

基于 SelectDB 构建 PB 级日志可观测系统

引领大模型 DevOps 最佳实践

香克斯

MiniMax 可观测架构师

2025/04/19

CONTENT 目录

01 MiniMax 简介

人工智能领域最具潜力的独角兽企业

02 可观测日志系统的探索与挑战

Apache Loki 和 Apache Doris 两次选型尝试

03 日志系统最佳实践：阿里云 SelectDB

更好的架构、更好的管控、更高的性价比

04 业务价值 & 总结展望

实现性能和降本增效双重效果

01 MiniMax 简介

人工智能领域最具潜力的独角兽企业



MiniMax-01



Speech-02



Video-01



Music-01



MiniMax Chat

AI 智能助手



海螺AI

视频创作平台



星野

AI 原生应用



Talkie

AI 原生应用

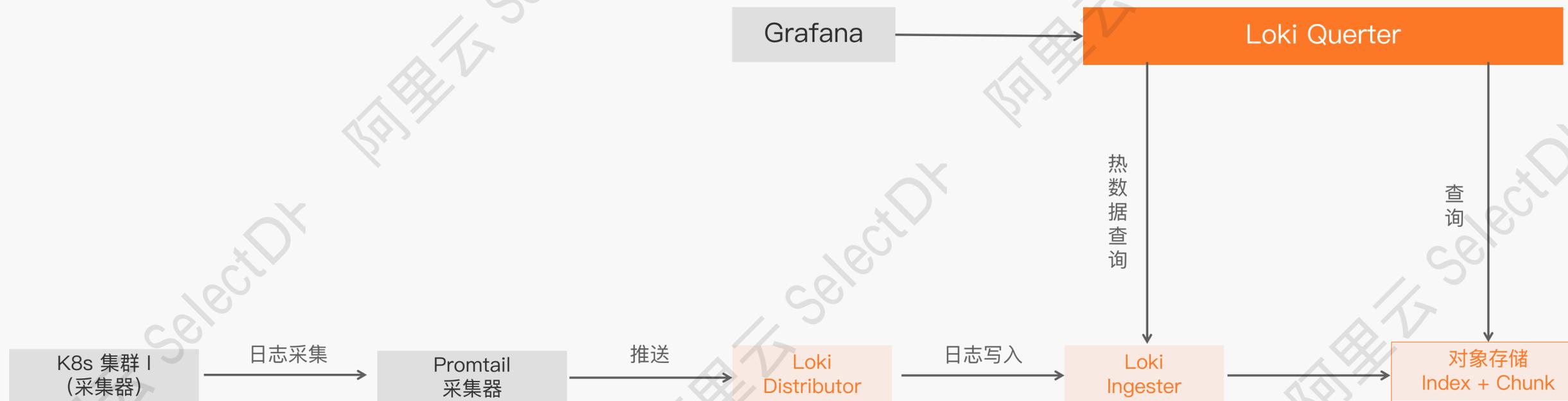
MiniMax 已然跻身人工智能领域最具潜力的独角兽企业之列



02 可观测日志系统的探索与挑战

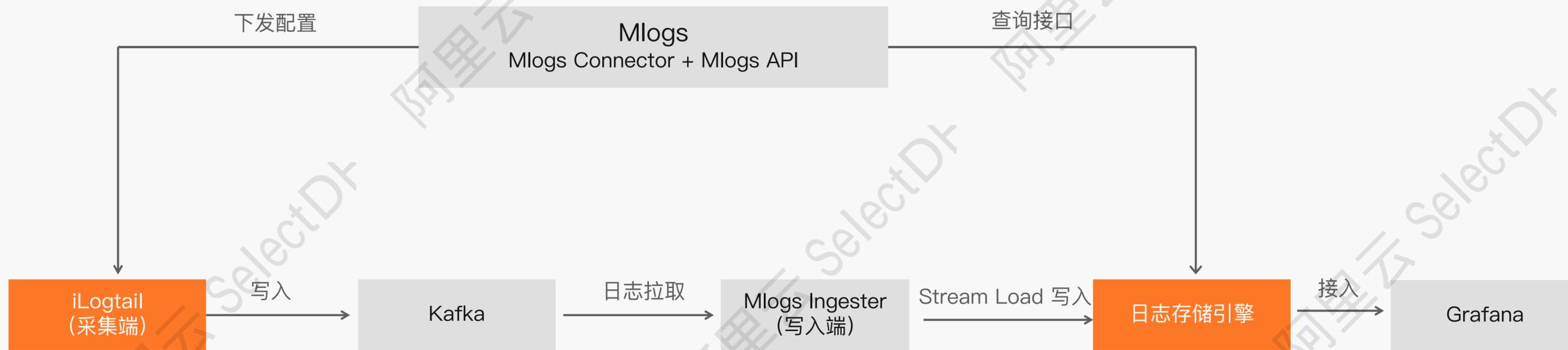
Apache Loki 和 Apache Doris 两次选型尝试

Loki 架构的尝试与局限性



- **MiniMax 初期架构**: 轻量化的 Grafana + Apache Loki, Promtail 负责采集日志给 Loki, Grafana 用以 UI 展示。
- **核心痛点**: 业务发展, 日志规模攀升至 PB 级之后, Loki 在资源消耗、写入性能和查询易用性等方面暴露出瓶颈。
- **核心诉求**:
 - ①更高的查询性能: 支持上亿条数据的秒级查询响应
 - ②更低的存储成本: 在 PB 级日志数据规模下, 实现更具性价比的日志采集与存储方案

基于 Doris 的日志系统架构升级



核心升级点一：采用阿里云开源的 iLogtail 作为日志采集工具，将日志数据推送至 Kafka 消息队列

核心升级点二：Doris 作为核心存储与查询引擎，实现了全量日志数据的**统一管理**，数据总量 **PB 级**，基于**倒排索引**实现日志检索**秒级响应**

自建 Doris 集群的痛点

扩容数据迁移耗时

业务快速扩张导致数据和查询量激增，频繁的集群扩容需要进行数据迁移，因数据规模较大，迁移过程繁琐耗时。

业务之间资源抢占

日志可观测系统负责多个业务的数据分析，单实例多业务并发时，内部资源竞争和干扰导致实例稳定性和查询性能。

运维成本较高

自建 Doris 随着集群规模不断变大，运维成本较高，参数调优和集群管理耗费了大量的人力物力，人效成本较高。

问题解决效率不可控

在遇到 Apache Doris 内核相关问题时，社区支持有时候需要花费比较多的时间，解决问题的时间成本和风险不可控。

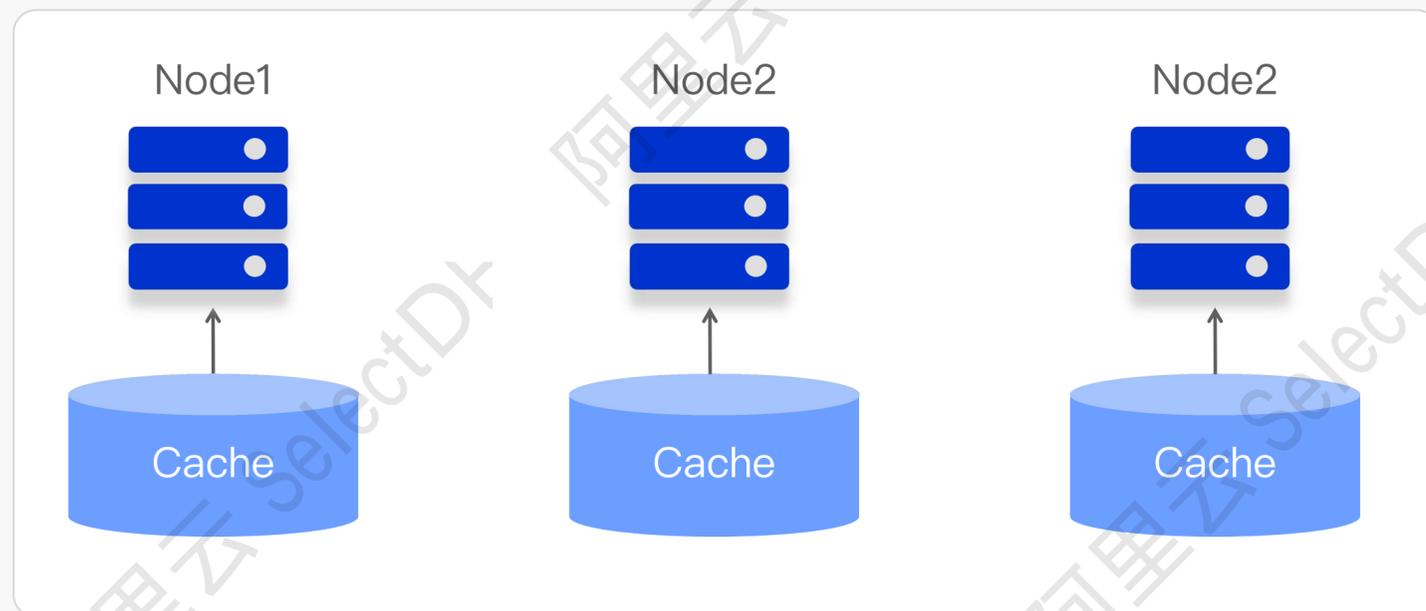
03 日志系统最佳实践：阿里云 SelectDB

更好的架构、更好的管控、更高的性价比

阿里云 SelectDB 技术方案优势

优势一：弹性伸缩，提高集群扩容效率

实例



存算分离

存储



01 兼容原有方案迁移成本低

SelectDB 沿用了 Apache Doris 的技术架构，100% 兼容 Doris 语法，并针对写入吞吐和查询性能等方面进行了深度优化。

02 解决Doris扩容瓶颈

Apache Doris 采用 MPP 架构，基于分桶逻辑进行数据的物理水平拆分，单个数据分桶节点可能会达到资源瓶颈。

03 弹性分离架构轻松应对业务流量

阿里云 SelectDB 采用存算分离的云原生架构，将计算与存储分层解耦，支持独立扩缩容。整体集群扩容时间可达到分钟级别，大大降低了运维成本，能够通过弹性伸缩能力迅速应对突发业务流量。

阿里云 SelectDB 技术方案优势

优势二：存算分离，提升吞吐并降低存储成本

A

Apache Doris

数据量大，在成本和性能之间很难两全

MiniMax 在使用 Apache Doris 集群时，为了实现数据高可用，生产环境采用Doris 的**两副本**模式，导致存储资源消耗和集群写入压力均增至单副本的两倍。此外，考虑到过高的存储成本，MiniMax 在 Doris 数仓中仅保留 15 天的业务数据，其他数据通过冷归档的方式存储；而需要对这部分归档数据进行查询分析时，则临时从归档库中解压加载后才能进行分析，极大降低了数据查询的效率。

VS

B

阿里云 SelectDB

更低的成本，更高的性能，更多的承载量

阿里云 SelectDB 采用存算分离的设计，存储层基于阿里云对象存储 OSS 提供存储服务。MiniMax 在使用 SelectDB 后，利用 OSS 的数据高可用能力，**计算引擎仅需单份数据写入，存储资源需求减少至 Doris 的 50%，实际业务写入吞吐能力提升超 20%**。此外，由于整体存储成本的降低，**SelectDB 支持对历史全量数据的实时查询分析**，大大提高了数据查询效率。

阿里云 SelectDB 技术方案优势

优势三：资源隔离，提高并发读写效率

实例



存储

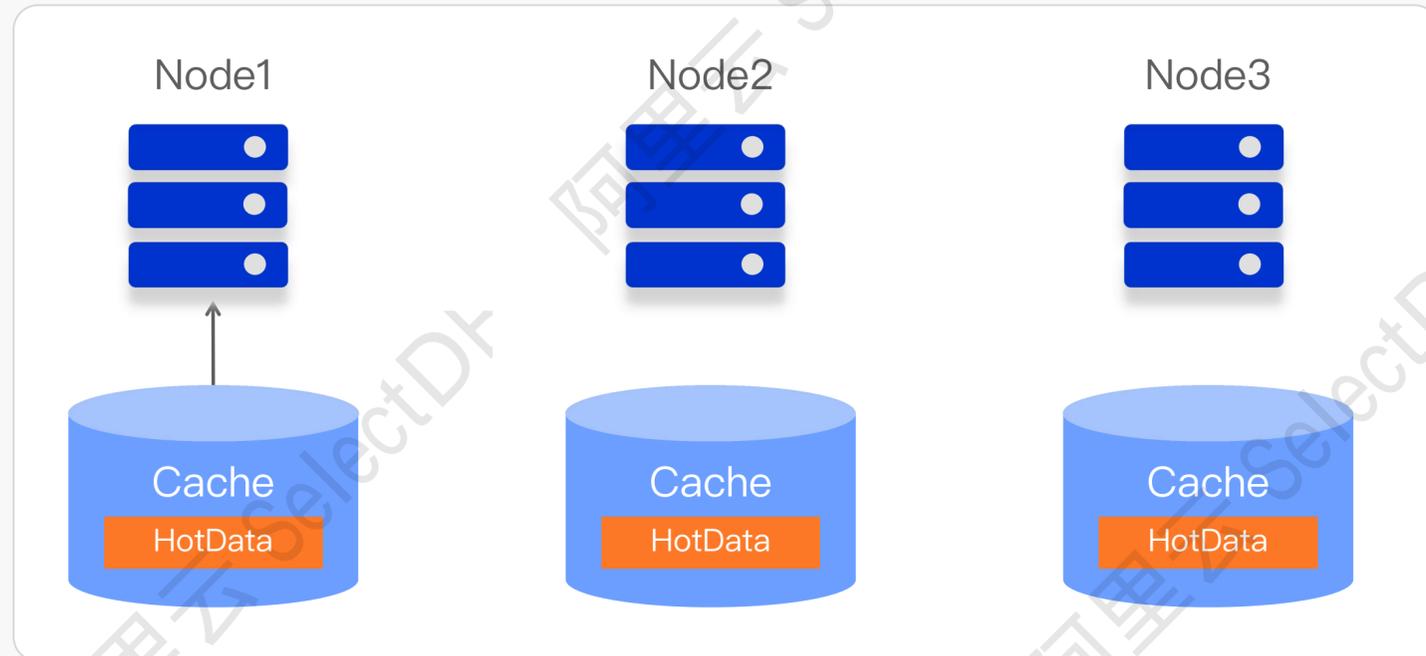
多集群资源隔离

阿里云 SelectDB 支持云原生多集群硬隔离能力，用户可以将单个实例的计算资源划分为多个逻辑集群，不同集群之间的分配独立的计算资源，实现不同集群的严格**物理资源隔离**和**数据共享**，很好的解决负载隔离问题。此外，SelectDB 还支持**读写分离**能力，进一步提高了并发查询稳定性。MiniMax 在使用了 SelectDB 后，采用了 SelectDB 多集群隔离能力，并将读写集群分开，避免了读写资源抢占带来的实例稳定性问题，大大提高了并发读写效率。

阿里云 SelectDB 技术方案优势

优势四：缓存加速，提供高吞吐与低延迟

实例



存算分离

存储

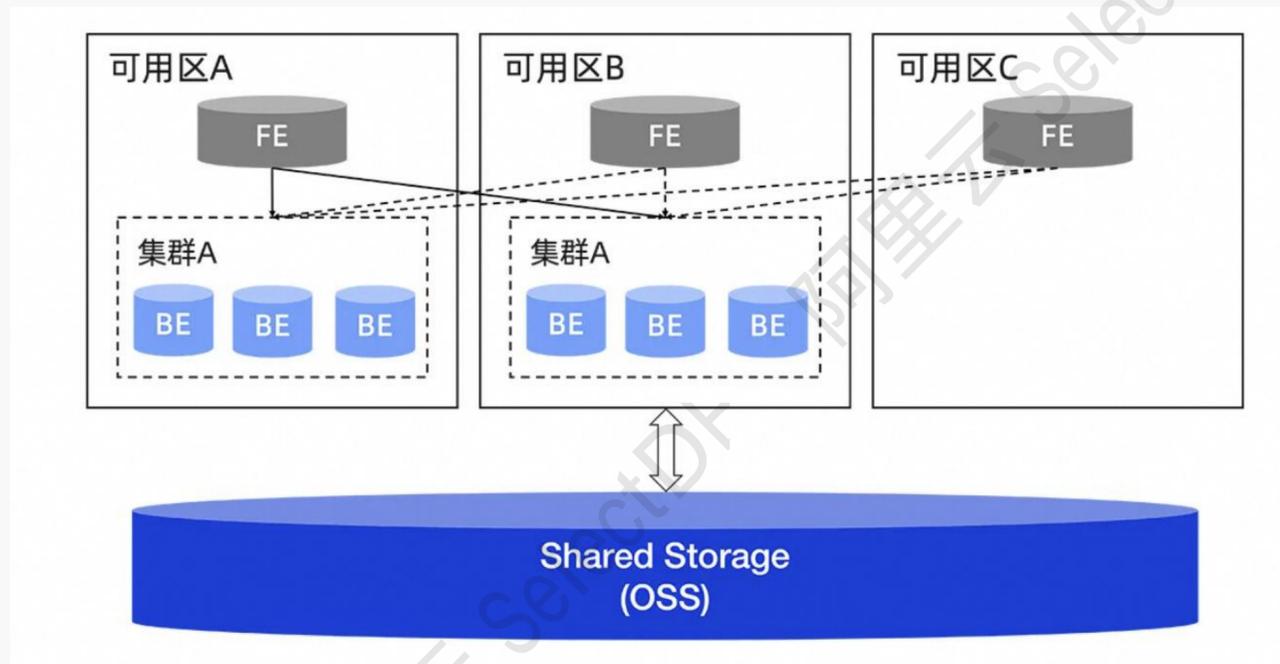


灵活的缓存机制保证查询效率

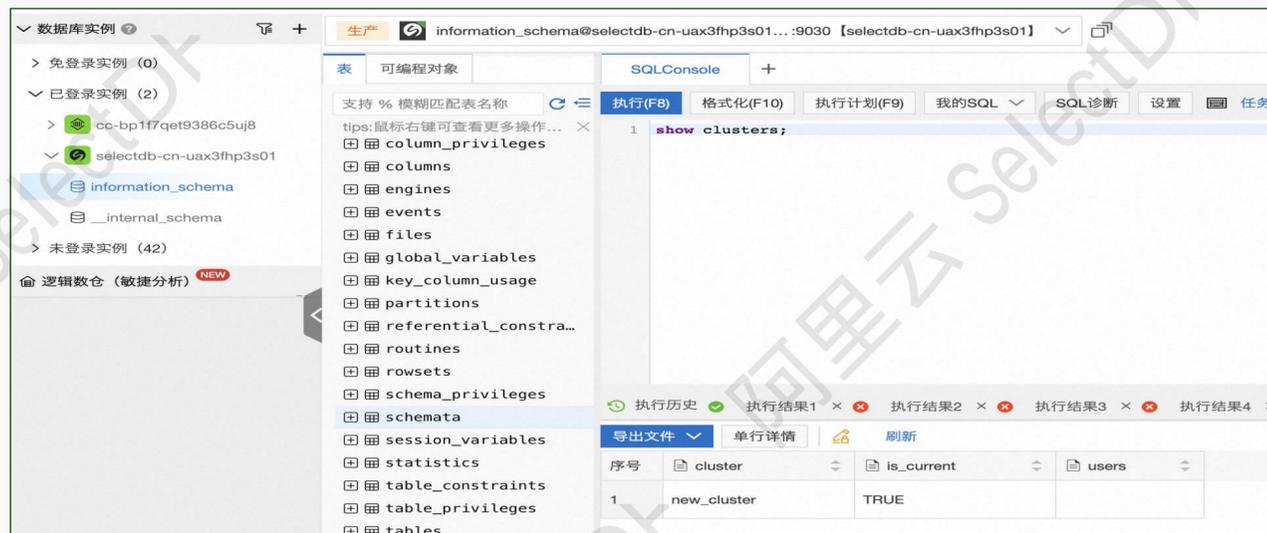
阿里云 SelectDB 通过单副本本地读写缓存、智能数据淘汰策略、高效列式存储格式和先进压缩算法，显著提升了海量数据的读写效率。业务进行数据查询时，依据 LRU 的读缓存策略，保证业务对于实时写入数据和高频查询热数据的查询性能。当发现缓存命中率低和查询性能不及预期时，可以进行实时的缓存空间扩容，以提升缓存命中率，PB 级数据 P95 查询可以在 3 秒内响应，提高了数据查询效率。

阿里云 SelectDB 技术方案优势

优势五：高 SLA 和运维白屏化



DMS: 可视化数据管理



监控告警



白屏化参数管理

参数名	参考值	类别	当前值	可修改参数值	描述
disable_storage_row_cache	true	查询	true	[true false]	
doris_remote_scanner_thread_pool_threa...	96	查询	96	[0-10000]	
doris_scanner_thread_pool_thread_num	48	查询	48	[0-10000]	
es_http_timeout_ms	5000	查询	5000	[0-100000]	
es_scroll_keepalive	5	查询	5	[0-14400]	
file_cache_max_file_reader_cache_size	100000	查询	100000	[0-2000000]	
jsonb_type_length_soft_limit_bytes	1048576	导入	1048576	[0-2147483640]	
load_process_max_memory_limit_percent	50	导入	50	[0-100]	
load_process_soft_mem_limit_percent	80	导入	80	[0-100]	
max_base_compaction_threads	4	后台任务	4	[0-100]	
max_cumulative_compaction_threads	10	后台任务	10	[0-200]	
max_segment_num_per_rowset	5000	导入	5000	[0-1000000]	

04 业务价值&总结展望

更好的架构、更好的管控、更高的性价比

业务价值 & 总结展望

业务价值

1 业务规模

总体数据规模超过数 PB，日均新增日志写入量达数百 TB。

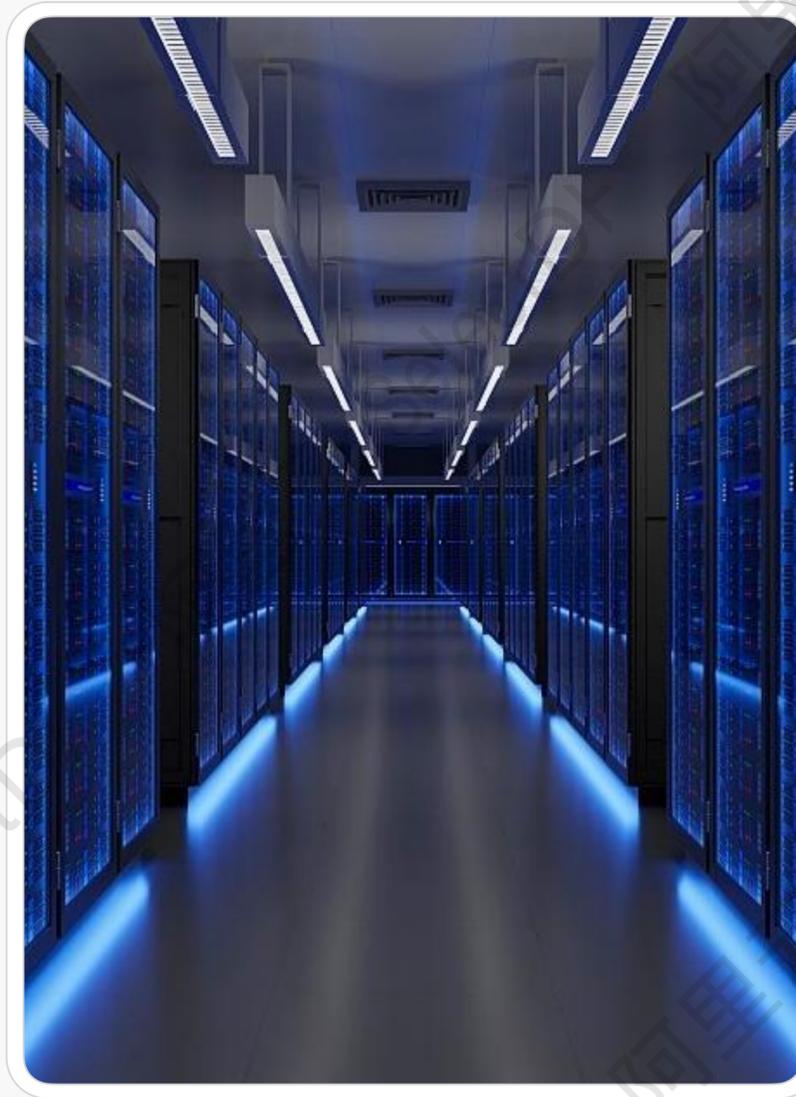
2.性能指标

日志系统查询 P95 分位延迟小于 3 秒

峰值写入吞吐超过 10GB/s, 1000万条/秒

3.降本增效

计算资源用量降低 40%，热数据存储用量降低 50%



总结展望

- 回顾 MiniMax 可观测系统的演进历程，从初期的 Loki 架构到 Apache Doris 的引入，再到 SelectDB 的全面升级，每一次技术迭代都体现了 MiniMax 对业务需求的深刻理解和对技术创新的不懈追求。
- 未来，随着 MiniMax 业务的持续高速发展，日志可观测系统将继续作为洞察系统运行状态和优化性能的核心工具。阿里云将与 MiniMax 携手，进一步挖掘日志数据的潜在价值，为业务创新提供更强有力的支持。

THANKS



计算,为了无法计算的价值