale:** A single CDS-OSS cluster can store more than 200 petabyte of data and an OSS bucket can store trillions of documents. Tens of millions of files can be queried within 10 minutes. A single node support a throughput of 300 MB/s and 5,000 QPS.
- **Multiple replica policies:** CDS-OSS supports three-replica mode and EC mode, such as 2+2, 8+2, or 20+4 configuration.
- **Version management:** After you enable the version management feature for a bucket, data that is overwritten or deleted in the bucket is saved as a previous version. After an object is overwritten or deleted, the object can be restored to any previous version.
- **Comprehensive access control:** You can configure access methods such as RAM and STS based on users. You can also configure access policies such as a bucket policy based on resources.
- **Data integrity:** MD5 verification is implemented for uploaded and downloaded objects, and CRC verification is implemented for internal data.
- **Disaster recovery:** CDS-OSS supports data replication in the same zone and across zones, and zone- and geo-disaster recovery with a RPO less than 15 minutes.
- **Data encryption:** Before you upload a file to OSS, you can use an OSS SDK to encrypt the file on the client. After the file is uploaded to OSS, OSS encrypts the uploaded file on the server.
- **Support for WORM:** CDS-OSS supports time-based retention policies. After a retention policy is configured for a bucket, objects stored in the bucket cannot be deleted or modified during the retention period.
- **Automatic lifecycle management:** You can configure lifecycle settings for a bucket to achieve automatic deletion of expired objects and parts.
- **Intelligent image and video processing:** CDS-OSS allows you to convert image formats, zoom in or out, crop, and rotate images, add watermarks to images, and take video snapshots.
- **Multi-chip architecture in one cloud:** CDS-OSS supports Intel, Hygon, Kunpeng and Phytium CPUs.



Architecture



Cloud Interconnection

- Provides custom VPC network environment.
- Provides Internet-facing and internal-facing access to business.
- Provides high-concurrency and high-elasticity services by using SLB.
- Provides secure links by using IPsec and SSL VPNs.

On-premises Collaboration

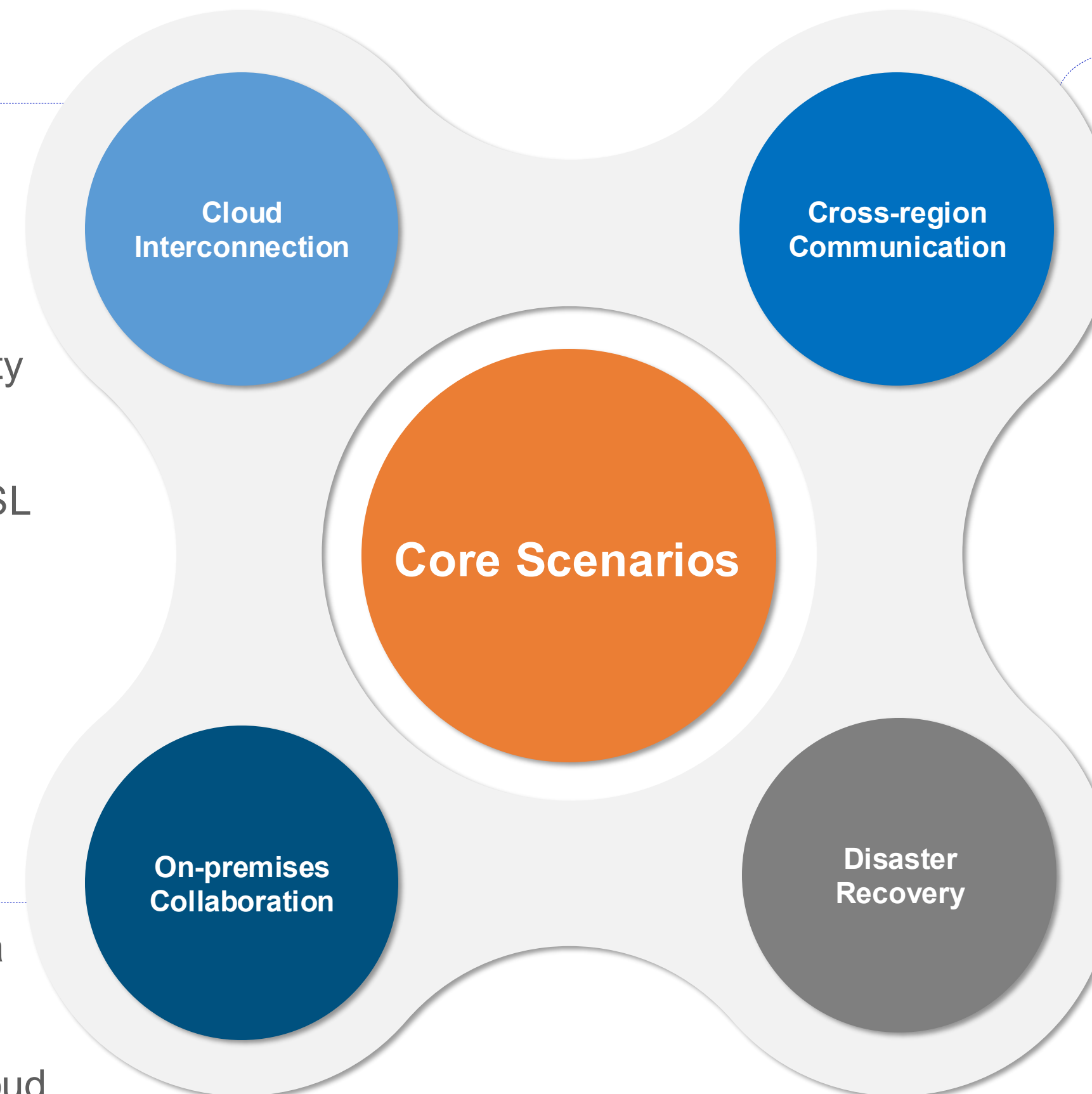
- Provides secure access to on-premises data centers over Express Connect.
- Supports security protection provided by Cloud Firewall on a leased line.

Cross-region Communication

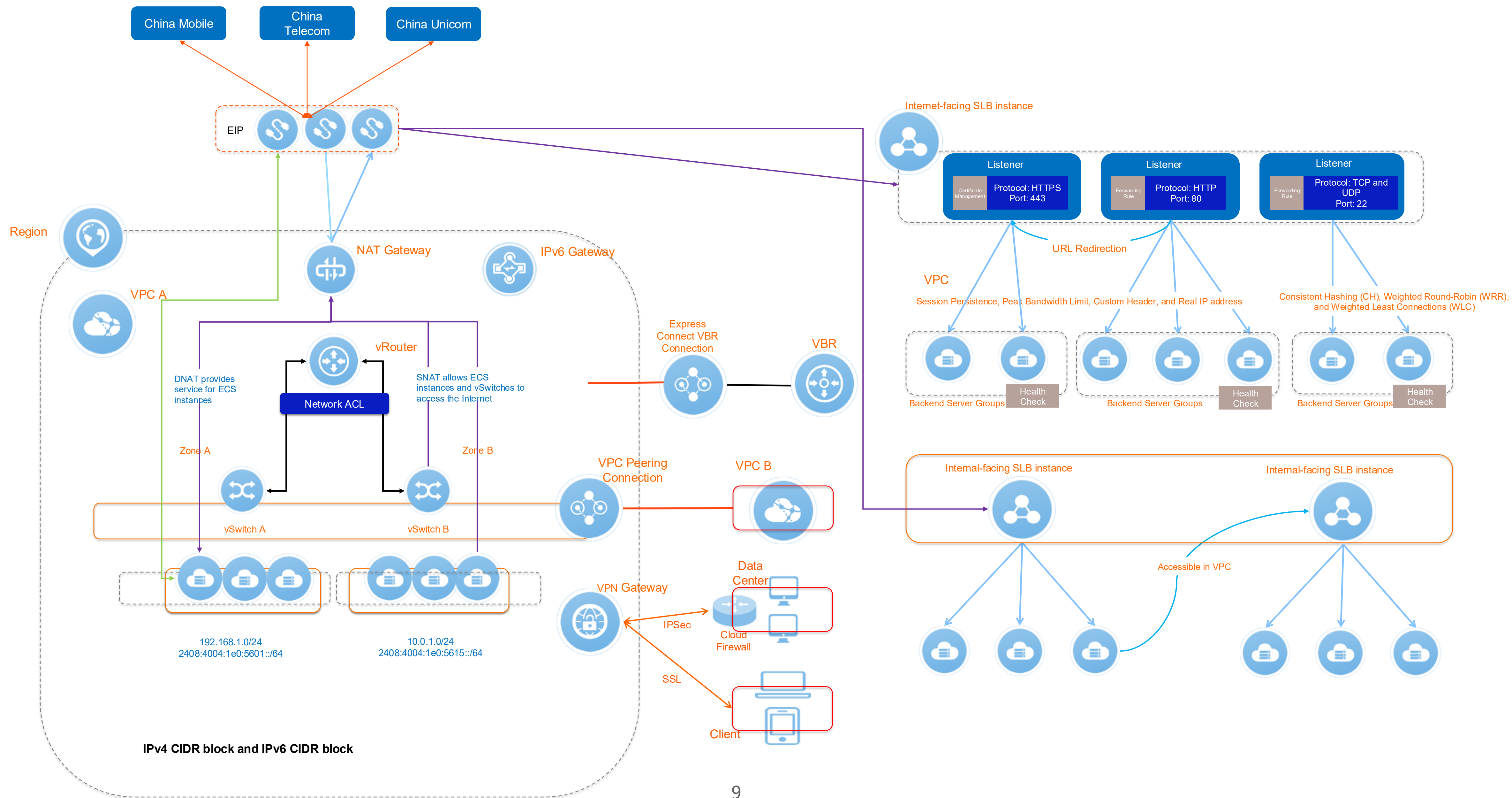
- Provides access to resources in multiple regions over VPC.
- Provides reverse access for cross-cloud communication to meet the requirements of database and big data migration.

Disaster Recovery and High Availability

- Allows business deployment across data centers. This way, the requirements of zone active-active application deployment are met.
- Supports active-active failover in a single region.



Panorama of Networking Services



Virtual Private Cloud (VPC)

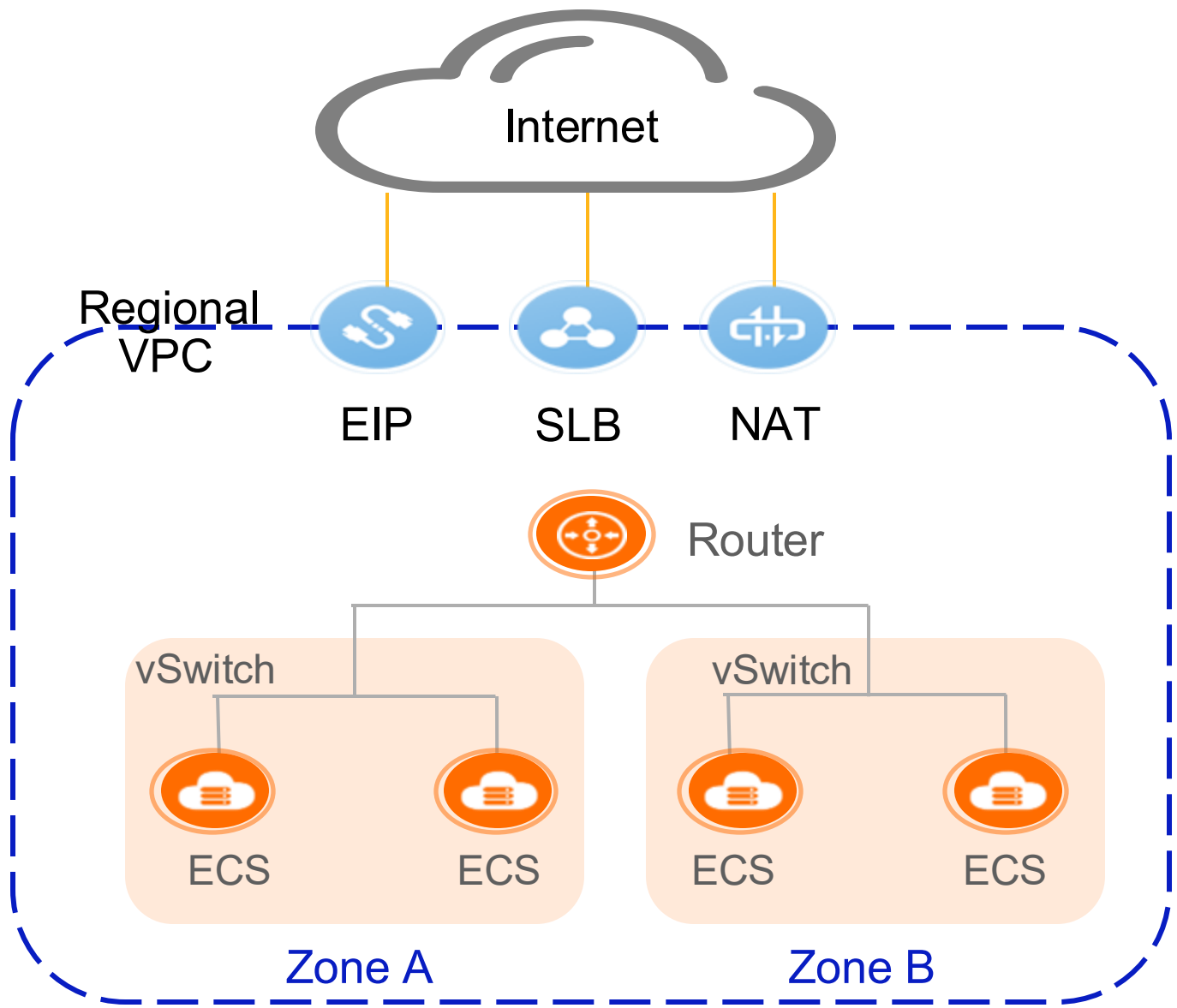


Application Scenarios

- **Isolate business systems:** VPCs are logically isolated.
- **Connect a VPC to the Internet:** You can associate an EIP with a VPC and configure a NAT gateway to connect the VPC to the Internet. This way, cloud services in the VPC can communicate with the Internet.
- **Implement cross-zone disaster recovery:** vSwitches in the same VPC can communicate with each other over the internal network. You can deploy resources in vSwitches in different zones to implement cross-zone disaster recovery.
- **Connect a VPC to another VPC:** You can create a pair of router interfaces to connect a VPC to another VPC to implement high-speed and secure communication over private network.
- **Connect a VPC to an on-premises data center:** You can connect an on-premises data center to a VPC by using an Express Connect circuit to smoothly migrate on-premises applications to the cloud.



Architecture



Overview

- VPCs are private networks deployed on Apsara Stack. VPCs are logically isolated from each other. You can manage your VPCs and Apsara Stack resources in your VPCs.



Benefits

- **Layer 2 isolation between VPCs:** The IP packets of an ECS instance are encapsulated by using the tunneling technology. Therefore, information at the data link layer (the MAC address) of the ECS instance is not transferred to the physical network. ECS instances in a VPC use security groups and firewalls to control inbound and outbound traffic at Layer 3.
- **Large-scale deployment:** By default, 3,000 VPCs can be created in a single region. Each user can create 2,000 VPCs.
- **Support for DPDK:** SLB uses the Data Plane Development Kit (DPDK) network driver to improve network transmission performance by using technologies such as user mode, replacing polling by interrupts, and zero-copy.
- **Automatic route learning:** vSwitches learn routing on demand. The first packet is forwarded to vSwitches. Distributed Layer 3 gateways are supported.
- **Topology visualization:** You can visualize the topology of the tenant network and subnet addresses to view the usage details of your networks.
- **Shared VPC:** You can share VPCs by tenant or resource set. The Alibaba Cloud account that owns a VPC can share the VPC with other tenants of the same organization.
- **Intelligent ENI:** ENIs can be bound to ECS instances and migrated between ECS instances. You can associate an EIP with an ENI.
- **High-availability virtual IP address:** A high-availability virtual IP address (HAVIP) is a private IP address that can be independently created and released. You can use HAVIPs with high-availability (HA) software such as Keepalived to deploy services in active/standby mode. This improves the availability of your services.
- **NAT Gateway:** VPC supports enterprise-class Internet gateways that provide the SNAT and DNAT features to support cross-zone disaster recovery.
- **IPv4 and IPv6:** VPC supports IPv4 and IPv6.

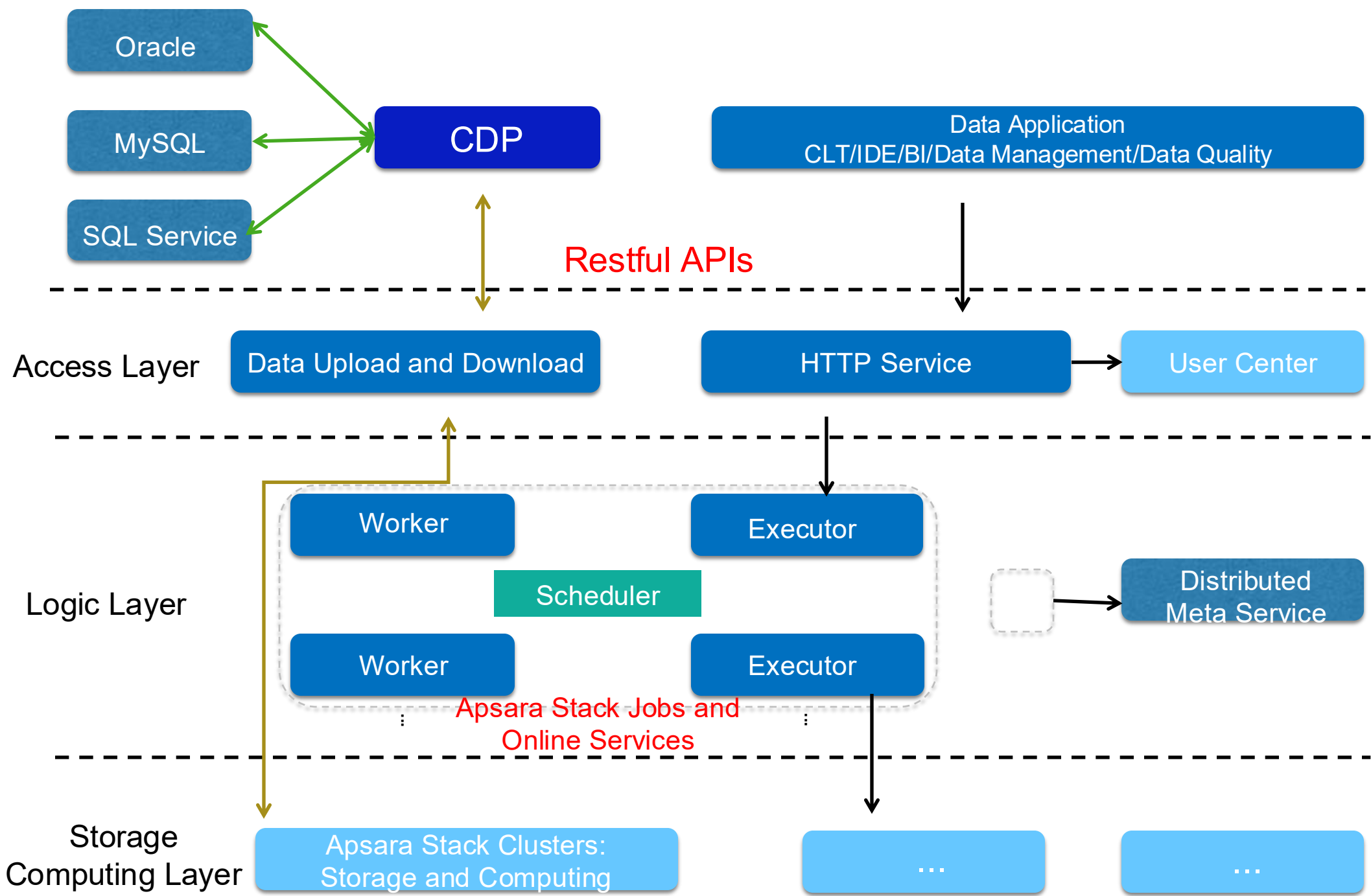


Application Scenarios

- MaxCompute is an enterprise-level cloud data warehouse that provides features for data classification, analysis, and mining. It helps enterprises make business decisions, reduce costs, and improve efficiency.



Architecture



Overview

- MaxCompute is an enterprise-level cloud data warehouse that uses the SaaS model and is suitable for scenarios that require data analysis. MaxCompute provides a fast, fully managed online data warehousing service in a serverless architecture, eliminates the constraints of traditional data platforms in terms of resource scalability and elasticity, minimizes O&M costs, and allows you to efficiently process and analyze large amounts of data in a cost-effective way.



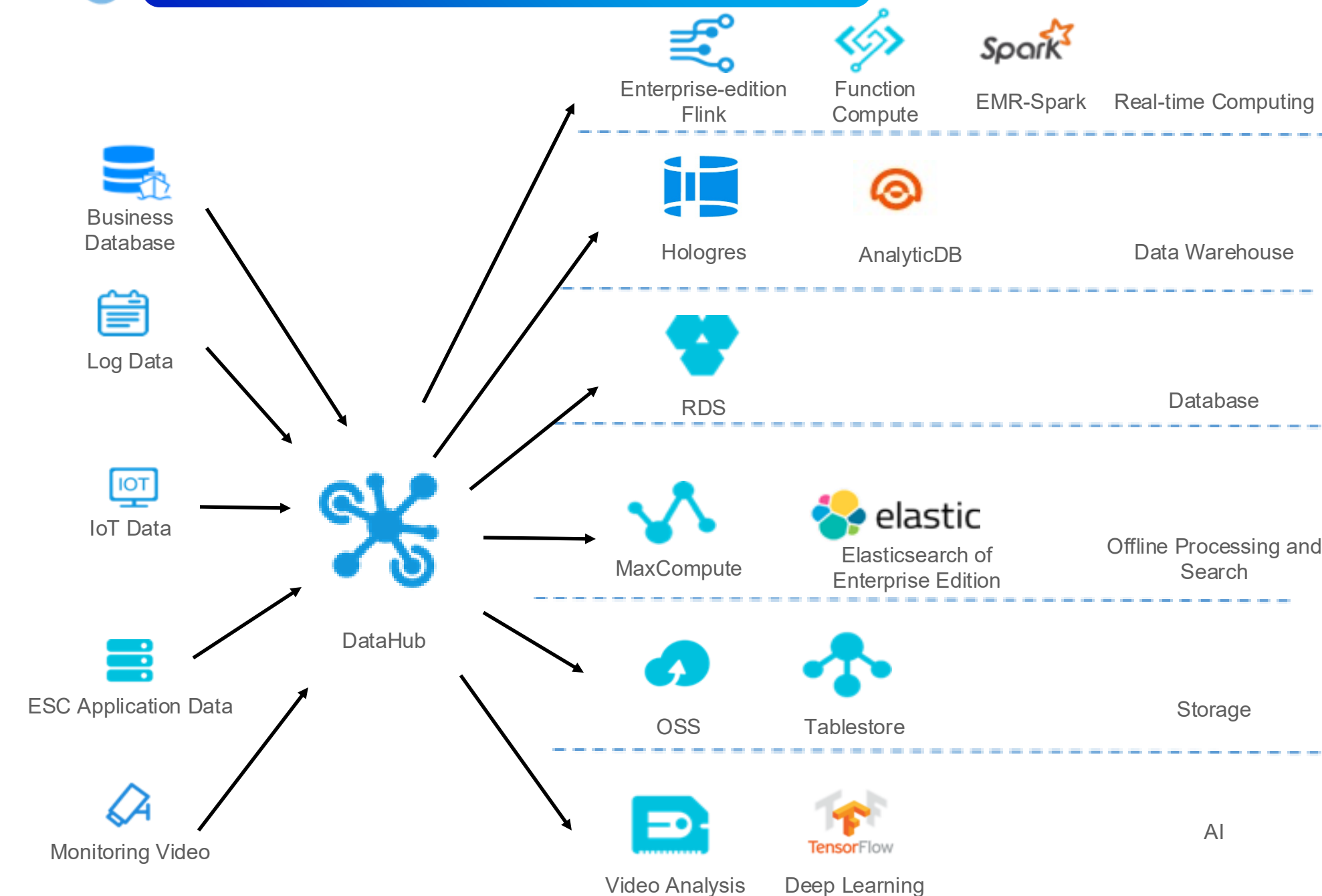
Benefits

- Ultra-large scale:** A single MaxCompute cluster can schedule more than 30,000 nodes.
- Excellent performance:** In the TPCx-BB test, MaxCompute scored 22,000 in processing 100 TB of data, maintaining the leading position in the industry.
- Lakehouse:** MaxCompute can be integrated with the open source Hadoop and supports federated computing.
- Cloud-native multi-tenant deployment mode:** MaxCompute supports fine-grained access control and data isolation and sharing across tenants. MaxCompute allocates resources to multiple tenants, which is different from node allocation.
- EC policy:** MaxCompute supports the three-replica mechanism and EC policy. You can manually archive data or implement automatic data archive in EC storage mode to save storage space.
- Nearline queries:** MaxCompute allows you query data within seconds based on the compute engines, SQL syntax, and an intelligent tiered storage policy for cold, hot, warm, and archived data in the same cluster.
- Multi-chip architecture in one cloud:** MaxCompute fully supports Intel, Hygon, Kunpeng, and Phytium chips.

Application Scenarios

- **Data collection:** DataHub provides various SDKs, APIs, third-party extensions, and Kafka protocols for data collection. You can import heterogeneous data that is generated by applications, websites, IoT devices, or databases to DataHub in real time.
- **Data storage:** DataHub allows you to flexibly set the storage duration to ensure that downstream systems can consume data again. DataHub supports data backup, storage encryption, and cross-data center disaster recovery, which is secure and reliable.
- **Data distribution:** You can synchronize data to downstream systems in real time based on simple configurations.

Architecture



Overview

- DataHub is a data distribution platform designed to process streaming data. You can publish and subscribe to applications for streaming data in DataHub and distributed data to other platforms. DataHub allows you to analyze streaming data and build applications based on streaming data.

Benefits

- **Ease of use:** DataHub serves as a unified graphical control platform. It is compatible with the Kafka protocol, supports SDKs for mainstream languages such as Java, Go, Python, and C++, and plug-ins that are used to collect data, such as OGG, Logstash, and Flume.
- **Stability and reliability:** DataHub supports multiple data replicas and service availability of 99.99% that is guaranteed in the SLA. It supports hot upgrade and does not affect users during cluster upgrade. It also supports data transmission encryption and enterprise-level authorization and authentication.
- **Powerful performance:** DataHub supports the processing of petabytes of data each day, data writing from 150,000 clients each day, QPS up to 2.55 billion during Double 11.

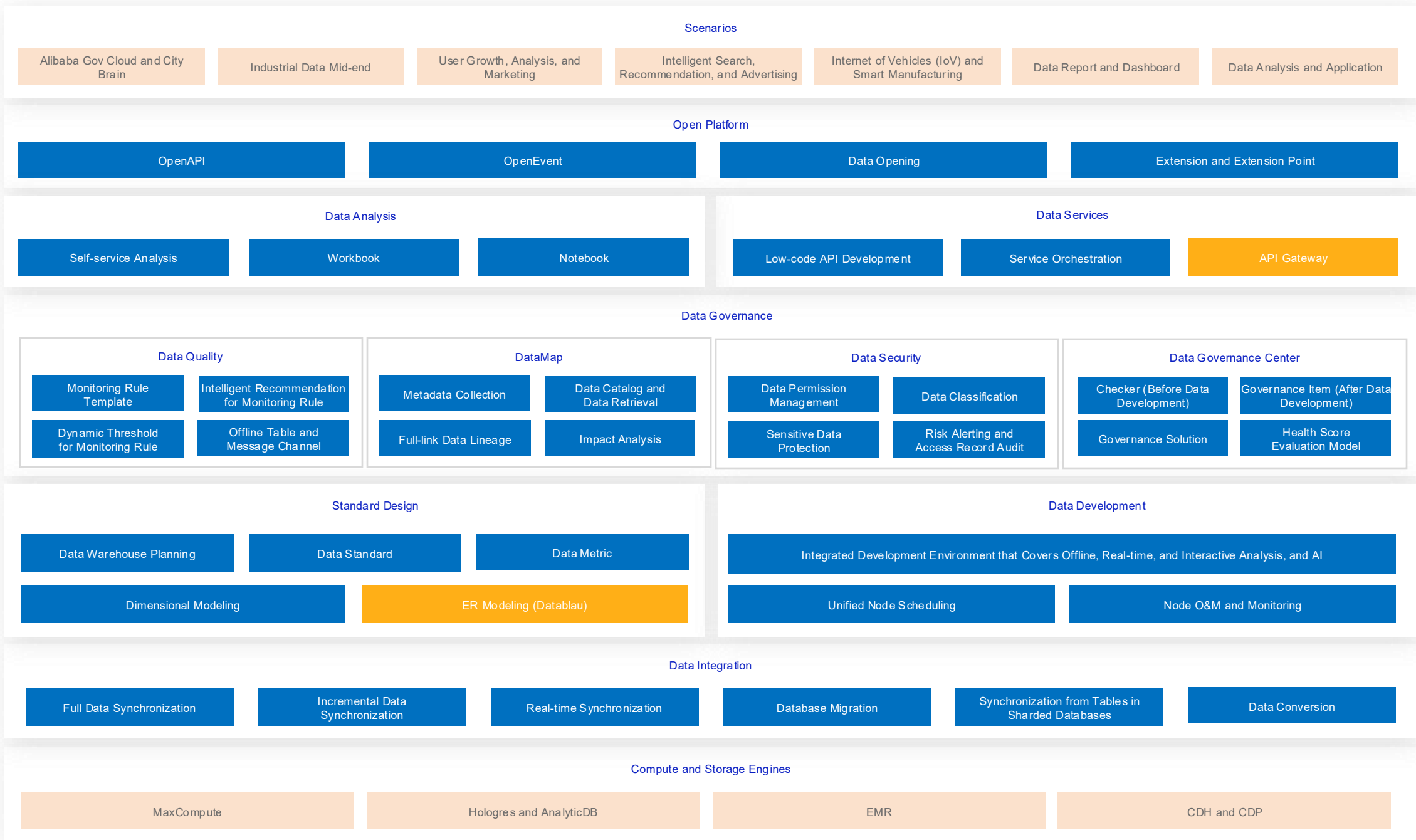


Application Scenarios

- DataWorks is used to construct data mid-ends and data warehouses that integrate offline and real-time processing and analysis.
- DataWorks is used to collect data from multiple data sources and synchronize data in real time.
- DataWorks is used for ad hoc query and online data analysis.
- DataWorks is used to migrate data from on-premises data centers to the cloud.



Architecture



Overview

- DataWorks is an end-to-end big data development and governance platform that provides data warehousing, data lake, and data lakehouse solutions based on big data compute engines, such as MaxCompute, Hologres, EMR, AnalyticDB, and Cloudera Data Platform (CDP).



Benefits

- **Support for multiple engines:** DataWorks supports various compute engines, such as MaxCompute, EMR, Hologres, and AnalyticDB.
- **Intelligent SQL-based IDE and visualized workflow orchestration:** DataWorks provides AI-powered SQL editor, supports intelligent prompt and visualized SQL structures, and allows you to construct workflows by performing drag-and-drop operations.
- **High-performance and high-stability scheduling system:** DataWorks provides AI-powered SQL editor, supports intelligent prompt and visualized SQL structures, and allows you to construct workflows by performing drag-and-drop operations.
- **Intelligent monitoring and O&M diagnostics:** DataWorks provides end-to-end diagnostics suggestions for node running and helps you quickly identify the causes of exceptions.
- **Workbook:** DataWorks supports data uploads from on-premises machines or online data sources, data pivoting and analysis, and data profiling.
- **SQL query:** DataWorks supports various types of data sources and serves as a more convenient and lightweight tool that is used to obtain data from SQL query results.
- **Self-service analysis:** DataWorks allows you to analyze and gain insights into data metrics based on dimension-based models.

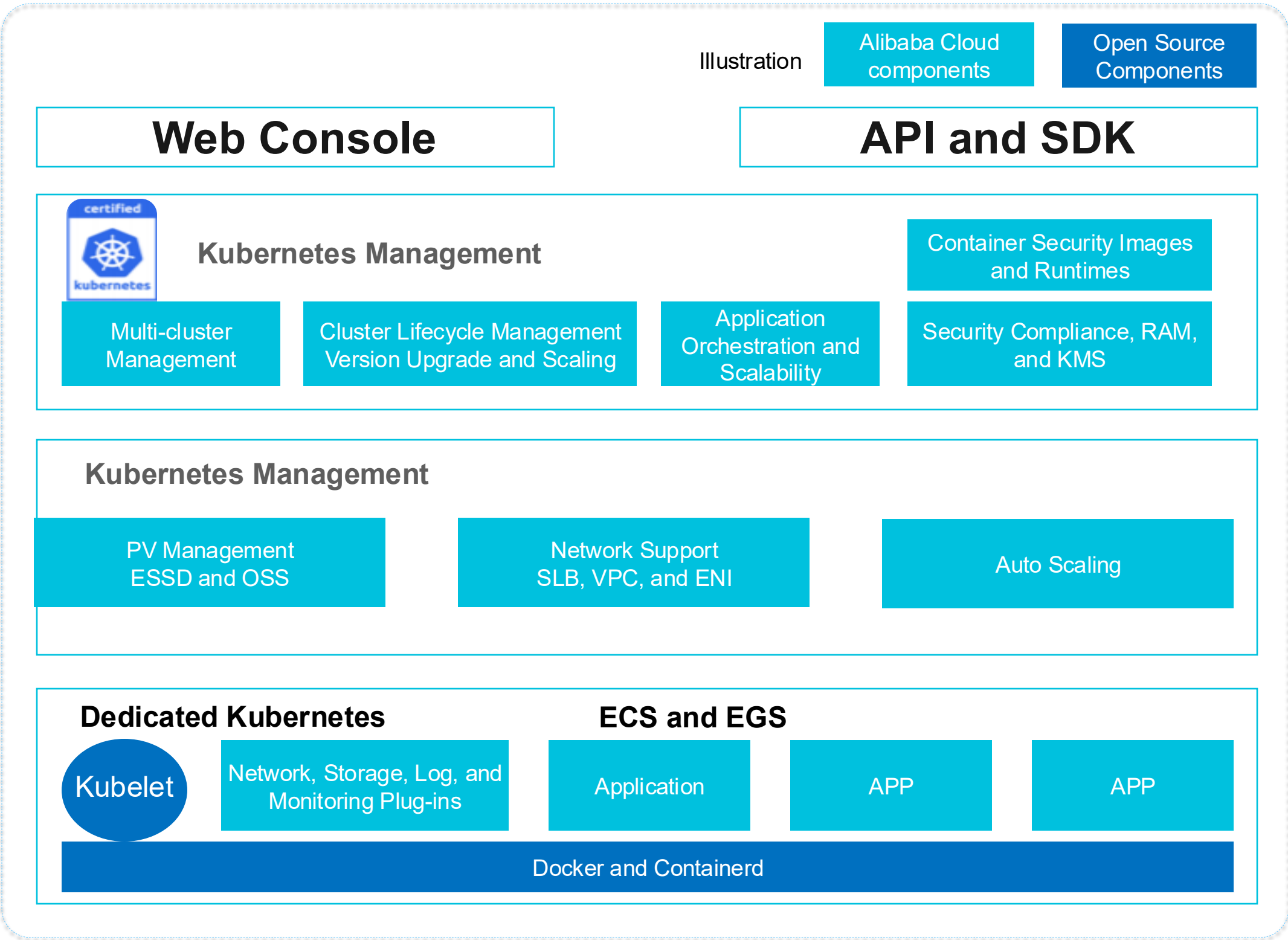


Application Scenarios

- DevOps and continuous delivery
- Machine learning based on Cloud-native technologies
- Microservices, hybrid cloud, and auto scaling architecture



Architecture



Overview

- ACS provides high-performance, scalable, and enterprise-level management service for Kubernetes containerized applications throughout their lifecycles. ACS simplifies O&M such as cluster building and scaling, and integrates the virtualization, storage, networking, and security capabilities of Alibaba Cloud to create the best runtime environment for Kubernetes containerized applications in the cloud. Alibaba Cloud is a Kubernetes Certified Service Provider (KCSP). ACS is one of the first services to participate in the Certified Kubernetes Conformance Program in the world. ACS provides high performance services for the management of containerized applications.



Benefits

- **Edge cluster management:** ACS allows you to create edge Kubernetes clusters and edge node pools and implement edge deployment.
- **Namespace and quota management:** ACS allows you to manage resource quotas for namespaces, such as CPU, memory, storage, and pods.
- **Workload management:** ACS provides secure containers that are equipped with dedicated kernels. This provides better isolation capabilities. ACS supports GPU in pass-through mode and virtualization capabilities.
- **Container network management:** You can resolve private domain names for VPCs in an ACS cluster. Pods remain unchanged after the fixed IP addresses of the pods are rebuilt. ACS also allows you to create policies for controlling inbound and outbound traffic of clusters.
- **Route management:** You can configure weights for an Ingress to proportionally distribute traffic. You can configure canary release policies for an Ingress to split traffic based on request headers, cookies, and query parameters. This way, multiple versions of an application can be released and coexist.
- **Monitoring and O&M:** ACS provides the log collection module. This module collects standard output logs and application log files from containers and allows you to customize tags to facilitate log analysis and filtering. You can log on to a container terminal in the web console to view logs and perform O&M operations. ACS supports the security audit feature. You can use this feature to check the overview of events in clusters, and view information about important events and computing and storage network resources.