



让数据串流不息

阿里数据传输服务揭秘

程实 (时勤) @ 阿里巴巴
2015-12-19

Geekbang >

极客邦科技

整合全球最优质学习资源, 帮助技术人员和企业成长
Growing Technicians, Growing Companies

InfoQ^{ueue}

专注中高端技术人员的技术媒体



EGO EXTRA GEEKS' ORGANIZATION NETWORKS

高端技术人员学习型社交网络



StuQ^{ueue}

实践驱动的IT职业学习和服务平台



GiT GEEKBANG INTERNATIONAL TRAINING 极客邦培训 东京

一线专家驱动的企业培训服务



旧金山 伦敦 北京 圣保罗 东京 纽约 上海
San Francisco London Beijing Sao Paulo Tokyo New York Shanghai

QCon

全球软件开发大会

2016年4月21-23日 | 北京·国际会议中心

主办方 **Geekbang** & **InfoQ**
极客邦科技

7折 优惠 (截至12月27日)
现在报名, 节省2040元/张, 团购享受更多优惠

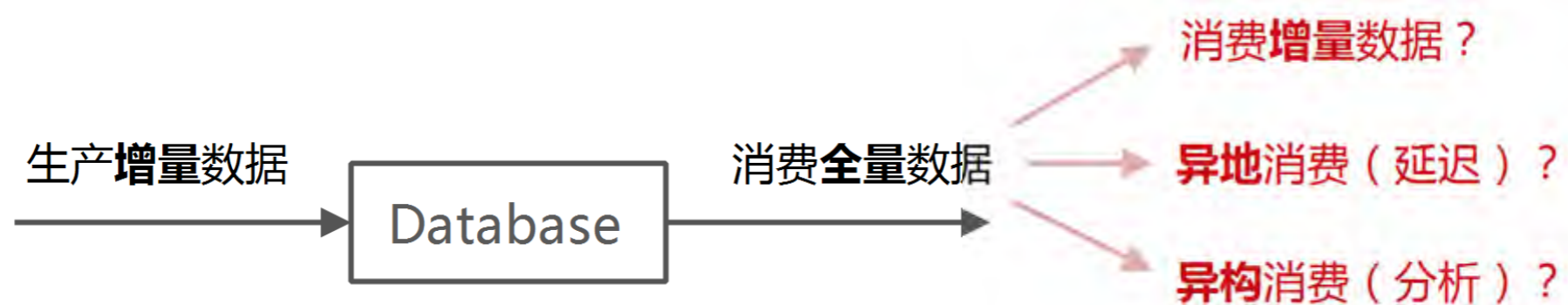
www.qconbeijing.com



扫描获取更多大会信息

DT时代——利用数据

- **生产数据不再是难题**
- **利用数据产生价值，才是关键**
 - 实时分析（实时商业决策、搜索、安全审计）
 - 典型案例：数据从DB实时同步到Storm，要求低延时
 - 离线分析（大数据分析）
 - 典型案例：数据从DB周期性导入到Hadoop，要求高吞吐率
 - 异地容灾，异地读写
 - 典型案例：云数据库多地容灾，本地读写
- **传统数据服务架构，对于利用数据并不友好**





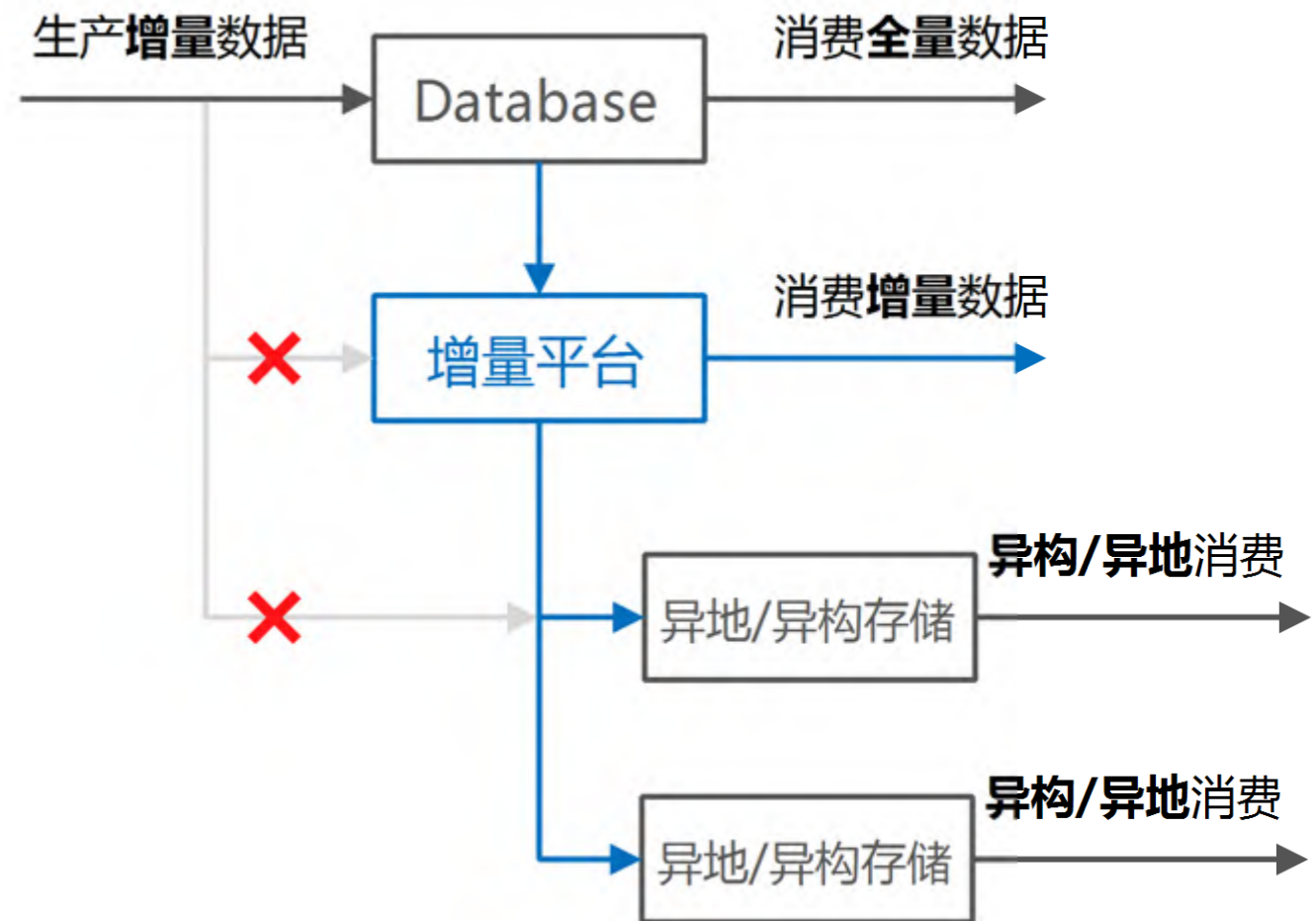
DT时代——增量数据服务化

• 业界已有的一些方案

- 统一的OLTP和OLAP
 - Databases on Hadoop
 - SAP Hana
- 分布式DB: 异地问题
- 迁移工具
- 业务同时写入到多个DB/平台

• 增量数据必须服务化

- 降低DB查询压力
- 降低业务方开发难度和耦合性
- 多种DB统一服务
- 多种下游统一服务
- 支持异地多活架构

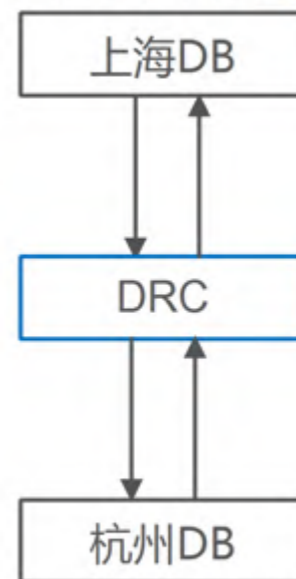


DT时代——阿里内部案例

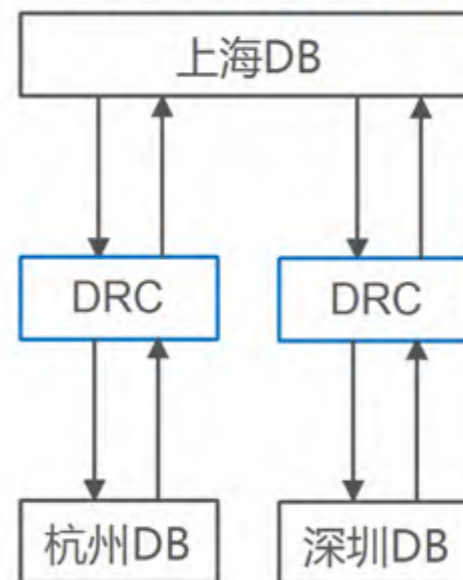
• 淘宝核心交易库

- 双十一数据量
 - 全天同步总量：百TB级别（全网6倍）
 - 高峰增量流量：Gbps级别
- 增量消费的下游
 - 5个实时下游业务
 - 实时媒体大屏
 - 实时商业分析
 - 实时搜索
 - 实时备份
 - 1个离线分析平台
 - 含多个子任务
 - 2个异地实时备份链路（异地多活）
 - 上海到杭州
 - 上海到深圳

2014年阿里异地双活架构



2015年阿里异地多活架构





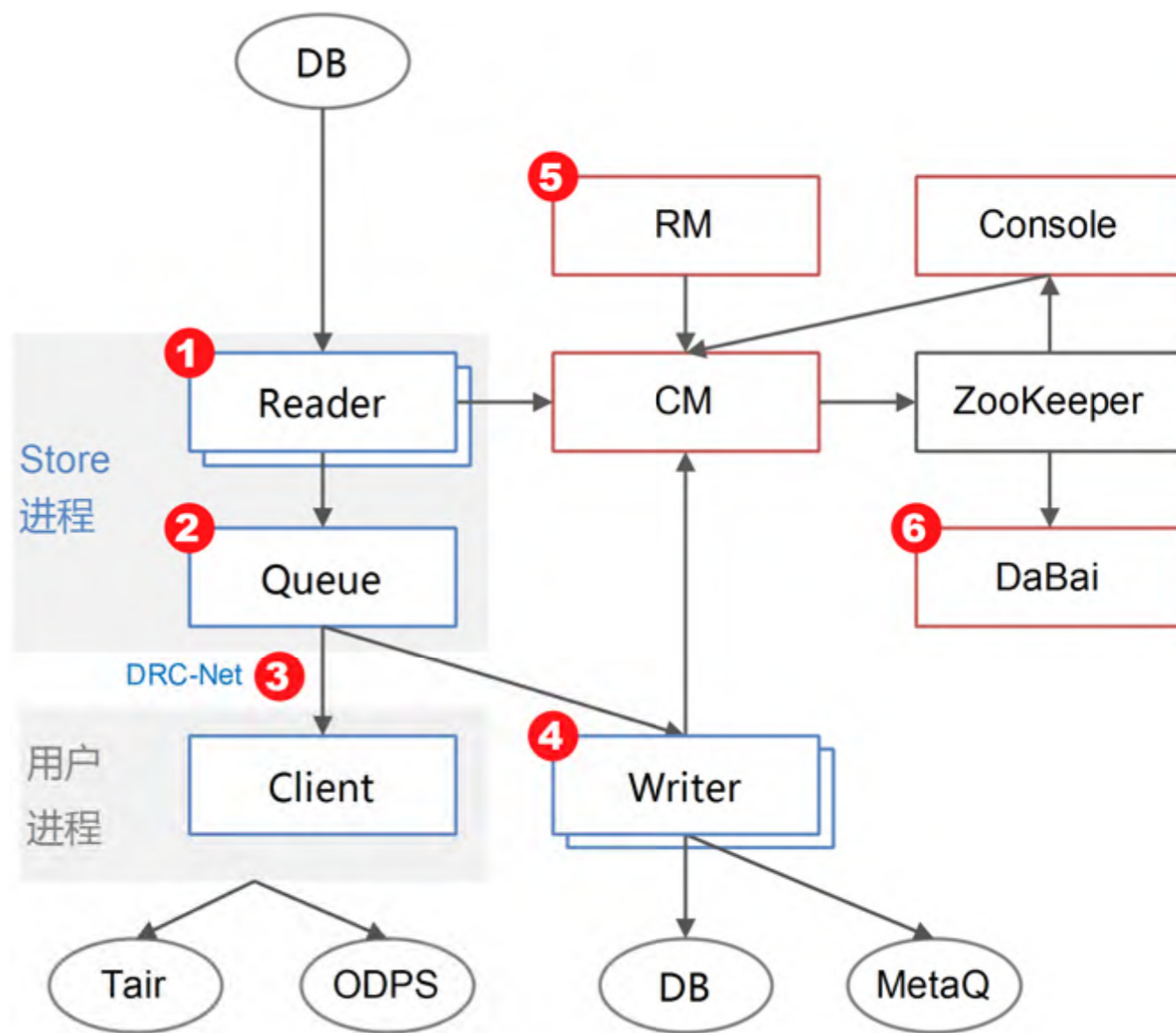
DRC和DTS

- **DRC——增量同步基础设施**
 - 增量抓取和存储
 - 增量同步和订阅
 - DRC在阿里广泛使用
- **DTS——云上传输数据服务**
 - 上云数据迁移
 - 云内外混合部署
 - 跨地域部署
 - 订阅服务

DRC模块构成

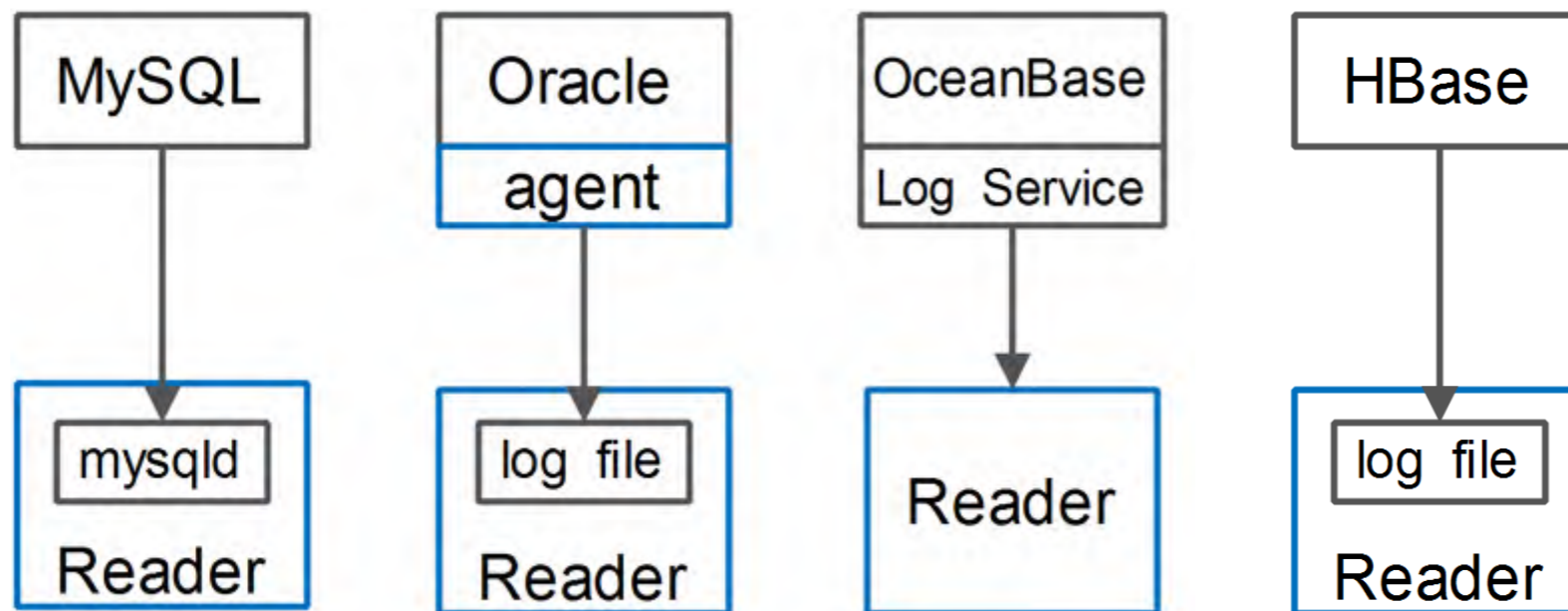
• 模块构成

- Reader: 抓取日志
- Queue: 存储
- Client: 订阅SDK
- Writer: 同步到DB
- CM: 集群管理
- RM: 自动化接入
- Console: 运维控制台
- DaBai: 全链路监控大屏



DRC抓取日志 (Reader)

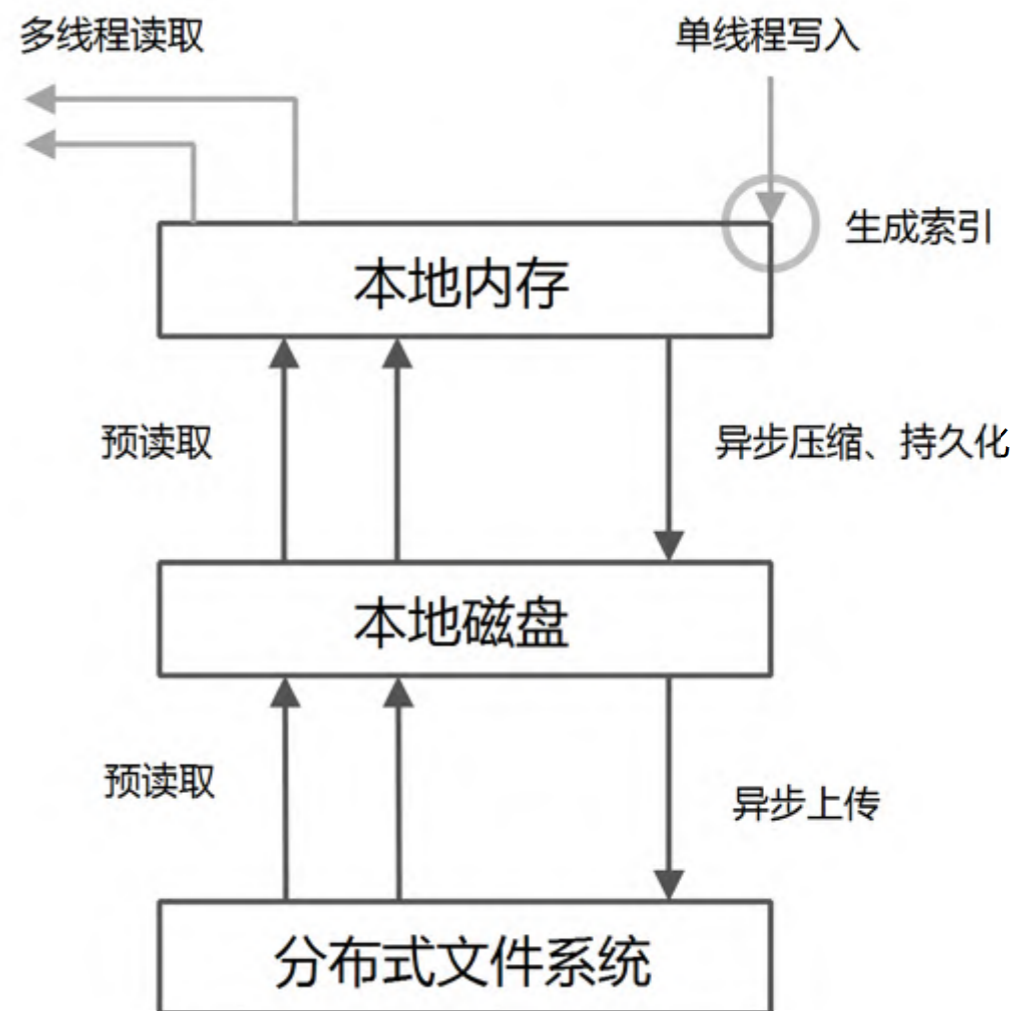
- 不同类型DB采用不同Reader设计
 - MySQL (5.1+): 基于mysqld主备同步机制
 - Oracle: 解析AchieveLog (非实时) 和RedoLog (实时)
 - OceanBase: 提供日志服务
 - Hbase: 解析HLog



DRC的队列存储 (Queue)

• DRC Queue队列存储

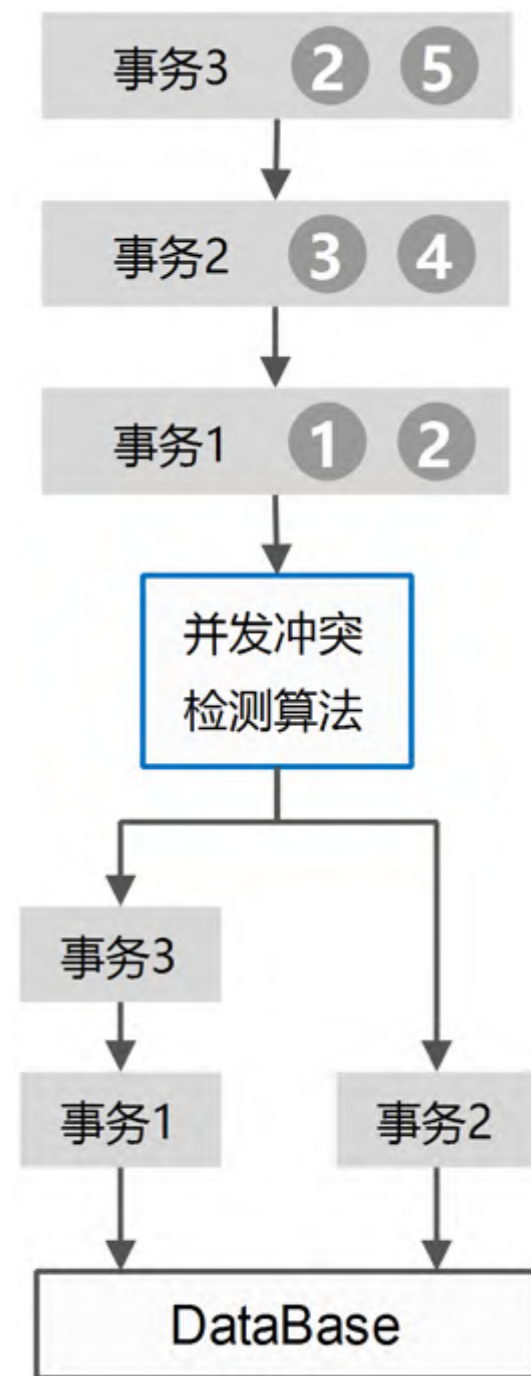
- 高性能
 - 实时读写28万rps，读历史90万rps
 - 支持1K连接并发读
- 压缩持久化
 - 压缩率30%
- 缓存和预读
- 位点/时间索引
 - 快速定位起始位置
- 自定义(库表名)索引
 - 大幅降低过滤开销
- 主备模式
 - 增加备节点的额外开销低于25%
- 分布式容灾



DRC同步增量 (Writer)

• Writer组件

- 高性能
 - 低冲突场景10万rps
- 支持多种数据库
 - JDBC模式: RDBMS
 - K/V模式: NoSQL、MQ
- 并发问题
 - 性能需要: 串行数据流->并发事务
 - 实时检测事务之间是否冲突
 - 数据库同步API带来的问题
- 防数据回流
 - 多向同步时, 避免重复写入
- 支持DDL同步





DRC网络传输

- Net组件

- 实时压缩

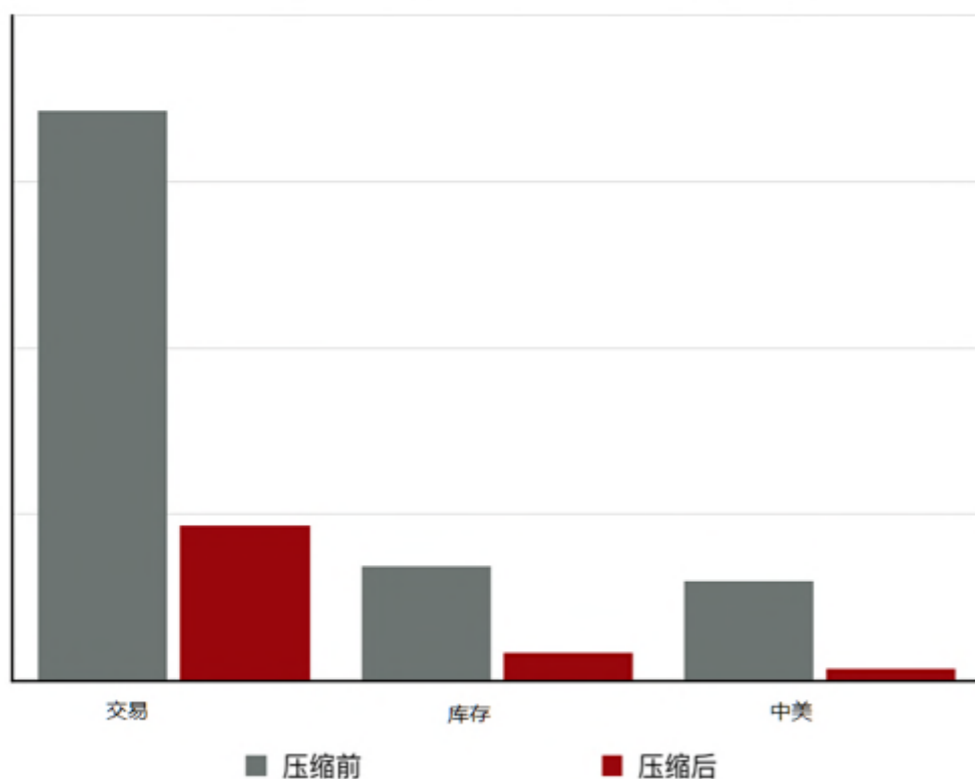
- 解决骨干网、主机网卡资源瓶颈问题
 - 压缩后延时增加不高于5ms，CPU增加不高于单核100% (占满2G带宽时)

- 多TCP协同传输

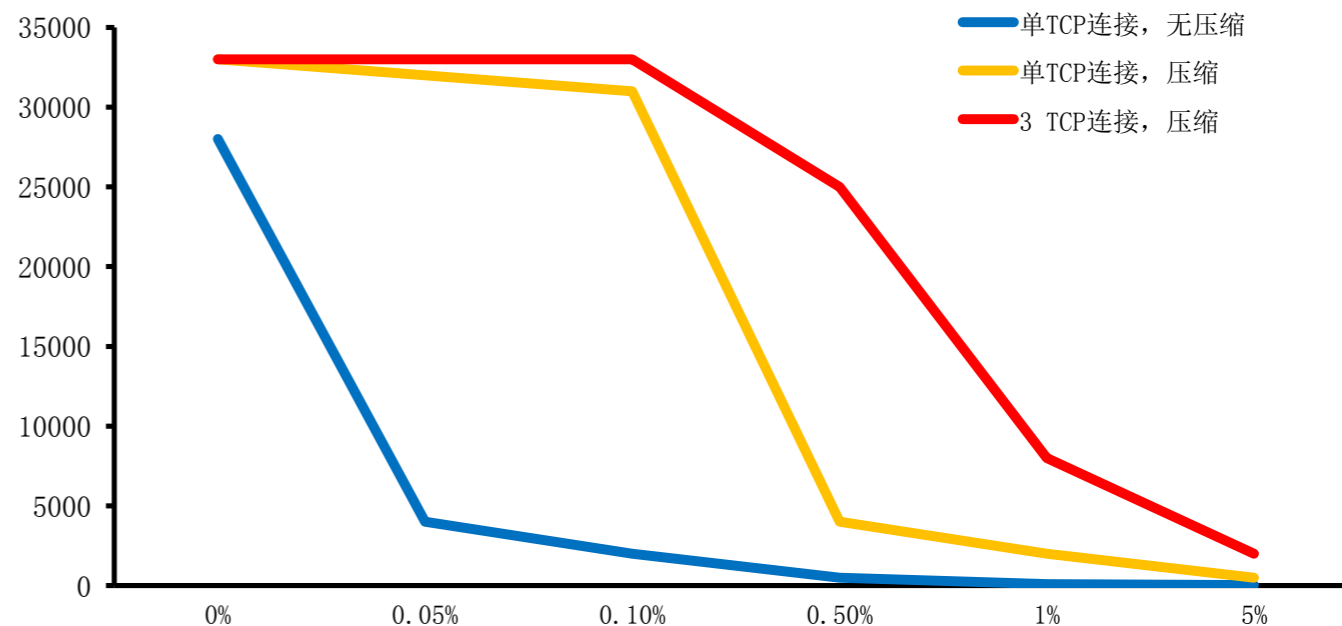
- 降低丢包对TCP传输带宽的影响

- SSL加密

DRC带宽优化及压缩组件DRCNet带宽压缩效果

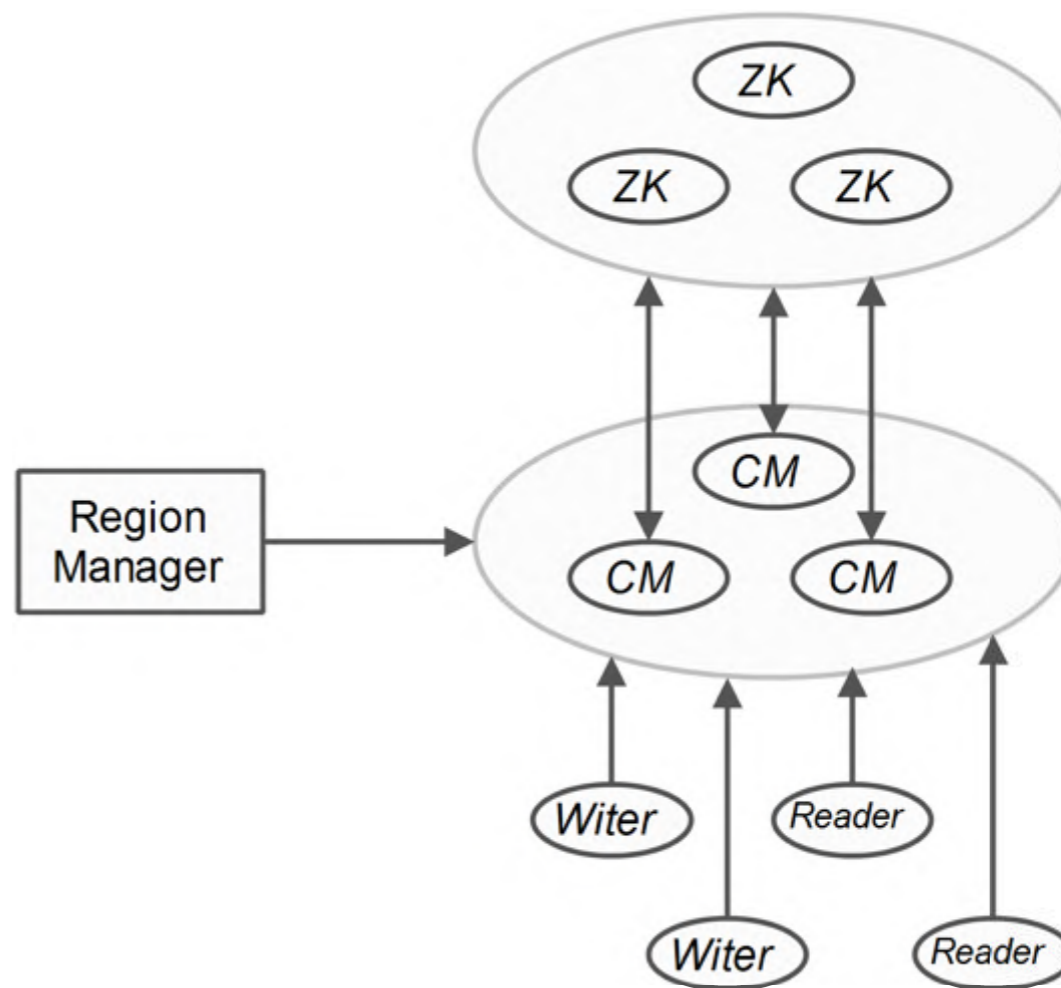


中美环境网络测试 (RT=121ms)



DRC集群管理

- Cluster Manager
 - 跨机房多节点部署
 - 无状态
- ZooKeeper
 - 跨机房高可用
 - 云信息存储
 - 任务元信息
 - 用户、进程、主机状态
- Region Manager
 - 自动化接入
 - Web控制台+API
 - 接入任务已超过1万，成功率90%+
 - 跨集群的调度



从DRC到DTS

- 云上用户更加需要数据传输服务
 - 一次性迁移上云
 - 全量+增量同步
 - 部分上云（云内外DB混合部署）
 - 技术过度平滑
 - 利用现有的IT资产
 - 数据可靠性
 - 订阅云上DB的增量
 - 支持在线查看订阅数据
 - 异地容灾
 - 异地读写
 - 迁移评估
 - 帮助用户决定合适容量的RDS服务

DTS迁移服务

• 用户迁移到RDS

- 跨数据库类型
 - 兼容不同类型数据库之间的规则差异
- 多种迁移方式
 - 支持在线连接数据库，或离线文件复制
- 全量+增量
 - 全量迁移期间新写入的数据不丢失
- 高级特性
 - 支持用户自定义过滤表名、字段、行

数据库类型	Schema	全量	增量
Oracle->MySQL	✓	✓	
MySQL->MySQL	✓	✓	✓
SQL Server->SQL Server	✓	✓	
Oracle->PGSQL	✓	✓	

	DTS	GoldenGate	SSIS
产品形态	云服务	商用软件(Oracle)	商用软件(微软)
系统架构	高可用集群	单链路，无容灾机制	单链路，无容灾机制
功能	迁移、订阅、同步	迁移、同步	迁移、数据分析
数据源	MySQL、SQLServer、Oracle、PostgreSQL、PPAS、OceanBase	MySQL、SQLServer、Oracle、DB2、Sybase、teradata、timesten(只能为目的端)	源：MySQL、Oracle、Sybase、access、DB2 目的：SQLServer、SQL Azure
易用性	WEB服务	命令	客户端软件
运维体系	性能展示、监控	无	无



DTS迁移服务Web控制台

DTS

数据传输服务DTS

数据迁移

上云评估

迁移任务列表 (共 492 个, 其中运行中 0 个, 异常 103 个, 暂停 1 个, 未启动 187 个, 完成 201 个)

搜索 排序: 默认排序

<input type="checkbox"/>	名称: [redacted]	状态: 完成	查看任务配置 创建类似任务 查看迁移详情
2015-08-31 12:03:57 创建		2015-08-31 12:05:52 完成	
结构迁移 100%		全量数据迁移 100% (已迁移1行)	一致性校验 100%
<input type="checkbox"/>	名称: [redacted]	状态: 完成	查看任务配置 创建类似任务 查看迁移详情
2015-08-12 23:55:33 创建		2015-08-12 23:57:37 完成	
结构迁移 100%		全量数据迁移 100% (已迁移80000行)	一致性校验 100%
<input type="checkbox"/>	名称: [redacted]	状态: 完成	查看任务配置 创建类似任务 查看迁移详情
2015-08-12 00:24:34 创建		2015-08-12 00:25:09 完成	
一致性校验 100%			
<input type="checkbox"/>	名称: 汇报1	状态: 完成	查看任务配置 创建类似任务 查看迁移详情

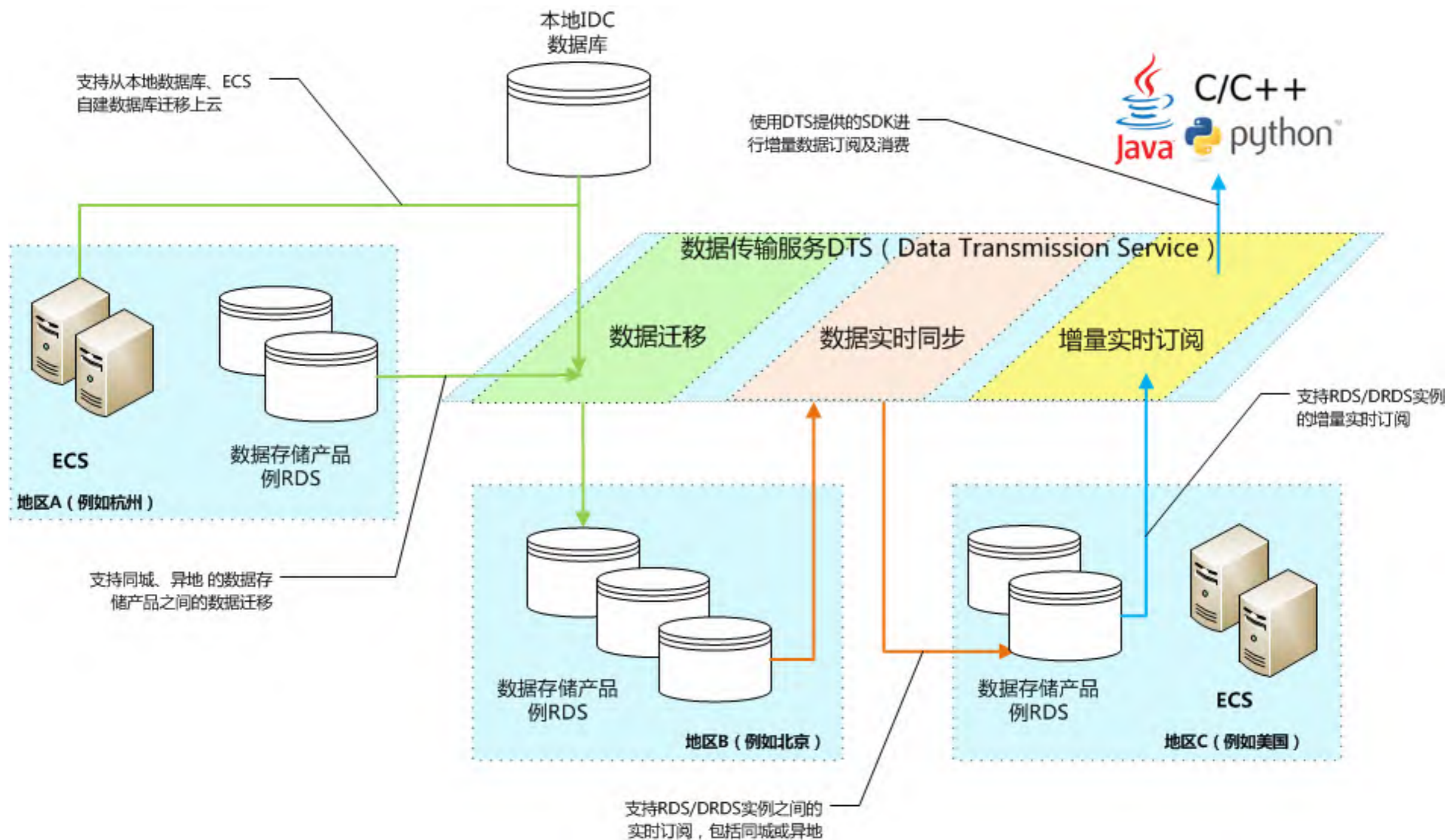
预检查 启动 暂停 结束 删除

共有201条, 每页显示: 20条

« « 1 2 3 » » GO



DTS实时同步服务



Thank you

Thank you



 @阿里技术保障



 @阿里技术保障