AnalyticDB : Real-time OLAP Database System at Alibaba Cloud

Chaoqun Zhan, Maomeng Su, Chuangxian Wei, Xiaoxiang Peng, Liang Lin, Sheng Wang, Zhe Chen, Feifei Li, Yue Pan, Fang Zheng, Chengliang Chai
Alibaba

Abstract: In this paper, we introduce AnalyticDB, a real-time OLAP database system developed at Alibaba. AnalyticDB maintains all-column indexes in an asynchronous manner with acceptable overhead, which provides low latency for complex ad-hoc queries. Its storage engine extends hybrid column layout for fast retrieval of both structured data and data of complex types. To handle large-scale data with high query concurrency and write throughput, AnalyticDB decouples read and write access paths. To further reduce query latency, novel storage-aware SQL optimizer and execution engine are developed to fully utilize the advantages of the underlying storage and indexes. AnalyticDB has been successfully deployed on Alibaba Cloud to serve numerous customers (both large and small). It is capable of holding 100 trillion rows of records, i.e., 10PB+ in size. At the same time, it is able to serve 10m+ writes and 100k+ queries per second, while completing complex queries within hundreds of milliseconds.

Lambda architecture

- Support for insert
- Block index for incremental data
- Multi-version
- Block level data at insert
- Copy-on-write for backup
- Support snapshot read
- Support data and update
- Merge
- Incremental index build
- Version-based merge
- Merge in background
- Data vacuum

Execution and Optimization

- Pipelining
- Codegen
- Vectorsized execution
- Memory pool/cache

Storage Architecture

- Read/Write decoupling
  - High throughput query
  - High-throughput write
  - High scalability
- High availability
- Fault-tolerant
- Self-healing
- High availability
  - Up to 1024 nodes/DB
- High scalability
  - High-throughput query
  - High-throughput write
  - Read/Write decoupling

Index and Data Store

- All-columns indexing
- Multi-dimensional analysis
- Multi-version
- Time/size based merge
- Incremental index build
- Support delete and update
- Support snapshot read
- Copy-on-write for dedup
- Mark for delete with bitsets
- Lambda architecture
- Block index for Incremental data
- Support fast insert

Example:

- TPC-H 22 Queries
- Full scan query
- Point lookup query
- Multi-join query
- Complex query
- Live updates
- 10k+ TPS
- 10+ OPS

Capabilities

- Multi-dimensional analysis
- Complex long computing
- Semi-structured, large fields
- Live Read/Write

Accuracy

- Full text and JSON
- int/varchar/time/date/
- Various data type

Concurrence

- K-way merge for indexing results
- Wide table

Agility

- Arbitrarily join
- Arbitrarily filter

Low latency

- 95% 50ms
- 10K+ QPS

High availability

- 100% 95% 10ms