

MySQL/Oracle 到 Apache Doris 数据实时同步技术实践

叶正盛 NineData 创始人 & CEO

个人简介-叶正盛

叶正盛

■ NineData 创始人 & CEO

- 资深数据库与云计算领域专家
- 曾担任阿里云数据库产品管理与解决方案部总经理，阿里云技术架构与产品决策委员会核心成员
- 阿里巴巴去 IOE、异地多活、云计算多次技术变革核心成员
- 构建阿里巴巴&蚂蚁集团数据库 DevOps 体系
- 创立云计算数据传输 DTS、数据管理 DMS、数据库备份 DBS、数据库自动驾驶服务 DAS 等多款云计算数据库产品



目录

01 数据库分类与数据复制简介

02 NineData 数据复制技术架构

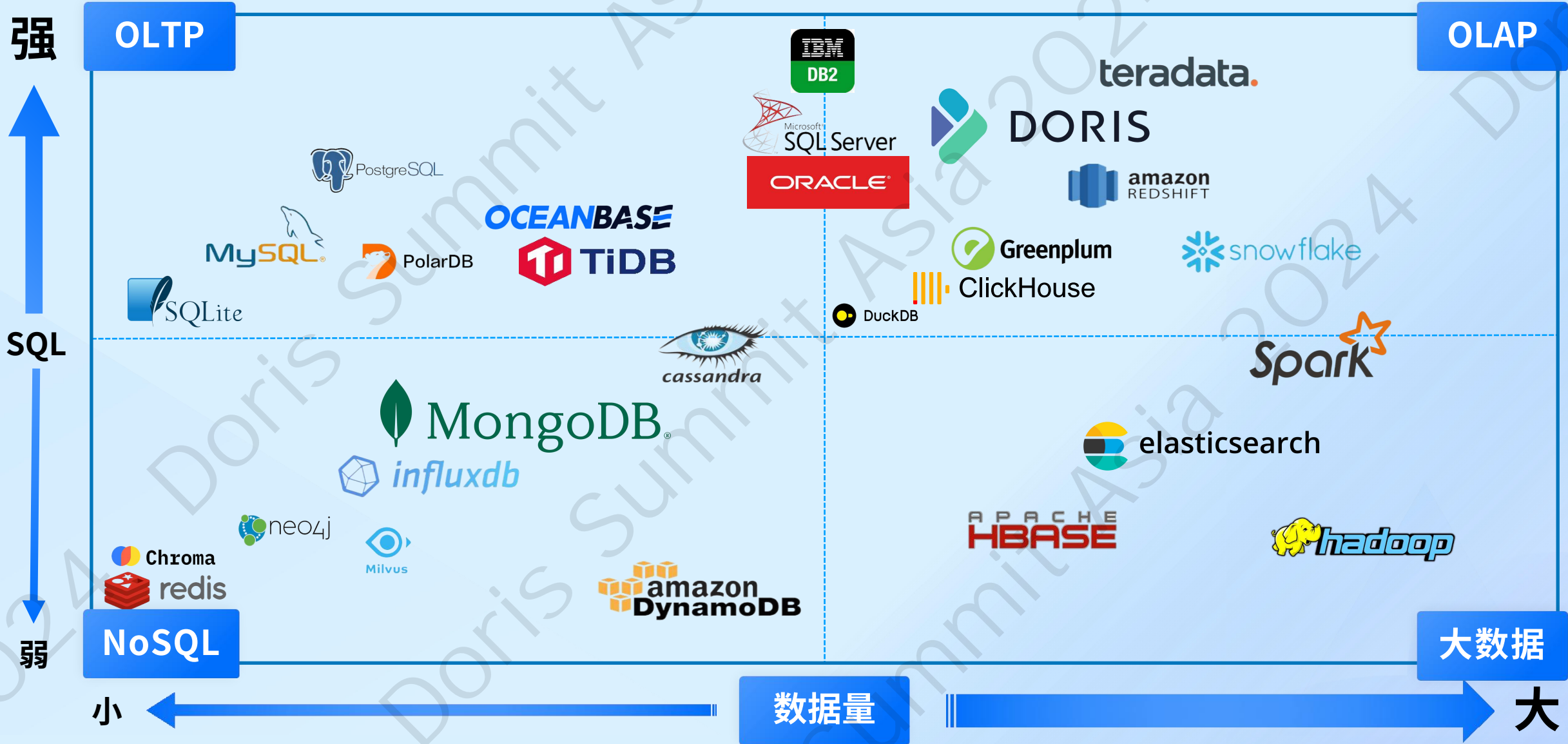
03 技术测试

04 客户实践

01

数据库与数据复制简介

数据库系统分类大图



国产数据库百花齐放

老牌关系型数据库

达梦

人大金仓

南大通用

神舟

...

分布式关系型数据库

OceanBase

TiDB

GoldenDB

AntDB

...

云数据库

PolarDB

GaussDB

TDSQL

GaiaDB

...

数据仓库

Doris/SelectDB

StarRocks

星环大数据

OushuDB

...

NoSQL

TDengine

巨杉数据库

NebulaGraph

Milvus

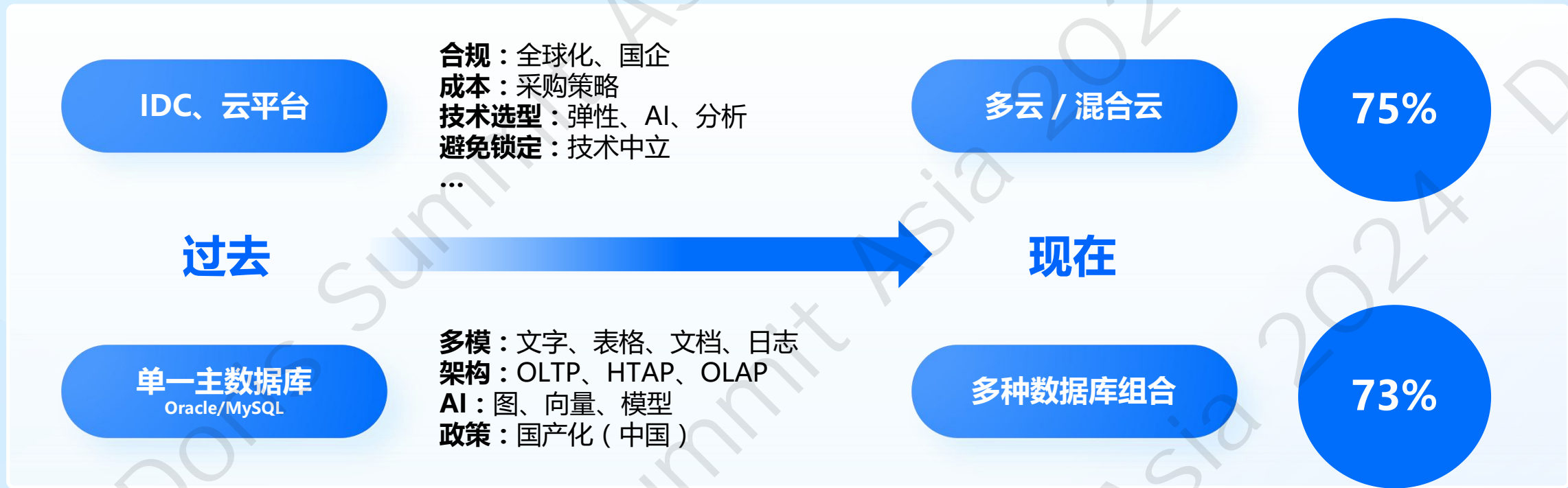
...

墨天轮收录：227 个，2024 年通过国家安全评测 17 个
 (*截至 2024 年 12 月)

数据库与生态产品



【多云多源】数据基础架构发展趋势

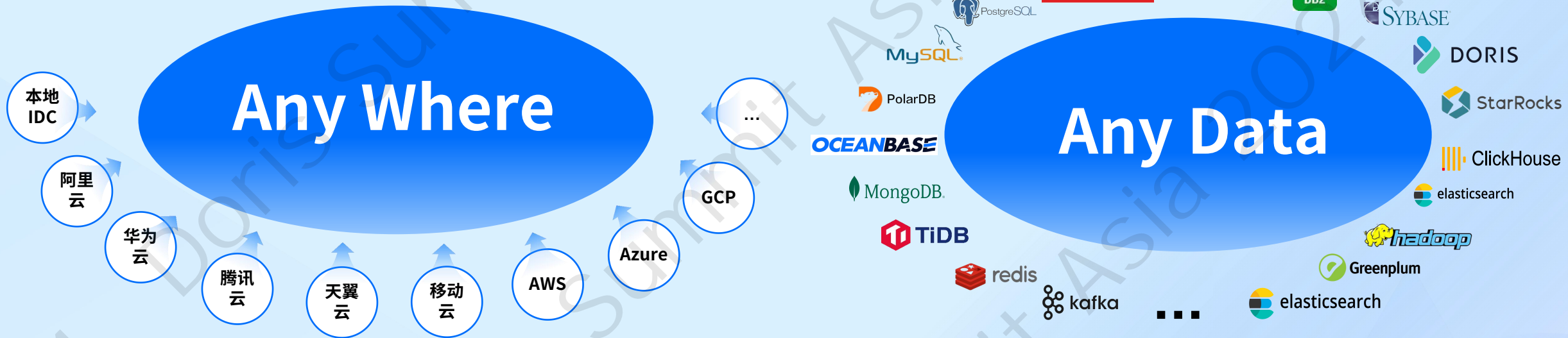


云计算（多云/混合云） + 数据库组合解决方案（多源）

调研企业： 字节跳动、小红书、米哈游、B站、小米、理想、小鹏、格力、美的、海尔、Minimax、Volvo
 中国移动、中国联通、中国电信、杭州银行、建设银行、光大银行、国家电网、中国石化
 OpenAI、Airbnb、Netflix、宝马、沃尔玛、Spotify、Dropbox、X(Twitter)

NineData：安全高效的管理无处不在的数据

跨云/IDC私网连接



NineData 云原生智能数据管理平台



数据复制典型场景

实时数据迁移
上云/国产化

MySQL->RDS
Oracle->Doris
Oracle>PostgreSQL

数据复制容灾

上海<->北京
阿里云<->AWS
云<->IDC

ETL

Oracle->Doris
MySQL->Doris
MySQL->TiDB

数据对比

结构对比
数据对比
差异数据订正

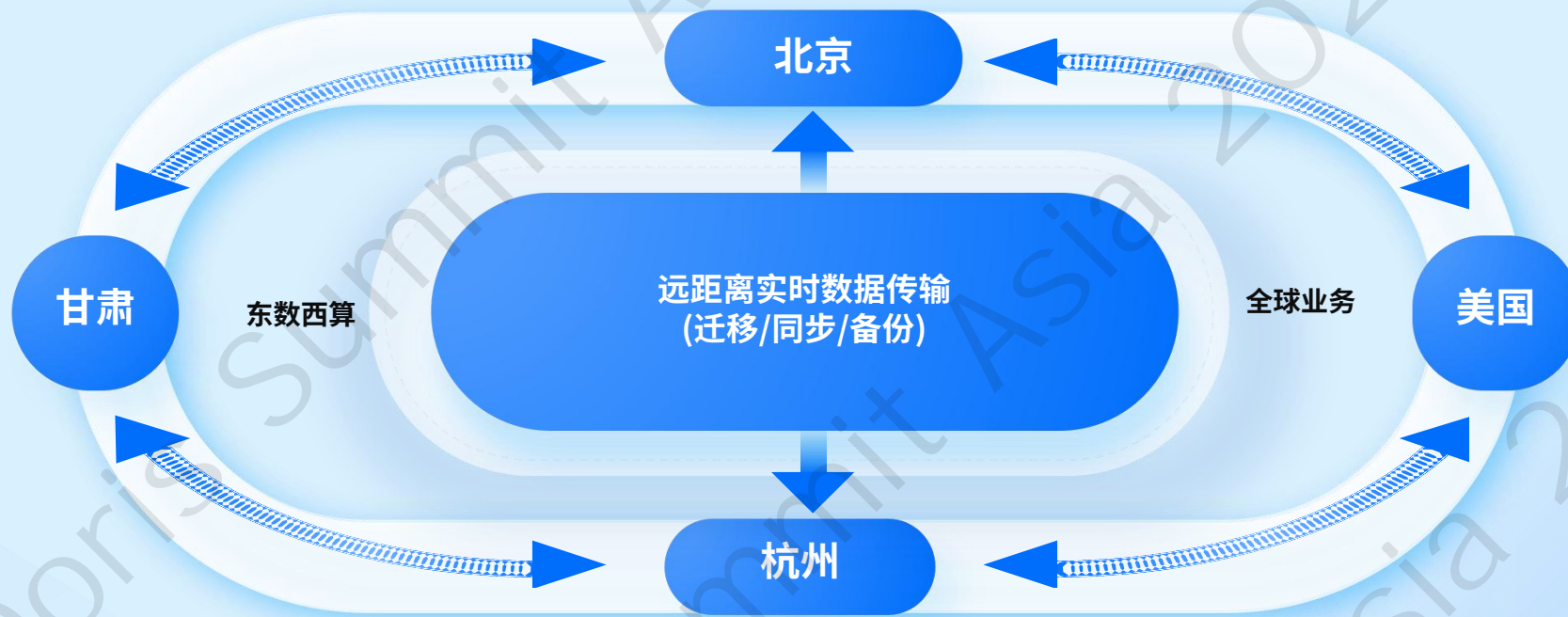
02

NineData 数据复制技术原理

NineData 数据复制架构图



NineData 数据复制：让数据自由流动，创造更多数据价值



丰富数据源

60+

强劲性能

10万 RPS
业界领先

跨云内网传输

业界首创

超远距离实时传输

>10000公里
业界领先

03

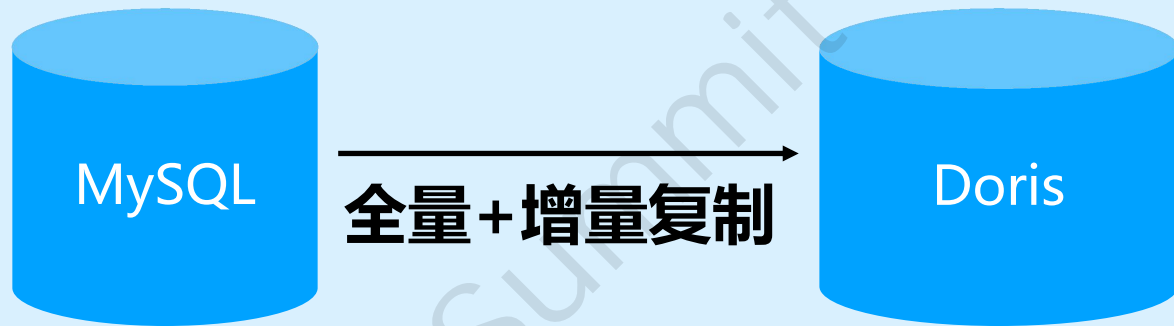
数据同步性能测试

NineData 数据复制对比

对比维度	NineData	Canal	FlinkCDC	某云 DTS
核心定位	多云多源实时数据复制	MySQL 增量复制	大数据实时复制	数据迁移上云
劣势	商业产品	仅支持 MySQL 需要二次开发	易用性、稳定性差 需要二次开发	非云中立，优先支持 自家云数据库
数据源	60+ 种	类MySQL	9 种	18 种
全量数据迁移	33MB/s	32MB/s	31MB/s	27MB/s
增量数据迁移	2.7万记录/s	0.83万记录/s	不稳定	0.55万记录/s
结构迁移	全量 + 增量	全量（仅表对象）	全量（仅表对象）	全量
数据对比	支持	×	×	-

场景：MySQL -> MySQL 数据复制
 数据库硬件：2台ECS 8c16G，ESSD

高性能实时复制 (MySQL > Apache Doris)



测试数据库硬件

- MySQL : 2c16G / SSD
- Apache Doris : 2c16G / SSD

性能数据

- 全量迁移 : ~15.2万 rows/s
- 实时同步 : ~3.1万 rows/s
- 同步延时 : < 1s

数据复制核心技术

- CDC 日志解析 (binlog)
- 分片读取与写入
- 数据压缩
- 多线程 Streaming
- DDL 支持 (自研 SQL Parse)

高性能数据对比 (MySQL & Apache Doris) (百万记录/秒)



测试数据库硬件

MySQL : 32c 256G / SSD

Apache Doris : 32c 256G / SSD

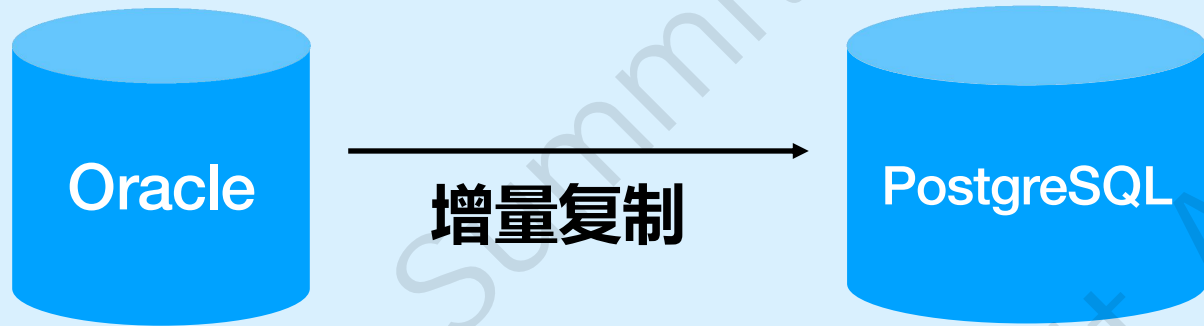
性能数据 : 429 万行/s

数据对比核心技术

- 异构 Hash
- 批量 Hash
- 梯度 Group

高性能实时复制 (Oracle->PostgreSQL)

国产化去 Oracle



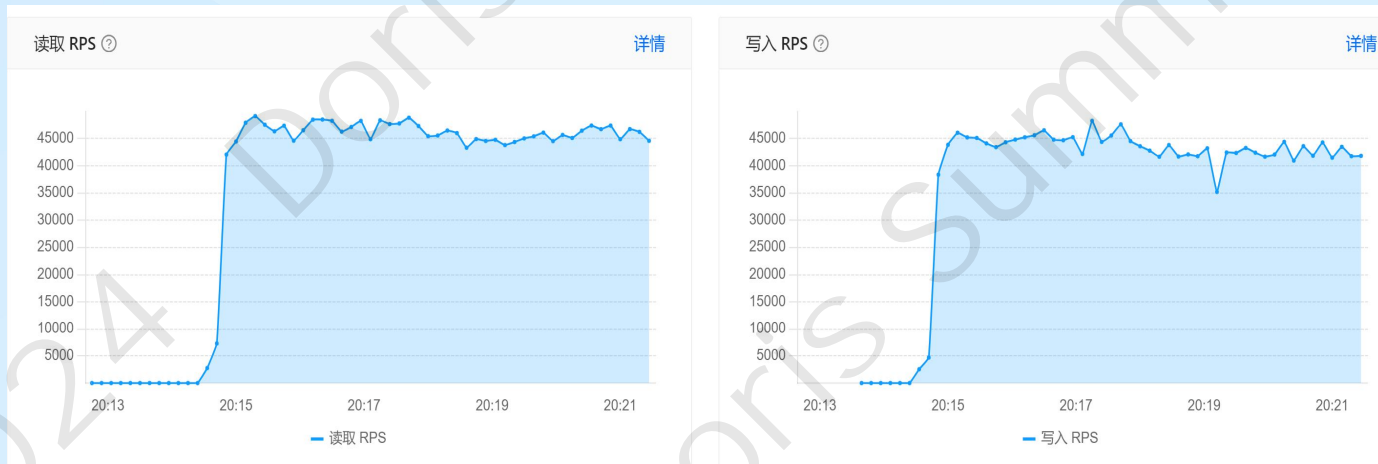
性能数据

- 实时同步 : 4.5万 rows/s
- 日志消费 : 100~200GB/小时

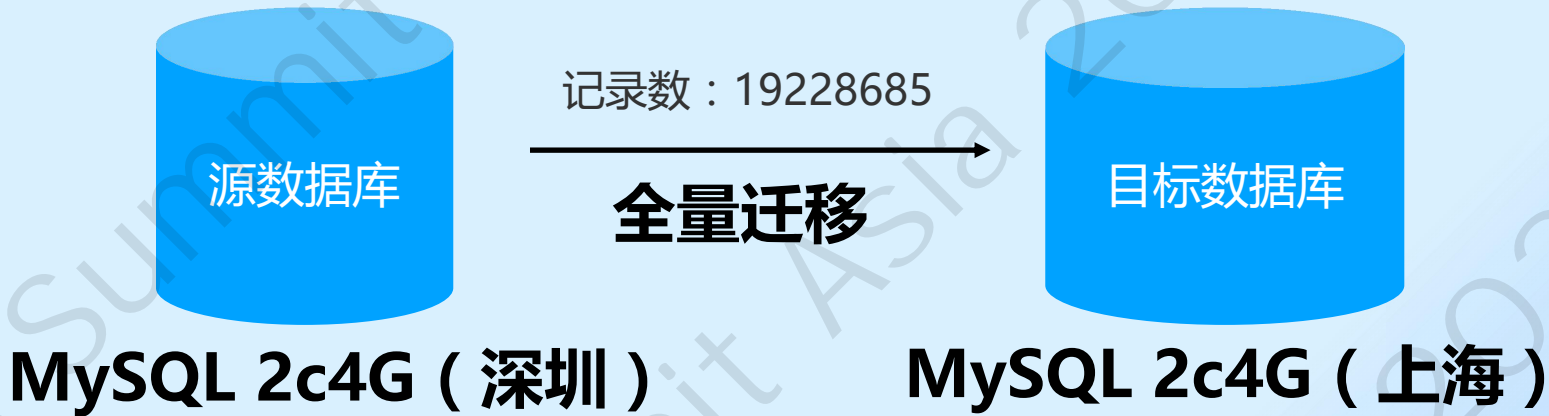
测试数据库硬件

Oracle : 4c32G / SSD

PostgreSQL : 4c32G / SSD



跨城远程全量数据复制（云 DTS vs NineData）

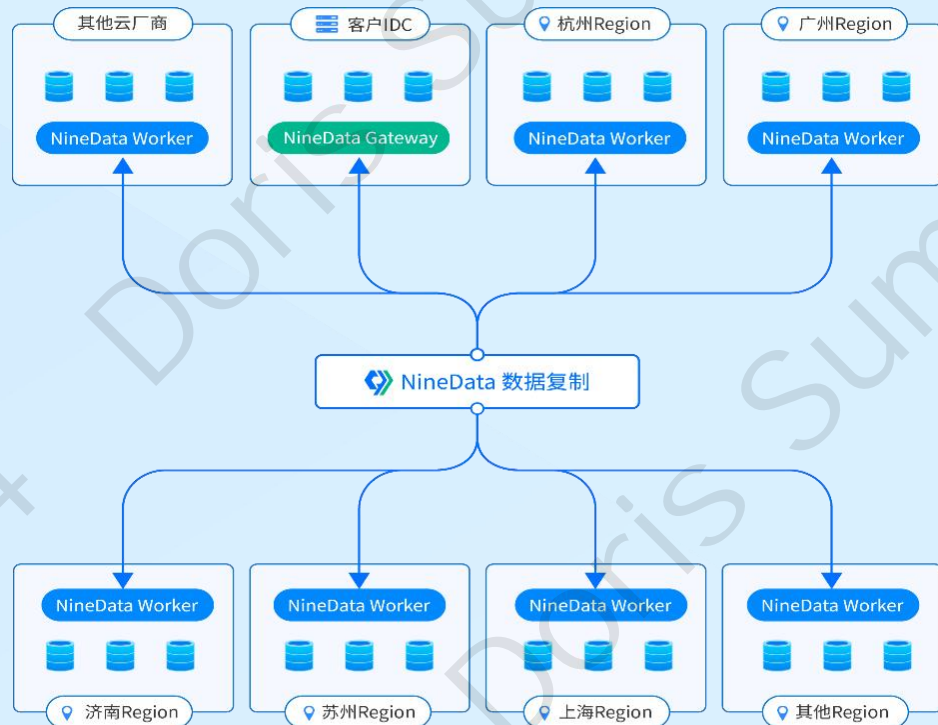


对比维度	A云DTS	B云DTS	C云DTS	D云DRS	NineData
数据复制时间	511秒	315秒	失败	20分钟没响应 后端日志显示失败	245秒
数据复制性能	37630 rows/s	61043 rows/s	-	-	78484 rows/s
数据对比时间	216秒 部分表失败	不支持	-	-	113秒

04

客户实践

客户案例：中国移动云数据复制



❓ 主要问题与挑战

- 客户本身有很多数据需要同步，同时还要提供数据复制产品给他的客户使用，比如该云的客户从其他云厂商或自建系统中迁移上云。
- 应用场景复杂：包括迁移上云、跨云迁移、跨区域迁移、数据容灾、异地多活等业务场景。
- 网络环境复杂：Region 内部/ Region 之间，和其他云厂商与客户自有系统之间等各种链路。

👍 使用成效

- 每天稳定运行的数据链路数百条。
- 通过 NineData 双向复制构建了超长距离（超过 1000 公里）的异地多活集群，支撑业务高稳定运行。

客户实践：某医疗服务企业

核心需求：MySQL > Apache Doris 实时同步



主要问题与挑战

- **实时性**：为了保证业务的连续性，数据同步过程需要尽可能地实时。需要高效的数据传输和处理机制。
- **数据与结构同步转换**：MySQL 和 Apache Doris 拥有不同的数据类型和结构，在同步数据时，涉及到复杂的数据类型和结构映射，来确保迁移后数据的准确性。MySQL结构变化，Doris也需要同时变化
- **缺乏监控告警机制**：在数据同步过程中，可能会遇到各种不可预见的问题，导致同步中断。需要有效的告警机制来及时进行人工干预，以确保数据同步的稳定性。



使用成效

- 稳定运行 1 年
- 边缘部署架构，安全高效、高性能
- DDL 实时同步

客户实践：某国有银行数据库国产化迁移同步

核心需求：去 O、国产化数据库选型、数据同步质量与性能...



PostgreSQL



TiDB



DORIS



AnalyticDB

...

客户实践：沃尔沃

每天自动化完成数百次变更发布



世界500强

“以其安全性和可靠性而闻名”

数百数据库实例

近千研发人员

阿里云+华为云

从 DMS 全部迁移到 NineData

如何

保障不同的经验/习惯的研发人员，
依旧能够构建规范的生产环境

如何

保障生产环境的变更，符合稳
定性规范

如何

保障不同的数据库、不同的环
境，差不多的使用和访问体验

多种功能支持

- ✓ 统一访问与权限管理
- ✓ SSO
- ✓ 数据变更与备份
- ✓ 规则规范
- ✓ 数据库变更管理
- ✓ Online DDL/DML

多种不同的数据库支持



GaussDB



客户实践：绿城中国

管理数十家软件供应商数据库操作

数百数据库实例

IDC / 阿里云/华为云

数十家应用集成厂商

数十个业务应用系统



*股票代码 03900.HK

“中国领先的优质房产品开发及生活综合服务供应商”

数据库

变更随意，多次导致生产环境业务连续问题

开发人员

习惯和经验不同，生产环境字段、索引、注释混乱

数十家

供应商、数百研发人员的数据库权限管理混乱

多种功能支持

- ✓ 统一访问与权限管理
- ✓ 数据库变更管理
- ✓ 数据变更与备份
- ✓ 规则规范
- ✓ 可视化结构设计
- ✓ Online DDL/DML

多种不同的数据库支持

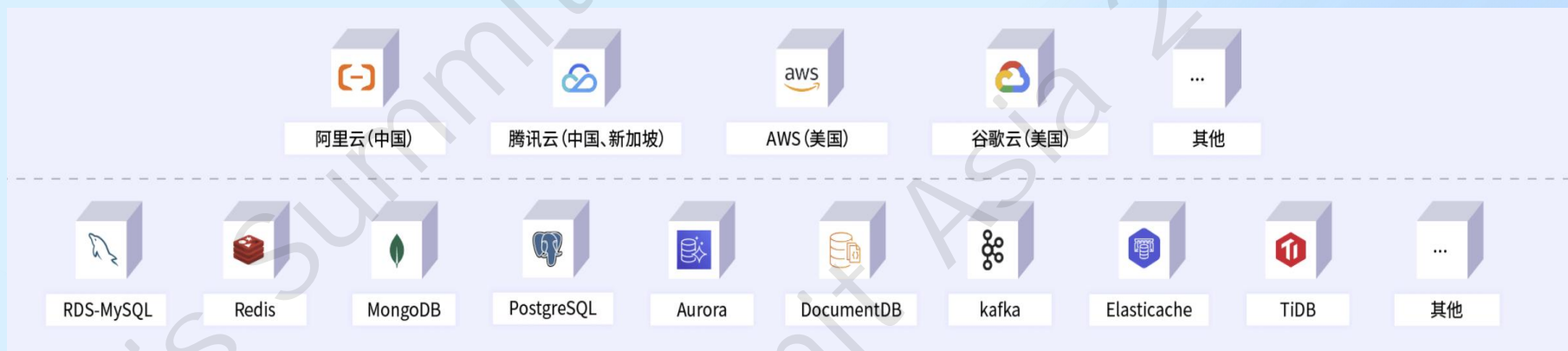


客户实践：MiniMax（全球多云多源）



“AI 大模型独角兽”

- 1 个 DBA 管理数百个数据库
- 4 朵云，数百个数据库接入
- 1 个月完成数据库 DevOps 在公司全面上线



欢迎使用 NineData

www.ninedata.cloud

让每个人用好数据和云

扫码关注 NineData 公众号



Thanks for Watching !