

Apache Doris 助力企业实现数据智能化应用

赵书贤

启高科技 联创&方案架构师

目录

01 关于启高科技

02 零售业数据特征

03 基于 Apache Doris + 启数道的数据平台

04 Q & A

分享嘉宾 - 上海启高信息科技有限公司



赵书贤

联合创始人&方案架构师

超过 18 年企业大数据 BI 咨询规划与实施落地经验，为超过百家大中型企业提供大数据BI咨询实施服务，对零售、制造行业有丰富经验，对 SAP BW/HANA、大数据各组件、数据开发平台均拥有丰富实践经验。

启高科技，企业数据领域的持久伙伴与坚强后盾



让数据成为驱动力

总部上海虹桥
辐射全国

专注于大数据 BI 产品与实施服务

1个产品 启数道 DataOps 平台

2个行业 聚焦零售、制造行业

100+名顾问 技术团队规模

50+家客户 近三年服务的企业

服务客户

鞋服零售、消费品行业



汽车零部件、制造行业



近年服务过的客户还包括：华为2012实验室、银联云、国泰君安、永祥股份、华油信通、方太、河南日报、德力西、淮海控股、顾家家居、澳优乳业、富奥股份、德赛西威、东方国际、红豆、阿吉豆、顺丰-星巴克、大金空调、维他奶、科迈化工、瑞人堂药业、Majnoon油田、上海电信、良信电器、绝味食品、兴业银行、卡斯柯、NB、贵人鸟、欧莱雅、元祖食品.....

荣誉与资质证书



启高科技陪跑企业数据价值化全过程

数据价值化方法论

前期准备

明确数据价值化的目标、现状、线路图

采集与整理

建立完善的数据采集和整理机制，确保数据的准确性、完整性和及时性。

分析和挖掘

运用数据分析和挖掘技术，对数据进行深度挖掘和分析，发掘出潜在价值，并为业务决策提供依据。

应用和创新

将数据应用于不同场景和业务领域，开展数据驱动的业务创新和变革，推动企业发展。

保障与安全

加强数据保障和合规管理，确保数据安全和合法使用，避免因数据泄露等问题带来的风险

人才与组织

通过人才培养和组织变革，建设具备数据思维和能力的团队，推动企业实现从传统到数字化转型

- 数据战略咨询
- 数据治理咨询
- 数据应用规划

- 数据治理落地
- 数据平台建设
- 数据资产管理

- 指标体系
- 数字大屏
- 报表
- BI

- 门店提效
- 电商提效
- 成本监控
-
- 预测维修
- 车间无纸化
- 车间目视化
- 质量优化
-

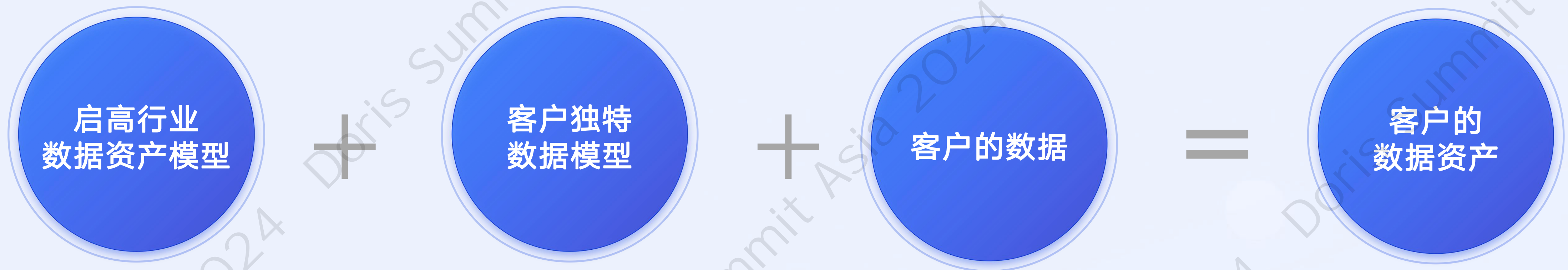
- 权限管控
- 数据API
- 数据脱敏

- 数据人才培养
- 数据文化
- workshop

启高科技服务内容

深耕行业，专注企业数据深加工，聚焦数据模型

行业数据模型



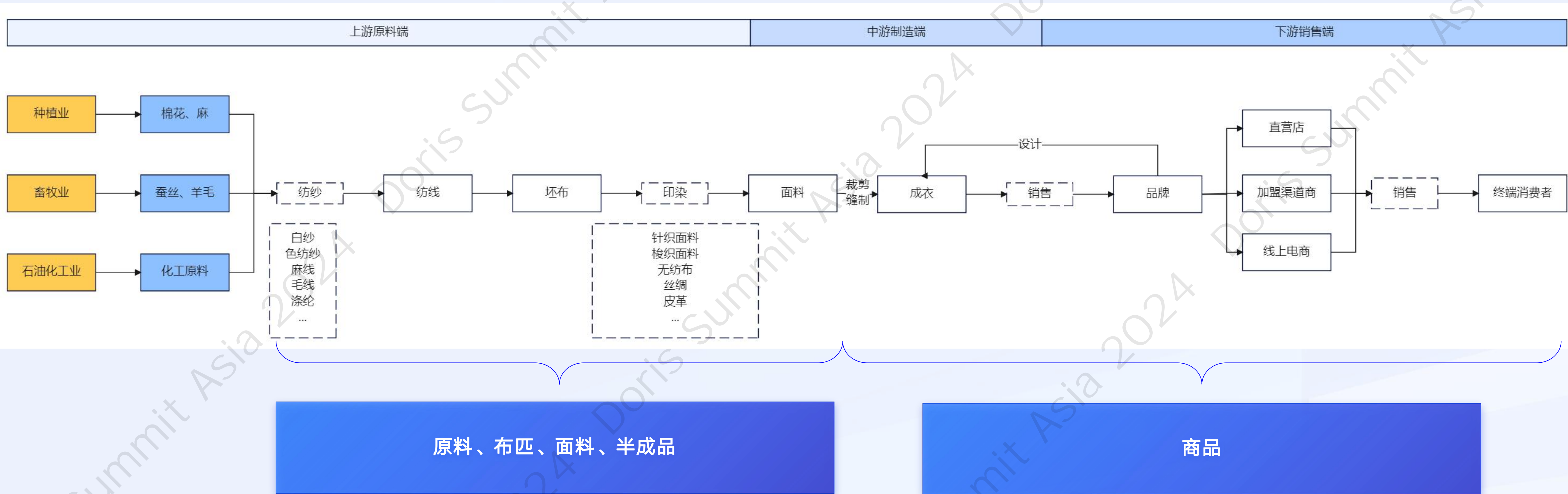
有参考、有验证、
兼顾企业特点

门槛要求低、
快速匹配企业业务流程

项目中就定义最终指标，快速交付

优势

纺织服装企业全方位数字化转型，消除牛鞭效应



牛鞭效应（Bullwhip Effect）是指在供应链中，较小的需求变化或波动逐渐放大，从而在上游供应链环节引发更大的波动和不稳定现象。

服装企业数字化转型过程的困境

数据量庞大、数据多元且最复杂

场景复杂、响应要求高

DRP

CRM

OMS

O2O

ERP

WMS

SRM

POS

PLM

MES

商城

Wechat

电商平台

.....

沉淀多年的信息

渠道多元化

商品多样化

经营链路长

多品牌矩阵

柔性快返

流量分配

DTC

千变万化的商业模式

充分利用大数据将供需实时连结在一起，实现全链路拉通

终端需求驱动

大数据技术掌握终端需求，并以需求为始点进行研产销存全链路经营活动

效率提升

全链路信息互联互通，实现信息的高效传递，避免长鞭效应

协同作业

订单跟踪，实现全环节协同，减少成本降低，提升协同能力

精准定位

借助大数据AI技术，实现人货场的精准定位

- 需求预测、快速迭代、终端与工厂制造端联动，减少了链条中的无用环节，将高质量的流行商品快速交付给了客户；
- 通过算法从海量数据中挖掘时尚趋势，将数据反馈给庞大的设计和打板团队进行后续生产；数据中台管理库存、生产能力数据，直接连结工厂车间与客户；借助强大的数据和算法支持，实现快速增长

用户偏好与行为

线下门店电商平台

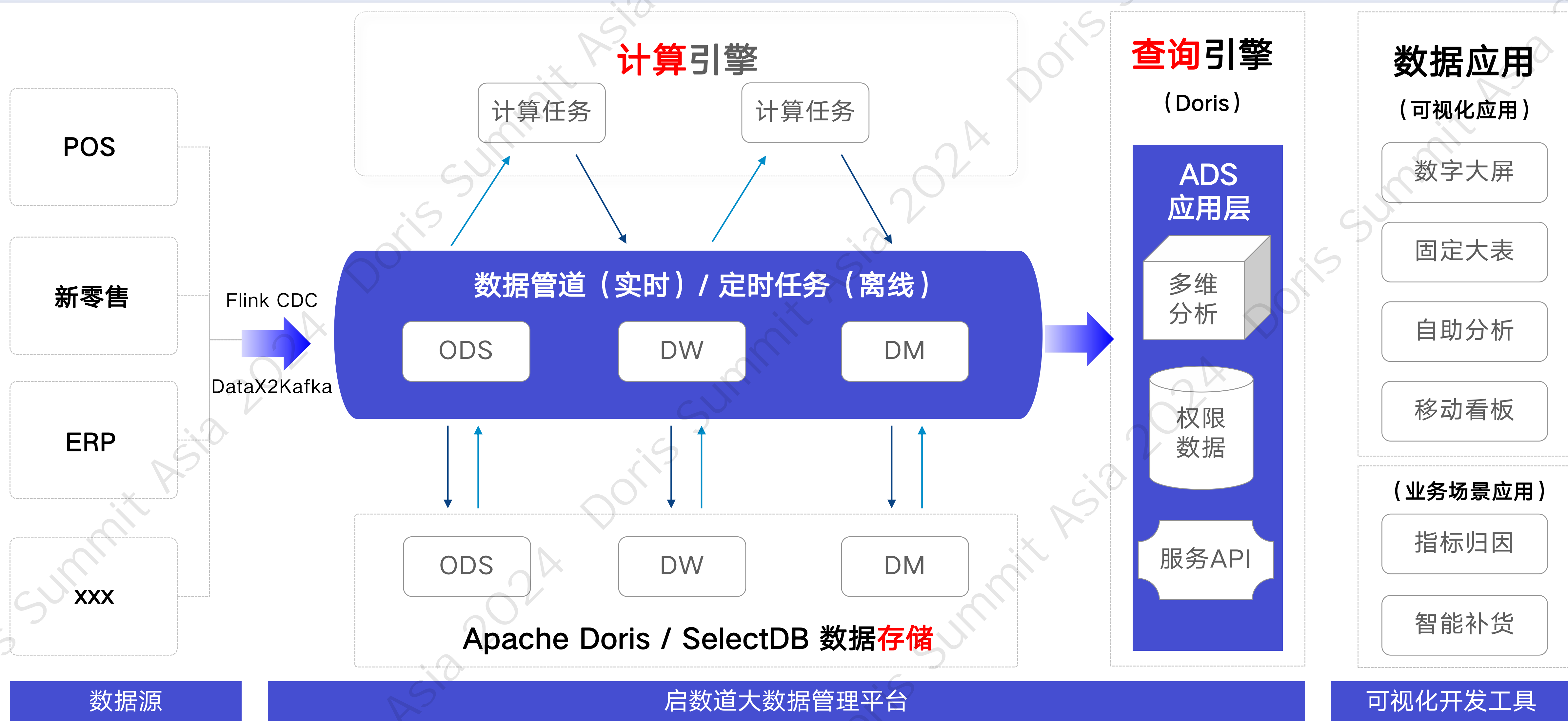
品牌商信息系统

供应链信息系统

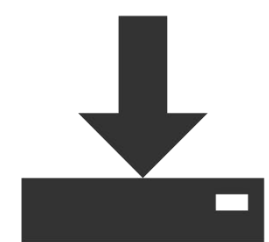
服装生产工厂车间

全链路数据拉通，实时反馈

存、算、查询分离的大数据平台，解决复杂零售场景数据应用

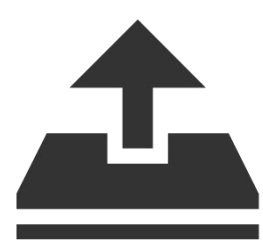


Doris作为企业大数据存储查询引擎的几大亮点



数据写入效率

1. 支持多种方式的高效写入，尤其是Stream Load性能非常高
2. 单表导入性能，在并发数为1，单次导入数据量达到 10 MB，导入吞吐量能达到 160 MB/秒，或者 10W 行/s。



数据查询效率

1. Doris 的一些特性，如分区分桶，数据模型，物化视图，优化器等提升查询效率
2. 高并发



数据更改能力 (随机写)

1. Doris提供的Unique Key数据模型，根据 key 列，修改 value 列。

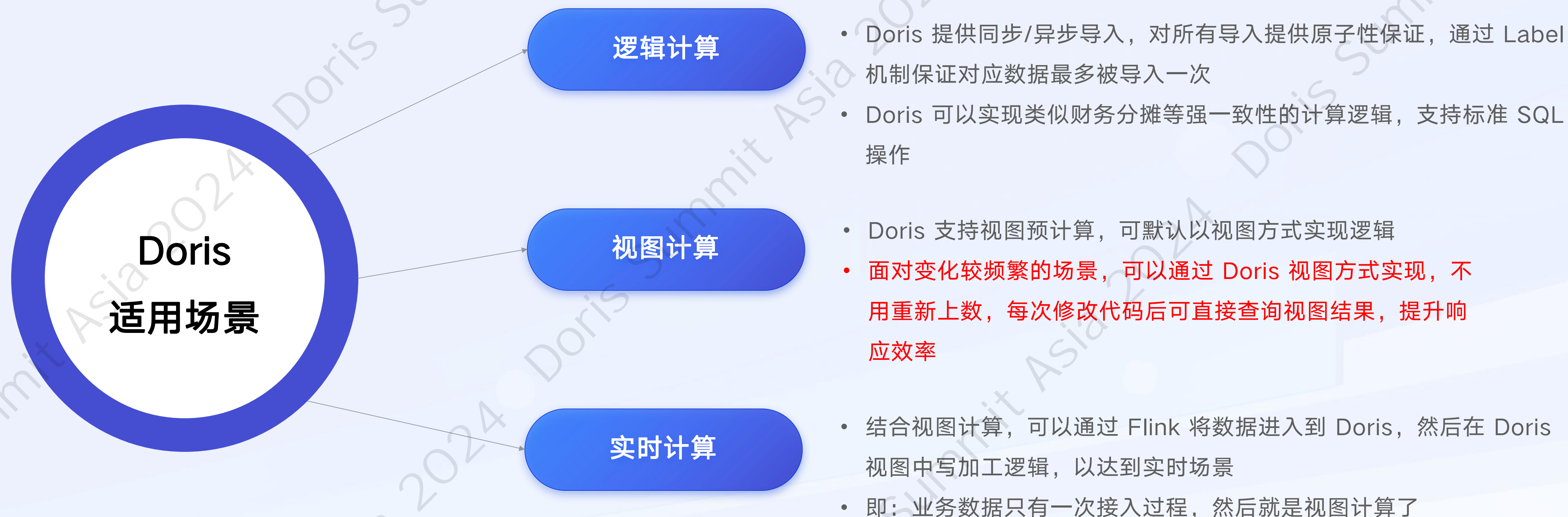


数据拉取能力 (增量数据流)

1. Doris 2.1 版本之前，默认使用 Merge_on_read 机制
2. Doris2.1 版本开始，默认使用 Merge_on_write 机制：在更新数据时，完成所有数据合并的工作，以便提高查询性能

Doris 在零售行业数据应用适用场景

Doris 是一款分布式查询数据库，在拥有较高的查询性能的同时能够保持数据一致性、并支持 Join 及标准 SQL，简单运维等特性，适合集数据仓库与数据查询于一体的应用场景。



启高科技打造的覆盖全链路的大数据服务平台：启数道

实时接入、调度管控、数据查询加速与可插拔的 OLAP 引擎架构

统一数据接入、加工、建模、治理及服务发布

与 BI 可视化、数据产品集成

架构开放

屏蔽大数据架构复杂性

DolphinScheduler

DataX

Flink

APACHE Spark

APACHE kafka
A distributed streaming platform

Doris

大数据平台门槛越来越高
实时场景开发代价太大

实施简单

加速数据中台实施部署

	数据源	数据对象
	数据建模	数据开发
	任务编排	调度管理
	指标管理	数据质量
	数据标准	数据资产
	元数据	数据血缘
	API定义	训练集
	API监控	数据标签
	算法库	智能场景

图形化的便捷的数据功能管理页面是加速实施、降低运维的基础

高效数据应用

组件化可拖拽式数据应用

数据API

数据门户

自服务

极致易用

移动跨屏

KPI 可视化不是唯一的数据应用
通过 API 集成业务场景才是数据价值的发现

Thanks for Watching!

上海启高信息科技有限公司

赵书贤 15821001961