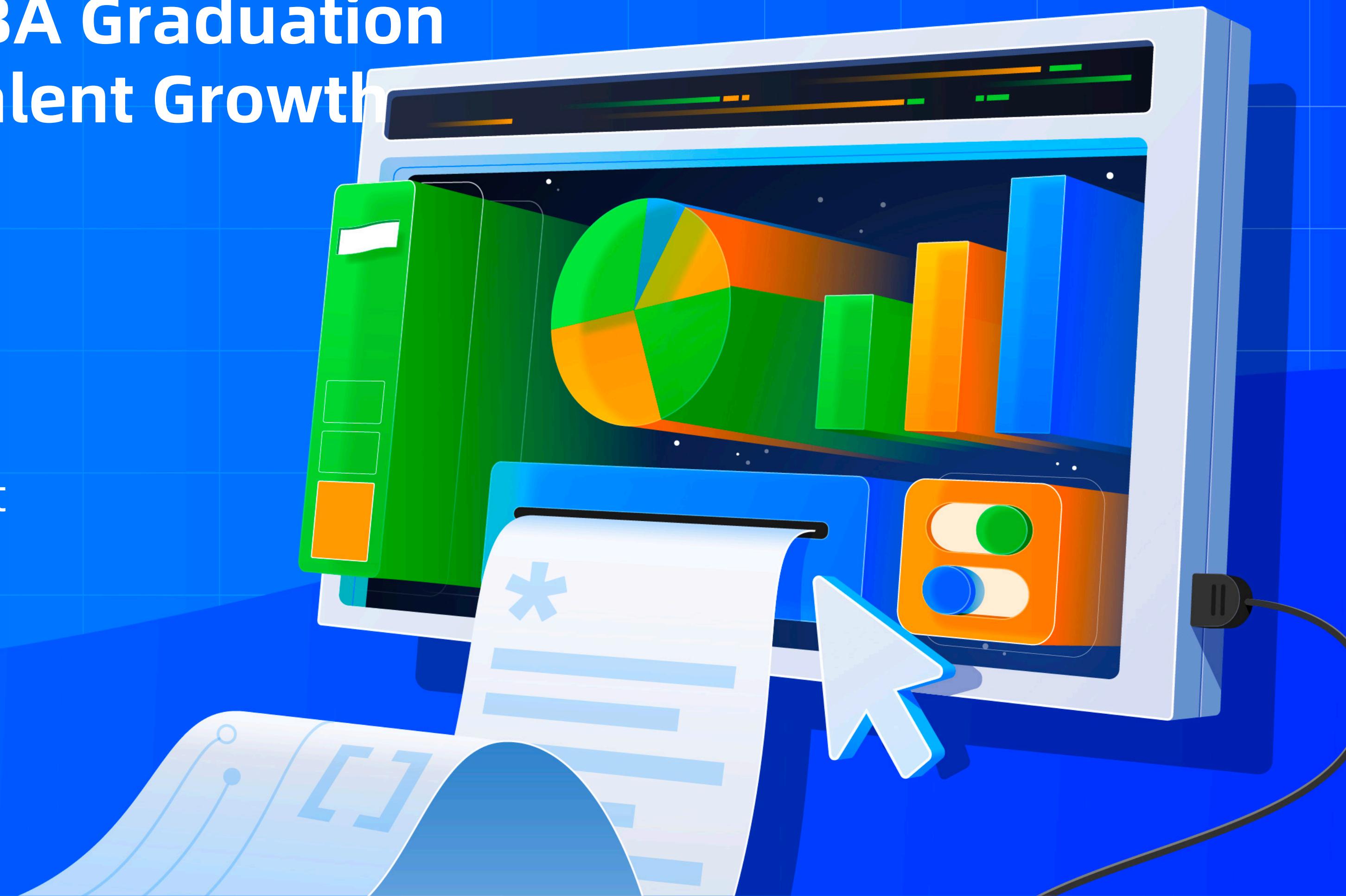


FROM INTRODUCTION TO PRACTICE

Lesson 10: OceanBase DBA Graduation
- High Availability and Talent Growth

Peng Wang
OceanBase Global Technical Evangelist



Agenda



- OceanBase 4.X High Availability
Architecture - Primary and Standby
Database
- OceanBase Database Talent Training
Program

Oceanbase 4.X High Availability Architecture - Primary and Standby Database

Physical Standby Database Solution Based On Redo Log Asynchronous Replication

Solution Introduction

This solution is similar to the primary-standby replication solution of traditional databases. Between two or more clusters, it is allowed to build a tenant-level primary-standby relationship by asynchronously replicating Redo logs at the tenant level, providing two disaster recovery capabilities: lossless switching within the plan and lossy switching during failure.

Disaster recovery capability

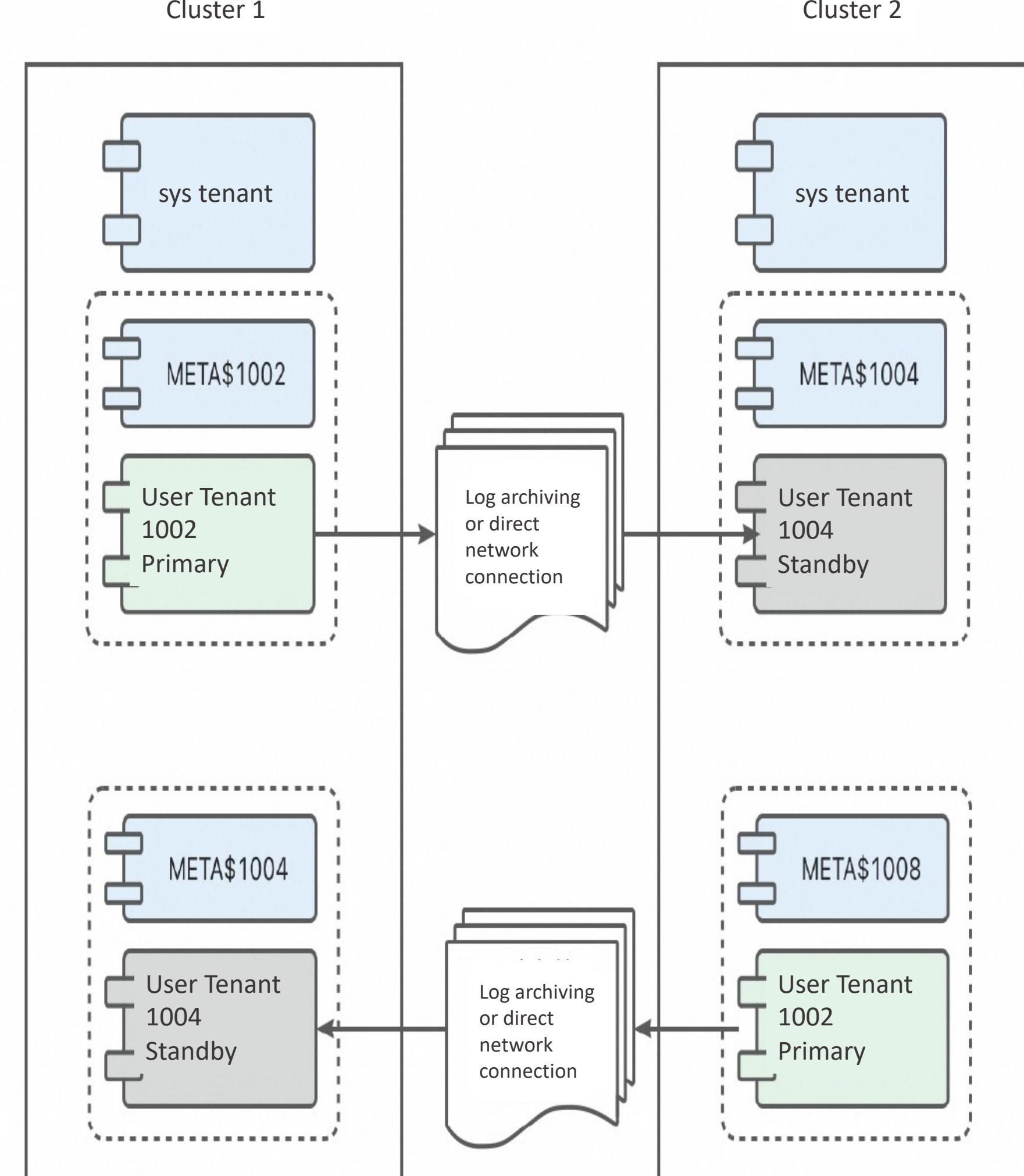
Lossless switching: When executing a lossless switch within the plan, the primary tenant and the standby tenant swap roles without data loss ($RPO = 0$), and the switching time is in seconds (RTO is in seconds)

Lossy switching: When the cluster where the primary tenant is located fails, a lossy switching can be performed to switch the standby tenant to the primary tenant. At this time, data loss cannot be guaranteed, RPO is larger than 0, and the switching time is in seconds (RTO is in seconds)

Oceanbase 4.X High Availability Architecture - Primary and Standby Database

Architecture Changes

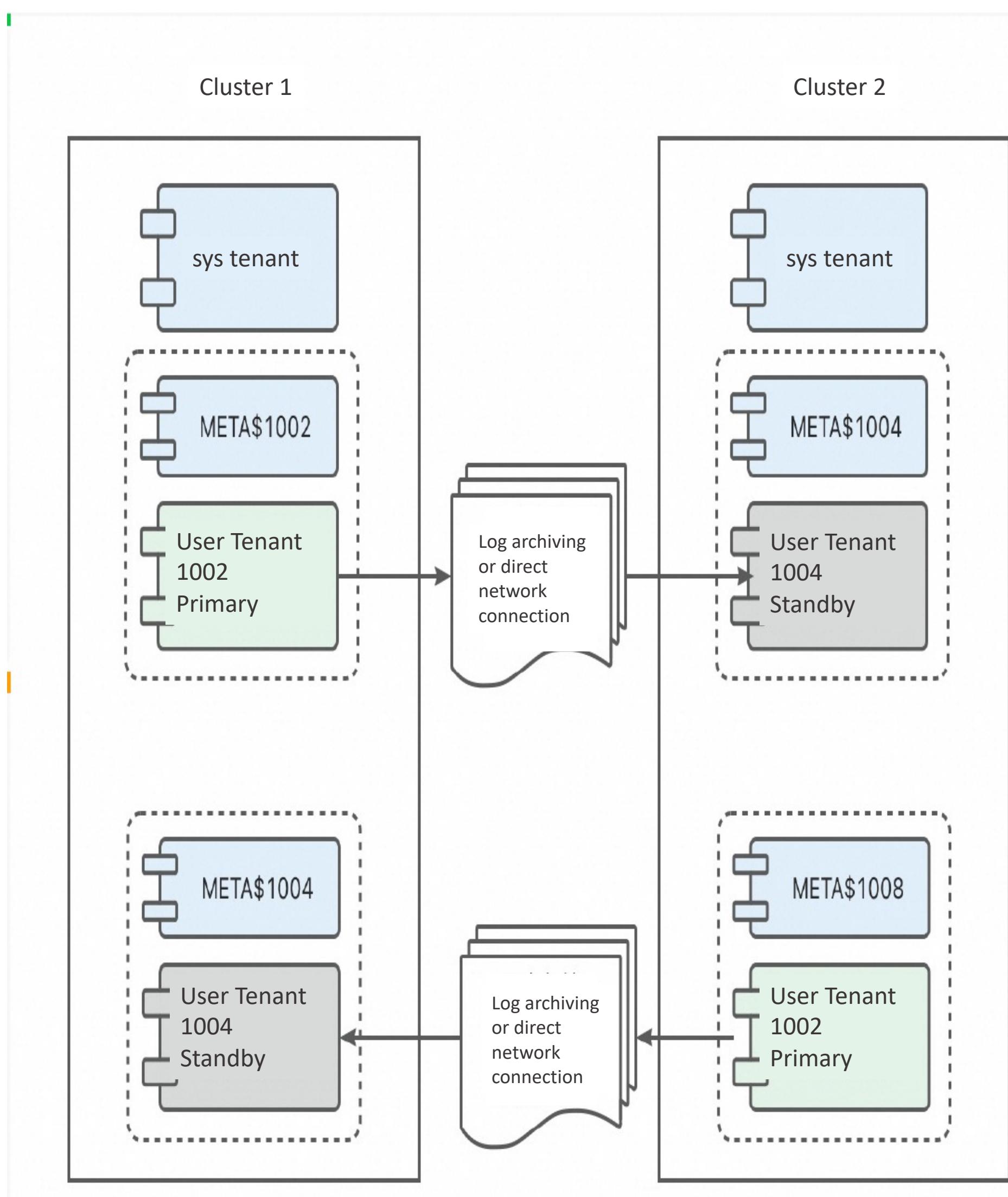
- In OceanBase V4.2.0, the physical standby database adopts an independent primary-standby database architecture, and the primary-standby relationship exists at the tenant level. Different from the centralized architecture (cluster-level master-slave) of previous versions, under the independent primary-standby database architecture, each cluster is independent of each other, and users can manage clusters more flexibly.
- The OceanBase database physical tenant-level primary-standby only provides asynchronous mode, that is, it only supports the maximum performance mode and does not support the maximum protection and maximum availability modes.
- The tenant names, resource specifications, configurations, and locality of the primary and standby tenants are not required to be the same.
- The log transport service provides two different usage modes, which determine two different deployment schemes of physical standby databases: log archiving-based physical standby database and network-based physical standby database.
- The primary and standby tenant replicas are decoupled, and the primary and backup tenants run independently. There is no longer a mandatory direct connection to maintain membership relationships.



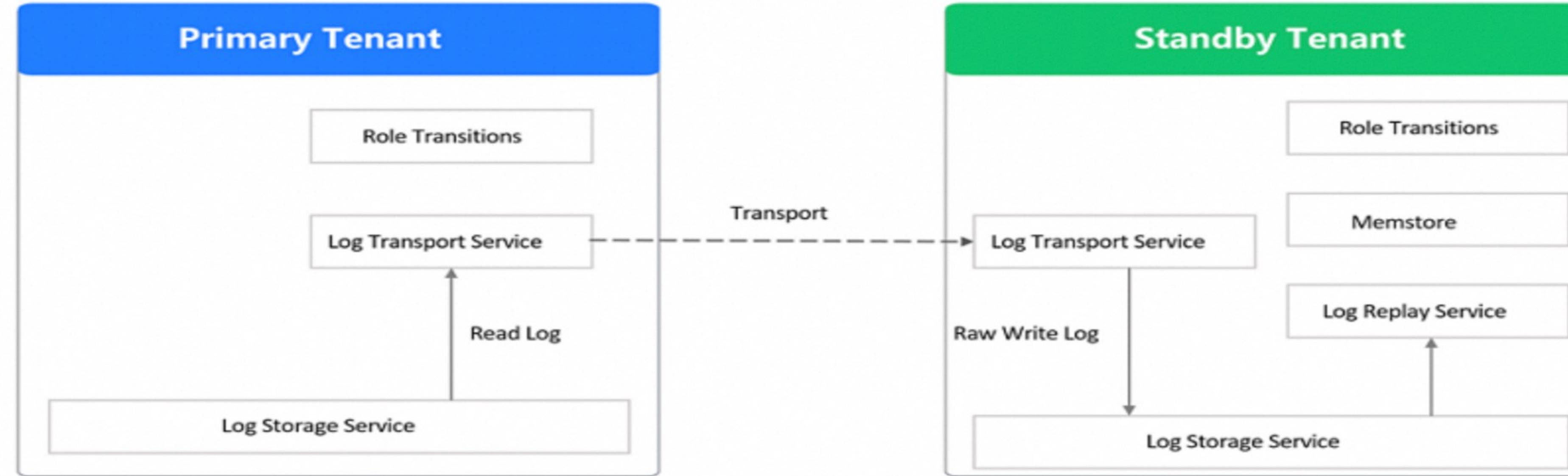
Oceanbase 4.X High Availability Architecture - Primary and Standby Database

Solution Benefits

- The primary-standby synchronization infrastructure requirements are more flexible and can be synchronized through direct network connection or shared disk log archiving
- The primary and standby databases synchronize data based on the log-based consumption system. Due to the introduction of single-machine log streams, partitions are no longer a concern, log stream site advancement is more lightweight, and RPO and RTO are improved
- The complexity of the primary and standby database architecture is simplified, supporting independent operation and maintenance, making it more stable and reliable
- V3.X to maintain activity between the primary and standby databases, heartbeats (keepalive) need to be sent periodically. When the number of partitions increases, CPU and network resources are likely to become bottlenecks. V4.X The bandwidth and other resource requirements between the primary and standby databases are lighter
- The log synchronization mode is changed from out-of-order synchronization to sequential synchronization. The switchover and failover switching takes seconds



Core Components of the Primary and Standby Databases



LOG Transport Service

- The standby database synchronizes the Redo logs between the primary tenant and the standby tenant in real-time through the log transmission service. In particular, the primary tenant will not actively push logs to the standby tenant, but only rely on the standby tenant to pull logs from the primary tenant.
- The log transmission service will automatically address log location information, handle log lag scenarios, and handle high availability issues such as failures of the cluster nodes where the primary tenant is located. The backup tenant can obtain logs through the log archiving service of the primary tenant or directly connect to the cluster where the primary tenant is located through the network.

Log Storage Service

- The log storage service provides highly available and reliable log storage and read/write capabilities for standby databases. The log storage service can be either single-copy or multi-copy and uses the Paxos protocol to achieve high availability.
- The log storage service for the primary tenant and the standby tenant uses two completely different working modes. The primary tenant provides the Append mode, which receives data written by upper-level modules such as transactions and DDL. The standby tenant provides the Raw Write mode, which only allows the writing of logs obtained from other tenants or log archives through the log transmission service.

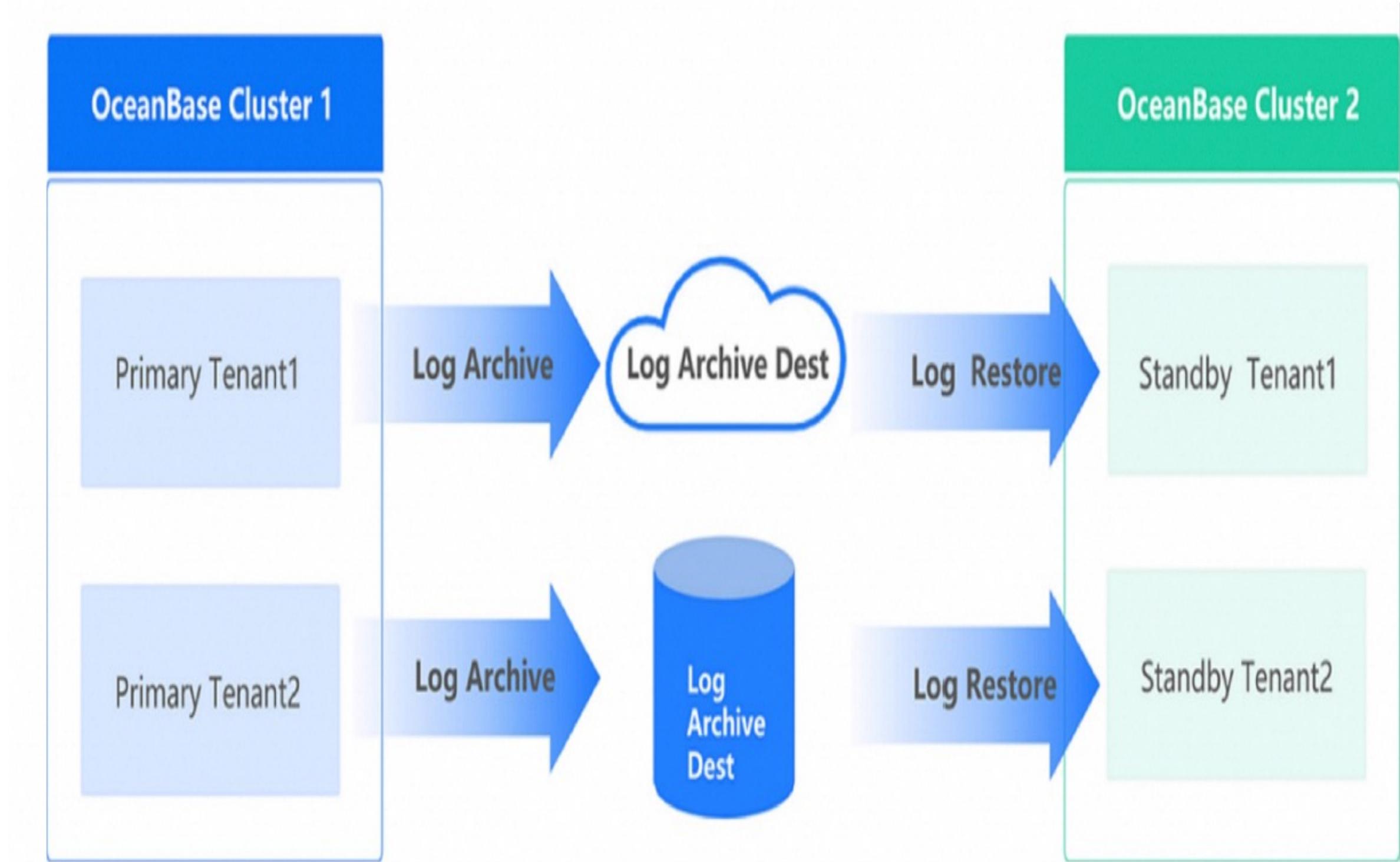
Log Replay Service

- The logs written by the standby tenant to the log storage service will be applied to the Memstore in memory in real-time or in a delayed manner through the log playback service, thus ensuring that the data of the primary tenant and the standby tenant are completely consistent.

Primary and Standby Database Log Transport Service

Physical backup database based on log archiving

- In a standby database based on log archives, the Redo logs of the standby database come from the log archives of the primary tenant or other standby tenants, similar to Far Sync in the Oracle database. The standby tenant only interacts with the log archives and does not have any other form of interaction with the upstream primary tenant or standby tenant.
- In this deployment mode, the standby tenant and the upstream tenant do not need network connectivity, but their synchronization performance and availability will be affected by the log archiving media.

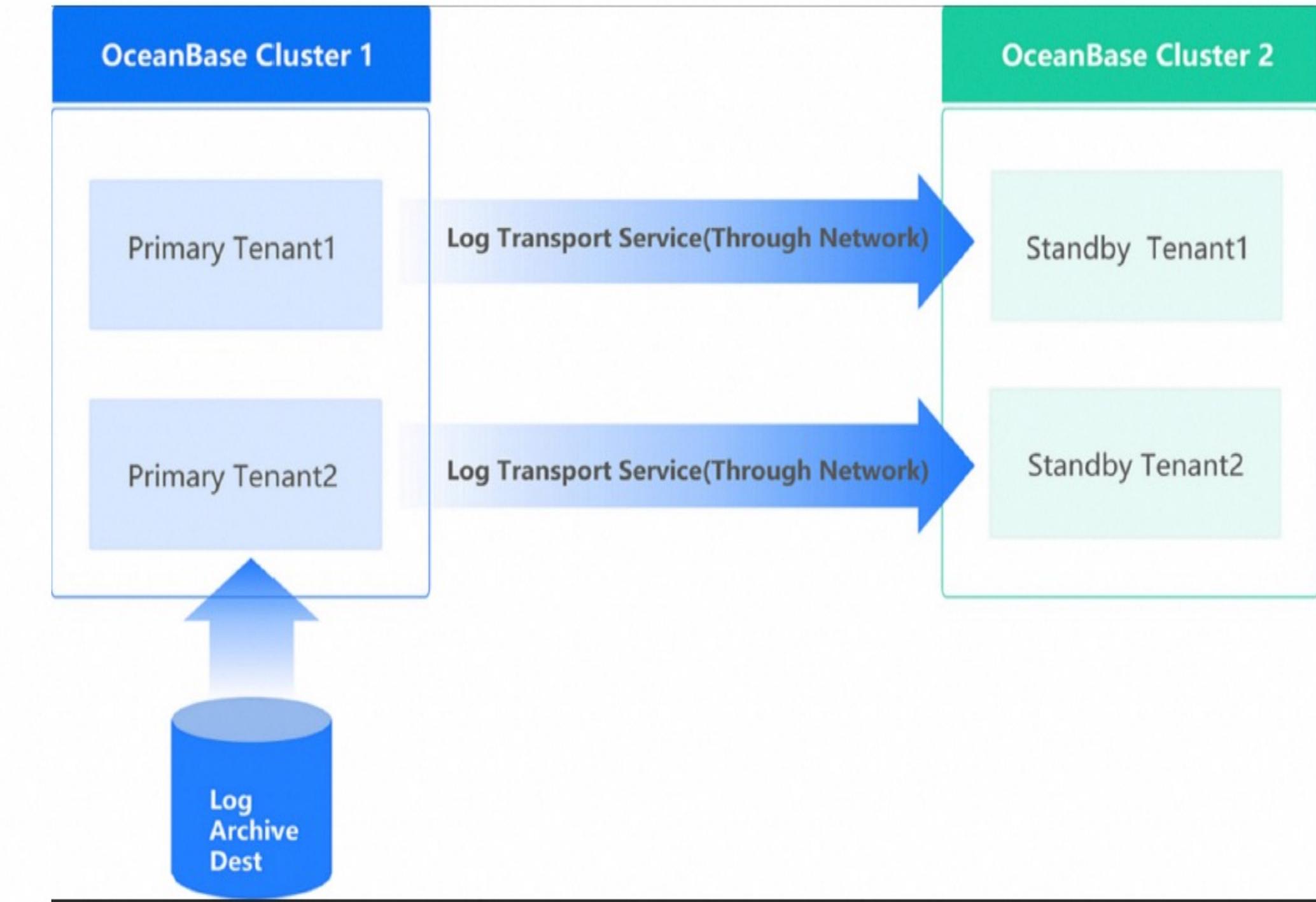


Log Archive, Log Archive Dest, Log Restore

Primary and Standby Database Log Transmission Service

Network-based physical backup database

- In a network-based physical standby database, the standby tenant directly connects to the primary tenant or other standby tenants through the network to read logs, similar to the Replication of MySQL database.
- In this deployment mode, the standby tenant and the primary tenant need to be connected. The standby tenant will send an RPC request through the network to read the Redo log in the primary tenant cluster.
- The logs read by the standby tenant from the primary tenant can be either the primary tenant's online logs or the primary tenant's archived logs (assuming the primary tenant has enabled the log archiving mode). The two log sources support automatic switching, which is transparent to the backup tenant and business users.

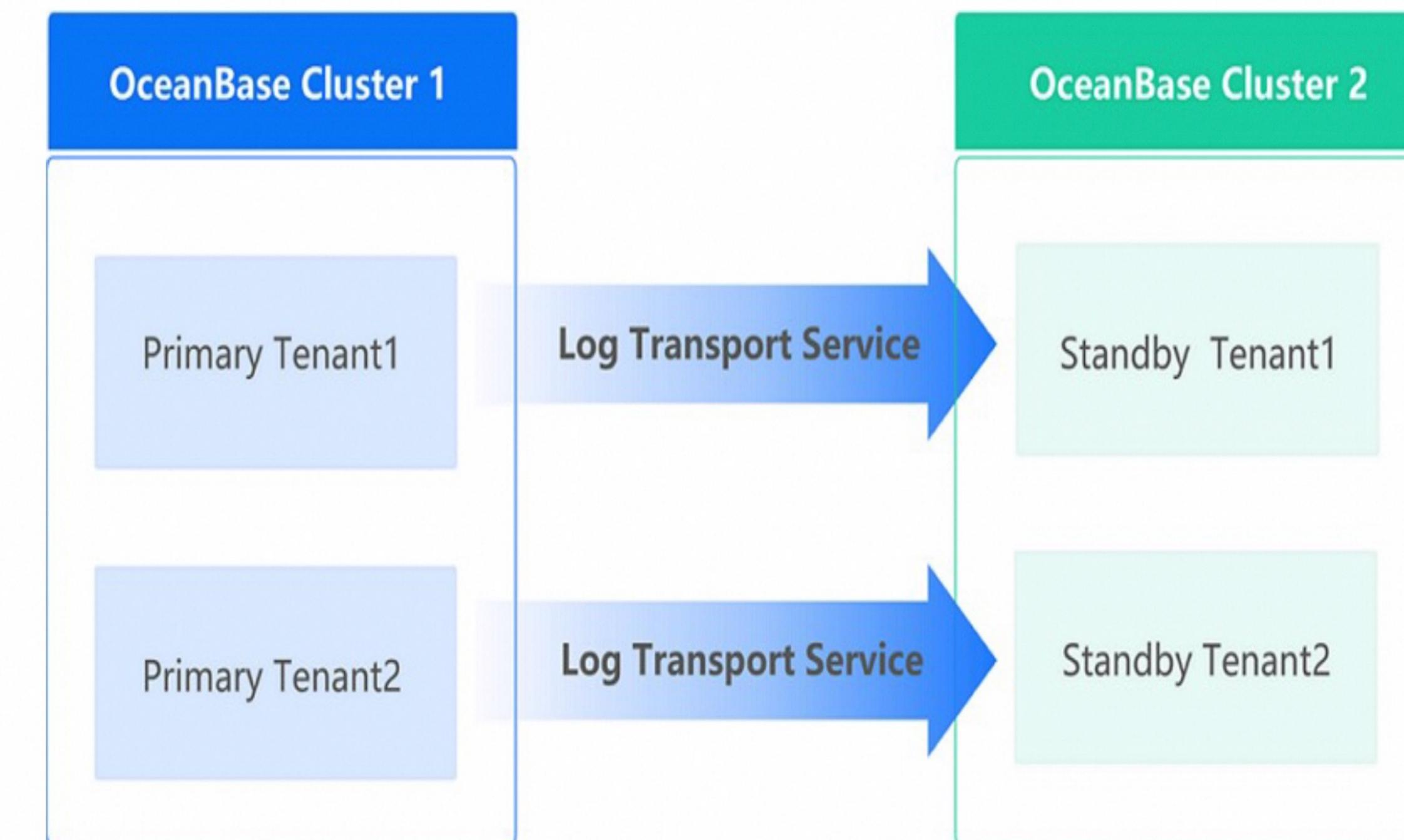


Primary Tenant1 does not have log archiving enabled, and Standby Tenant 1 only synchronizes the online logs of Primary Tenant 1 through the network; Primary Tenant2 has log archiving enabled, and Standby Tenant 2 can synchronize the online logs of Primary Tenant 2, and can also automatically switch to synchronizing the archived logs of Primary Tenant2 after the online logs of the primary tenant are recycled.

Primary Standby Database Scenario

Physical standby database in dual data centers in the same city or different locations

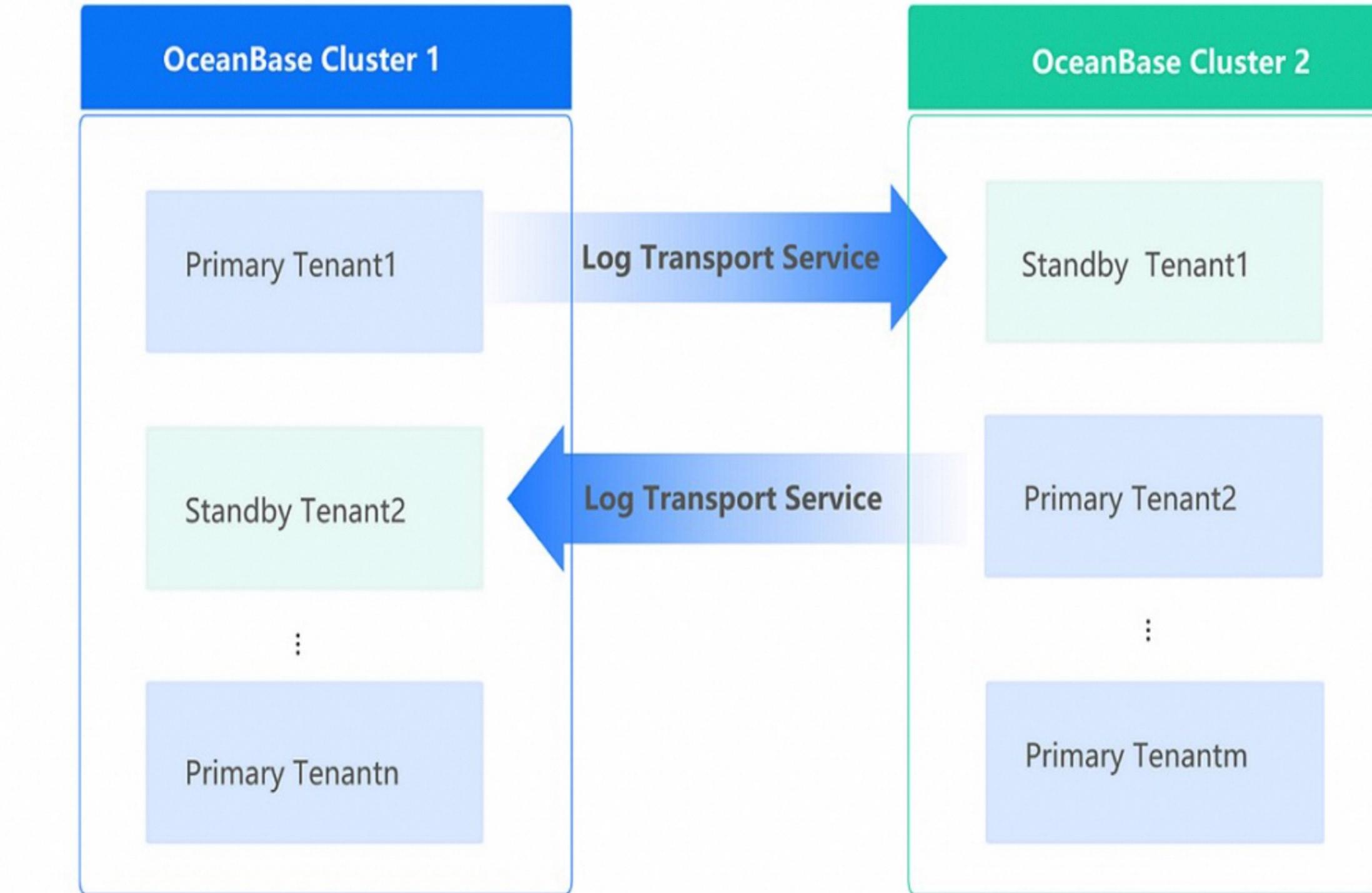
- In the most typical deployment mode, users can use the physical standby database function to meet various requirements, such as remote disaster recovery. Use management tools such as OCP to achieve the master-slave role switching of all tenants at the cluster level
- If there are only two computer rooms in the same city and you want to achieve computer room-level disaster recovery, or if there are two computer rooms in different locations and you want regional disaster recovery, you can use a physical standby database and deploy a cluster in each computer room. When the main computer room is unavailable, the standby computer room takes over the business services, and RPO > 0



Primary Standby Database Scenario

Active-active and disaster recovery in the same city or across different regions

- In the usage scenario, the business has read/write and remote disaster recovery requirements in two different regions, so the primary and standby databases are required in the two regions. When using other database-based primary/standby solutions, users need to deploy two (or more) clusters in each region, and the clusters in the two regions serve as primary and standby for each other.
- In the V4.X solution, you only need to deploy one cluster in each of the two regions, and the tenant-level master-slave can meet business needs, greatly simplifying the complexity of database cluster management.



Primary and Standby Database Management

Physical Standby Database Usage Restrictions

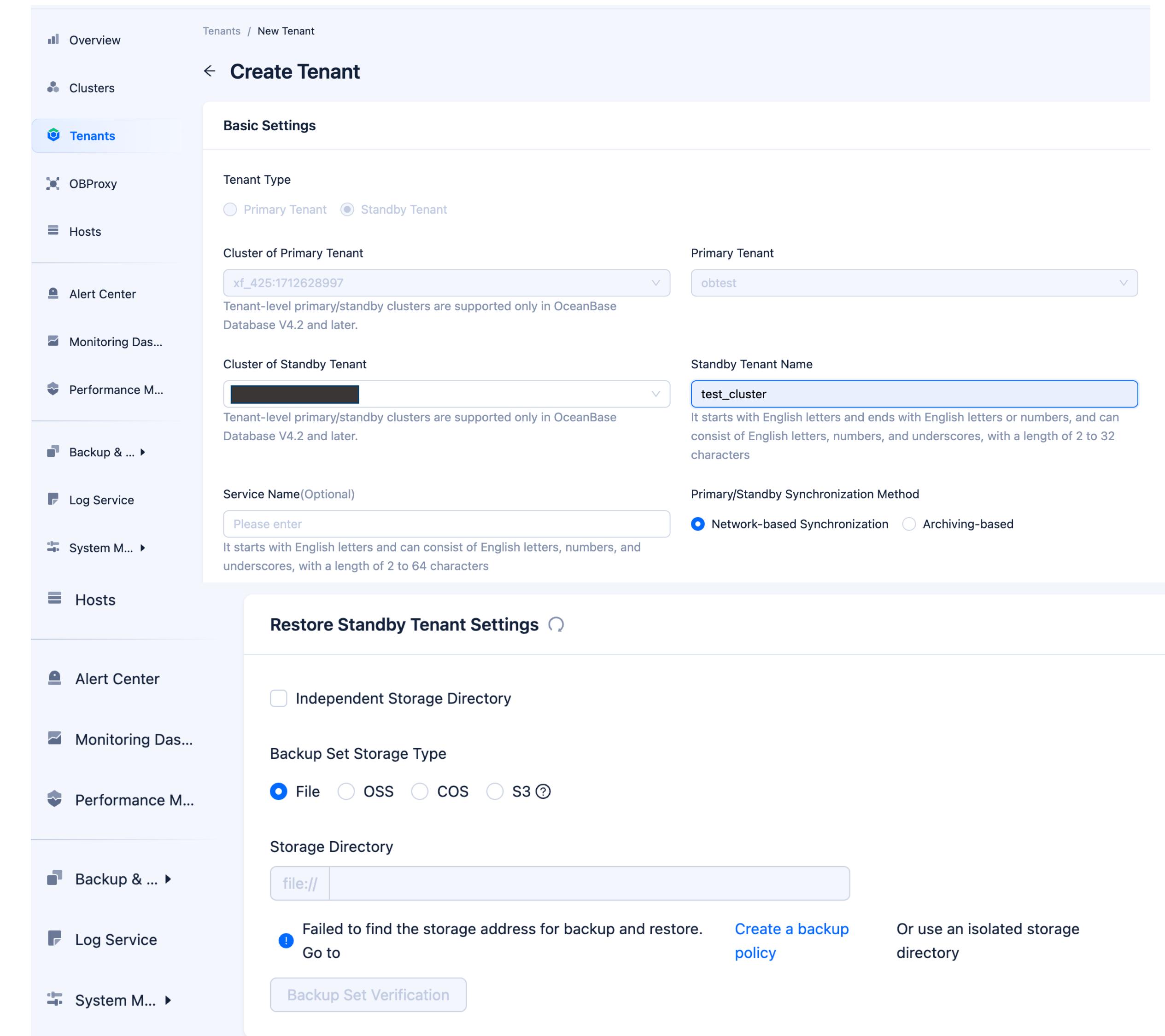
Restrictions	Specific Description
Maximum number of standby tenants supported by one primary tenant	No restrictions
Do the primary tenant and the standby tenant require the same architecture?	Not required. It is just recommended that the primary tenant and the standby tenant use the same resource specifications.
Configuration Items	The configuration items of the primary tenant and the standby tenant are independent of each other and will not be physically synchronized. If you modify the configuration items of the primary tenant, you need to evaluate whether you need to modify the same configuration items of the standby tenant.
System Variables	The system variables of the primary tenant and the standby tenant are physically synchronized. If the system variables are modified on the primary tenant, the system will synchronously modify the same system variables of the standby tenant.
User and Password	Only the primary tenant can create users and modify user passwords. The updated information will be synchronized with the standby tenant.
Read and Write Operations	The standby tenant supports read operations but not write operations.
Minor Compaction & Major Compaction	The dumps of the primary tenant and the standby tenant are relatively independent. The standby tenant synchronizes and merges information from the primary tenant, and independent merging is not supported.
Switchover Limitations	Requires all copies of the standby tenant log stream to be online
Failover Limitations	Requires all copies of the standby tenant log stream to be online

Primary and Standby Database Management

Create a standby tenant

- The primary and standby tenants can be located in the same OB cluster or different clusters.
- Three ways to create standby tenants
- Master/Slave Log Synchronization Mode:
 - Web-based
 - Archive-based
 - The same master-slave relationship only supports the same synchronization mode
- Based on the storage media currently supported by the archive:
 - NFS
 - OSS
 - COS
 - S3

* Currently, the OceanBase database physical standby database only provides asynchronous mode.



Primary and Standby Database Management

Method 1 for Creating Standby Tenants

Create standby tenants based on direct network connection + subsequent network-based synchronization

Applicable to the scenario where the primary tenant has just been created or the primary tenant's Redo log is complete (This can be determined by querying the view GV\$OB_LOG_STAT)

Primary Tenant obtest

Standby Tenant Name test_cluster

It starts with English letters and ends with English letters or numbers, and can consist of English letters, numbers, and underscores, with a length of 2 to 32 characters

Primary/Standby Synchronization Method

Network-based Synchronization Archiving-based

Service Name(Optional) Please enter

It starts with English letters and can consist of English letters, numbers, and underscores, with a length of 2 to 64 characters

Select a load type based on the actual loads to avoid compromising the runtime performance metrics of the database.

Select a value

-

The REDO logs of the primary tenant are complete. You do not need to restore a tenant. [Restore by Using Backup Set](#)

Primary and Standby Database Management

Method 2 for Creating Standby Tenants

Create standby tenants based on backup and recovery functions + subsequent network-based synchronization

Applicable to scenarios where log archiving and data backup exist, and the primary and standby databases are interconnected over the network.

The screenshot shows the OceanBase management console interface for creating a standby tenant. The left sidebar includes options like Overview, Clusters, **Tenants**, OBProxy, Hosts, Alert Center, Monitoring Datasources, Performance Metrics, Backup & Restore, Log Service, and System Management. The main area has sections for Service Name (Optional), Primary/Standby Synchronization Method (Network-based Synchronization selected), Load Type (Load Type dropdown), and Restore Standby Tenant Settings. The 'Restore Standby Tenant Settings' section is highlighted with a red box and contains fields for Independent Storage Directory (unchecked), Backup Set Storage Type (File selected, OSS, cos, S3 options available), Storage Directory (file:// input field), and error message: 'Failed to find the storage address for backup and restore.' It also includes links to 'Create a backup policy' and 'Or use an isolated storage directory'. At the bottom are 'Backup Set Verification' and 'Submit' buttons.

Tenant-level primary/standby clusters are supported only in OceanBase Database V4.2 and later.

Service Name (Optional)
Please enter

It starts with English letters and can consist of English letters, numbers, and underscores, with a length of 2 to 64 characters

Primary/Standby Synchronization Method
 Network-based Synchronization Archiving-based

Load Type
Select a load type based on the actual loads to avoid compromising the runtime performance metrics of the database.

Load Type
Select a value

Restore Standby Tenant Settings

Independent Storage Directory

Backup Set Storage Type
 File OSS cos S3

Storage Directory
file://

Failed to find the storage address for backup and restore.
Go to [Create a backup policy](#) Or use an isolated storage directory

Backup Set Verification

Cancel Submit

Primary and Standby Database Management

Method 3 for Creating Standby Tenants

Create a standby tenant based on the backup and recovery function + subsequent synchronization based on archive logs. Applicable to the scenario where log archiving is enabled on the primary database, data backup is done, and the standby database can access the backup.

The screenshot shows the OceanBase management console interface for creating a standby tenant. The left sidebar navigation includes: Overview, Clusters, **Tenants**, OBProxy, Hosts, Alert Center, Monitoring Datas..., Performance M..., Backup & ..., Log Service, and System M... ▶.

The main configuration area has the following sections:

- Service Name (Optional):** A text input field with placeholder "Please enter". Below it is a note: "It starts with English letters and can consist of English letters, numbers, and underscores, with a length of 2 to 64 characters".
- Primary/Standby Synchronization Method:** A radio button group with two options: "Network-based Synchronization" and "Archiving-based" (which is selected and highlighted with a red box).
- No backup set available?**: A message with a link to "Back Up Now".
- Load Type:** A dropdown menu with placeholder "Select a value". A note below says: "Select a load type based on the actual loads to avoid compromising the runtime performance metrics of the database."
- Restore Standby Tenant Settings:** This section is enclosed in a large red box.
 - Independent Storage Directory:** An unchecked checkbox.
 - Backup Set Storage Type:** A radio button group with four options: File (selected), OSS, COS, and S3. The "File" option is highlighted with a blue circle.
 - Storage Directory:** An input field containing "file://".
 - Failed to find the storage address for backup and restore.**: A note with a warning icon and a link to "Go to".
 - Create a backup policy**: A link.
 - Or use an isolated storage directory**: A note.
 - Backup Set Verification**: A button.

At the bottom right are "Cancel" and "Submit" buttons. The footer of the page includes the OceanBase logo and the URL "oceanbase.github.io/".

Primary and Standby Database Management

Daily Operation and Maintenance

- Daily Switchover
 - Only primary-standby daily switchover is supported by the primary tenant.
 - If the standby tenant's log synchronization delay is larger than 5000ms, it is not allowed to switch to the primary tenant.
 - The primary-standby switchover is not supported if the standby tenant does not enable log archiving. You need to enable archiving for the standby tenant first.
- Disaster Recovery Failover
 - Only supports disaster recovery switchover from the standby tenant, provided that the primary tenant corresponding to the standby tenant cannot operate normally.
- Primary-Standby Decoupling
 - Only the standby tenant can initiate the primary-stanby decoupling. After decoupling, the standby tenant will exist as an independent tenant and will no longer synchronize data with the original primary tenant.
- Pause Sync
- Turn on Sync
- Upgrade
 - It is recommended to upgrade the standby tenant's OB cluster first and then upgrade the primary tenant's OB cluster; otherwise, if you upgrade the primary tenant first and then the standby tenant, the standby tenant synchronization will be temporarily stuck.

Primary and Standby Database Management

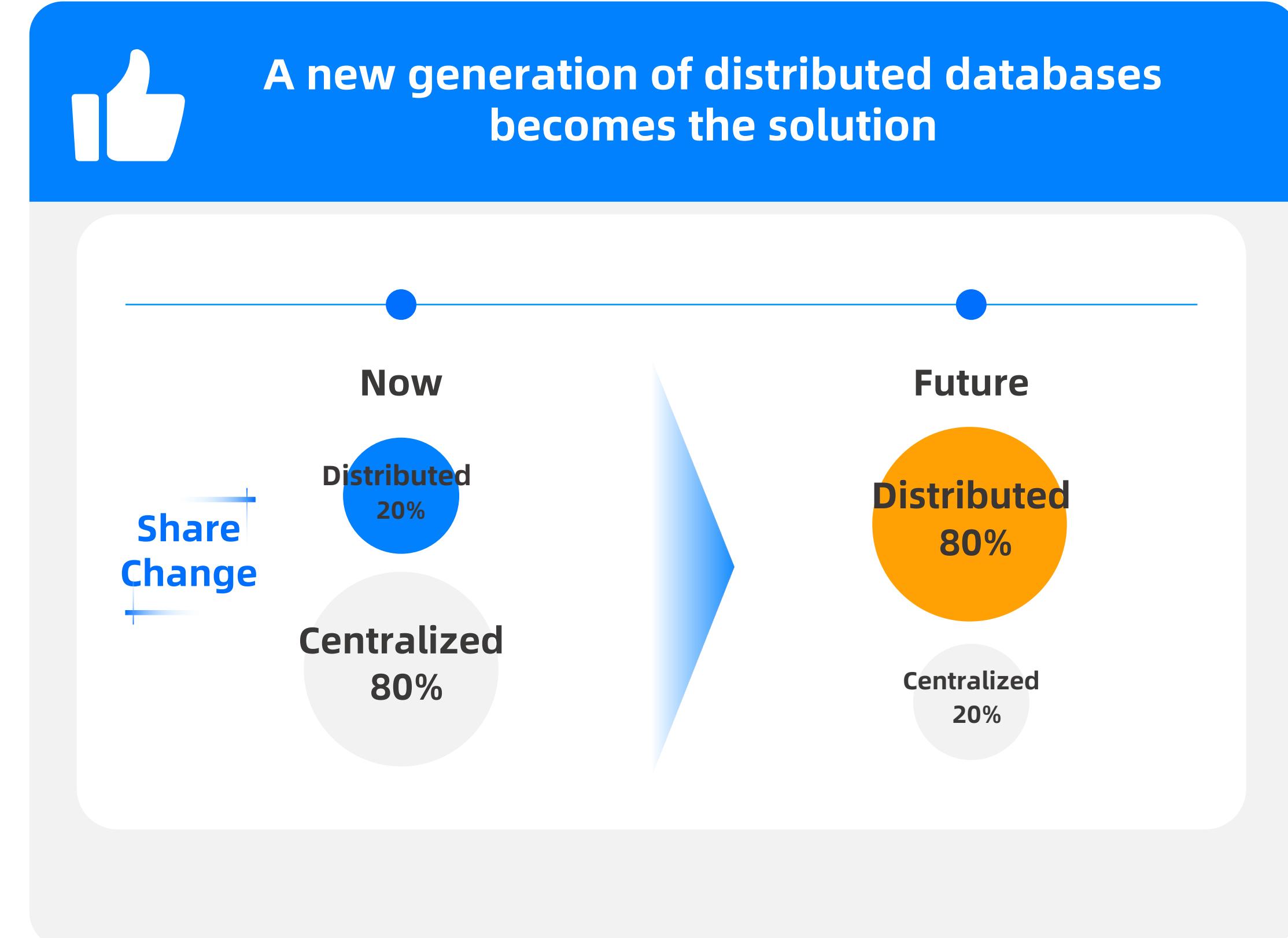
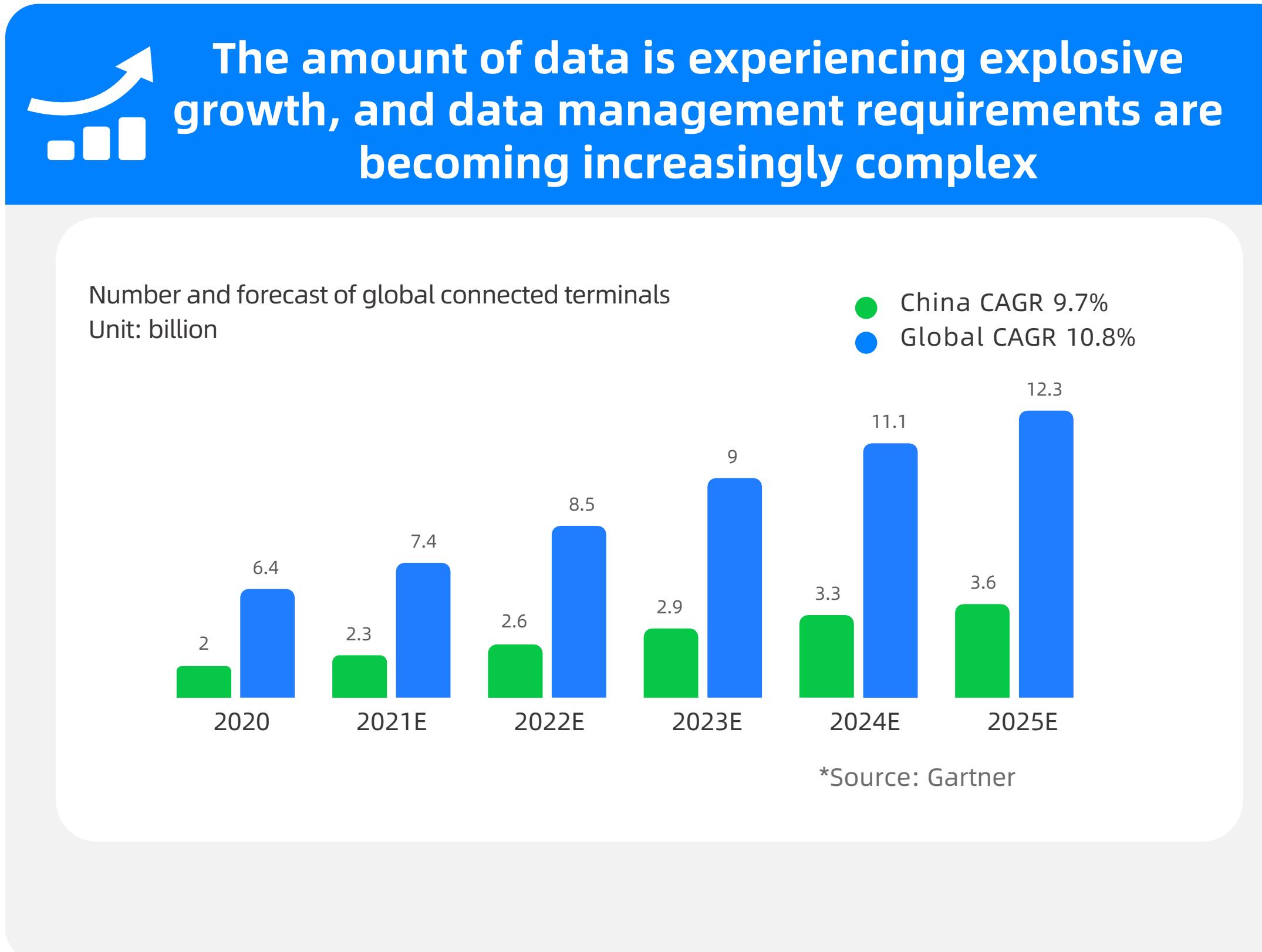
	Synchronize the standby database based on log archiving	Synchronize the standby database based on direct network connection
Whether Switchover is supported	Support	Support
Whether to support Failover	Support	Support
Whether it supports connecting one primary database to multiple standby databases	Support	Support
Whether to support cascaded standby database	Support	Support
Whether to synchronize asynchronously	Yes	Yes
Whether to support maximum availability/protection mode	Not supported	Not supported
Whether to support speed limit	Not supported	Support, cluster-level rate limit
Data source of the standby database	Archive logs	Primary database online log or archive log, support automatic switching
Whether to require the main library to enable archiving	Required (otherwise, standby tenants cannot be created)	Not required (but it is recommended that log archiving be enabled on the primary database)
Whether to enable archiving in the standby database	Required (otherwise Switchover cannot be performed)	Not required (but it is recommended to enable log archiving for the standby database)
Real-time	Second-Level to Minute-Level	Second-Level
Storage media supported by log archiving	OSS/NFS	Not involved

Agenda

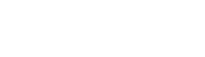
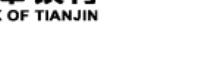
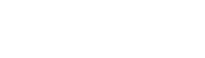


- OceanBase 4.X High Availability
Architecture - Primary and Standby
Database
- OceanBase Database Talent
Training Program

The New Generation of Distributed Databases is a Requirement for Industry Change



Oceanbase Milestones and Architecture Evolution

1.0 Era Distributed Engine		2.0 Era Native Distributed Database		3.0 Era HTAP Hybrid Engine		4.0 Era Integrated Architecture Multi-cloud Deployment	
Distributed KV Storage		SQL engine, multiple copies, high availability		Compatibility, transparent expansion, multi-active disaster recovery		Batch processing, enterprise-level features, HTAP	
2010 Product project establishment	2013 Expanding the scope of use	2014 Core transaction launched	2016 Full business coverage	2017 Multiple financial clients	2019 Breaking world records	2020 Independent Commercialization	2021 Large-scale promotion
OceanBase was born First customer	Serving the business systems of dozens of e-commerce platforms of Alibaba Group	Supporting Alipay's core transaction system Carrying 10% of the transaction traffic during "Double 11" Achieving RPO=0 and RTO < 30s for the first time	Launched Alipay core account and payment system Supports a payment peak of 120,000 transactions per second and a transaction peak of 175,000 transactions per second	Completed the replacement of the last Oracle core of Ant Group's core system Nanjing Bank became the first online user "Double 11" created a database processing peak record of 42 million times/second	Oracle compatible Public cloud service TPC-C 60.88 million tpmC topped the list "Double 11" created a database processing peak record of 61 million times/second	TPC-C 7.07 tpmC, breaking its own world record, 23 times faster than Oracle Launched on the core system of top customers	Released HTAP engine TPC-H 1,526 tpmC topped the list, becoming the world's only distributed database to top TPC-C and TPC-H Community version released, 3 million lines of core code open Pilot overseas customers
 Taobao	 Taobao	 支付宝	 花呗  借呗  余额宝  芝麻信用  网商银行	 四川农信 SICHUAN RURAL CREDIT  南京银行 BANK OF NANJING  顺德农商银行 SHUNDE RURAL COMMERCIAL BANK  广东农信 GDRC  苏州银行 BANK OF SUZHOU  天津银行 BANK OF TIANJIN  浙商证券	 招商证券 China Merchants Securities  常熟农商银行 CHINA COMMERCIAL BANK OF CHENGSHU  中国人民保险 PICC  西安银行 BANK OF XI'AN  中国移动 China Mobile  交通银行 BANK OF COMMUNICATIONS  中国工商银行 INDUSTRIAL AND COMMERCIAL BANK OF CHINA  DANA  中国石化 SINOPEC CORP.  中华保险 CHINA INSURANCE	 江西省人力资源和社会保障厅 JIANGXI PROVINCIAL HUMAN RESOURCES AND SOCIAL SECURITY DEPARTMENT  中国移动 China Mobile  交通银行 BANK OF COMMUNICATIONS  中国工商银行 INDUSTRIAL AND COMMERCIAL BANK OF CHINA  DANA  GCash  中华保险 CHINA INSURANCE  数字江西 DIGITAL JIANGXI	 中国人寿 CHINA LIFE  太平洋保险 CPIC  平安银行 PING AN BANK  北京银行 BANK OF BEIJING  中国联通 China unicom  国家电网 STATE GRID  中国联合航空 CHINA UNITED  翼支付  easypaisa  Touch n Go  easypaisa  palmpay  海底捞 POP MART 用友 yonyou 理想 Midea 快手 Kuaidi 贝壳 Beike 携程 Ctrip vivo vivo Inkeverse 映宇宙

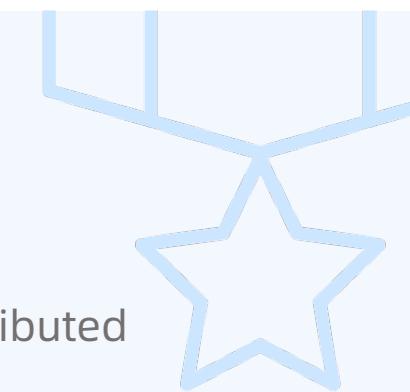
Who I Am

OceanBase is a native distributed database developed entirely independently, dedicated to providing a stable and reliable database for enterprise core systems. It creatively develops distributed technology, achieves financial-grade high availability on ordinary PC servers, and solves the database scalability, ease of use, and high-cost issues.

In 2020, Beijing OceanBase Technology Co., Ltd. was registered in Beijing, and OceanBase began independent commercial operations. At present, OceanBase has become a leader in distributed databases, ranking first in the database market in the financial industry, and has moved from finance to the national economy and people's livelihoods, overseas. It has now served the Industrial and Commercial Bank of China, Bank of Communications, China Life, China Pacific Insurance, China Life, China Merchants Securities, Shandong Mobile, Sinopec, Haidilao, Ideal Auto, GCash, etc., helping to upgrade key business systems in multiple industries such as finance, government affairs, operators, retail, and the Internet.

NO.1

It broke the TPC-C test world record twice and became the first distributed relational database on the list



11 Year

Stable support for "Double 11"
Created a peak database processing record of 61 million times per second

25000+

Database talent certification helps the development of the database industry

300w+

Fully self-developed distributed database, open source core code

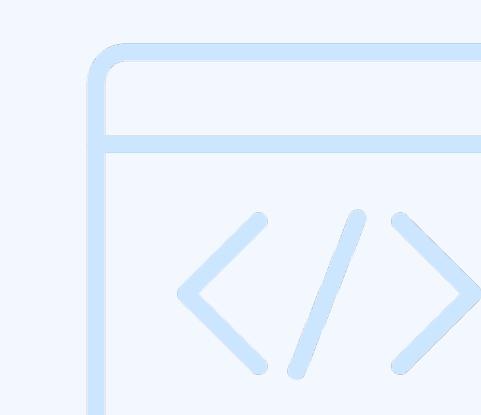
The Only One in The World

Breaking world records in both TPC-C and TPC-H tests



13 Year+

Core R&D team's experience in distributed database industry

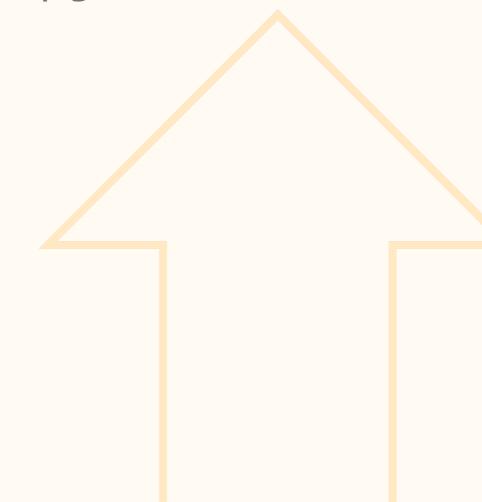


500+

Database Patent

1000+

Industry customers achieve key business system upgrades

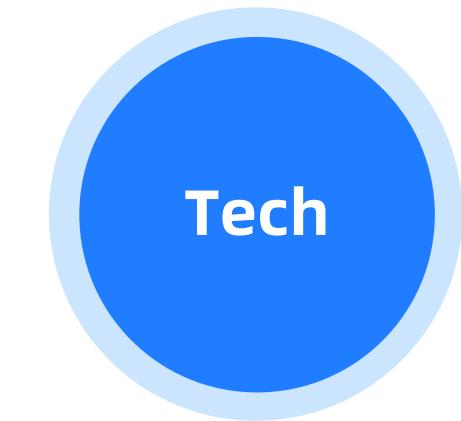


Challenges in Database Talent Training



The talent gap is widening

Favorable policies and high demand from the capital market will promote the database industry into a period of vigorous development, and the demand for the number and quality of practitioners will increase in the next few years.



Rapid technological evolution

Most enterprise database practitioners face the challenge of upgrading existing databases to distributed architectures to help implement innovative businesses.

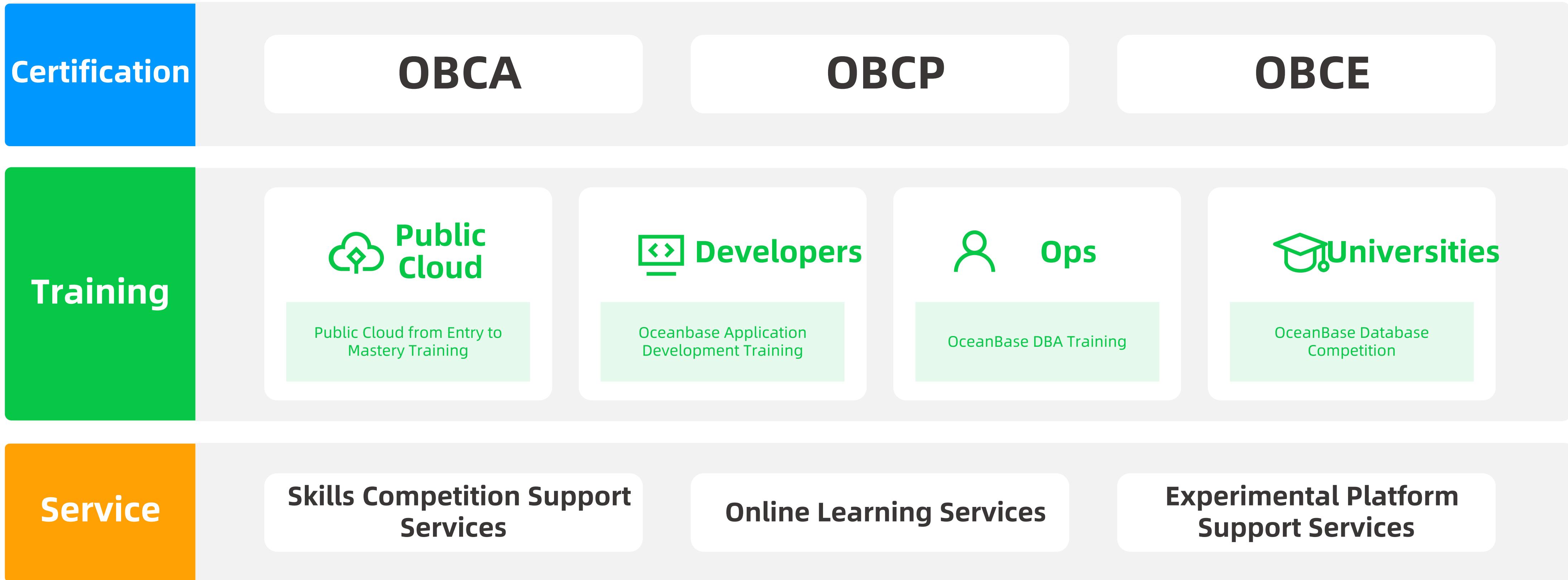


Uneven talent pool

The differences in scientific and technological capability reserves are becoming more prominent, there is a skill gap between industries, and there is an uneven distribution of talent between regions and cities.

Oceanbase Database Talent Training Framework

OceanBase's talent training philosophy: Innovation, Openness, and Cooperation



Ops: Operation and maintenance personnel

OceanBase Database Technology Certification



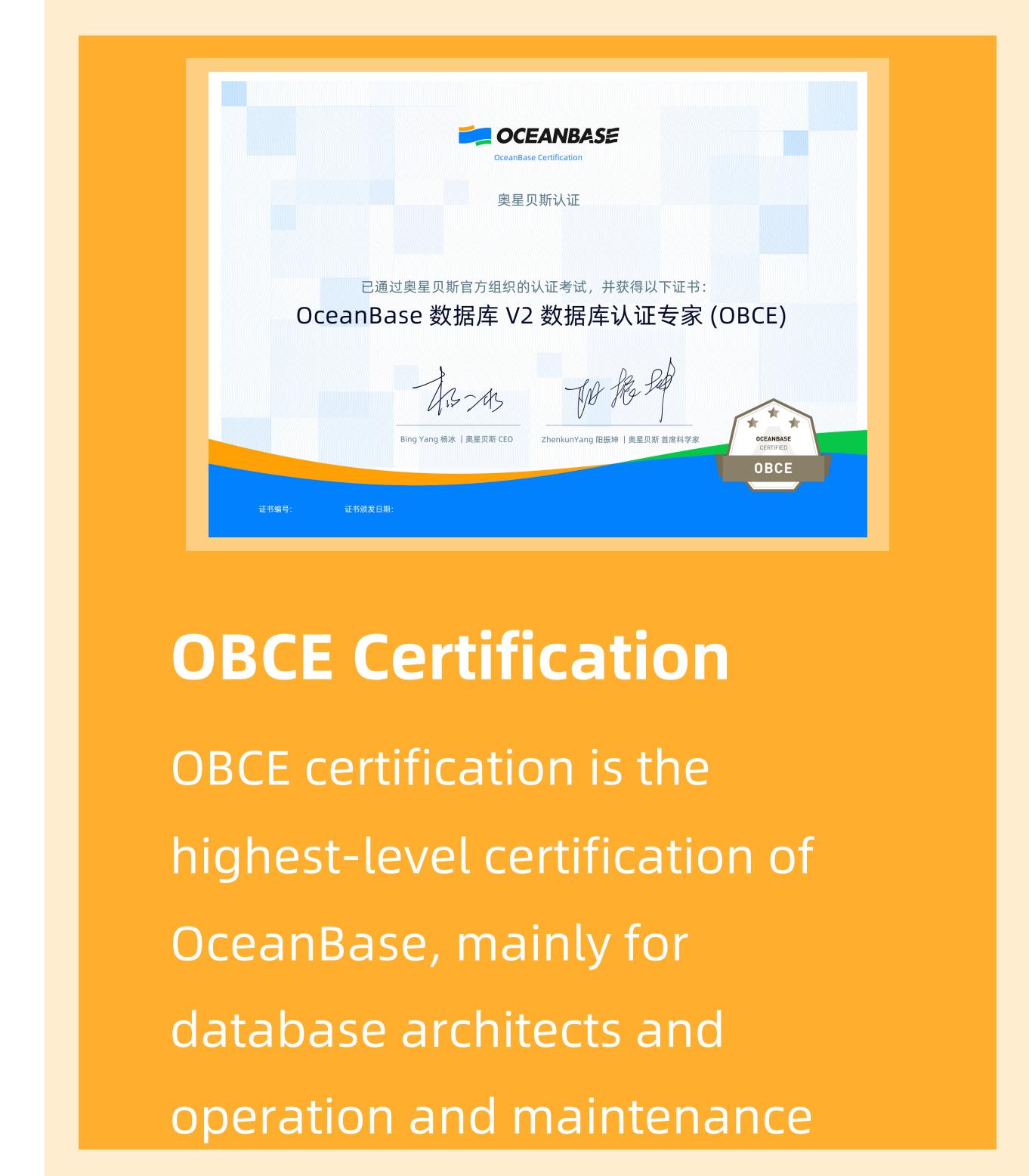
OBCA Certification

OBCA certification is suitable for junior database administrators, junior application developers, partner on-site service personnel, etc.



OBCP Certification

OBCP certification is suitable for architects, senior database administrators, senior application developers, and other senior after-sales service



OBCE Certification

OBCE certification is the highest-level certification of OceanBase, mainly for database architects and operation and maintenance

OceanBase Certification Benefits

| Learn leading distributed database technology

| Obtain authoritative professional database capability certification

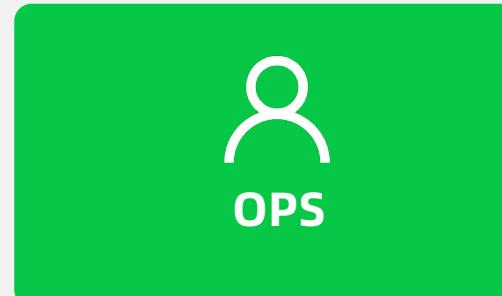
| Priority recommendation to enter the OceanBase industry chain

OceanBase Database Product Course

Based on completely self-developed code-level teaching materials

- 100% self-developed native distributed relational database. The core content of the course comes from self-developed research results, rich core knowledge of distributed autonomous and controllable databases;
- Keep up with the version evolution and functional features, and update teaching materials in real-time, through OceanBase training, students can experience the latest product features and master practical operations in the first place;

Design a post-based and hierarchical curriculum system based on different job responsibilities



- Public cloud-independent course system, mainly in the form of MOOC
- Help developers and DBAs get started quickly

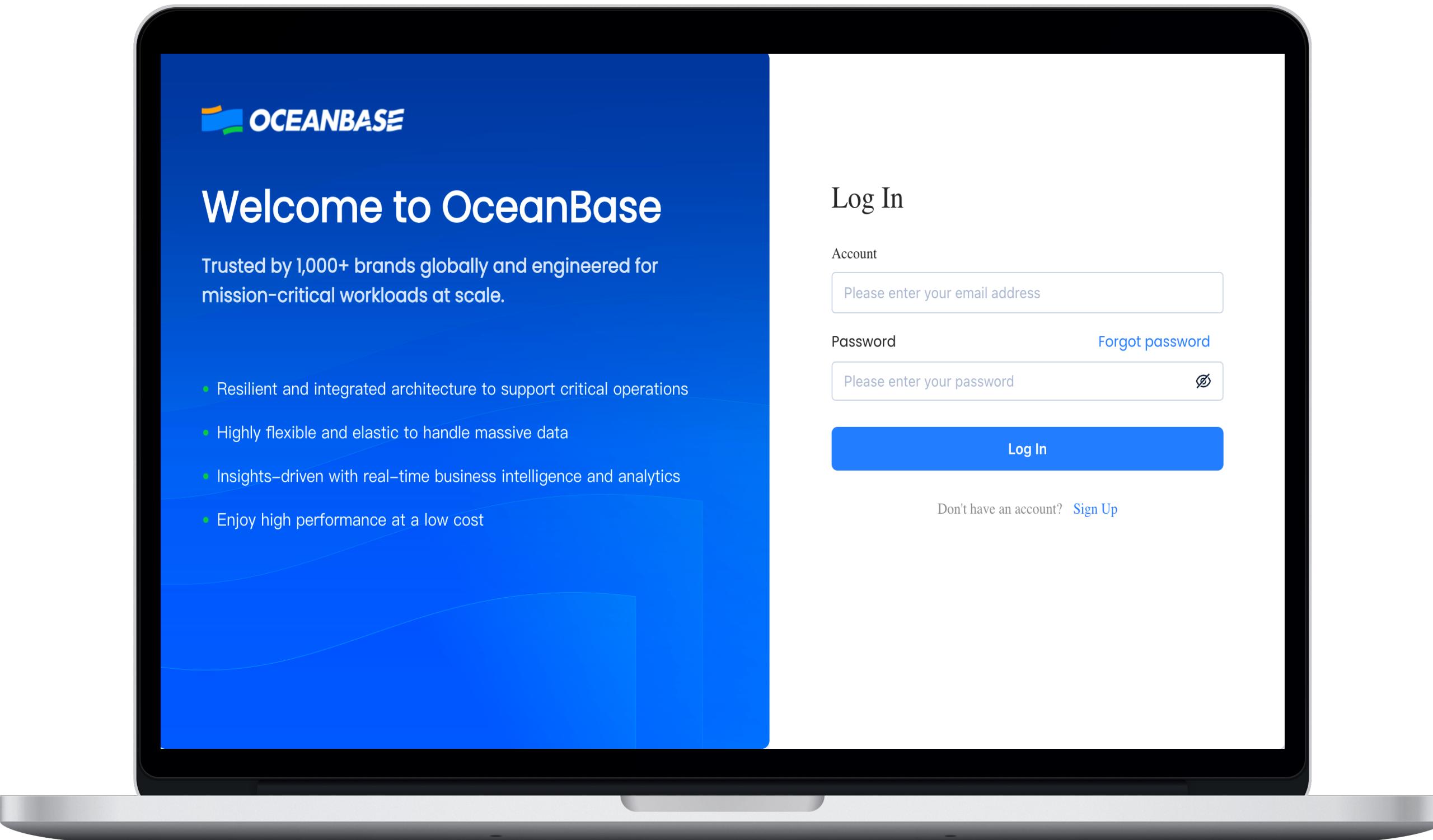
- Development based on distributed database;
- SQL performance tuning;

- Design and plan database construction plan;
- Deploy, install, manage and optimize database system;

DEV: Developer

OPS: Operation and maintenance personnel

Oceanbase Database Talent Training Course Sample (Public Cloud)



OceanBase Public Cloud: Getting Started to Mastery

Target Audience:

- Public cloud database DBA/application development engineer

Learning Goal:

- Quickly master the use of public cloud databases
- Have data migration capabilities
- Application interface adaptation and optimization

Course Outline:

- Cluster creation and resource management
- Database migration
- Application access and SQL programming
- Database object design and development specifications
- SQL monitoring and tuning
- System operation and monitoring

Oceanbase Database Talent Training Course Sample (DEV)



OceanBase Development Engineer Training (DEV)

Target Audience:

- Application Development Engineer

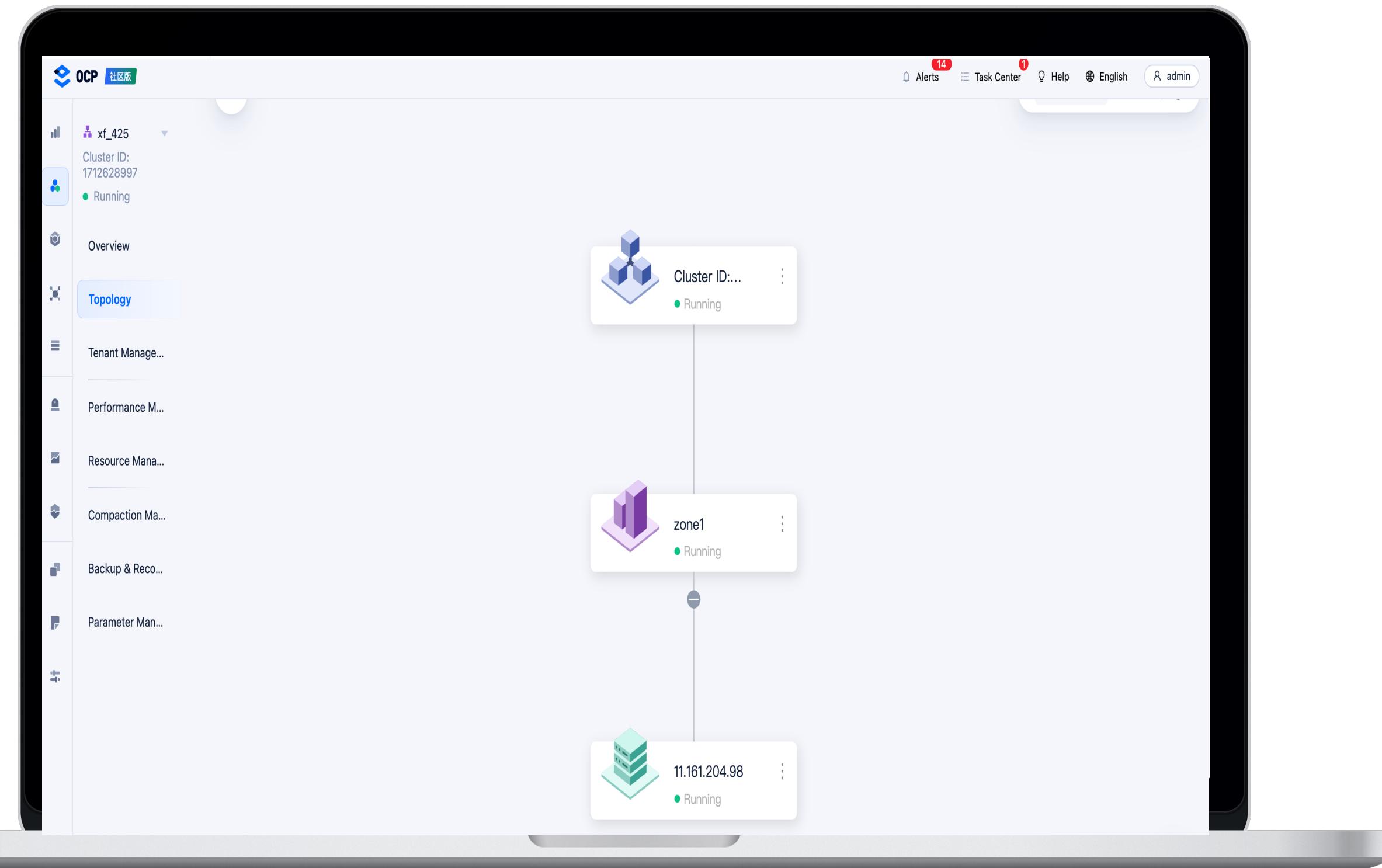
Learning Objectives:

- Connect and access OceanBase
- Basic OB database development code
- Development based on OB distributed database features

Course Outline:

- Introduction to OceanBase Database
- Access OB Database
- OB Database Common Object Management (Table)
- OB Database Common Object Management (Index)
- Distributed Database SQL Statements
- SQL Statement Performance Optimization
- OB Data Export & Import

OceanBase Database Talent Training Course Sample (Operation and Maintenance)



OceanBase Operation and Maintenance Engineer Training (DBA)

Target Audience:

- Database operation and maintenance engineer

Learning Objectives :

- Understand OceanBase infrastructure, core features and key technologies
- Install and deploy OBProxy and master usage skills
- Perform backup and recovery tasks
- Daily operation and maintenance to implement database performance monitoring and daily alarm processing

Course Outline:

- OceanBase cluster technology architecture
- OceanBase SQL and storage engine
- Operation and maintenance platform and basic management
- Data distribution, links, distributed transaction management
- Backup and recovery management
- User management, data monitoring and exception handling
- SQL basic tuning

OceanBase Training Advantages



Success Stories

- OceanBase training has served clients in multiple industries including finance, telecommunications, and energy.
- Train more than 10,000 students to obtain OceanBase certificates every year
- Offline training centers and certification test sites in 10+ cities across the country



中国工商银行
THE PEOPLE'S BANK OF CHINA



中国工商银行
INDUSTRIAL AND COMMERCIAL BANK OF CHINA



中国建设银行
China Construction Bank



中国人寿
CHINA LIFE



招商证券
China Merchants Securities



中国石化
SINOPEC



中国移动
China Mobile



国家电网
STATE GRID



住房公积金



Thank You!

 OceanBase Official website:
<https://oceanbase.github.io/>

 GitHub Discussions:
<https://github.com/oceanbase/oceanbase/discussions>

