

腾讯云TCHouse-D：基于Apache Doris 打造全托管企业级实时数仓平台

沈彪

腾讯云 TCHouse-D 产品负责人

目录

1. 腾讯云 TCHouse-D 产品概述
2. 客户需求及痛点简析
3. 打造稳定易用的全托管Doris服务

1 腾讯云 TCHouse-D 产品概述

腾讯云 TCHouse-D 的产品架构及定位

腾讯云数据仓库 TCHouse-D 基于 Apache Doris 内核构建，兼容 MySQL 协议，与云上大数据生态体系融合，提供丰富的集群管控能力，及完善的巡检告警体系，为客户提供简单易用、极简运维的云上全托管服务。



2 客户需求及痛点简析

腾讯云 TCHouse-D 在多个行业及场景均有落地案例

覆盖 20+ 行业的落地

- | | | |
|--------|------|-------|
| 游戏 | 汽车 | 教育 |
| 通用SaaS | 电商 | 智慧零售 |
| 社交娱乐 | 金融 | 医疗 |
| 能源 | 政法 | 工业云 |
| 交通运 | 智慧建筑 | |

覆盖多种应用场景

- | | |
|--|--|
|  实时报表 |  实时大屏 |
|  用户行为分析 |  广告归因分析 |
|  业务经营查询 |  实时在线业务 |

客户来源及主要诉求



OLTP数据分析

- 关注端到端整体链路的时效性（分钟/秒级）
- 数据实时写入、大数据量OLAP实时数据分析



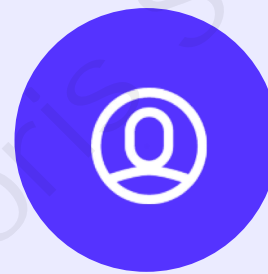
自建转化

- 专家支持，平台兜底
- 降低运维成本



Hadoop查询加速

- Hive查询加速
- 简化技术栈（替代Presto、Impala）
- 充分兼容现有大数据系统



一些资源降本诉求

- 友商迁移降本
- 日志场景降本

客户的常见痛点

实时OLAP分析场景

MySQL

- 字段格式兼容性问题
 - ✓ TIMESTAMP时区问题
 - ✓ VARCHAR精度问题等
- 数据迁移/接入效率
 - ✓ 成千上万张表的迁移
 - ✓ 分区、分桶设置等
- 联动稳定性问题
 - ✓ 失败重试机制
 - ✓ 企业级稳定性保障

自建 Doris

- 问题被动发现，缺乏主动预警及解决问题的能力
- 缺乏简便的数据备份和跨集群迁移方案
- 从其他系统迁移至Doris，缺乏应用改造的经验和专家支持

- Ranger兼容
- Hadoop版本兼容
- Kerberos
- Jar包兼容性

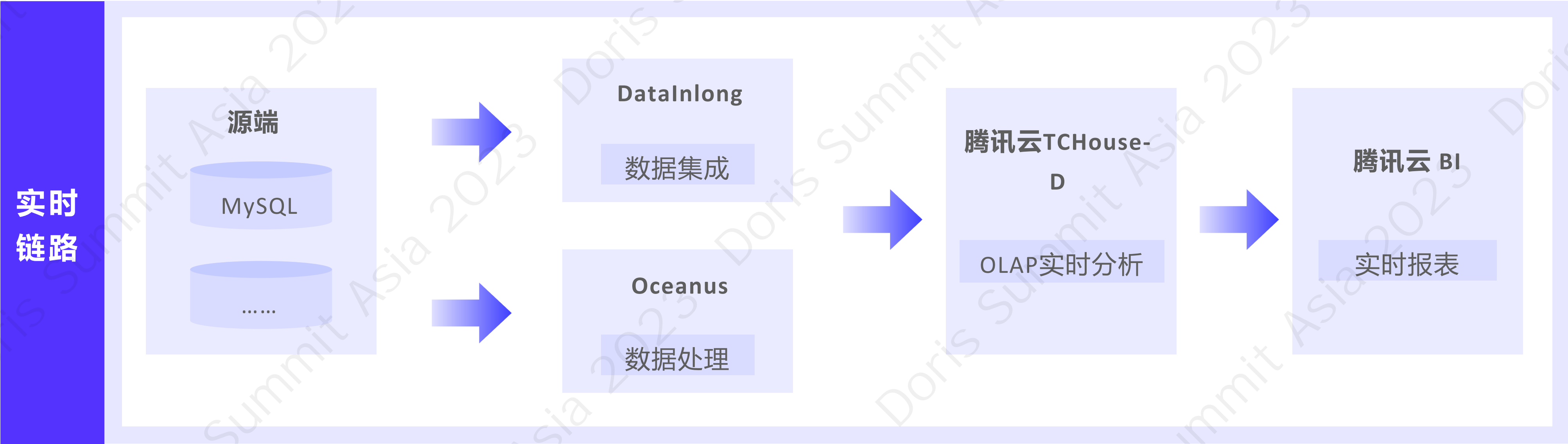
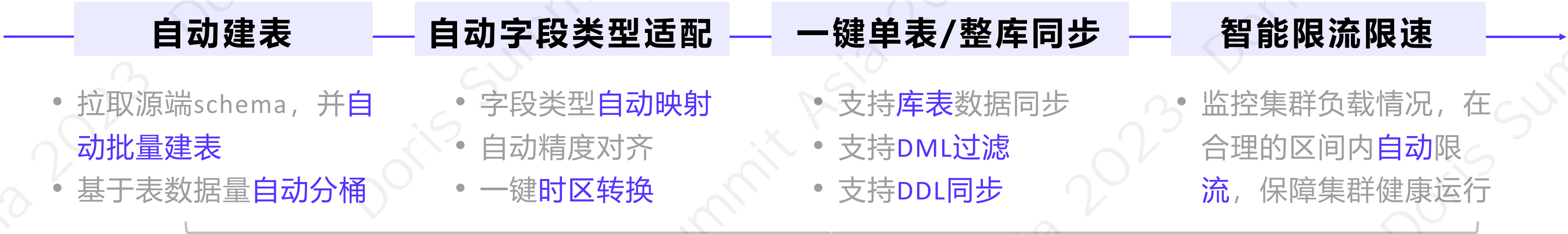
湖仓查询加速

Hive

Iceberg

3 打造稳定易用的全托管Doris服务

联动腾讯云上生态，快速接入数据并分析展示



一键数据备份恢复，一键跨集群迁移

■ 备份恢复

- 支持配置一次性、周期性任务
- 支持全量、按库、按表备份及恢复
- 支持精细化任务及实例管理
- 备份实例永久保留可追溯，备份进度可查看

■ 跨集群数据迁移

- 支持跨Doris集群一键数据/元数据迁移
- 全量、按库、按表迁移
- 支持COS数据迁移

备份恢复

备份进行中

备份服务

已开启

COS存储桶

carson-test-1301087413

数据备份或恢复过程中会因集群重启（如进行水平扩容，水平缩容，垂直变配，参数配置等操作）而导致数据丢失

支持对当前集群设置周期备份、一次性备份；一次性备份不影响周期备份策略，但同一个库表同一时间只能有一个正在执行的 BACKUP 或 RESTORE 任务，多个备份任务同时对一张表进行备份时，后续任务均会报错。为保障数据完整，待上个备份任务结束后，可通过“一次性备份数据”功能重新备份。

支持跨集群进行备份，即可以通过该方法将其他集群的数据通过备份和恢复的方式进行数据迁移。远端集群支持通过该功能进行数据迁移。

周期备份配置

一次性备份数据

数据迁移配置

备份/迁移任务管理

实例管理

任务名称	创建人	任务类型	任务状态
test-addtable	100025004453	周期备份	已完成
test-unstart-2	100025004453	周期备份	已完成
test-unstart-1	100025004453	周期备份	已完成
test-start-pause	100025004453	周期备份	进行中
test-unstart	100025004453	周期备份	进行中
test-im	100025004453	一次性备份	已完成

创建周期性备份任务

同一个库表同一时间只能有一个正在执行的 BACKUP 或 RESTORE 任务，多个备份任务同时对一张表进行备份时，后续任务均会报错。为保障数据完整，待上个备份任务结束后，可通过“一次性备份数据”功能重新备份。

调度配置

备份任务名称

backup-test

长度限制为6-36个字符，只允许包含中文、字母、数字、-、_、.

备份至COS桶

carson-test-1301087413

生效周期

2023-09-01 00:00:00 ~ 2023-10-31 00:00:00

调度周期

天

执行时间

天

备份范围

按表备份

备份粒度

按库备份

选择数据表

已选择 1 张表

确定

取消

8

backup succeed

2023-10-10 15:42:05

[查看实例](#) [编辑](#) [更多](#)

1

backup succeed

2023-09-21 20:07:59

[查看实例](#) [编辑](#) [更多](#)

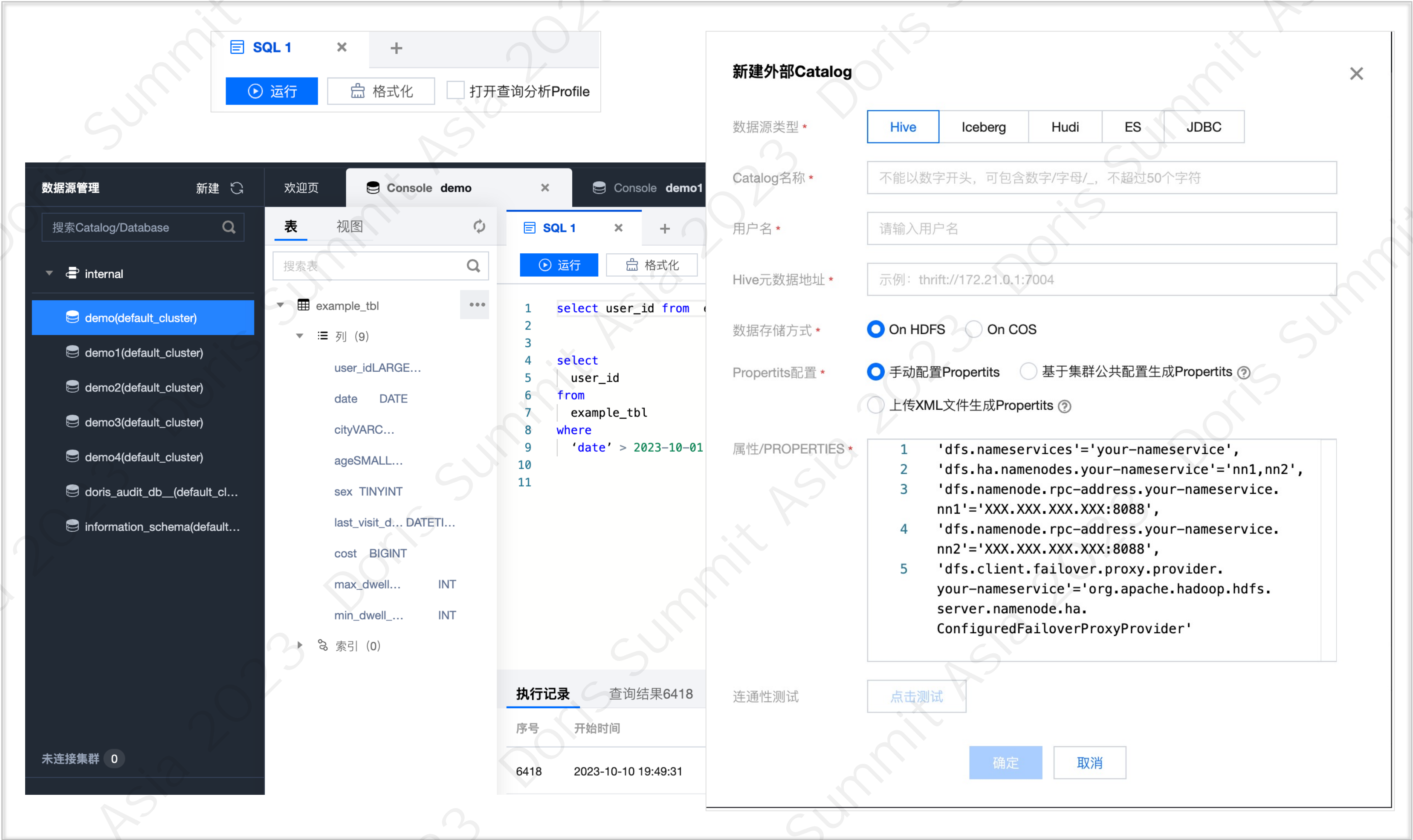
充分适配 Hadoop 生态，简化 Multi Catalog 配置

■ Catalog可视化注册

- 简单进行外部Catalog注册
- 支持on HDFS、COS多种存储方式
- 充分适配2.7、3.5的Hadoop版本
- 支持kerberos认证

■ SQL工作区便捷查询

- 支持源、库、表、视图预览
- 支持针对需要的SQL打开Profile



慢查询监控，可视化查询分析 Profile，高效业务调优

■ 问题感知

慢查询监控

慢查询告警

■ 初步分析

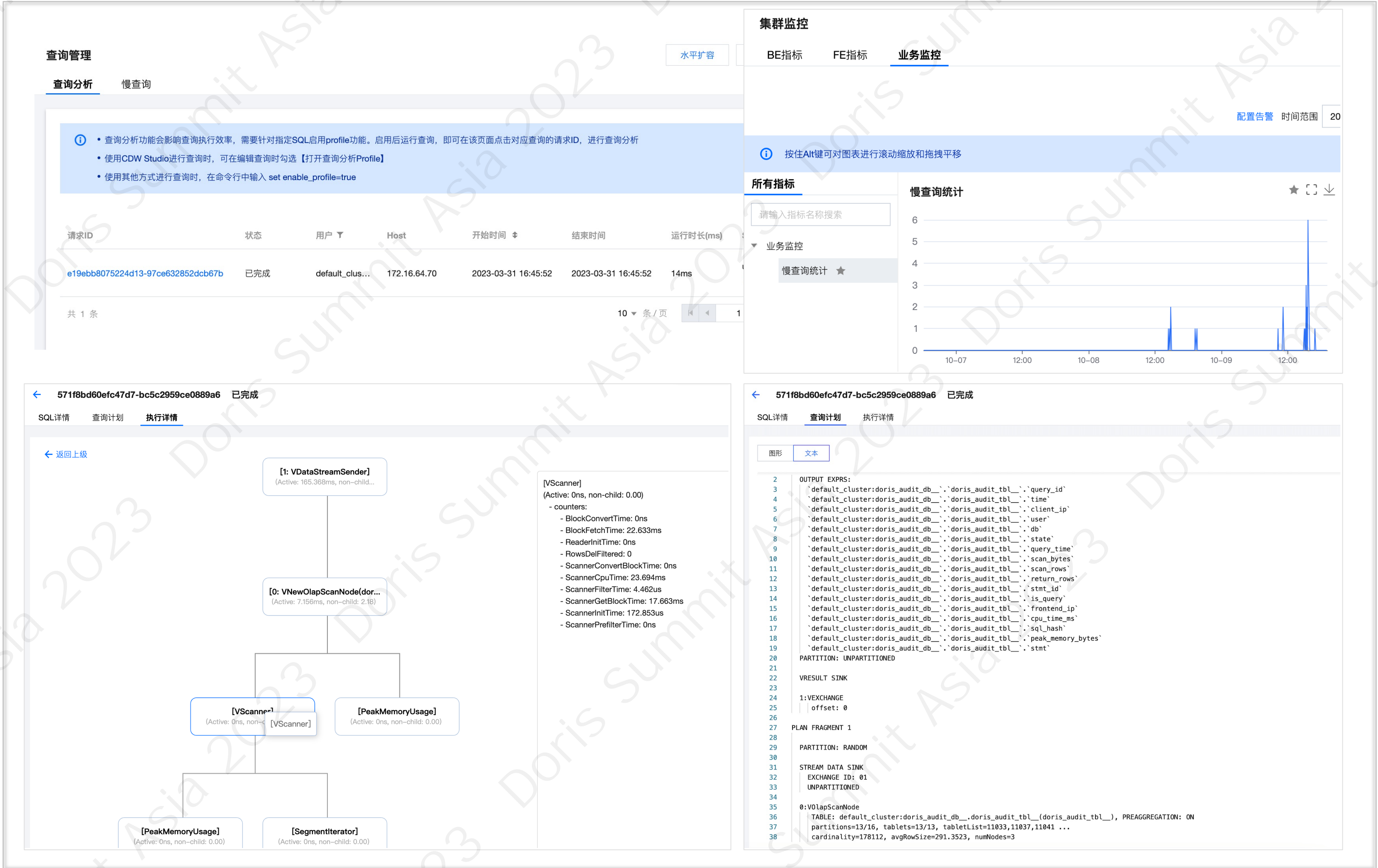
慢查询清单

慢查询日志下载

■ 深入分析

相关SQL打开Profile

查看查询计划&执行详情



丰富的集群+业务监控告警指标，常用指标一键关注



集群监控，直观定位问题



健康检查，风险高亮提示



实时巡检，问题提前发现



群聊/工单，问题快速解决



集群健康检查，主动进行预警，并给出处理建议



集群监控，直观定位问题



健康检查，主动提示风险



实时巡检，问题提前发现



群聊/工单，问题快速解决

资源预警：集群负载持续超过80%，磁盘使用率持续超过90%

状态告警：节点失活；指标预警：核心指标数值异常等

欢迎使用腾讯云数据库TCHouse-D [产品详情](#)

基于业内领先的 OLAP 数据库 Apache Doris 内核构建，为您提供方便易用、灵活稳定的实时数据仓库服务。只需要几分钟，便可完成数据仓库的搭建，简单轻松地实现对海量数据的实时查询分析。

集群概览

操作指引

地域总览

集群总数

8

运行中

7

变更中2个

创建中

0

异常

0

集群健康状态

运行中集群实时数据

7

服务不良集群

0

运行良好

ID/名称	集群运行状态	集群健康状态	异常概要
cdwdoris-mb1act07 studio12test	运行中	状态预警	3个FE节点资源预警、3个BE节点资源预警
cdwdoris-llod5pl2 tomas100801	运行中	状态预警	3个FE节点资源预警、3个BE节点资源预警
cdwdoris-anqy62nf liuming-es-test3fsdf	状态变更中	异常告警	3个FE节点不可用

网络信息

可用区

重庆一区

VPC ID

vpc-xxxxx

子网 ID

subnet-xxxxx

JDBC连接地址

10.0.12.188

HTTP连接地址

10.0.12.188

节点信息

序号	节点类型
1	FE
2	FE
3	FE
4	BE
5	BE
6	BE

集群私有网络IP

33	9.1.1.1
69	9.1.1.2
80	9.1.1.3
47	9.1.1.4
36	9.1.1.5

预警内容：近 5 分钟 BE 磁盘使用率频繁高于 1%，
问题影响：BE 写满，数据无法写入及更新，
处置建议：扩容 BE 磁盘，或删除不需要的表和分区(删除后表会进入回收区，磁盘真正的物理删除会默认在1天后，可提工单加速过程)

预警内容：近 5 分钟 BE 内存使用率频繁高于 1%，
问题影响：可能业务变慢或部分失败，
处置建议：建议降低读写频率或查询并发，或对 BE 进行垂直升配/水平扩容

预警内容：近 5 分钟 BE CPU 使用率频繁高于 1%，
问题影响：可能业务变慢或部分失败，
处置建议：建议降低读写频率或查询并发，或对 BE 进行垂直升配/水平扩容

预警内容：监控指标 DorisBeTabletBaseMaxCompactionScore > 0，
问题影响：数据可能部分写入失败，
处置建议：建议优化写入业务，降低写入频率，或对 BE 进行垂直升配/水平扩容

共 6 条

10 条 / 页

1 / 1 页

异常告警 + 实时巡检，降低 80%+ 问题升级，减轻对业务影响



集群监控，直观定位问题



健康检查，主动提示风险



实时巡检，问题提前发现



群聊/工单，问题快速解决



客户专项支持群 + 工单体系，保障问题解决效率



集群监控，直观定位问题



健康检查，主动提示风险



实时巡检，问题提前发现



群聊/工单，问题快速解决

小时级问题响应，资源不足一键扩容/升配，配置优化一键下发生效，Bugfix快速发版解决

VIP客户专项支持群

咨询建议

集群故障

群聊助手 + 值班人员支持

客户工单/巡检工单

客户咨询

调优建议

集群故障

SRE解答处理

疑难问题专家支持

一些后续规划



支持存算分离

- 支持直接读取对象存储的数据，并通过缓存等策略达到和本地盘相近的性能
- 支持冷数据自动写入对象存储



资源弹性扩缩容

- 产品容器化
- 支持分时弹性、按集群负载自动弹性



用户及作业资源隔离

- 支持用户配额、和SQL任务配额管理，资源隔离避免任务间互相影响



产品化支持集群读写分离

- 基于 CCR 跨集群数据同步能力，提升在线服务数据可用性，实现读写负载分离以及多机房容灾备份



获取更多社区动态与最佳实践

Apache Doris 官方平台:

- Apache Doris 官网: doris.apache.org
- Apache Doris GitHub: github.com/apache/doris/

获取更多峰会资料:

- Doris Summit 峰会官网: doris-summit.org.cn
- Doris Summit 峰会回放: <https://space.bilibili.com/1196172099/channel/collectiondetail?sid=1824324>