GeminiDB

Technical White Paper

Issue	01
Date	2021-12-30

HUAWEI

HUAWEI TECHNOLOGIES CO., LTD.

Copyright © Huawei Technologies Co., Ltd. 2023. All rights reserved.

No part of this document may be reproduced or transmitted in any form or by any means without prior written consent of Huawei Technologies Co., Ltd.

Trademarks and Permissions


and other Huawei trademarks are trademarks of Huawei Technologies Co., Ltd.
All other trademarks and trade names mentioned in this document are the property of their respective holders.

Notice

The purchased products, services and features are stipulated by the contract made between Huawei and the customer. All or part of the products, services and features described in this document may not be within the purchase scope or the usage scope. Unless otherwise specified in the contract, all statements, information, and recommendations in this document are provided "AS IS" without warranties, guarantees or representations of any kind, either express or implied.

The information in this document is subject to change without notice. Every effort has been made in the preparation of this document to ensure accuracy of the contents, but all statements, information, and recommendations in this document do not constitute a warranty of any kind, express or implied.

Security Declaration

Product Life Cycle

Huawei's regulations on product life cycle are subject to the Product End of Life Policy. For details about the policy, see the following website:https://support.huawei.com/ecolumnsweb/en/warranty-policy

Vulnerability

Huawei's regulations on product vulnerability management are subject to "Vul. Response Process". For details about the policy, see the following website:https://www.huawei.com/en/psirt/vul-response-process For enterprise customers who need to obtain vulnerability information, visit:https://securitybulletin.huawei.com/enterprise/en/security-advisory

Preconfigured Digital Certificate

Huawei has released the Huawei Preset Digital Certificate Disclaimer for the preconfigured digital certificates delivered with devices. For details about the disclaimer, visit the following website:https://support.huawei.com/enterprise/en/bulletins-service/ENEWS2000015789

Life Cycle of Product Documentation

Huawei released the Huawei Product Documentation Lifecycle Policy for after-sales customer documentation. For details about this policy, see the website of Huawei's official website:https://support.huawei.com/enterprise/en/bulletins-website/ENEWS2000017761

Contents

 Copyright © Huawei Technologies Co., Ltd.

1 Architecture Overview

GeminiDB is a distributed, multi-model NoSQL database service with decoupled compute and storage architecture. It is highly available, secure, and scalable and provides service capabilities including quick deployment, backup, restoration, monitoring, and alarm reporting. There are four types of GeminiDB products: GeminiDB Cassandra, GeminiDB Mongo, GeminiDB Influx, and GeminiDB Redis. They can provide high read/write performance at a low cost, and is well suited to sectors like IoT, the Internet, and gaming applications.

GeminiDB architecture has the following features:

- With decoupled compute and storage, compute and storage resources can be expanded separately.
- Compute nodes share full load.
- Strong consistency of multiple copies is supported at the storage layer.
- The distributed architecture and shared storage means that data migration is not required during scale up or down.
- The storage layer provides ultra-low latency data access.
- Fast and elastic scaling is provided.
- Efficient snapshot backup and recovery are provided.
- Strong consistency of multiple copies is provided in an AZ or across AZs.
- Horizontal expansion and storage pool sharing are supported.
- Ultra-low latency access (100 μs) is provided.

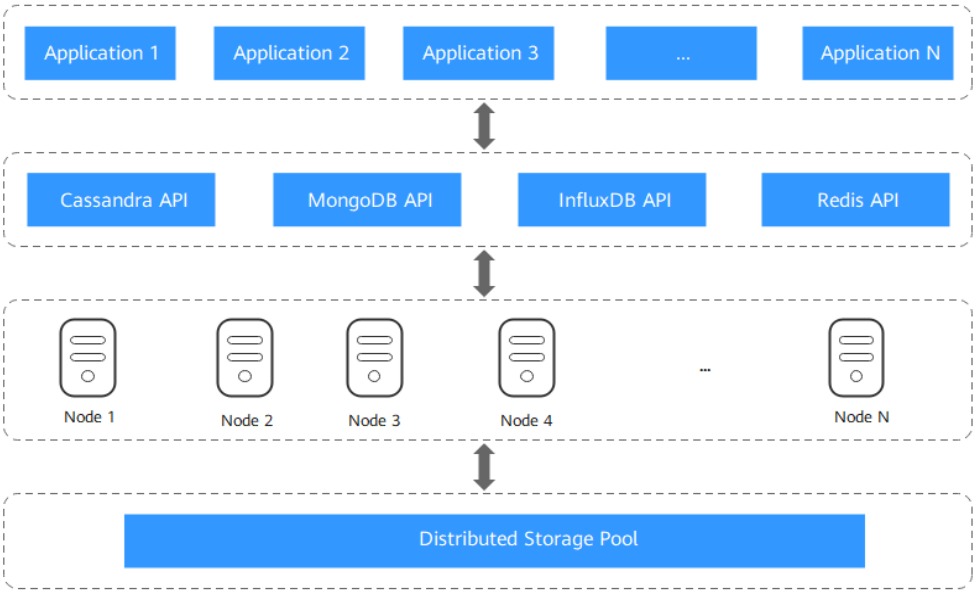
 Copyright © Huawei Technologies Co., Ltd.

2 Key Features

2.1 Database High Availability

GeminiDB uses the decoupled compute and storage architecture. Multiple DB instances in a cluster share the underlying distributed storage. The faults that may occur include compute node fault and storage node fault.

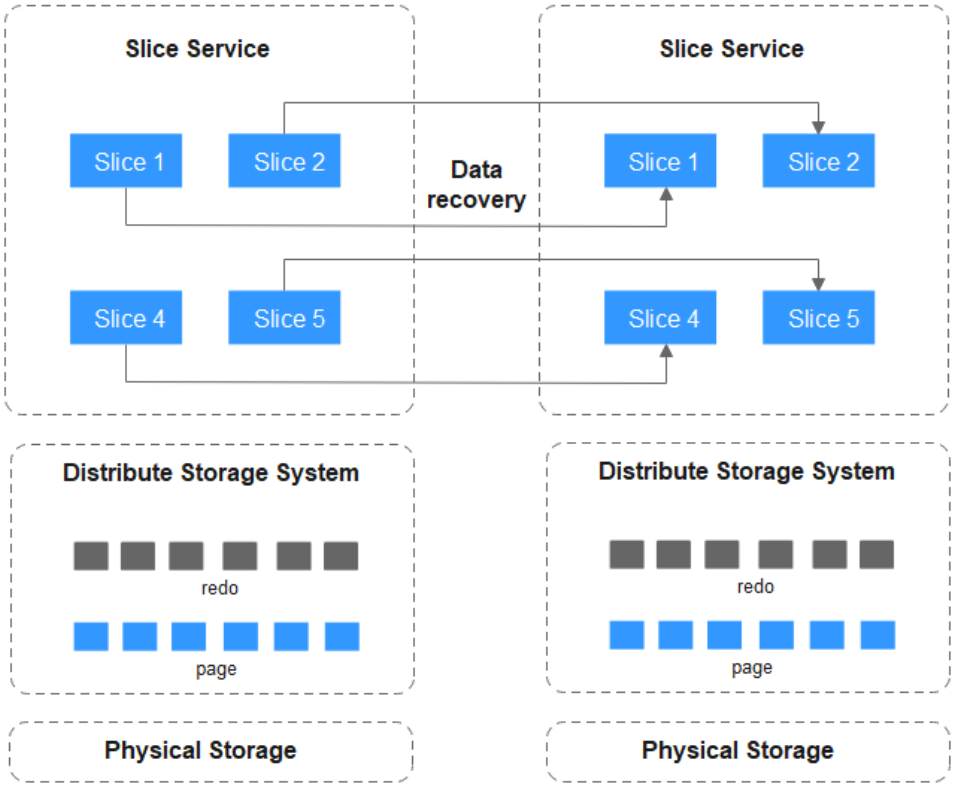
Figure 2-1 Diagram



- Thanks to the shared storage, data restoration is not required when a compute node becomes faulty. Once a node is faulty, other nodes that are running properly can directly take over services in seconds, and this process is unnoticeable to customers' applications due to the persistent connections provided by the upper-layer proxy.
- There are three data copies in the underlying storage. A single point of failure will not cause the entire system to be unavailable.

 Copyright © Huawei Technologies Co., Ltd.

Figure 2-2 Underlying storage diagram



2.2 Data High Reliability

All data of GeminiDB databases is stored in three copies on the underlying DFV distributed storage.

- If a storage node is faulty, the distributed storage mechanism ensures that reads and writes are routed to other available nodes.
- If a shard on an available storage node becomes faulty, the upper-layer service reports the faulty shard and route reads and writes to other available nodes.
- Faulty shards can be automatically repaired through the gossip protocol. The repaired shard can continue to provide services. If a shard fails to be repaired, a new shard needs to be rebuilt.

2.3 Online Elastic Scaling and Service Expansion

With the native decoupled compute and storage architecture, GeminiDB can use stateless compute nodes, making service expansion easily.

Service expansion refers to the expansion of compute nodes and storage capacity.

- GeminiDB provides shared storage, pay-per-use billing, one-click capacity expansion, and up to 96 TB of storage capacity. Services are not interrupted during the capacity expansion.

 Copyright © Huawei Technologies Co., Ltd.

- Compute nodes support both horizontal and vertical expansion.
  a. The number of shards can be increased by adding compute nodes because the compute nodes are stateless. You do not need to purchase additional storage devices or rebuild data. No matter how large the data volume is, it takes about 5 minutes to add shards.
  b. GeminiDB is deployed in containers. You only need to modify compute resources through adjusting container parameters, and no disk needs to be attached. In this way, specifications can be changed rapidly.

2.4 Efficient Backup and Disaster Recovery

Traditional database systems use file or block storage, which means that backup and restoration speed depends on the data volume. If there is a large amount of data, quick backup and restoration cannot be achieved, and data reliability is impacted.

GeminiDB uses the append-only storage in the DFV storage pool and optimizes the distributed parallel algorithm based on the database logic, improving the backup and restoration performance.

NOTE

Data Function Virtualization (DFV) is a high-performance and high-reliability distributed storage system that is vertically integrated with databases. Storage clusters are deployed in pools to improve storage utilization.

- Powerful data snapshot processing capability: Data is stored in append-only mode in multiple copies at multiple time points. Snapshots are generated within seconds, and massive amounts of snapshots can be processed.
- Parallel high-speed backup and restoration: Backup restoration logic is pushed down to storage nodes. Data is accessed locally and directly interacted with storage systems, improving database performance and concurrency.
- Quick instance recovery
- Full functions are available in minutes leveraging the asynchronous data synchronization and on-demand real-time data loading mechanism.

 Copyright © Huawei Technologies Co., Ltd.