

Apache Doris 在中国电信 及省公司的应用

尚书杰 (shangsj@chinatelecom.cn)

中国电信 大数据部门技术专家

目录

1. TDP-Doris 的技术背景与产品介绍
 - 新一代 MPP 技术背景
 - TDP-Doris 产品介绍
2. TDP-Doris 在电信集团和省公司的应用
 - TDP-Doris 迁移案例 1
 - TDP-Doris 迁移案例 2
 - TDP-Doris 迁移案例 3
3. Doris 社区互动

1 TDP-Doris 的技术背景与产品介绍

为什么我们需要新一代 MPP 数据库（1）

Hadoop + Hive 的数仓体系存在弱点

- 数据入库实时性不足，大部分天级，少部分小时级，实质是离线数仓
- 数据查询交互性不足，分钟级、小时级
- 组件众多，复杂度高，门槛高
- 数据湖组件是加强，但还不够
 - 仍然存在以上问题，实际生产情况下，入库最多分钟级，无法涵盖所有场景
 - 仍在快速发展期
- 复杂的 kerberos 配置

更全面地看待大数据

- 大数据不只有大表大文件，还有中数据，小数据
 - 要解决的是数据问题，而不只是大数据问题
- 大数据不只有 batch insert 和批量查询
 - 还有 append, delete, update
 - 点查询,点更新
 - 还有 Rullup
- Google 不只有三架马车，还有其它组件
 - 比如 Google Mesa
 - Google 也不是大数据的开始，也不是大数据的结束，只是大数据的插曲
- 新一代 MPP 蓬勃发展
 - Clickhouse/Doris 等

为什么我们需要新一代 MPP 数据库（2）

上一代 MPP 难以适应新变化

- 由单机数据库发展而来
 - 单机库偏OLTP
 - 主要基于行存，列存是补充，列存方面优化不足
 - 未使用CPU指令加速

电信省公司现状，易用好用够用是关键

- 人手少，要维护的集群规模大，压力大，任务重
- 绝大多数表，单天/月十亿级以内
- 组件众多，版本老旧
 - gbase, kudu+impala, phoenix + hbase等
 - 面临去oracle任务
- 国产化MPP替代，国产化机器适配

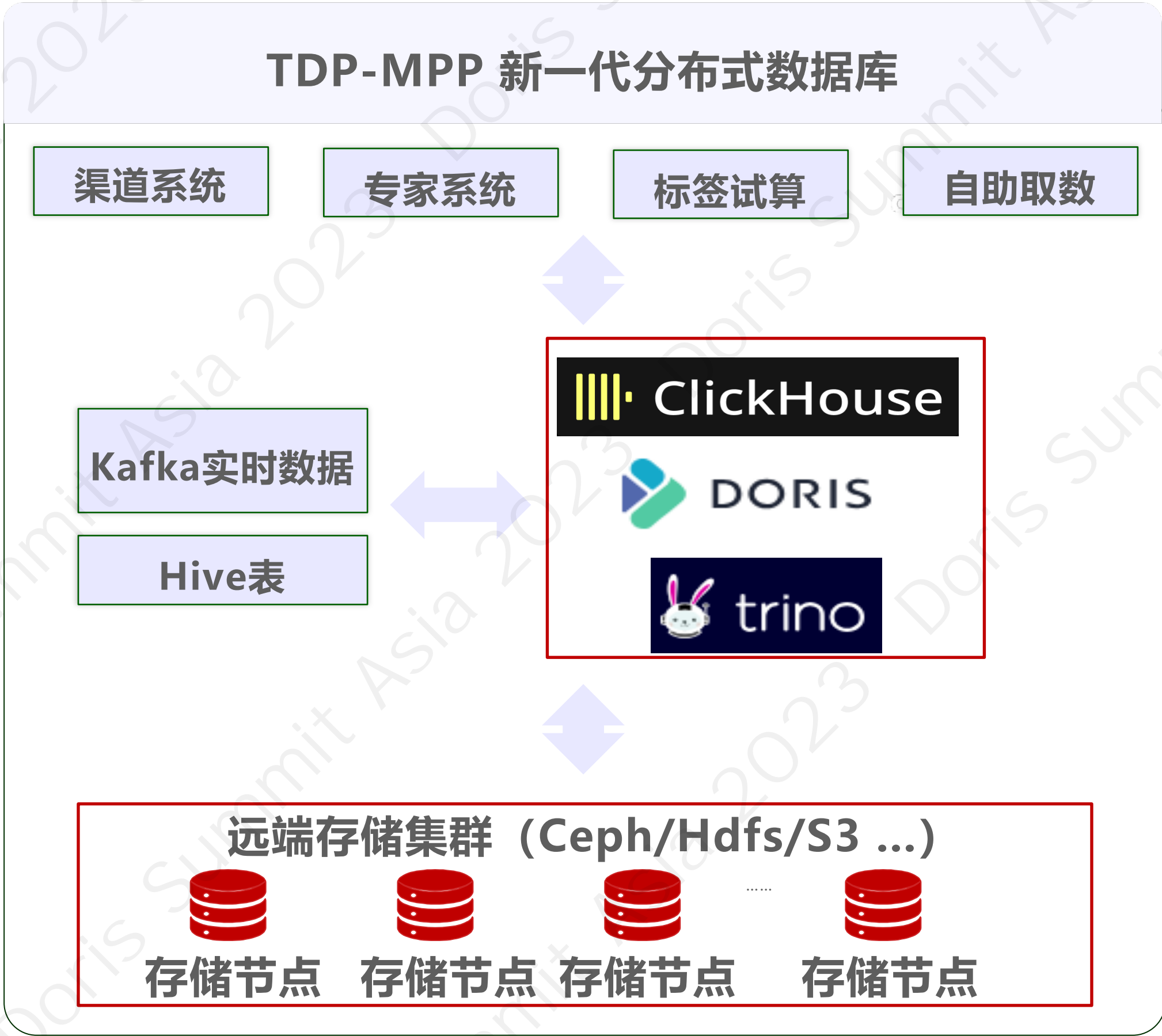
TDP-MPP 软件包

TDP: Telecom Data Platform

- 电信大数据平台，包含30多种组件
- 自研可视化部署、管理组件
- 自研数据中台、监控报警、日志采集系统

TDP-MPP 是电信二次开发形成的 MPP 数据库发行版本

- 包含Tdp-Doris(2.0), Tdp-Clickhouse, Tdp-Trino
- 完全兼容开源功能，不改变用户使用习惯
- 基于最新稳定版开发而成，并不断迭代升级，保证与社区版完全兼容
- 集成Tdp-manager集群管理工具
 - 支持可视化部署、配置修改一键下发、滚动升级
 - 监控、报警、日志采集分析
- 集成大数据中台组件
 - 支持可视化表管理、SQL编写与执行、可视化数据同步配置
- 自主编译，国产化适配，默认最佳参数配置
- 自研新功能，负载均衡，产品白皮书(案例集成)
- 支持物理机部署，容器化K8S部署，docker混部，冷热分层存储



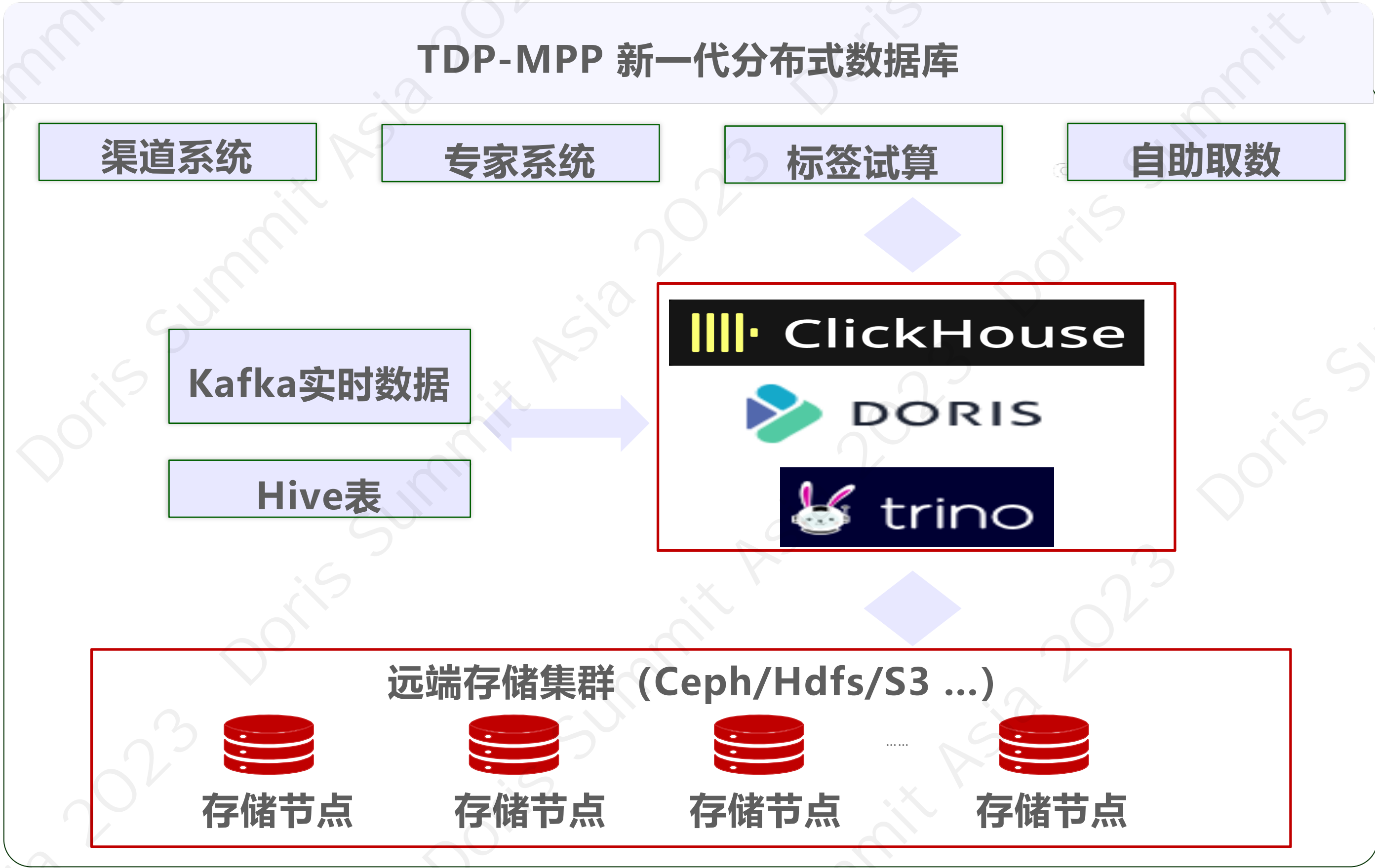
TDP-MPP 软件包

50人专家团队，社区密切合作

- 省公司问题第一时间响应
- 协助方案设计、架构优化、使用优化
- 省公司特殊需求优先开发
 - doris jdbc catalog for gbase
 - doris hbase迁移

其它

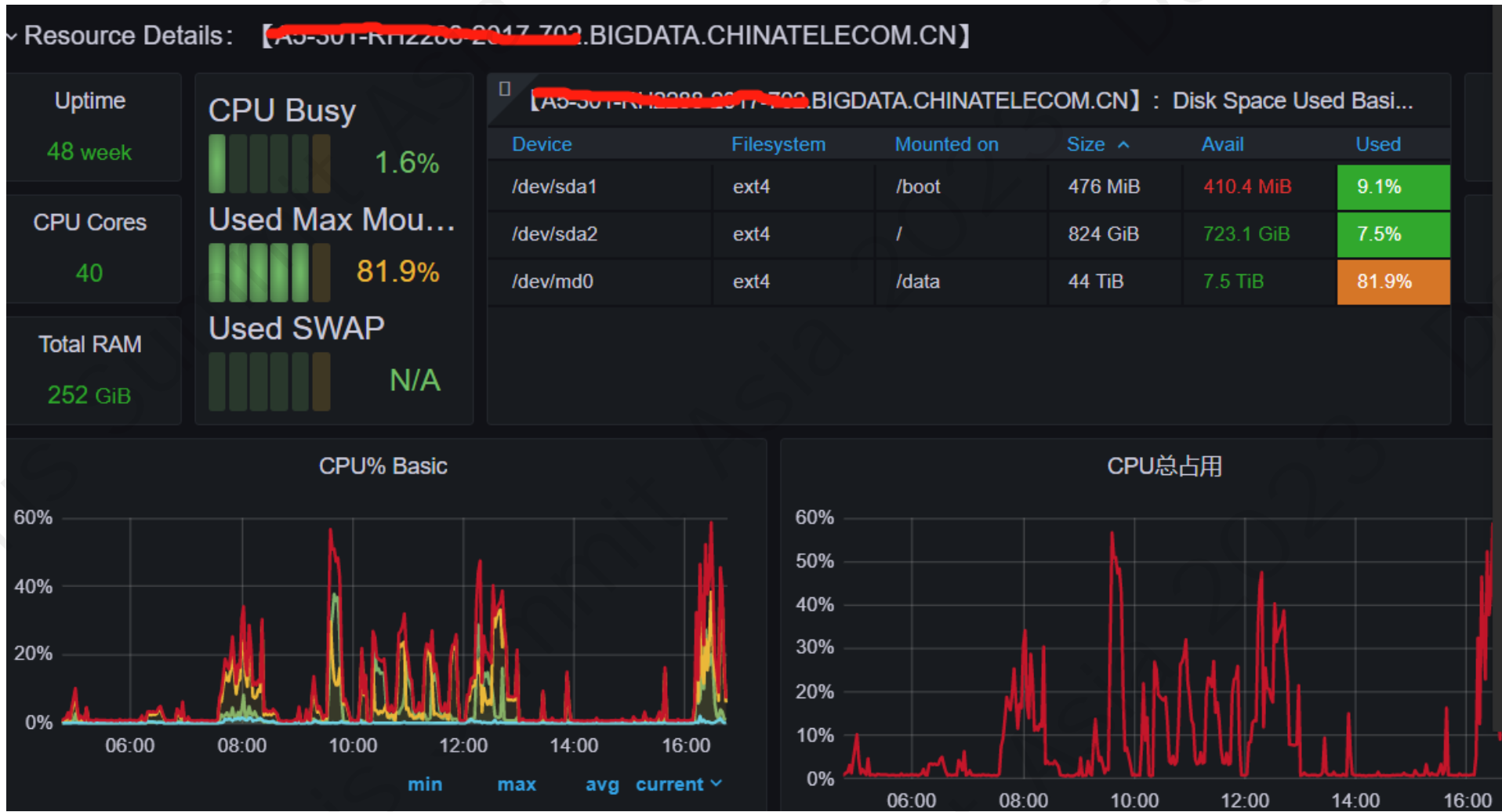
- Haproxy
- Milvus
- Kudu 1.15 + Impala 4.1
- TSDB
- GraphDB



TDP manager - 自研集群管理工具

可视化部署，简化运维

- 支持电信自研MPP引擎的自动化部署
 - 支持30+组件的自动运维
- 最优新版本,定制化,全网经验
- 一键集群部署,滚动重启，最优配置 (以doris fe为例)
- 进程监控,失败自动起
- 指标监控, 日志接入ELK



集群实例名称(英)	集群实例名称(中)	所属集群环境	备注	操作
doris	doris	Ha3_xty	doris	管理集群 管理服务 配置管理 ...
clickhouse	clickhouse	Ha3_xty	clickhouse	管理集群 管理服务 配置管理 ...
kafka_test	TDP2测试	Ha3_xty	TDP2测试	管理集群 管理服务 配置管理 ...
zookeeper_test	zookeeper_test	Ha3_xty	zookeeper_test	管理集群 管理服务 配置管理 ...
kerberos_test2	kerberos_test2	Ha3_xty	kerberos_test2	管理集群 管理服务 配置管理 ...
node_exporter	node_exporter	Ha3_xty	node_exporter	管理集群 管理服务 配置管理 ...
http_exporter	http_exporter	Ha3_xty	http_exporter	管理集群 管理服务 配置管理 ...

分配角色/选择主机

角色: 请选择角色

ip列表: clickhouse_servers 0/0

请输入关键字

10.142.104.176

10.142.104.180

10.142.104.177

10.142.104.181

192.11.22.33

10.142.104.183

10.142.104.161

请选择ip

取消

确定

自研数据中台

可视化数据中台

- 可视化sql查询界面
- 可视化表管理工具
- 可视化导入界面
- 创建数据源
- 字段映射
- 任务调度

常用菜单

请输入关键词搜索服务

查询

+

添加

数据开放目录

数据资产目录

指标目录

流程节点定义

业务流程设计

数据查询日志

菜单管理

角色管理

操作日志管理

参数管理

用户管理

团队管理

大数据开发场景

1 架构规划

2 数据建模

3 数据交换

4 数据研发

5 数据治理

6 数据服务

架构规划

对数据架构、流程、资源和团队进行规划和设计。

开始使用

数据建模

根据业务需要进行数据建模。

开始使用

数据交换

将不同数据源集成至大数据平台，实现数据入湖。

开始使用

数据研发

创建数据清洗、加工、编排任务，并发布。

开始使用

数据治理

数据加工全流程质量校验，发现数据问题并处置。

开始使用

数据服务

将高质量数据表，快速生成数据API，加速数据应用。

开始使用

个人基本信息

系统管理员

ID: admin

团队

数据开发...

角色

团队管理员

对应租户

成都数据演...

最近浏览

字段标准

2023-06-06 08:39:21

标准评估

2023-06-06 08:39:19

数据地图

2023-06-02 11:19:27

发起流程

2023-06-02 11:17:49

公告

3333333333333333

2023-09-13 10:40:41

222222222222

2023-09-12 16:29:08

11111111111111

2023-09-09 14:38:40

dfsdfsdfsdf

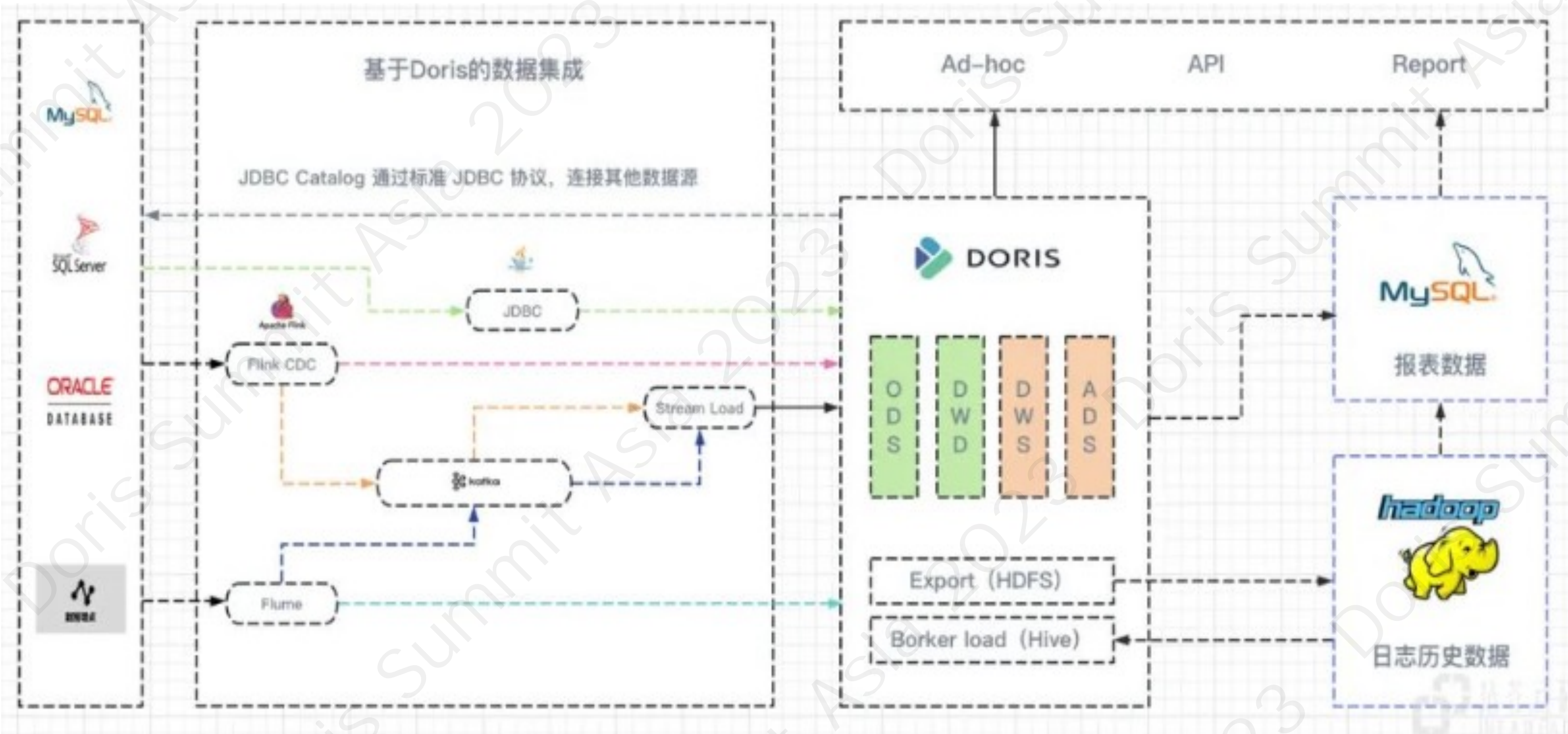
2023-09-07 00:05:51

查看更多

TDP-Doris 主要应用场景

1. 查询加速

- 应用层，汇总层导入MPP，做查询加速
 - 配置交互式报表，做多维查询
- 简化数仓体系
 - 将DWD,DWS层直接入库，直接查询，无须再加工



2. 实时应用、实时数仓

- Kafka数据直接入MPP，实入实查
- 保留历史数据,方便回查问题



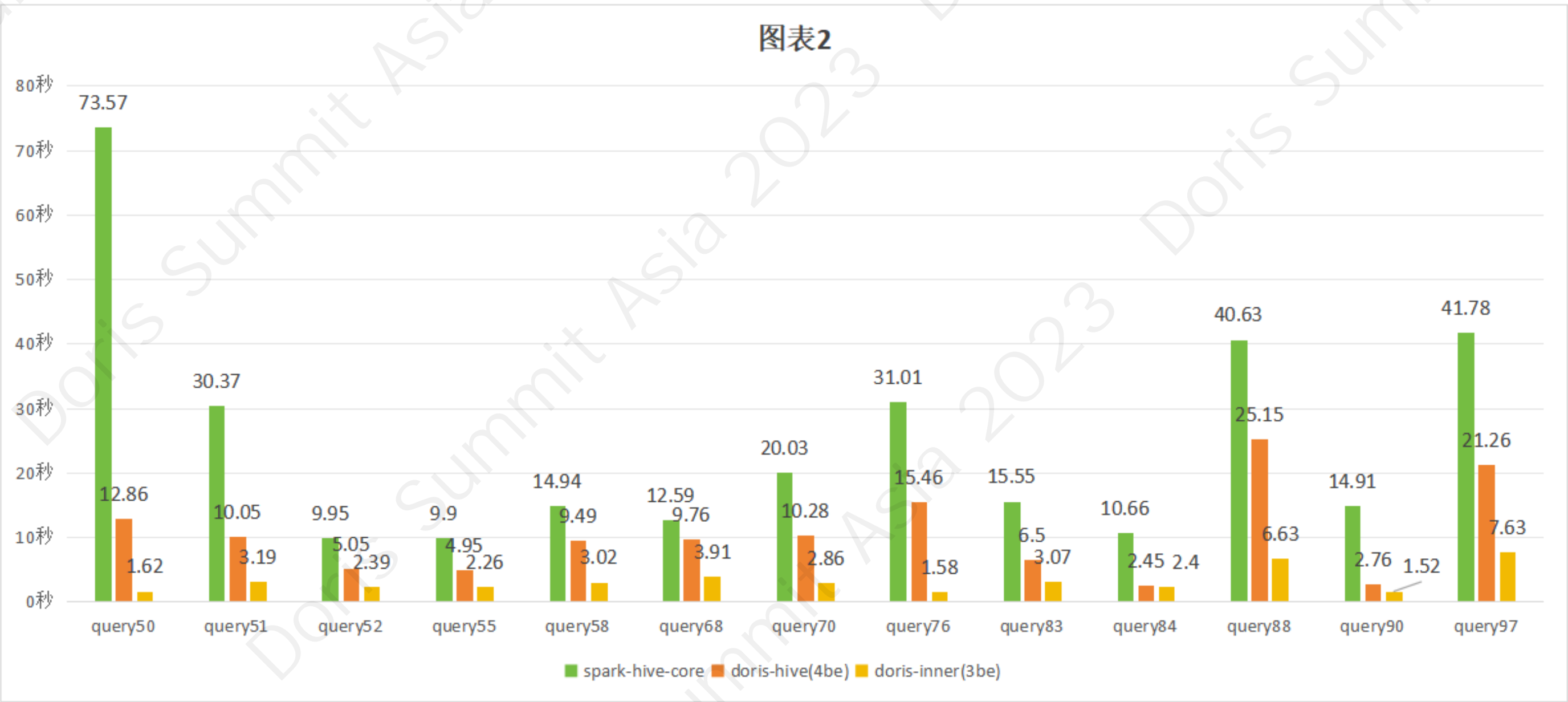
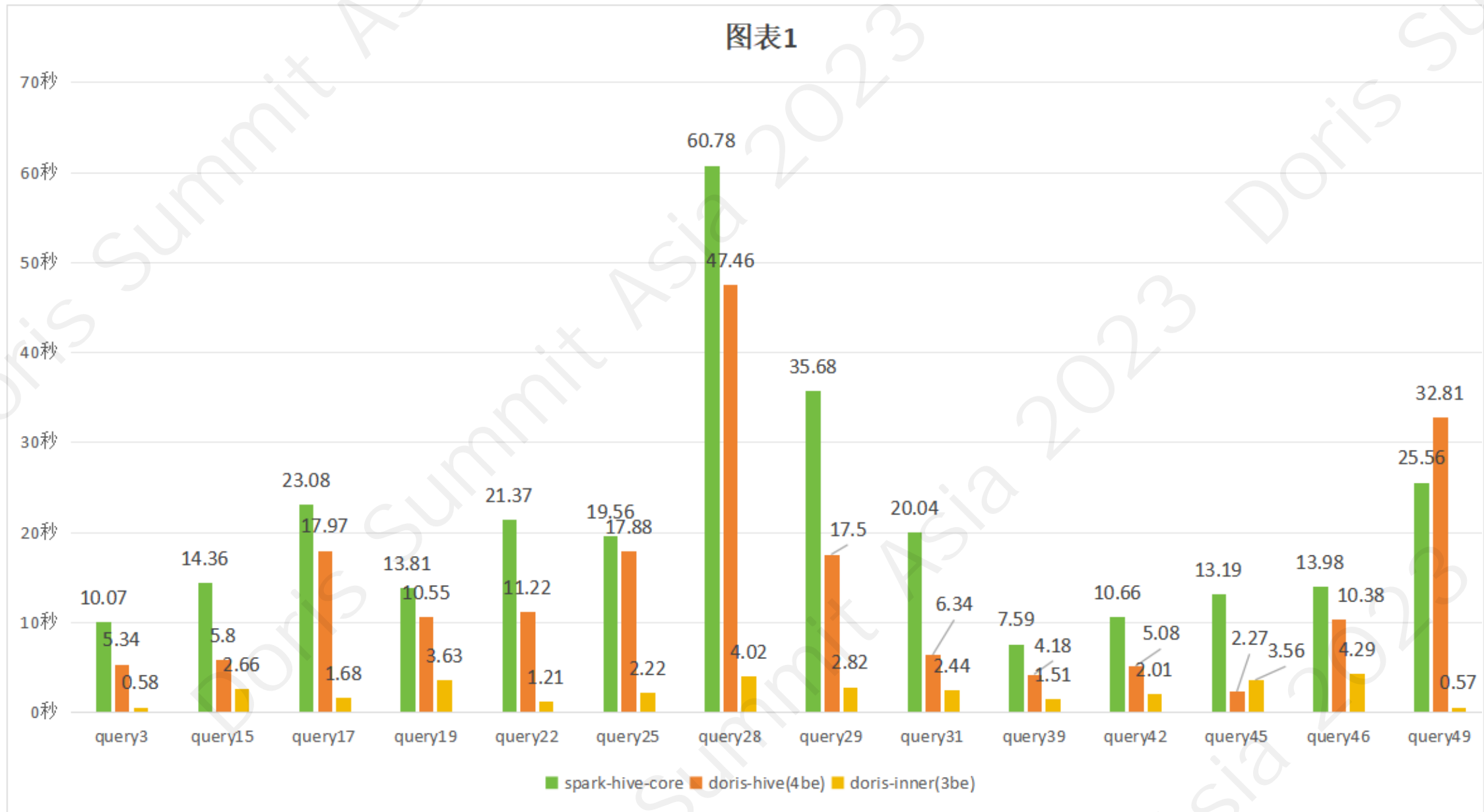
2 TDP-Doris 在电信集团和省公司的应用

TDP-Doris 性能对比测试

Doris / Spark 查询性能对比

TDP-DS 1T

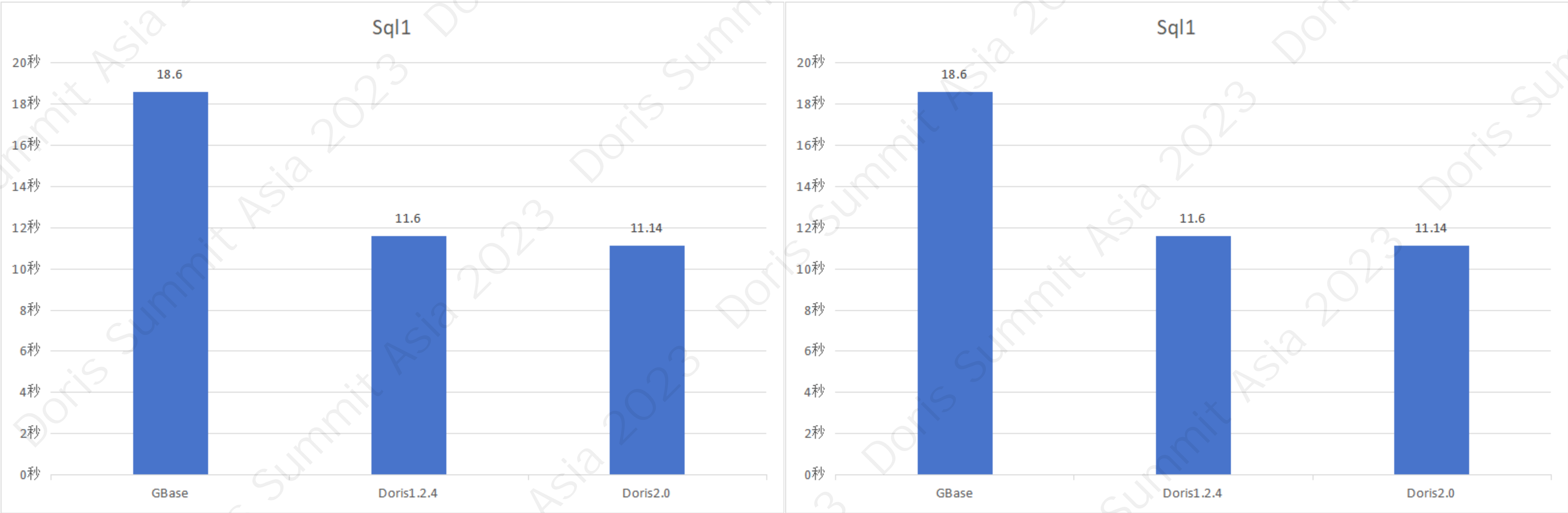
- spark on yarn: 32 executors、16G、2core
- doris-hive: 4be物理机、2.0.1版本,parallism设置为10
- doris-inner:3be物理机,2.0.1版本, parallism设置为10



性能测试: 10*doris > 30*gbase

对比测试 1

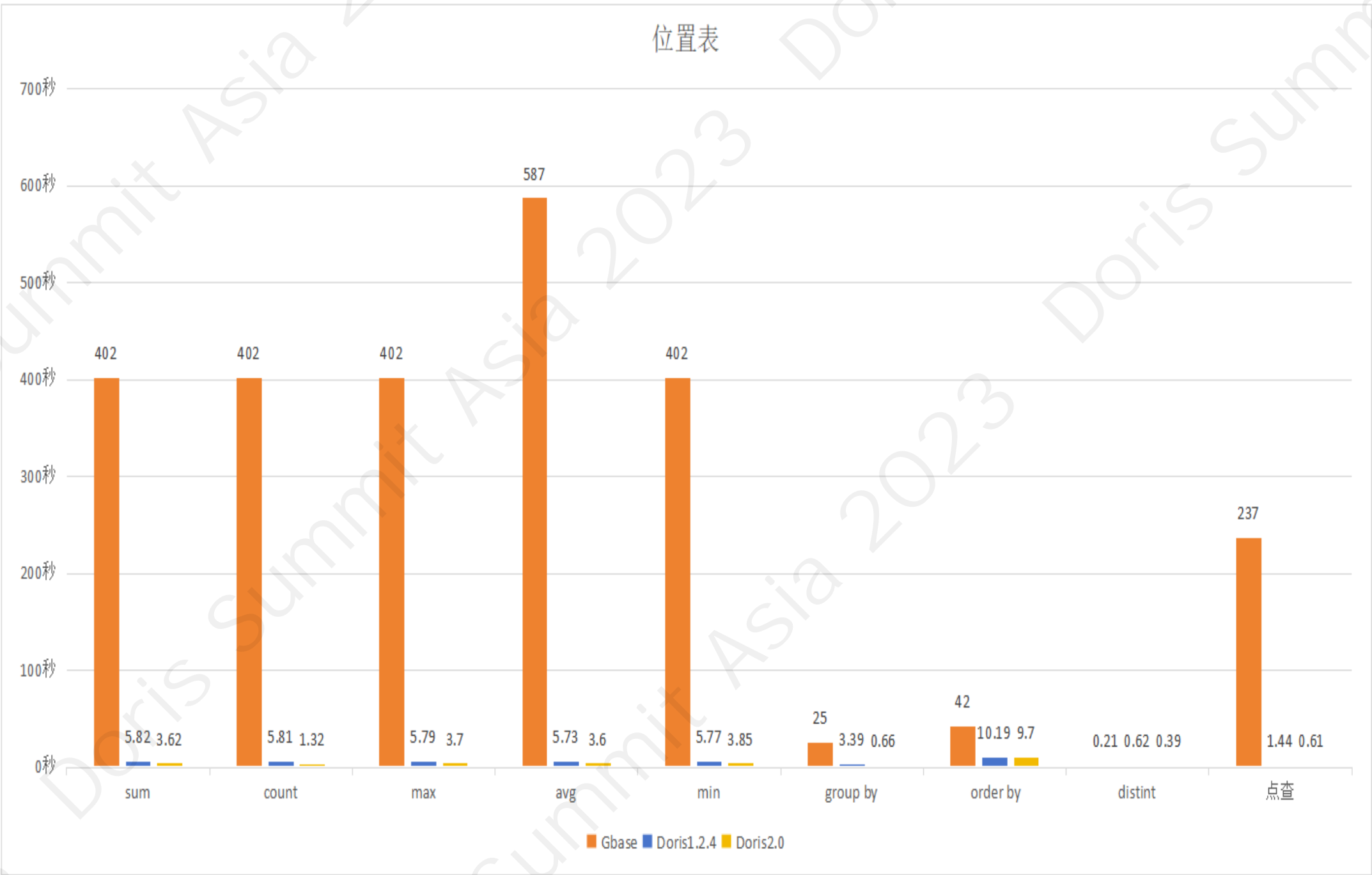
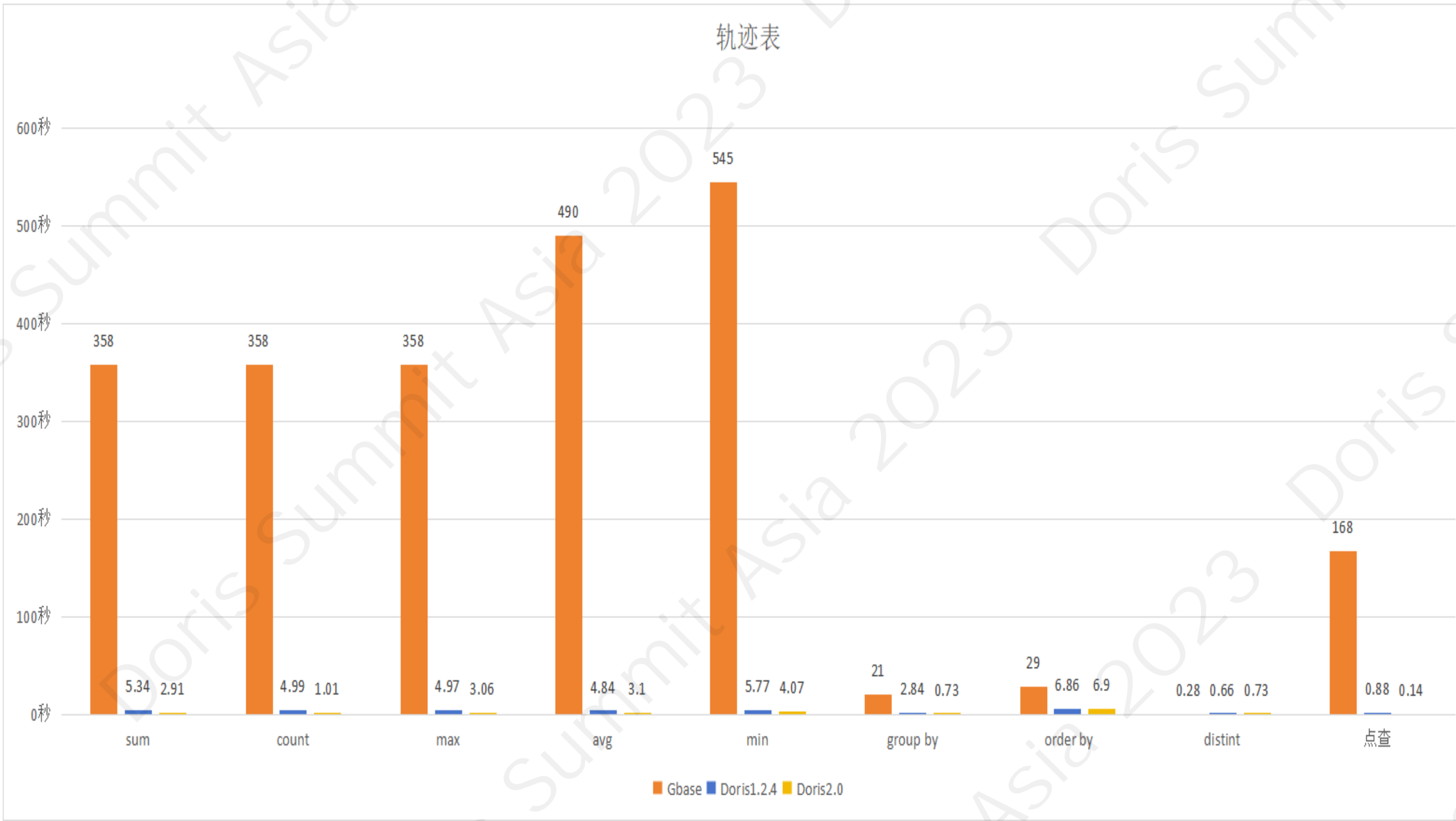
- 10台 doris be 和同等配置的30台 gbase
- 多表 join, t645 2亿, t17 5亿, t705 3千万



性能测试: 10*doris > 30*gbase

对比测试 2

- 10台 doris be 和同等配置的30台 gbase
 - 简单聚合查询
- gbase 对sum/count等作了预先统计
 - doris可用物化视图实现同样功能
- 添加 where 条件, gbase 预先统计失效
 - 下图均是添加 where 条件后的执行时间
 - 查询表数据量均在百亿级



导入性能测试

环境：3个fe, 10个be

- 经过调优，16个并行任务，20分钟内导入完成,整体脚本在30分钟完成
- 测试seatunnel
 - 任务出错时长时间无响应

不同测试条件 (资源、参数、模式等不同情况)	运行耗时	说明 (从操作易用性、特别注意事项方面补充)
broker load	3分钟	异步作业，需要后台查进度
seatunnel parallelism=16	3分钟	使用 spark 资源，减轻 doris 压力
spark load,20*5G	5分钟	并发数加大，单个内存需求降低，运行加快

Hive数据导入Doris性能测试

数据加工效率验证

测试场景	测试场景	数据量大小说明	Spark时长 (秒)	Doris 时长 (秒)	Oracle 时长 (秒)
决策支持 随翼发展 套餐清单 加工	单表加工	宽表1.7亿	367	71	782
	多表加工	1. 宽表1.7亿 2. 2表4000万 3. 中间过程表 1.7亿关联分析	155	83	623

环境配置

- ✓ **Doris集群环境说明：**
集群节点：3个fe、9个be，配置都为40核256GB内存
- ✓ **Oracle集群环境说明：**
集群节点：计算节点2个，72核512G内存，存储节点15台，16核64G内存

数据同步效率验证

测试场景	测试场景	Doris (broker load方式)	Oracle (sqoop方式)
销售品实例宽表 数据回写	地市1	18	18
	地市2	13	11
	地市3	14	12
	地市4	14	8
	地市5	9	9
	地市6	14	10
	地市7	15	8
	地市8	14	10
	地市9	8	9
	地市10	13	10
	地市11	7	7
	地市12	8	7
	地市13	14	8
	地市14	14	11
	地市15	7	7
所有地市并发数据同步		18	18

结论

- 1. doris在数据关联分析比Spark快3倍，比Oracle快8倍
- 2. 大数据平台数据同步至doris和oracle效率相当
- 3. 某省电信将大数据平台和Oracle加工流程迁移至Doris，其中大数据平台加工1小时缩短至20分钟，数据迁移0.5小时保持不动，Oracle数据加工2小时缩短至30分钟以内，总体流程由原来的3.5小时缩短至80分钟以内，应用展示提速2小时。

Doris KV 查询测试

用户标签场景

- 3 fe, 4 be, 128G/16核,Doris 1.2.4
- 测试语句：指定key, 查标签
 - select lab3,lab4,... from dm_awt_all_tran where key='xx';
- 要求：
 - 1亿行数据，800列，8000/min
- 还可支持非key字段的多维查询，join
- Doris 2.0 kv QPS 更高,达到8767

thread-num	平均响应时间/ms	QPS	Err率
10	9	1040.3	0.00%
20	10	1958.8	0.00%
50	12	3826.8	0.00%
100(5*20)	17.6	5534.1	0.00%

```
summary = 763220 in 00:03:20 = 3822.5/s Avg: 13 Min: 8 Max: 369 Err: 0 (0.00%)
summary + 115222 in 00:00:30 = 3840.7/s Avg: 12 Min: 8 Max: 128 Err: 0 (0.00%) Active: 50 Started: 50 Finished: 0
summary = 878442 in 00:03:50 = 3824.9/s Avg: 12 Min: 8 Max: 369 Err: 0 (0.00%)
summary + 115179 in 00:00:30 = 3839.3/s Avg: 12 Min: 8 Max: 117 Err: 0 (0.00%) Active: 50 Started: 50 Finished: 0
summary = 993621 in 00:04:20 = 3826.5/s Avg: 12 Min: 8 Max: 369 Err: 0 (0.00%)
summary + 114643 in 00:00:30 = 3821.4/s Avg: 13 Min: 8 Max: 126 Err: 0 (0.00%) Active: 50 Started: 50 Finished: 0
summary = 1108264 in 00:04:50 = 3826.0/s Avg: 13 Min: 8 Max: 369 Err: 0 (0.00%)
summary + 40138 in 00:00:10 = 3848.7/s Avg: 12 Min: 8 Max: 132 Err: 0 (0.00%) Active: 0 Started: 50 Finished: 50
summary = 1148402 in 00:05:00 = 3826.8/s Avg: 12 Min: 8 Max: 369 Err: 0 (0.00%)
Tidying up ... @ April 24, 2023 4:34:10 PM CST (1682325250430)
... end of run
```


2 TDP-Doris 在电信集团和省公司的应用

TDP-Doris 迁移案例分享

案例 1：将应用系统迁移 Doris 数据库

✓数据模型迁移：

整理网格全息系统、渠道支撑系统、企业健康度系统、财务智慧展示系统、企业数据门户等系统的模型，在doris数据库中新增需迁移的数据模型

✓存量数据迁移：

将存量数据从oracle和大数据平台迁移到doris数据库

✓加工脚本改造：

对照数据库的SQL语法差异，改造所有迁移的数据加工脚本，并确保运行效率

✓调度流程开发：

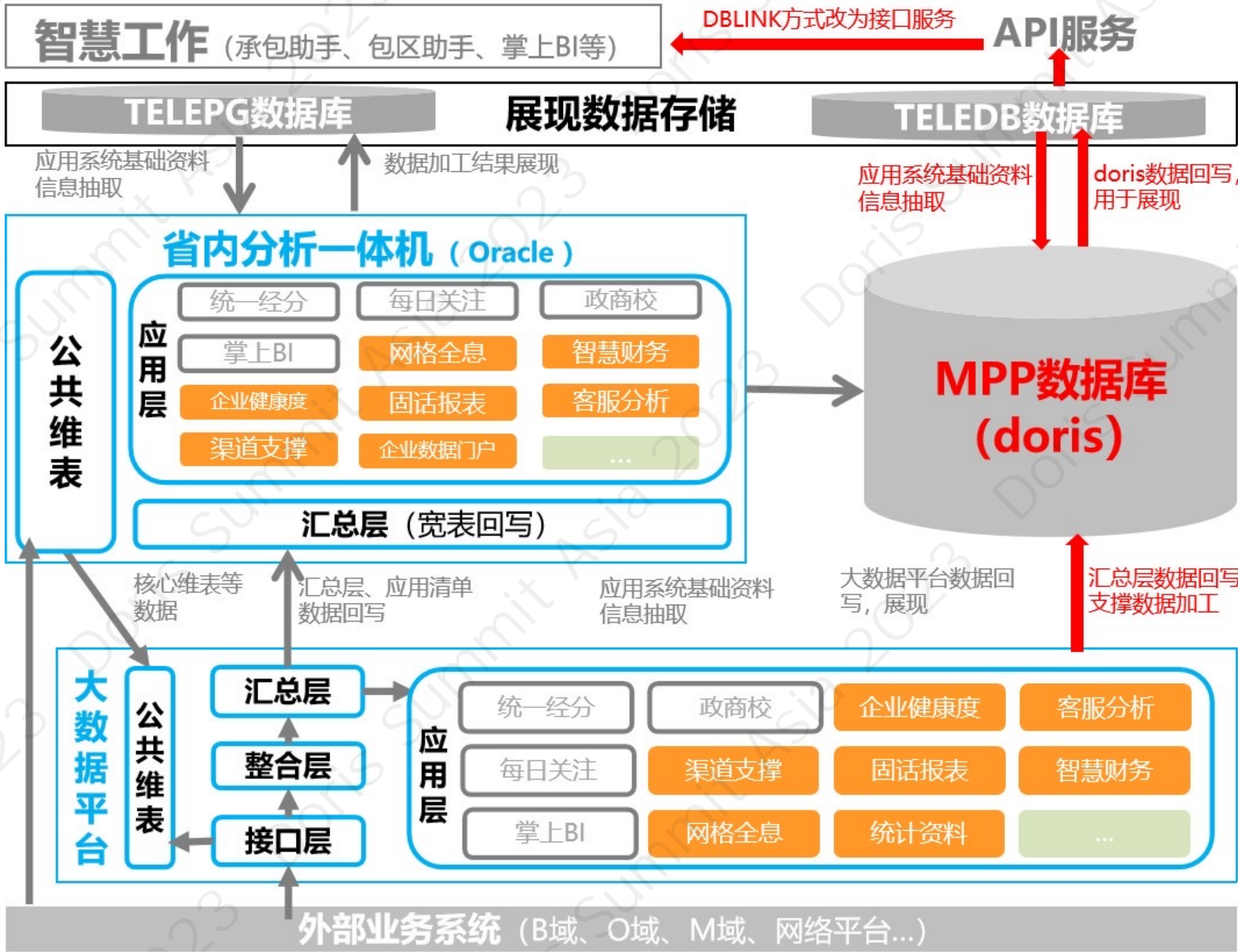
基于ETL平台，为所有数据加工脚本新配置对应调度流程（前置信号量、加工组件、输出信号量、质量校验与告警）

✓前置数据下发：

新增大数据平台到doris数据的数据下发流程、doris数据库到应用展现存储数据库的数据下发流程

✓新增API接口：

去DBLINK数据共享形式，新增API接口支撑承包助手、包区助手、掌上BI等系统用数



案例 2：持续探索大数据实时数据入湖，打造 Hudi+Doris 湖仓一体架构

✓实时数仓架构：

以大数据平台为主，MPP为辅，基于Hudi+Doris构建混搭架构的实时数仓：实时数仓在离线数仓的基础上引入实时整合层、准实时整合层、准实时汇总层、准实时集市层

✓接口层：

计费话单类数据实时入Doris库(文件方式)，其他实时数据入大数据集群（消息方式）

✓整合层：

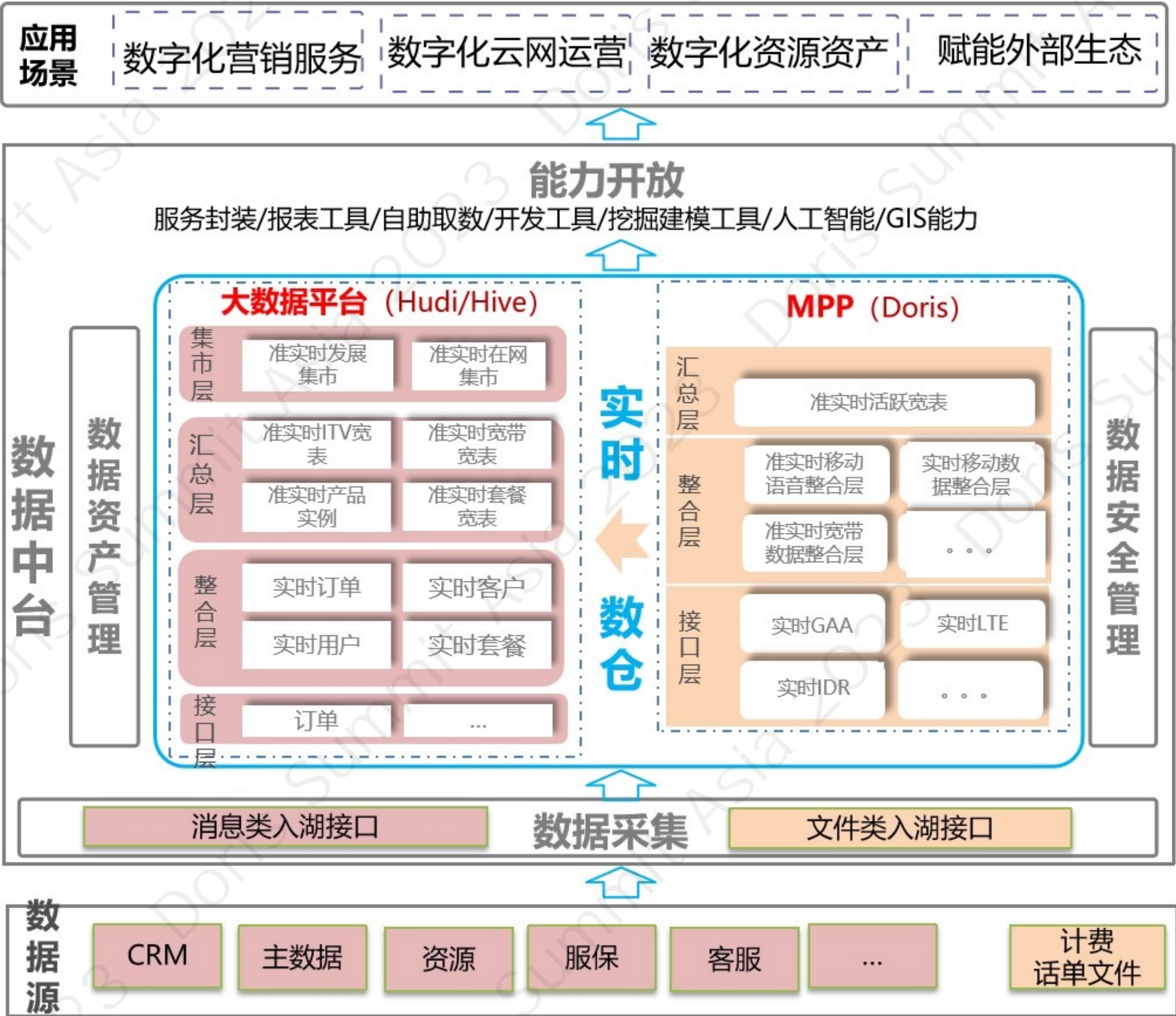
全面切换云化模型并落地原生数据、基于增量实时接口形成全量数据（涉敏数据加密入安全区）

✓汇总层：

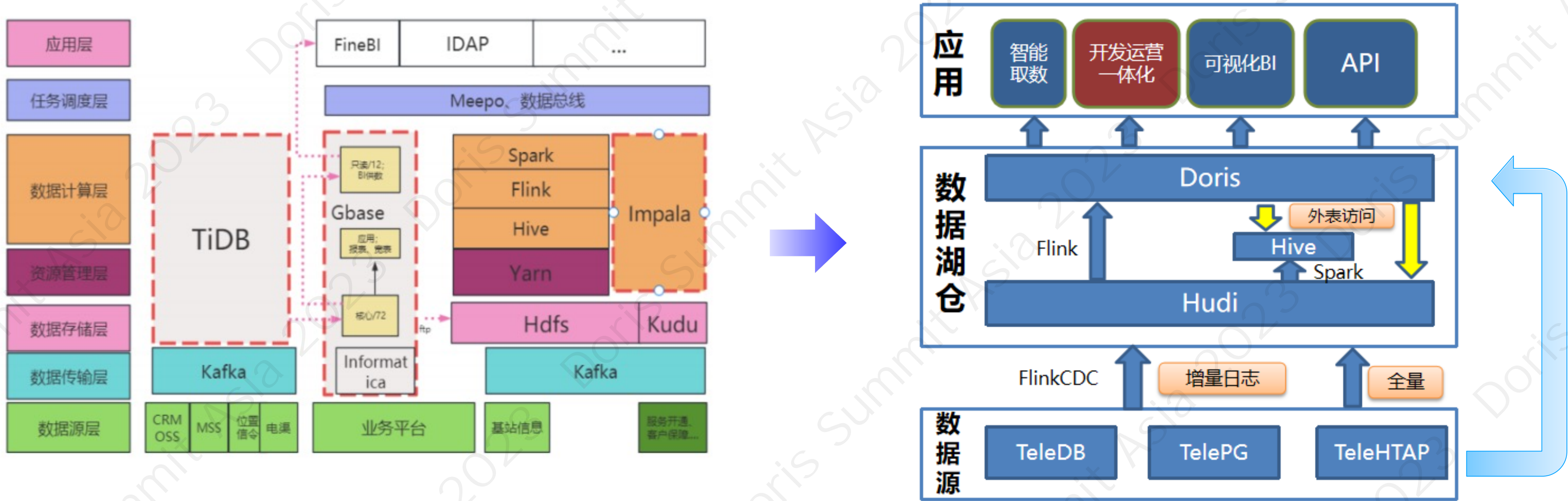
跟进业务发展热点，分主题构建天翼云眼、云业务、天翼视讯、随翼选等准实时宽表；实现用户订购、划配、位置、使用量类标签准实时化；

✓集市层：

构建实时/准实时的业务发展、产数标品、数字生活、客户感知等集市模型



案例 3：使用 Hudi+Doris 取代老旧组件



- 基于hudi + Doris的湖仓一体新架构
 - Hudi做为数据湖组件，使数据就绪时间从天级到分钟级
- Tdp-doris作为统一查询组件，替换Oracle, Kudu + Impala, Gbase，部分Hbase
 - 支持标签kv实时查询
 - 支持多维查询
 - 支持kafka数据直接入库
 - 支持联邦查询，读hive表，本地表，旧库表，方便数据迁移
- 新架构逻辑清晰，简化数据流，使用组件少，维护压力小

Oracle/Gbase替换-数据同步

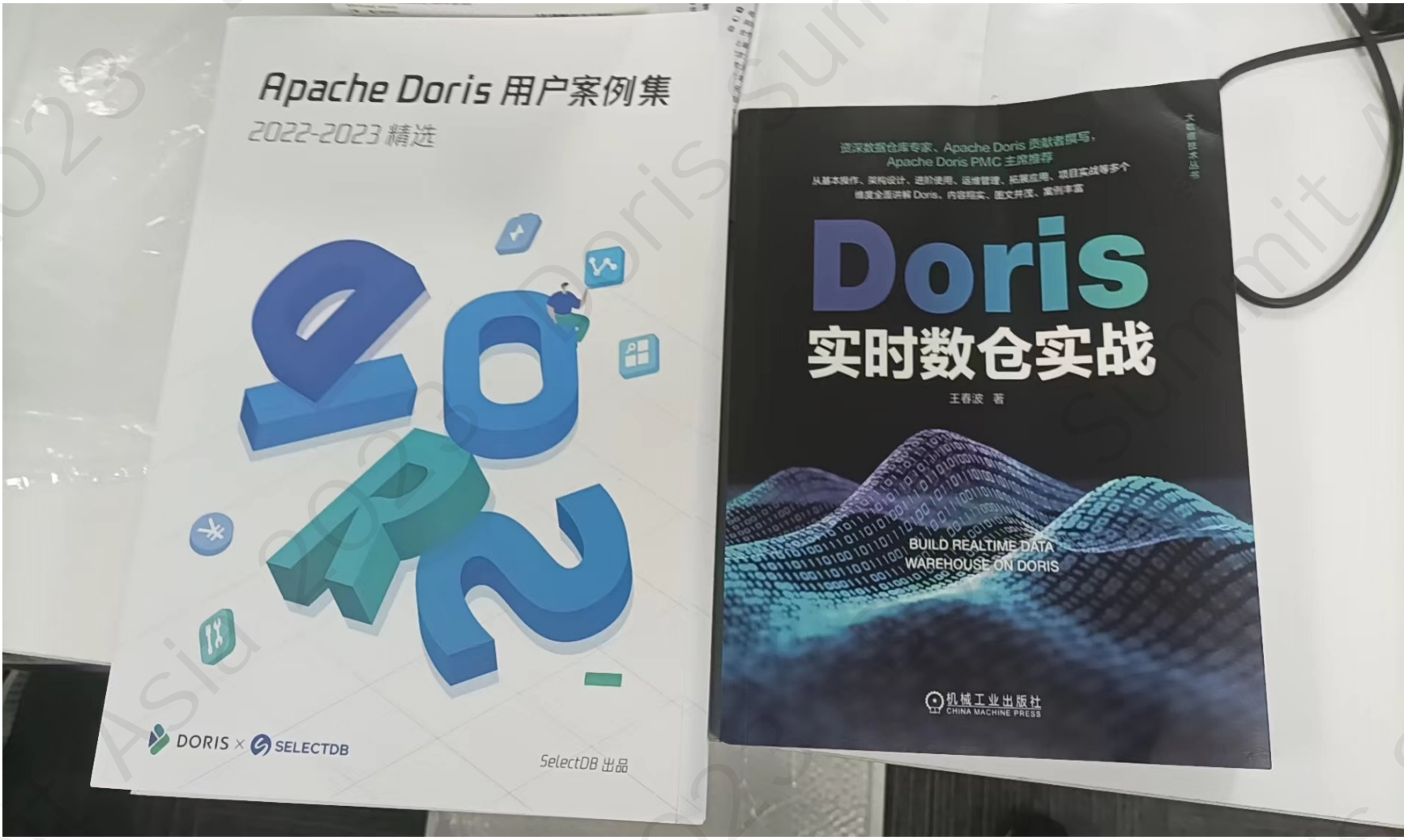
- 使用 jdbc-catalog的方式, 连接oracle/gbase
- 自研JDBC catalog
- 不需建外表, 建catalog后, 库、表自动可见
- 灰度上线, 逐步替换
- 直接查询
- `select * from mysql_catalog.db.table where k1 > 1000 and k3 = 'term';`
- 双向同步数据
- 使用insert into 本地表 `select mysql_catalog.db.table {where partition}` 的方式导入
- 新增Trino Gbase/Doris Connector
- trino worker 到 be, 高并发导入

```
CREATE CATALOG jdbc_oracle PROPERTIES (  
    "type"="jdbc",  
    "user"="root",  
    "password"="123456",  
    "jdbc_url" =  
    "jdbc:oracle:thin:@127.0.0.1:1521:helowin",  
    "driver_url" = "ojdbc8.jar",  
    "driver_class" = "oracle.jdbc.driver.OracleDriver"  
);
```

3 Doris社区互动

TDP-Doris 社区互动

- Doris创始人，SelectDB CEO 马如悦老师来电信访问
 - 连续3小时讲解Doris 2.0最新优化，用户使用情况
 - 协助处理线上问题
 - 技术会议、共享文档、现场教学等
- 近期开发计划
 - doris mixed operator
 - oracle语法兼容、存储过程自动化迁移
 - doris ranger适配





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>