



第十四届中国数据库技术大会

DATABASE TECHNOLOGY CONFERENCE CHINA

数智赋能 共筑未来



北京国际会议中心 | 2023/8/17-19



RedDTS：小红书异构数据同步解决方案

小红书 郑云龙

小红书

目录

CONTENTS

① 背景发展

② 平台建设

③ 应用案例

④ 未来规划

历史原因，小红书内遗存多个数据同步组件，用户和管理员叫苦不迭

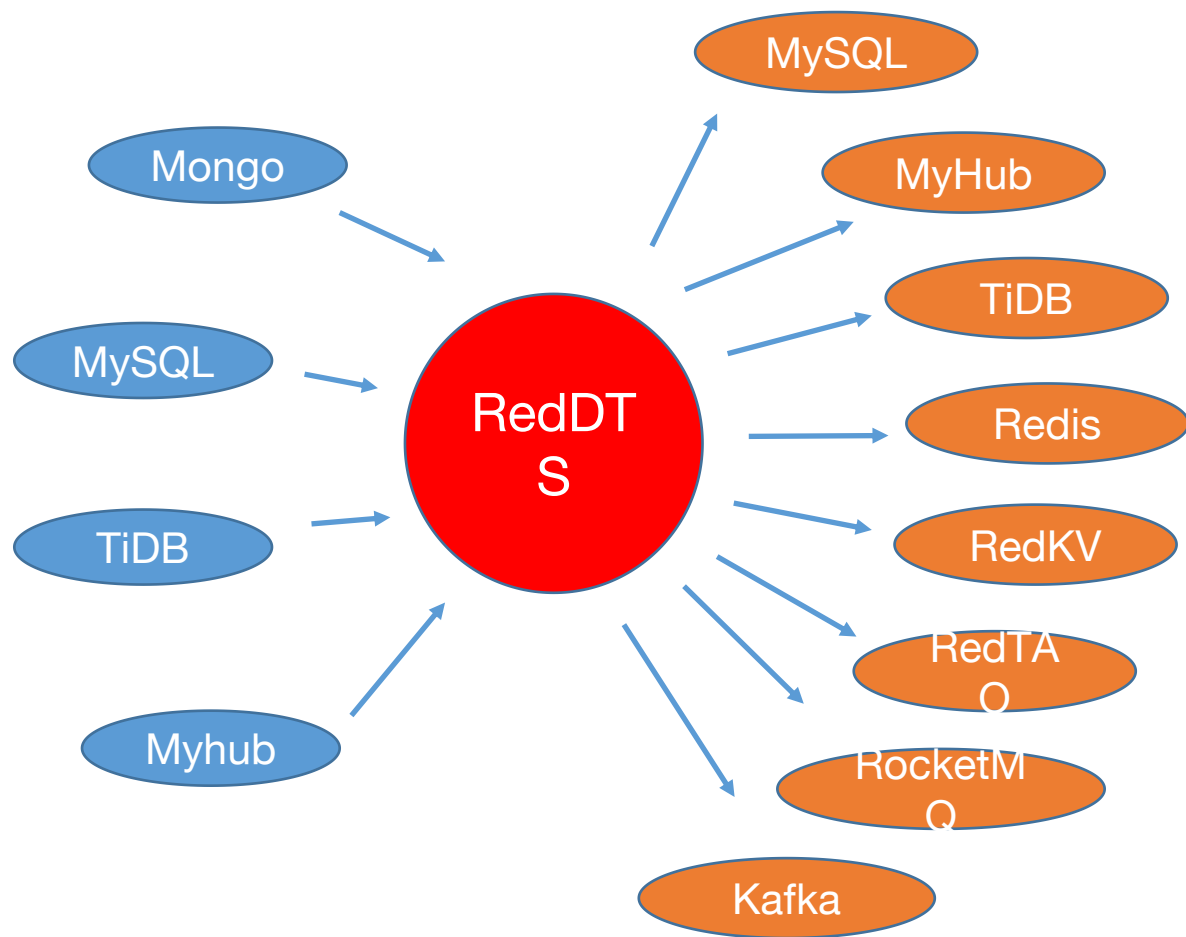
同步工具	源端	目标端	备注
DRC	MySQL	MySQL	自研，用于数据迁移 & 异地多活场景
Canal	MySQL	Kafka	开源基础上修改，主要用在数据订阅场景，将 Binlog 写入 MQ，用户消费做业务处理
Otter	MySQL	MySQL、TiDB、Redis	开源基础上修改，主要用在离线数仓同步
MongoBus	Mongo	Kafka、RocketMQ	Mongo OpLog 2.0
Mongo-Connector	Mongo	Kafka、RocketMQ	Mongo 数据同步
Warp-Prism	Mongo	Kafka、RocketMQ	Mongo OpLog 1.0
Ticdc	TiDB	Kafka	开源基础上修改，TiDB 数据同步

对于用户

1. 组件杂多，无从下手
2. 用户体验差，需求响应慢
3. 配置繁琐，关联平台多

对于管理员

1. 链路复杂，运维压力大，隐患多
2. 无法复用，重复开发，迭代慢



模型统一

- 统一架构、流程
- 能力扩展
- 支持更多的数据源

内部系统深度集成

- 隐藏上下游连接信息
- 屏蔽分库分表细节
- 默认参数场景定制

可视化

- 减少代码编写，直接通过平台可视化配置的方式实现转化逻辑
- 数据可视化展示，更加直观

工具化

- 提供数据工具，减少跨平台跳转

目标：打造集数据同步、校验、修复于一体的数据工具

2022.
02

项目启动

2022.
04

- 上线
- MySQL -> RDB
- MYSQL -> MQ

2022.
08

- MySQL -> TAO
- MySQL -> KV
- Mongo -> MQ
- Otter 下线

2022.
12

- 同构 DQC
- Mongo -> MySQL

2023.
04

- 异构 DQC
- Warp-Prism 下线

小红书

目录

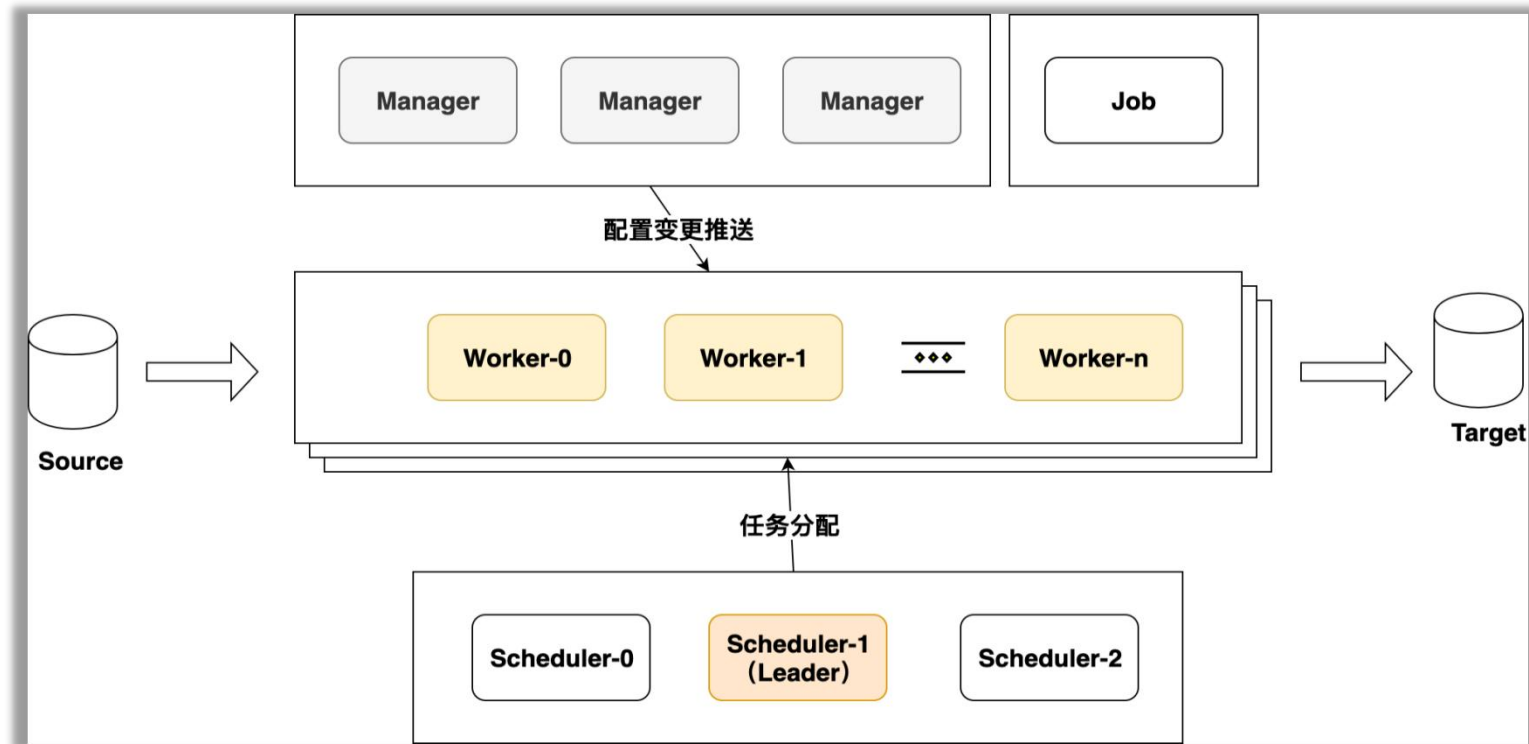
CONTENTS

① 背景发展

② 平台建设

③ 应用案例

④ 未来规划



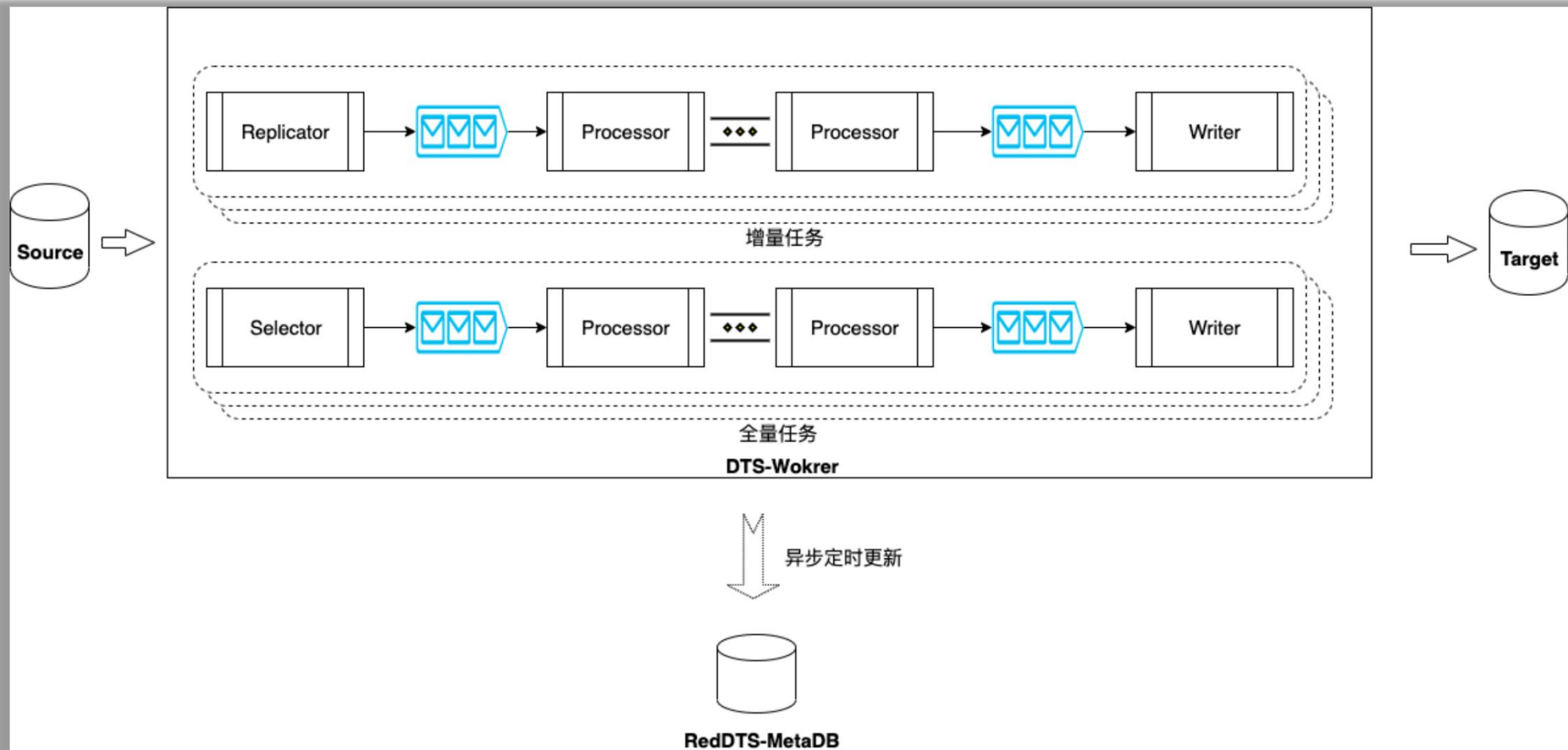
组件	名称	描述	部署
Manager	管控平台	用户交互入口，编辑任务配置，管理集群等功能	无状态
Worker	执行器	执行同步任务节点，获取上游数据，转化并写入下游	分片
Scheduler	调度器	动态调节 Worker 与 Task 的分配关系	主备
Job	定时作业	执行定时执行生成报表、开启校验任务等功能	单节点

逻辑层封装

- Job, 逻辑任务
- Task, 物理任务

Worker 核心组件

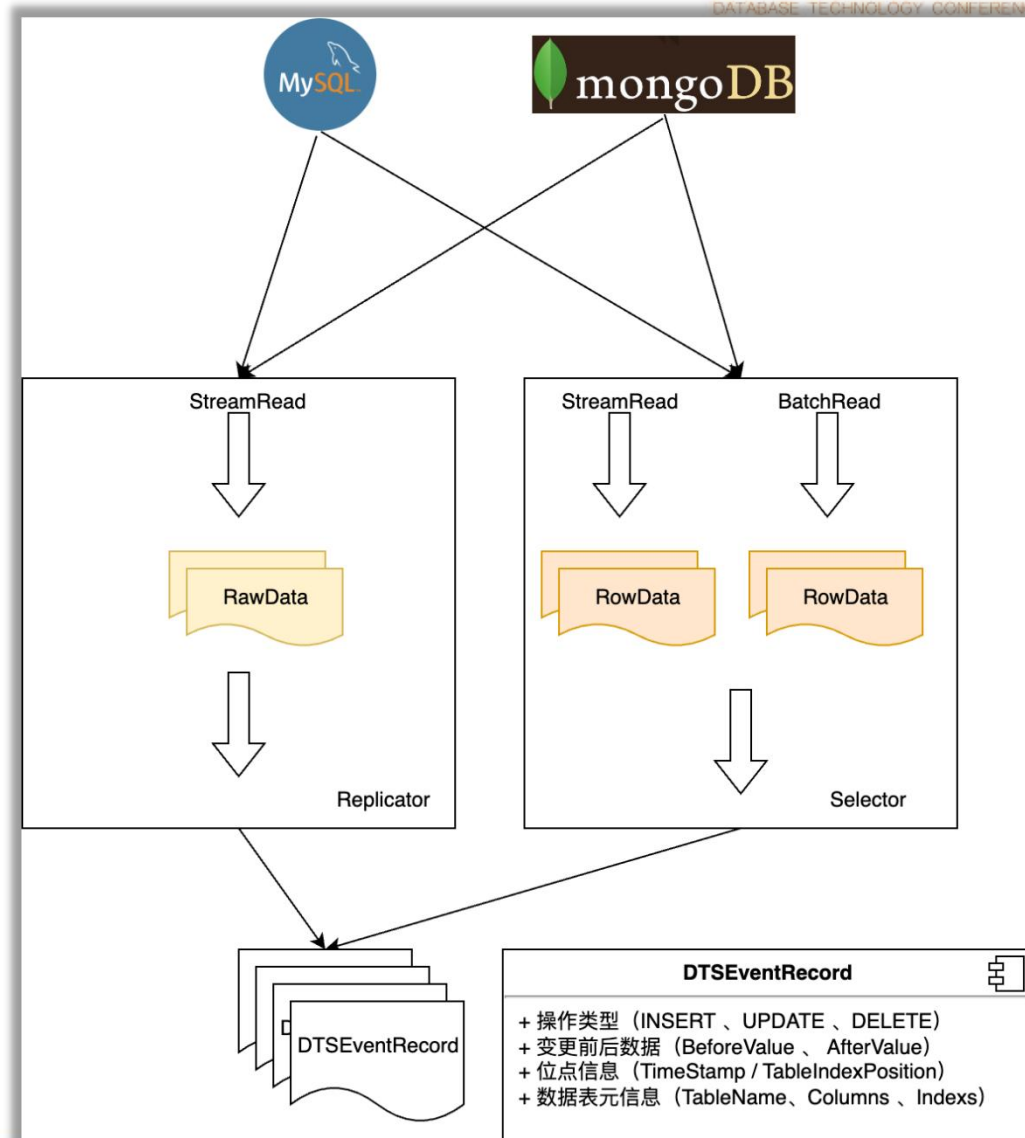
- Reader : 数据读取
- Processor : 数据处理
- Writer : 数据写入



平台建设 > 核心引擎

Reader：数据读取，转化成标准事件

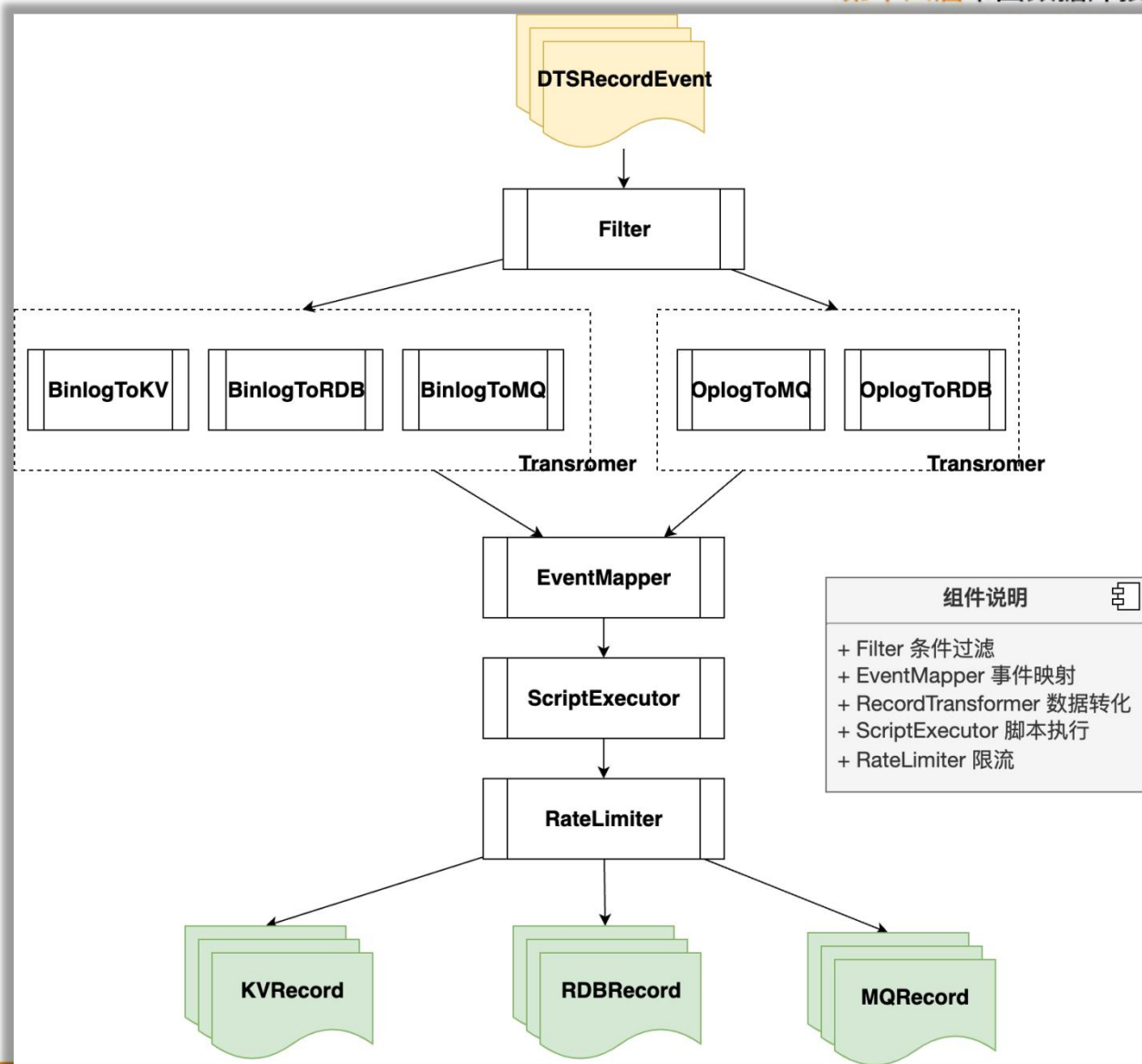
类型	组件	数据获取方式	线程模型
增量	Replicator	Binlog /OpLog	单线程
全量	StreamSelect or	流式读取	表级多线程
	BatchSelector	批式读取	表级多线程



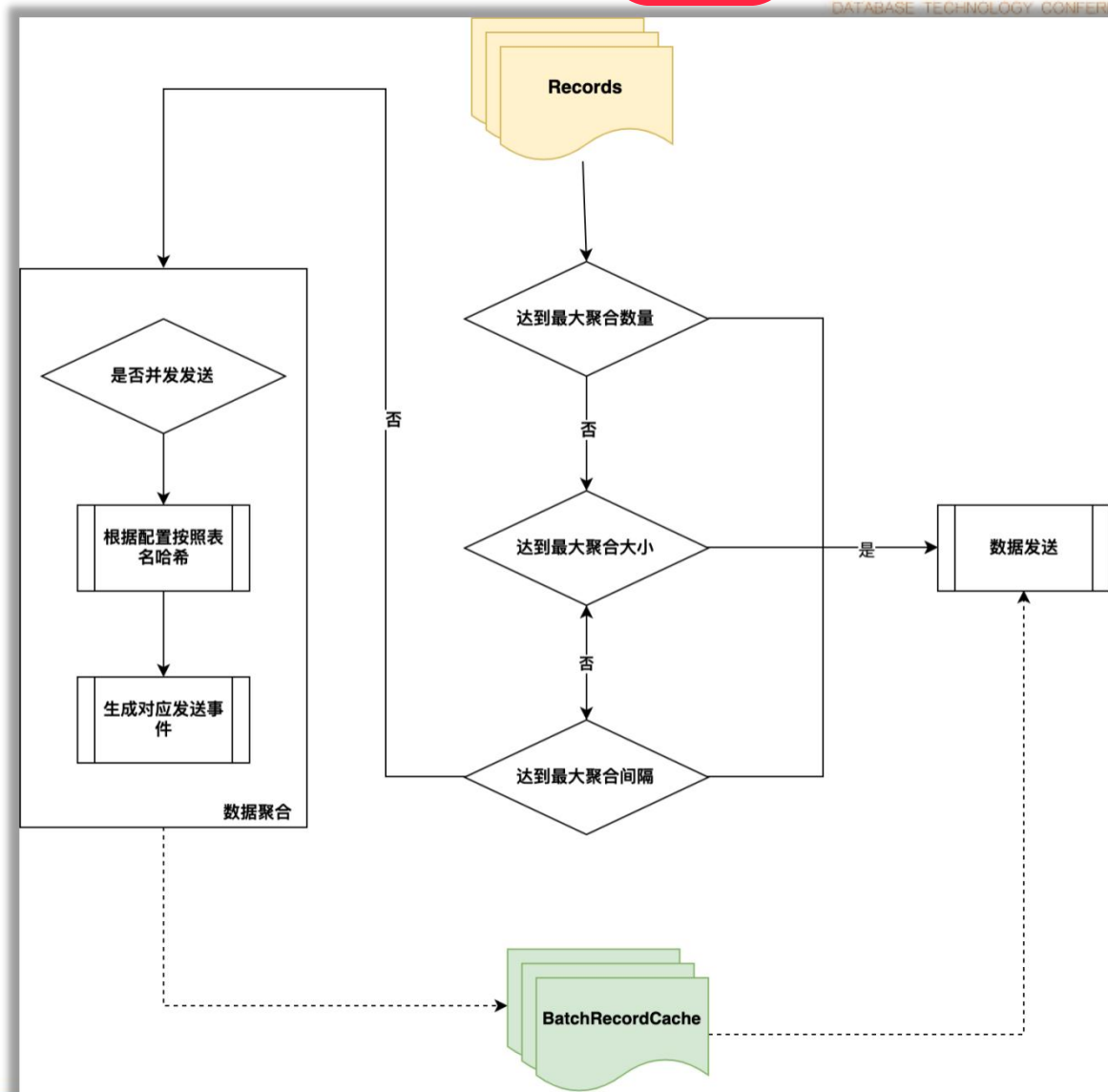
平台建设 > 核心引擎

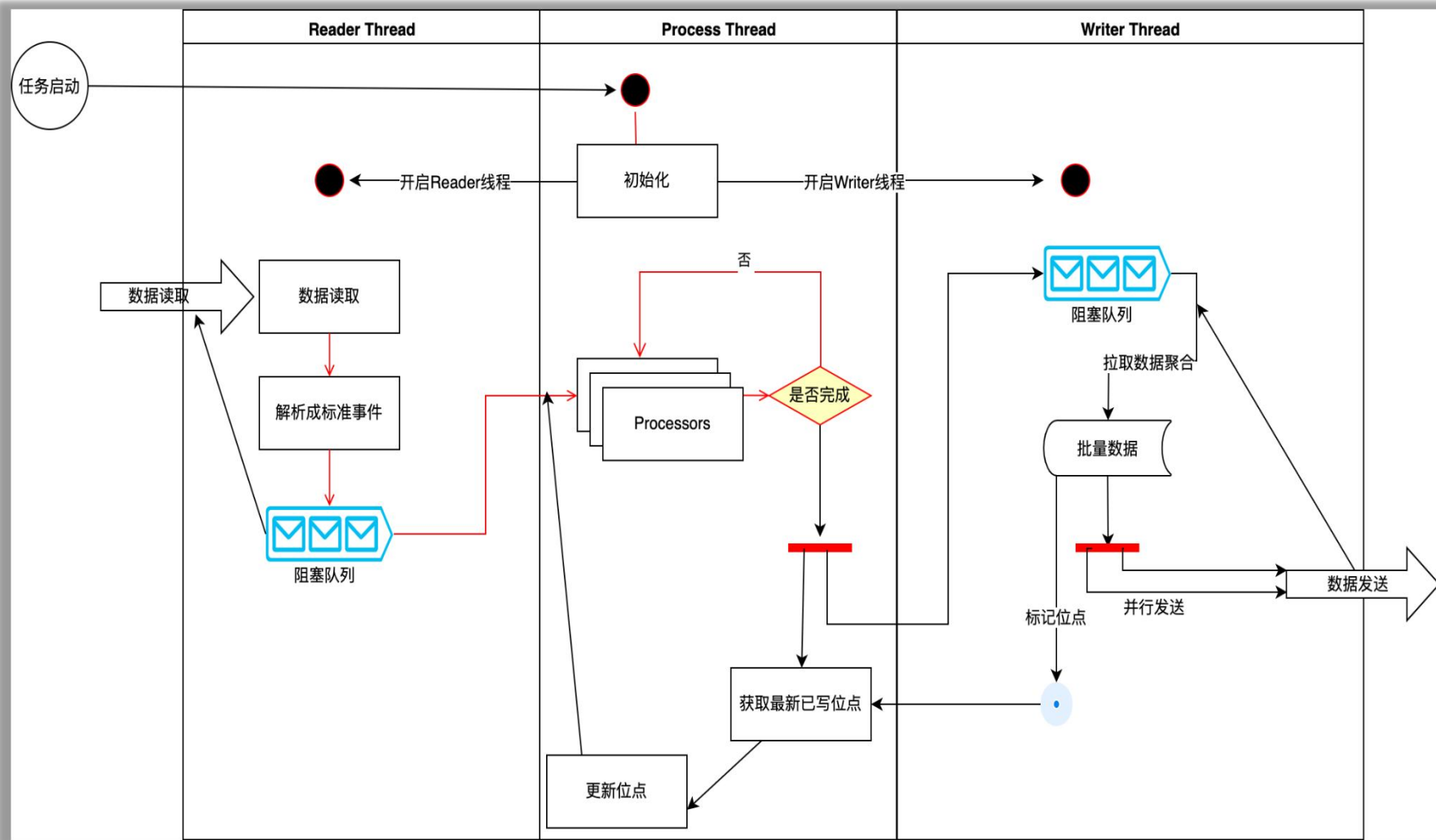
Processor : 数据过滤、转化成目标写入格式

- ? • 如何实现复杂的转换需求表达（拼接、裁剪）？
- 如何复用逻辑？
- ✓ • 支持过滤、转化、事件映射、限流等
- 定义专属数据映射模型
- 基于SpringEL表达转化关系
- 功能原子拆分，串联执行



- 如何解决弱网环境下写入性能问题？
- 如何保证数据一致性与幂等性？
- **批量聚合**：按照间隔，数据大小，数据量聚合
- **分桶并发**：源表名哈希，最小可达物理表级并发
- **批量发送**：采用批量发送语法，介绍 IO 开销
- **顺序写入**：分表之内单线程有序
- **幂等语法**：目标端操作幂等，保证最终一致，MQ 要用户消费幂等





线程模型

- 读写组件采用独立线程
- 组件之间通过阻塞队列解耦
- 由 Processor 线程驱动数据流转

任务启停与异常处理

- 启动时，Processor 完成其他线程初始化
- 运行时，Processor 线程检测其他线程是否结束，捕获异常
- 结束时，Processor 线程主动中断其他线程

平台建设 > 高可用与水平扩展

DTCC 2023

第十四届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2023

小红书

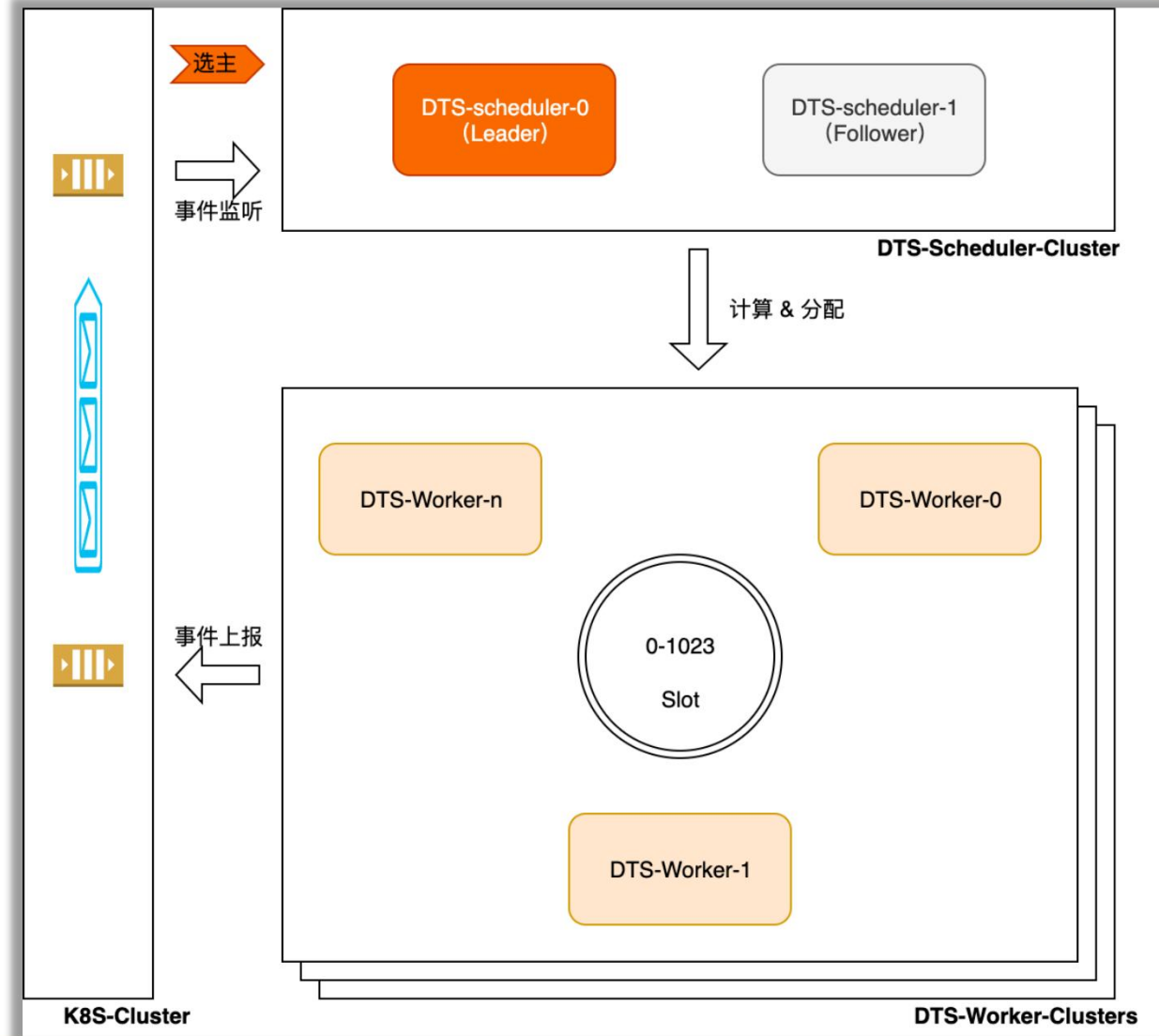
- ? • Worker 发版与宕机
- Worker 的负载均衡（扩缩容、运行时）
- ✓ • 组件以 StatefulSet 类型部署在 K8S 集群中
- 结合动态迁移实现故障转移以及负载均衡

•Worker

- 分片部署，可以水平扩展

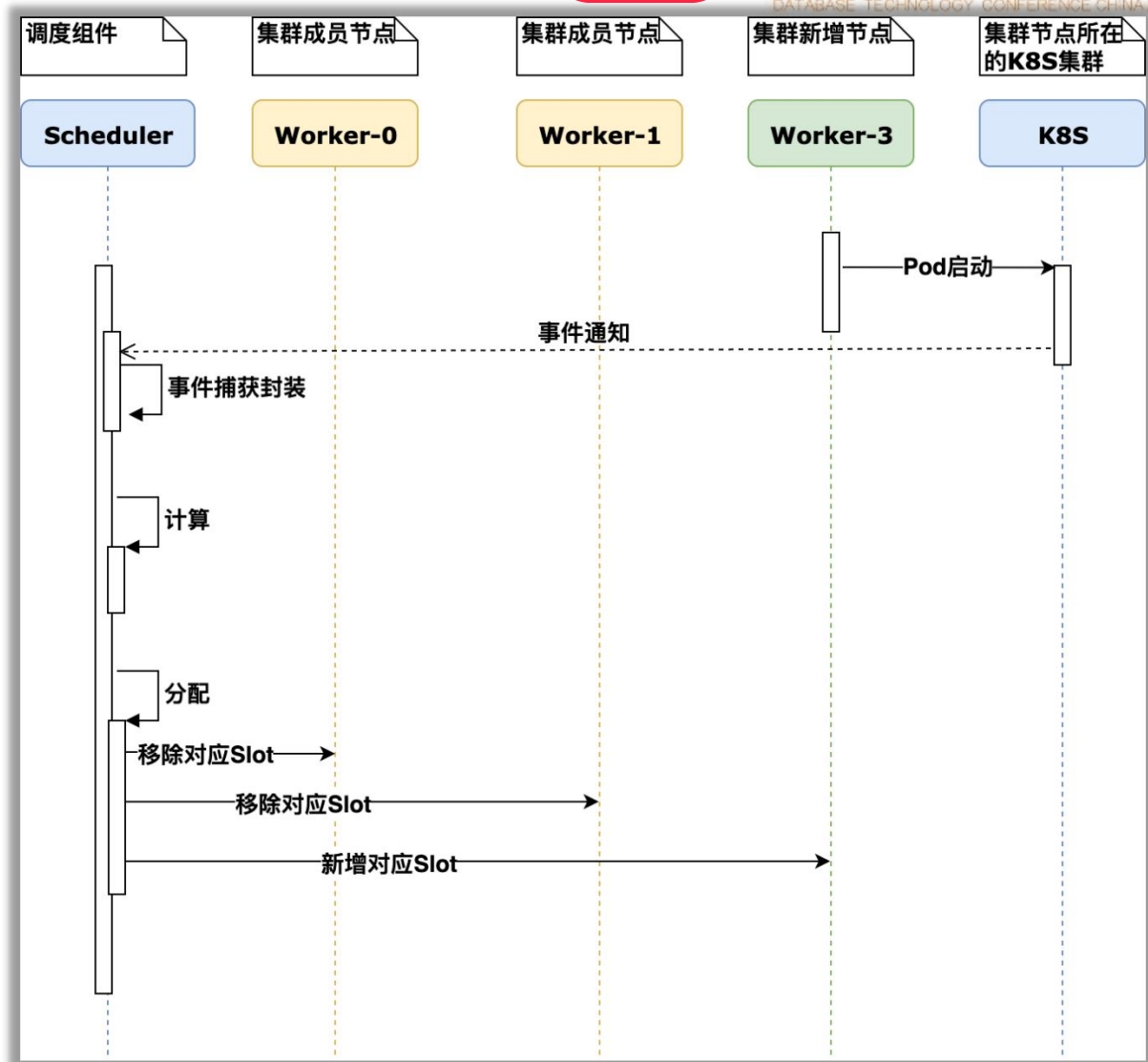
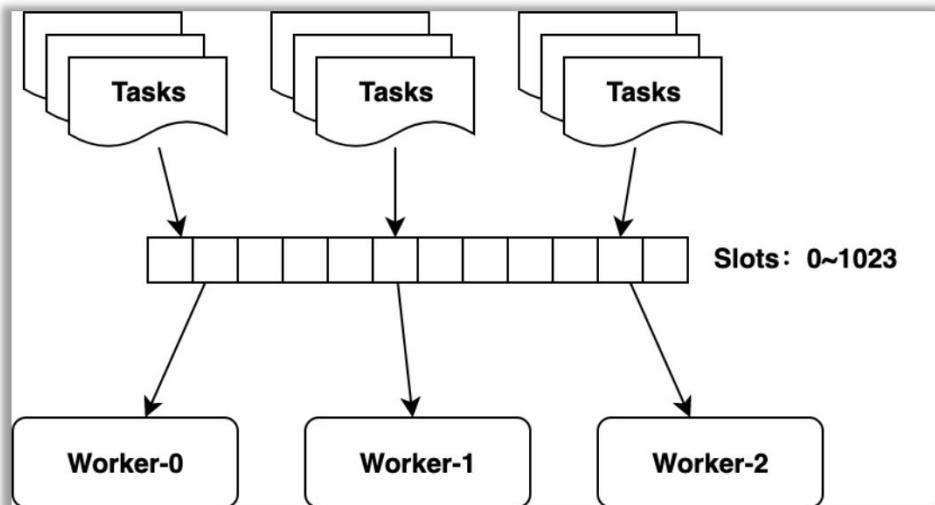
•Scheduler

- 监听 K8S Pod 变更事件，动态处理 Worker 与 Task 的映射关系。
- 主备部署，通过 K8S Lease 资源选主



平台建设 > 高可用与水平扩展

- 引入虚拟层 Slot 来解耦 Worker 和 Task
- Task 与 Slot 映射关系固定，Scheduler 动态处理 Worker 和 Slot 映射关系

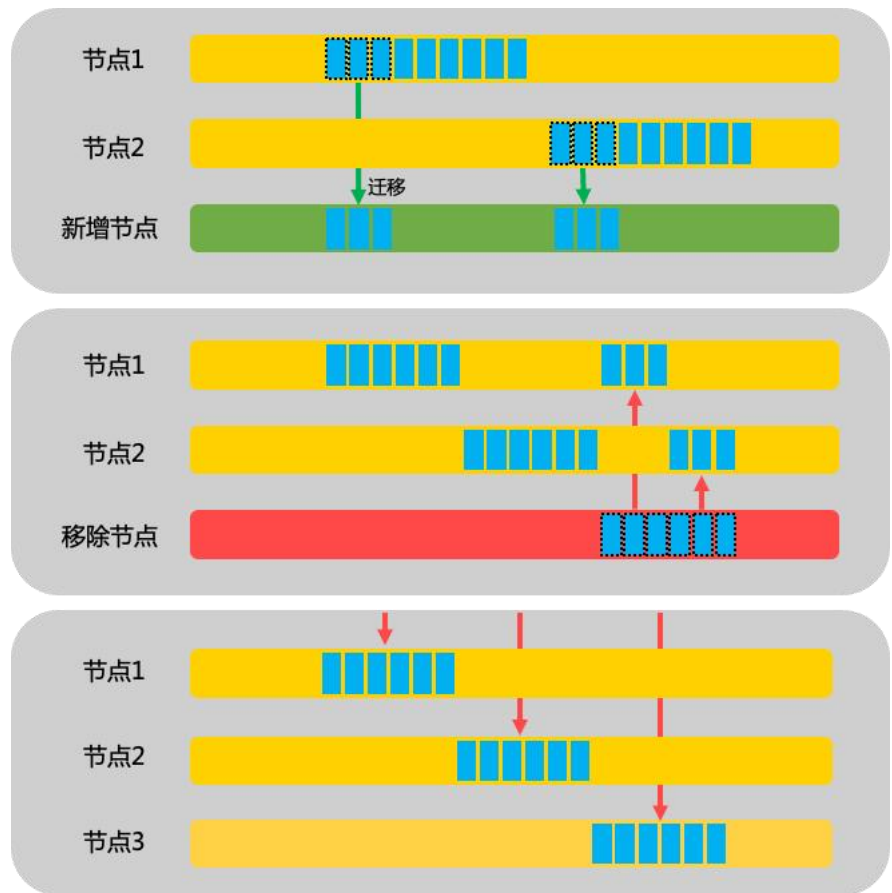


平台建设 > 高可用与水平扩展

小红书

DTCC 2023

第十四届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2023



均等分配

?

- 如何减小迁移规模
- 如何保证迁移之后负载均衡

✓

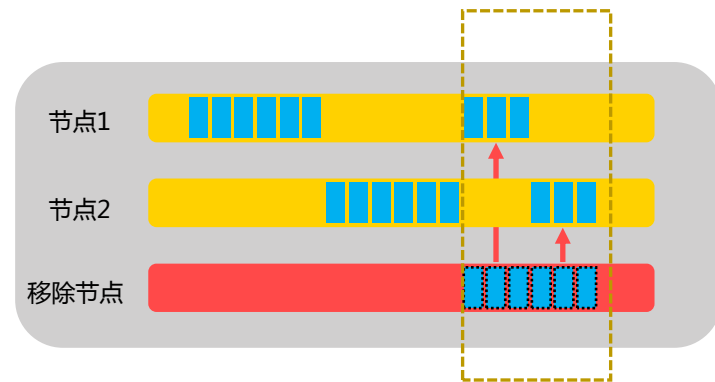
- 只关心 Slot 个数
- 最小粒度迁移
- 迁移之后个数尽量相等

粘性分配

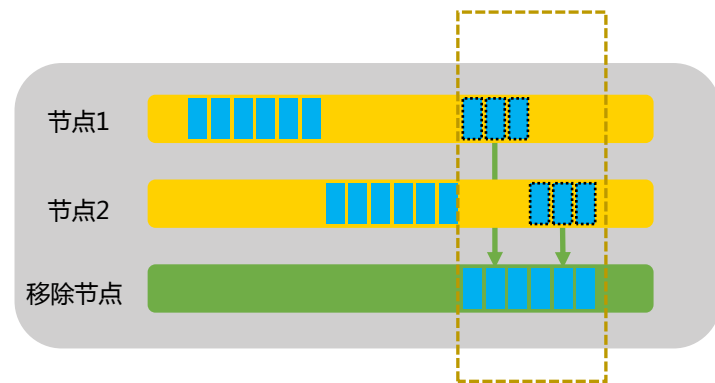
?

- 如何控制滚动升级场景下的任务迁移次数
- 记录上一次移除节点事件的分配关系，下一次当前节点加入按照记录关系还原
- 针对滚动发布场景优化，优先级高于均等分配

✓



移除节点时记录迁移关系



重新加入时按照记录还原

ID	Region	Zone	WorkerId	最后更新时间	Actions
4802	qc-sh	qcsh4	dts-worker-kv-0 ⓘ	2022-11-18 15:29:19	删除🗑
4803	qc-sh	qcsh4	dts-worker-kv-1 ⓘ	2022-11-18 15:29:18	删除🗑
4823	qc-sh	qcsh4	dts-worker-kv-2 ⓘ	2022-11-18 15:29:18	删除🗑
4824	qc-sh	qcsh4	dts-worker-kv-3 ⓘ	2022-11-18 15:29:18	删除🗑

基本信息

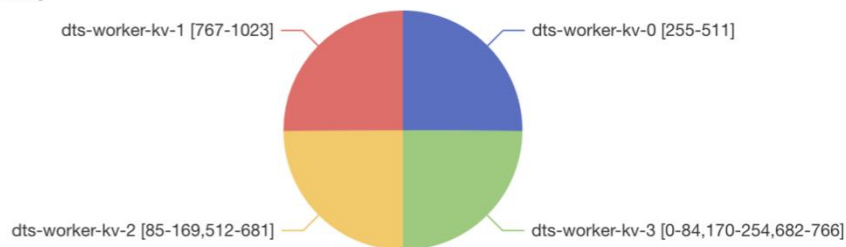
🔍 节点个数	4
🔍 分配中	false
🔍 是否合法	true

操作

启用 禁用 重置

- dts-worker-kv-0 [255-511]
- dts-worker-kv-3 [0-84,170-254,682-766]
- dts-worker-kv-2 [85-169,512-681]
- dts-worker-kv-1 [767-1023]

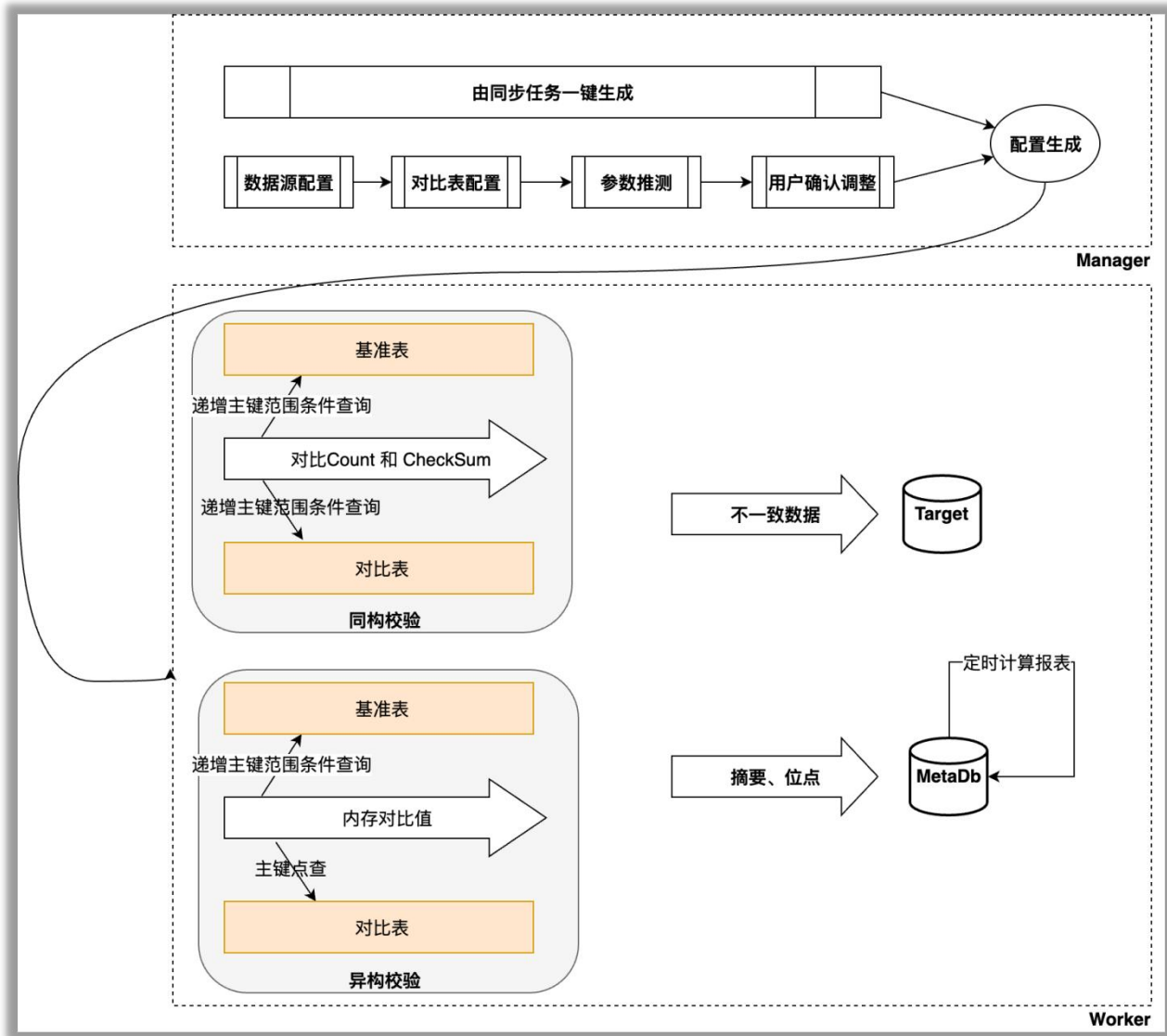
集群拓扑



发版延迟 10s 内

自动负载均衡，误差 10%

任务可迁移



配置时

- 可以根据同步任务一键生成，也可以选择数据源创建
- 自动根据表结构优化查询条件，选择索引
- 参数可选，满足不同场景
- 支持同构和异构校验两种模式

运行时

- 分块校验，减小慢查影响
- 自动按照索引递增，记录位点
- 记录不一致区块，方便快速定位
- 输出校验报告，推测异常原因



执行结果可视化

- 对比元数据
- 数据总数
- 不一致数据区块

报表分析统计

- 不一致原因推测
- 不一致表数统计
- 异常占比排名

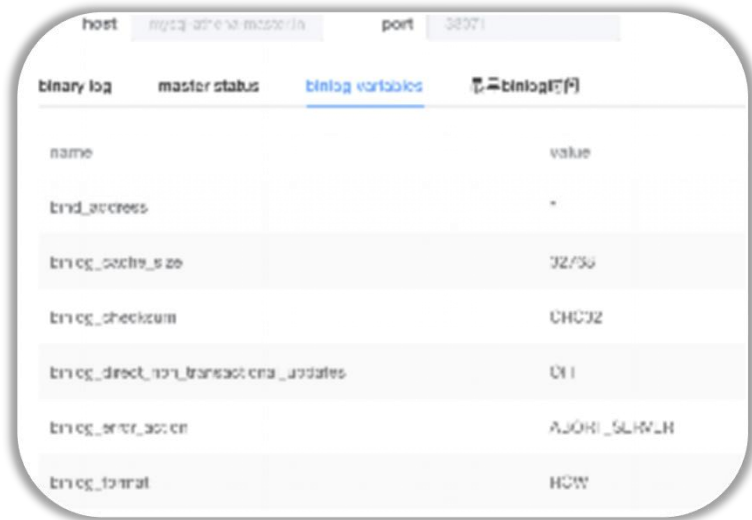


平台建设 > 数据工具

小红书

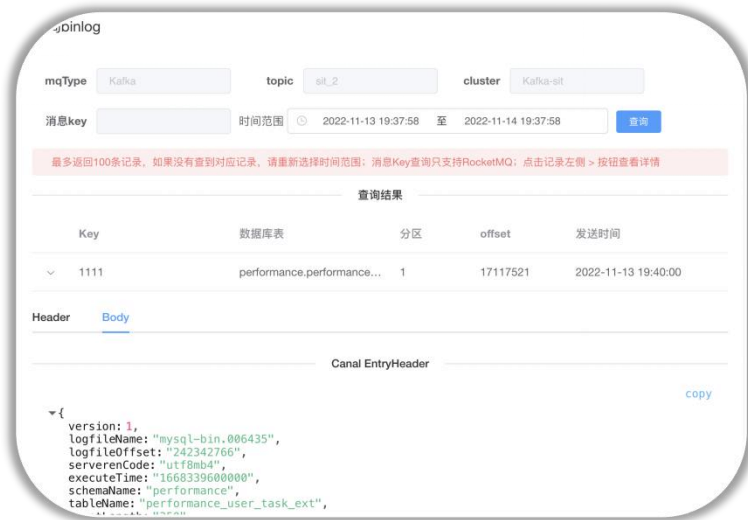
DTCC 2023

第十四届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2023



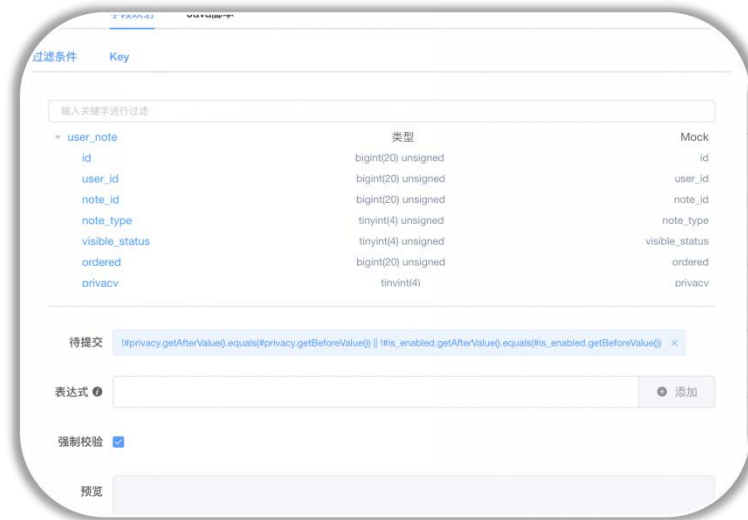
MySQL工具

- Binlog文件列表
- Binlog参数查询
- 库表结构查询
- 时序Schema管理
- Binlog内容查询【DOING】
- Binlog流量统计【DOING】
- 数据修复【DOING】



MQ工具

- Binlog消息查询
- 下游链路查询



映射编辑工具

- 映射规则编辑
- SpringEL表达式编辑调试
- Java脚本编辑上传

- 深度整合，自动建表 / Topic
- 逻辑层封装
- 断点续传、位点重置
- 进度预估
- 多种数据工具

方便易用

- 分类重试
- 写入限流

健壮安全

- 延迟写入
- Partition-Level 心跳
- 日期表

场景特性

- 动态调参
- 格式统一、组件复用
- 存算分离
- 自动负载、故障转移
- 链路查询、迁移

高效运维

小红书

目录

CONTENTS

① 背景发展

② 平台建设

③ 应用案例

④ 未来规划

DTCC 2023

第十四届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2023

DTCC

数智赋能 共筑未来

IT168.com

ChinaUnix.net

ITPUB

分类	上游	下游	功能描述	业务方
数据订阅	•Mongo •MySQL	•Kafka •RocketMQ	将数据库变更发送至 MQ 中 • Mongo Oplog订阅 • MySQL Binlog订阅	目前用途最广的任务类型，有实时数仓、以及各种核心业务使用
数据迁移	•Mongo •MySQL	•MySQL	通常是一次性的，完成所有数据的转移 • 数据库拆库、拆表 • 数据库自动化扩缩容 • 异构数据源迁移	用于 Mongo 迁移到 MySQL 工具以及 MySQL 数据库迁移、扩容等场景 目前 DBA 使用居多，也有业务在自助使用
数据同步	•MySQL	•MySQL	一般需要持续同步，比如 • 同步数据用于离线 OLAP 场景 • 构建二级查询索引 • 构建宽表	用于二级索引、DBA 元数据线上同步至线下，宽表整合等场景，一致性和实时性要求较高。 当前数据同步也承接了一些业务处理能力，比如互关需求通过同步任务来构建关系，无需业务逻辑实现关注和取关功能
缓存同步	•MySQL	•Redis •RedKV •RedTAO	MySQL 数据同步至 RedKV / Redis / RedTAO	用于笔记中台清除缓存，用户中台构建个人页缓存，RedTAO 更新缓存等场景
数据归档	•MySQL	•TiDB	• 分表汇聚成单表，通常用于 OLAP 场景 • 多个业务场景数据聚合汇总 • 历史数据归档	数仓、BI、中台等
数据清洗	•MySQL	•MySQL	对于表内数据进行加工并写回，比如 • 加密某个字段 • 新增字段补全 • 大表 DML	主要为全量任务，重要程度相对较低，个人信息数据加密，字段补充，拆解等

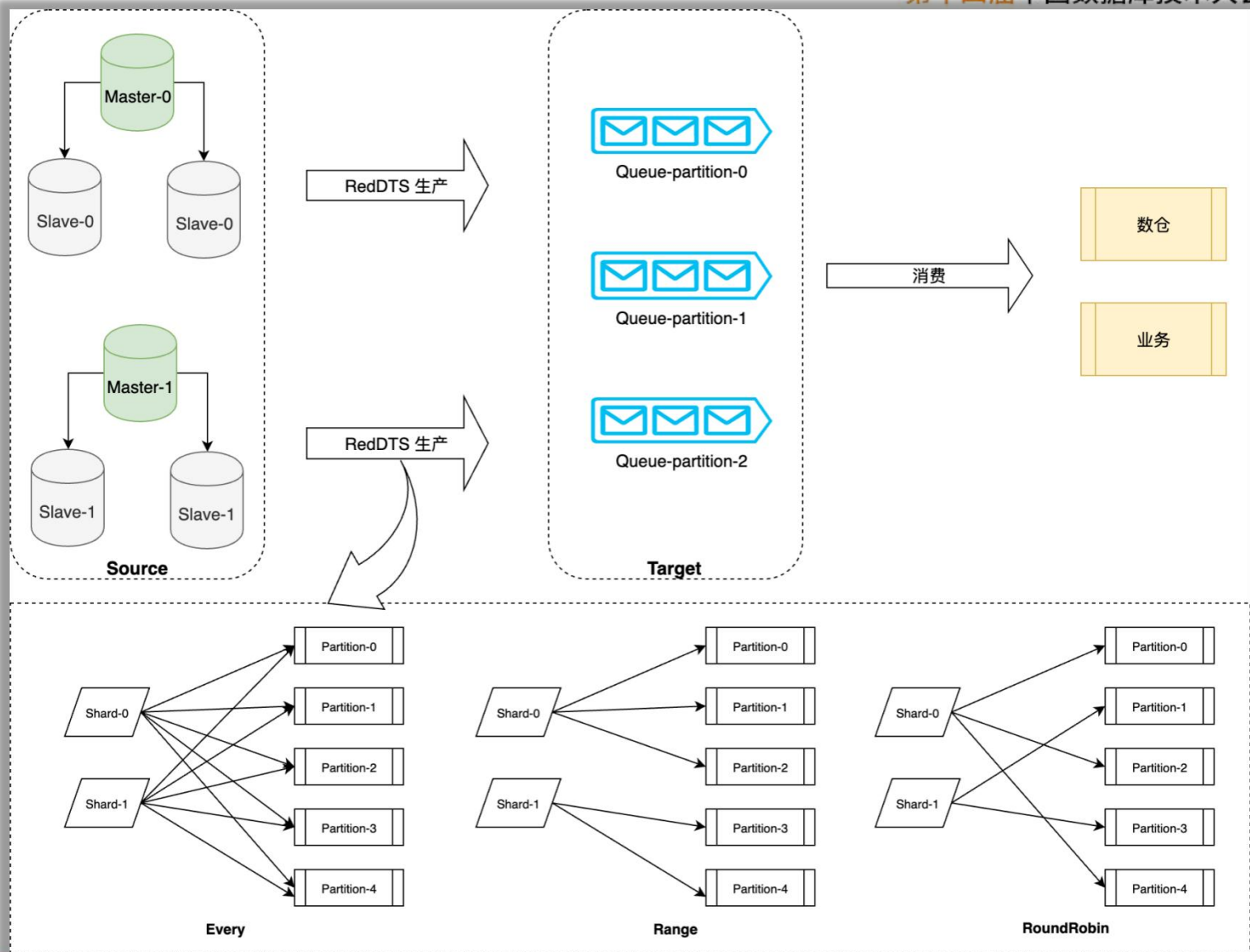
应用案例 > 数据订阅

?

- **重复创建**，存在大量的重复链路，浪费资源
- **健康检查**，下游需要实现Partition Level 心跳，感知源端是否正常

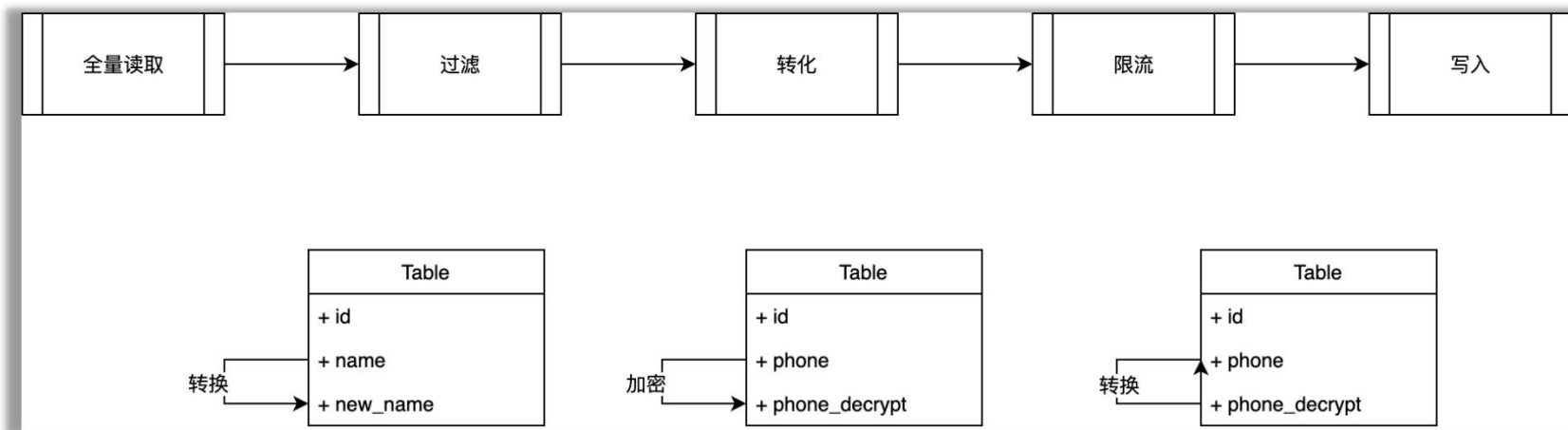
✓

- **格式统一**：以逻辑库为单位创建链路，全量增量格式相同，写入到不同的 Topic 中，增量由管理员维护，全量可由用户操作
- **链路复用**：复用已有链路，增量链路只允许加表，禁止用户修改，具备链路的检查、合并的能力
- **链路查询**：通过源端信息获取下游消费信息，方便迁移、通知等运维操作
- **局部有序**：增量数据通过主键 Hash 计算获取 Partition 编号，发往特定 Partition，保证相同的主键数据有序
- **心跳发送**：支持Every、Range 和 RoundRobin 三种路由模式，支持 Partition 级别的心跳发送





- 1. 字段改名或者类型
- 2. 字段值修改
- 3. 字段加密
- 4. 大表 DML



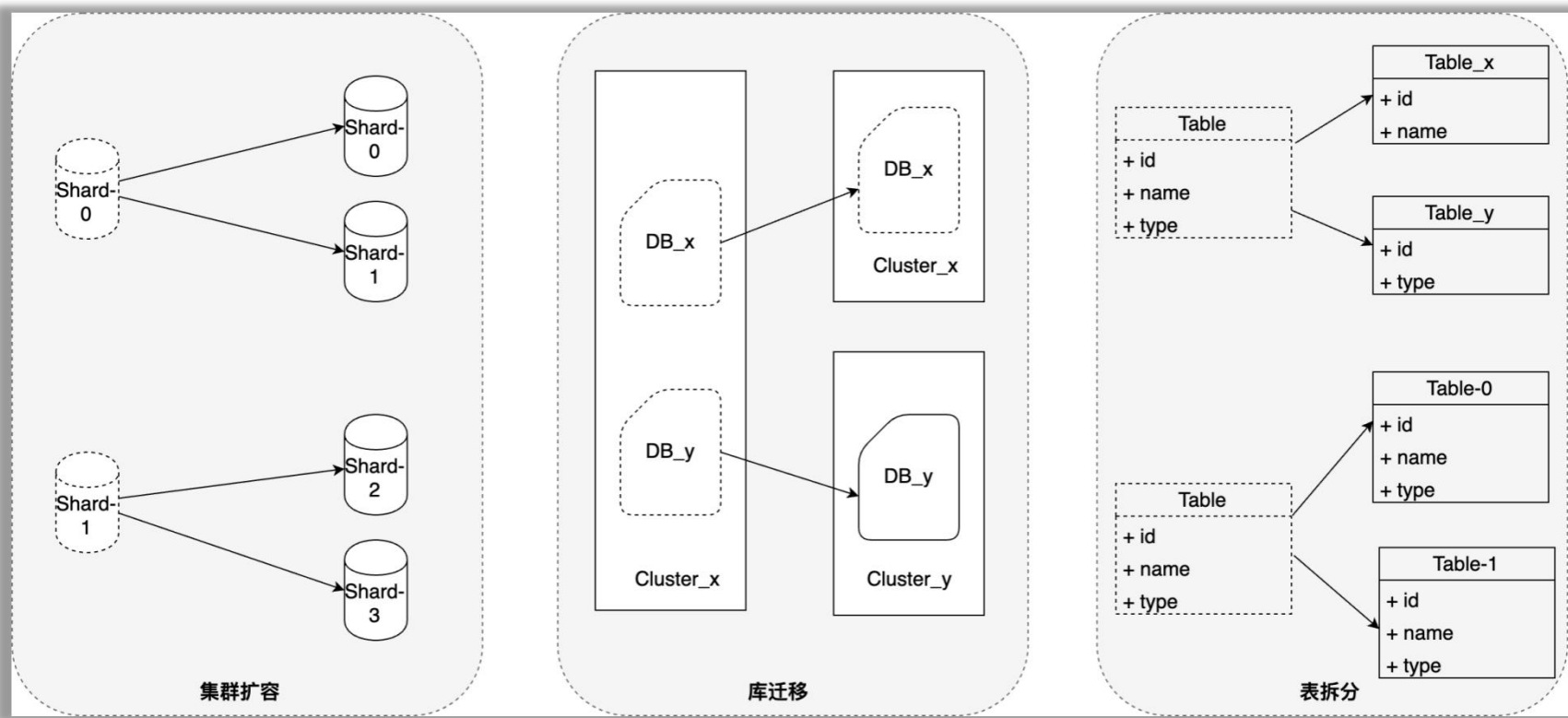
- 1. **流量控制**，未控制速度，导致流量过大，数据库抖动，产生大量变更，影响下游
- 2. **断点续传**，对于执行时间跨度较长时，需要具备任务中断，重新恢复的能力
- 3. **进度统计**：可以观察当前处理进度并估算预计完成时间
- 4. **转化过滤**：按照一定规则完成数据加工处理

?

- 1.数据库集群扩容
- 2.数据库库、表拆分以及合并
- 3.异构数据源迁移

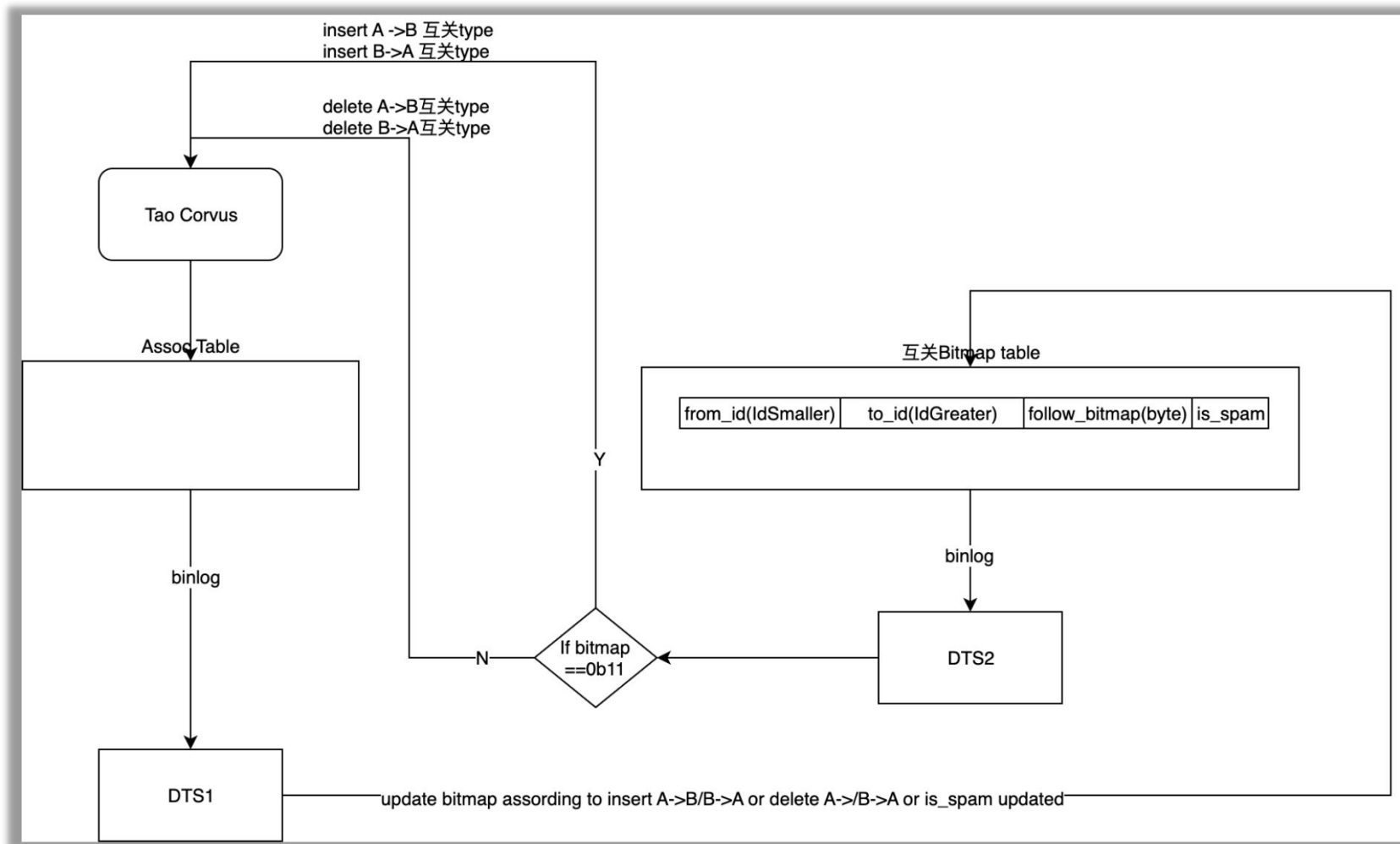
✓

- 1.数据同步，全量+增量
- 2.数据校验，校验 Count 和 CheckSum
- 3.链路迁移，一键查询下游链路，并迁移至目标库
- 4.反向链路，用于数据回滚



- 1. 多个业务场景数据聚合汇总
- 2. 历史数据归档
- 3. 同步数据用于离线 OLAP 场景
- 4. 构建二级查询索引

- 1. **自动建表**，自动识别分表键，根据场景映射到目标表结构，解决大多数数据冲突场景
- 2. **动态参数**，出现延迟等问题可以及时感知，调整参数提升性能
- 3. **数据校验**，一键创建校验链路，定时执行校验逻辑（Doing）



小红书

目录

CONTENTS

① 背景发展

② 平台建设

③ 应用案例

④ 未来规划

DTCC 2023

第十四届中国数据库技术大会
DATABASE TECHNOLOGY CONFERENCE CHINA 2023

小红书

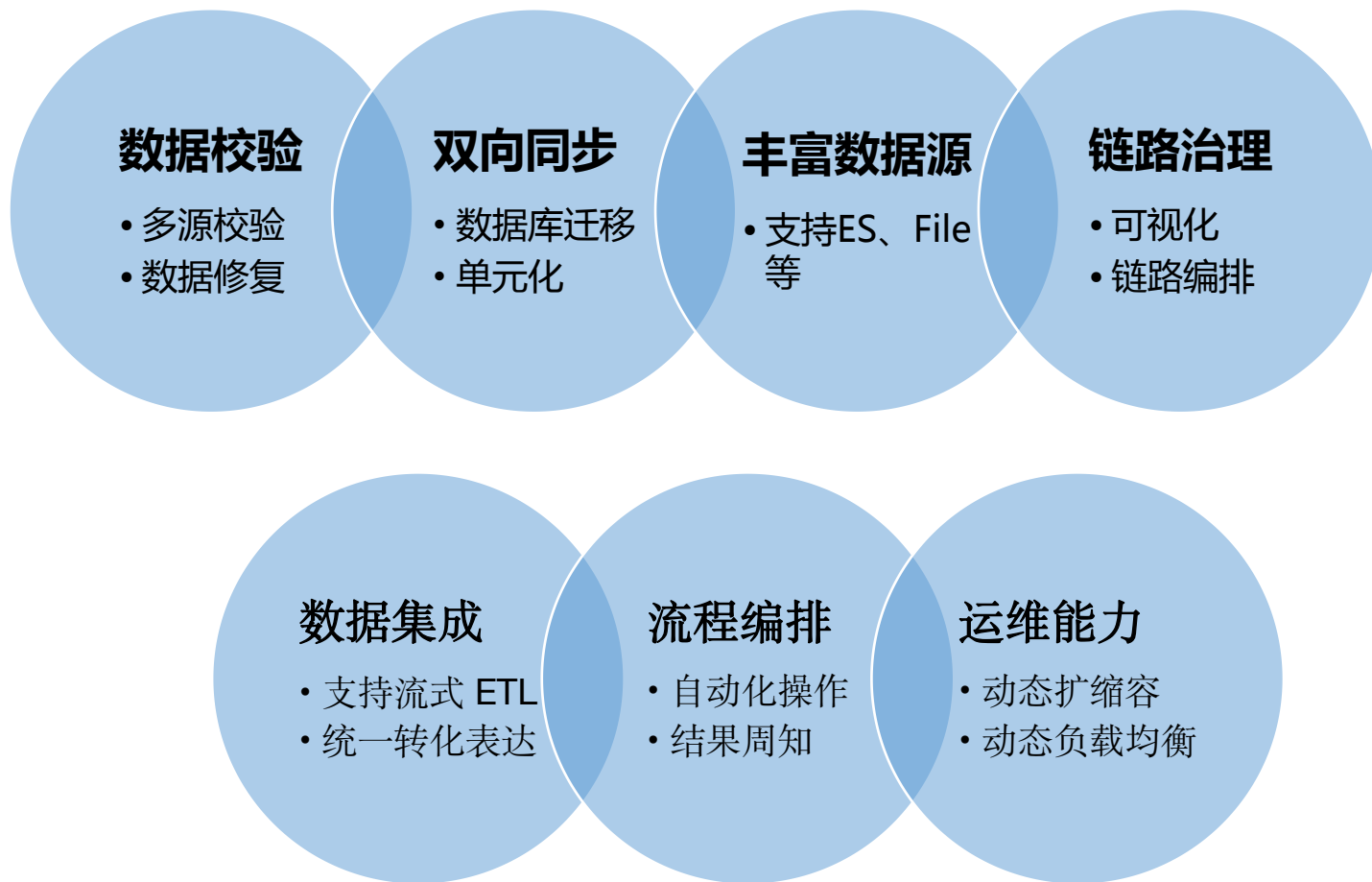
DTCC

数智赋能 共筑未来

IT168.com

ChinaUnix.net

ITPUB



小红书

THANKS



欢迎关注
小红书技术 REDtech