京东弹性数据库

基础架构部-吕信



目录

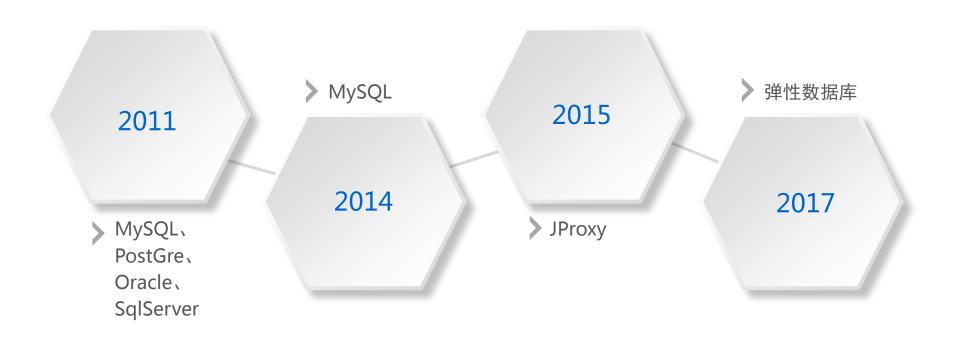


- 1 发展历程
- 2 功能特性
- 3 整体架构
- 4 实现细节
- 5 使用情况

Part

发展历程

诸侯混战->MySQL->Jproxy->弹性数据库





功能特性

京东弹性数据库是一个产品系列,包括:JED、BinLake和DBS,解决了数据库管理和使用过程中的动态伸缩、高可用、数据路由、日志采集和自动化运维等一系列问题



- 1 在线动态扩容
- 自动Fail0ver
- 兼容MySQL协议
 - 多源数据迁移
 - 数据库操作审计

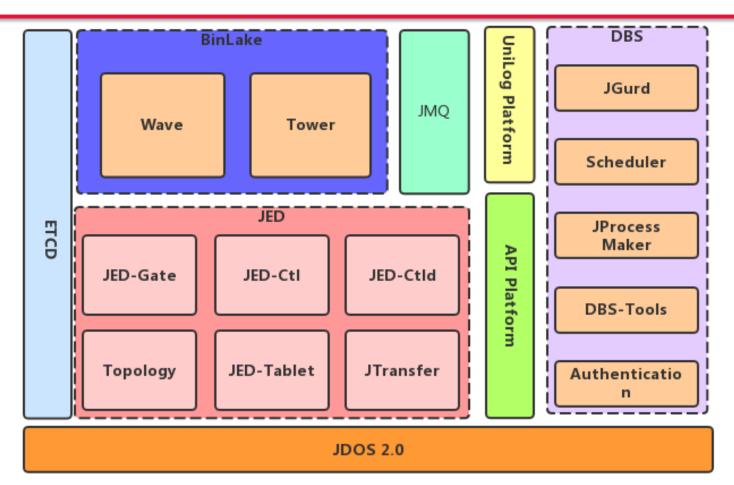
- 1 无状态集群化BinLog采集
- 高可用与自动故障转移
- 5 负载自动均衡
- 女持多种MQ
- 支持集群横向扩容

- 数据库服务自助化交付
- 2 数据库操作流程化管理
- 数据库健康度全面监控
- 4. 数据库自助备份与结转
- 作业任务多样化调度



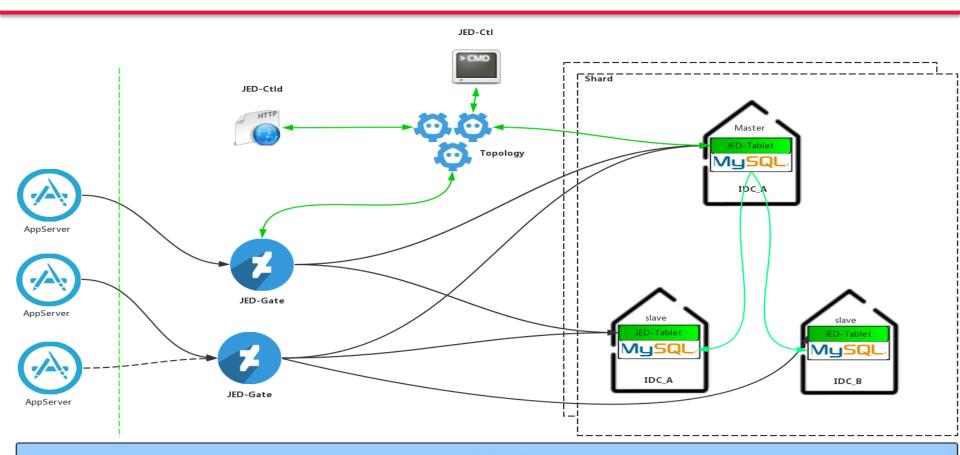
整体架构

弹性数据库各个产品模块的设计及相互协作



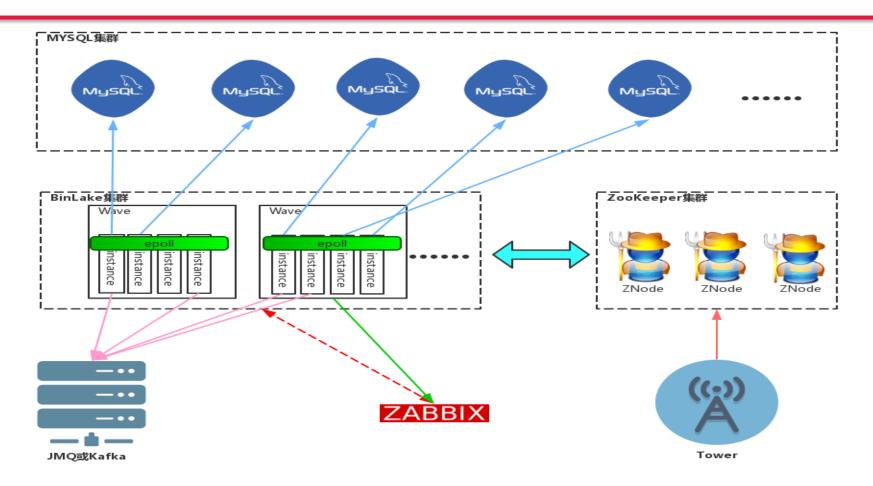
JED

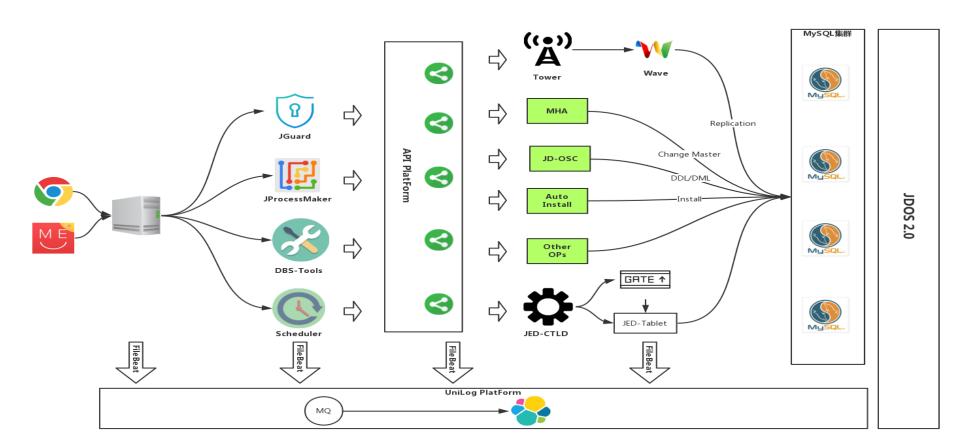
JD.COM 京东



JDOS 2.0

BinLake







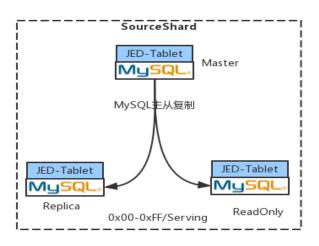
实现细节

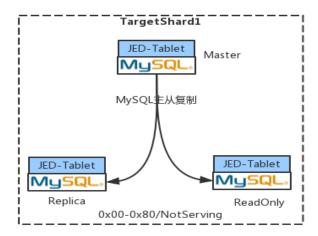
动态Resharding、流式排序、集群化BinLog采集与订 阅的实现细节

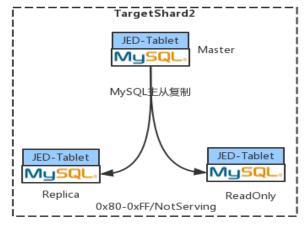


动态在线扩容

Step1:创建两个目标Shard







ReadOnly

0x80-0xFF/NotServing

My5Q Replica

动态在线扩容

Step2:全量数据过滤克隆 TargetShard1 (1)复制TableSchema信息 (2)将ReadOnly从Master复制关系摘下来,然后过滤copy全量静态数据,两种方式:两 JED-Tablet 次完成和一次完成 Master My5QL The state of the s MySQL主从复制 SourceShard JED-Tablet JED-Tablet JED-Tablet Master Mysc My5Q My5Ql Replica ReadOnly 0x00-0x80/NotServing MySQL主从复制 过滤Copy[0x80-0xF的数据] TargetShard2 JED-Worker JED-Tablet JED-Tablet JED-Tablet Master My5Ql MySQl Musq drain Replica MySQL主从复制 0x00-0xFF/Serving JED-Tablet JED-Tablet

动态在线扩容

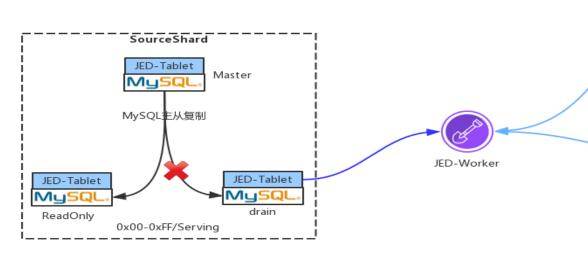
Step3: 增量数据过滤复制到两个目标Shard (1) JED-Worker通过JED-Tablet API建立SourceTablet与TargetTablet之间的复制关系 TargetShard1 (2) SourceTablet针对于每个TargetShard创建协程,并根据相应KeyRange发送BinLog JED-Tablet Master (3)每个TargetTablet中的BinLogPlayer重放BinLog Musq William Dado od Oddin Balling 创建复制关系 MySQL₺从复制 SourceShard JED-Worker JED-Tablet JED-Tablet JED-Tablet Master My5Q My5Ql – Replica ReadOnly 0x00=0x80/NotServing MySQL主从复制 TargetShard2 过滤复制[0x80-0xFF的增量数据] JED-Tablet JED-Tablet JED-Tablet Master My5Q N_USQ Replica ReadOnly MySQL主从复制 0x00-0xFF/Serving JED-Tablet JED-Tablet My5Q MuSQL Replica ReadOnly 0x80-0xFF/NotServing

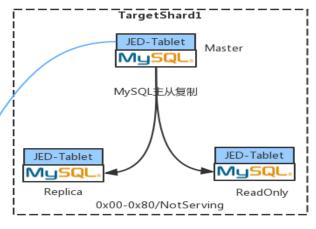
动态在线扩容

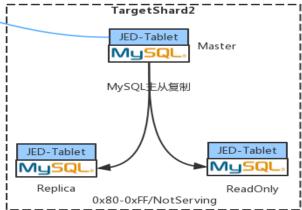
Step4:数据一致性校验

当Target与Source复制小于5S时,开始进行校验

- (1)将Replica从Master复制关系摘下来,待过滤复制无延时
- (2) JED-Worker执行逐行校验,校验通过之后,将Replica重新挂载到SourceMaster下





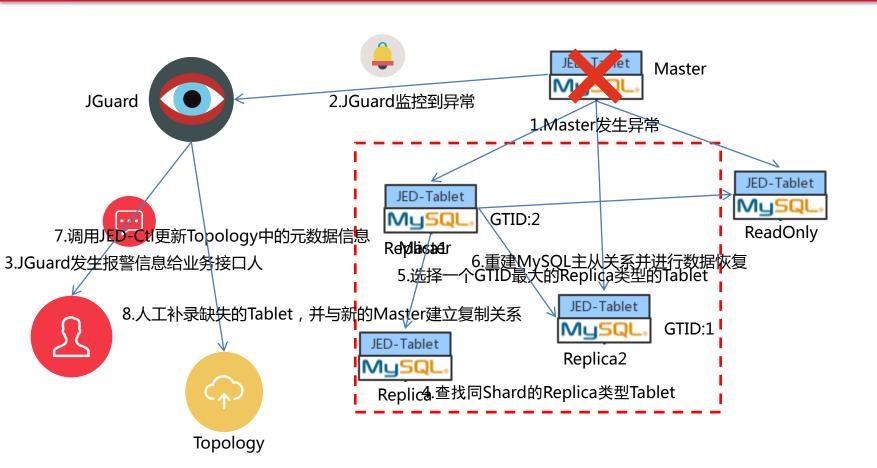


0x80-0xFF/Serving

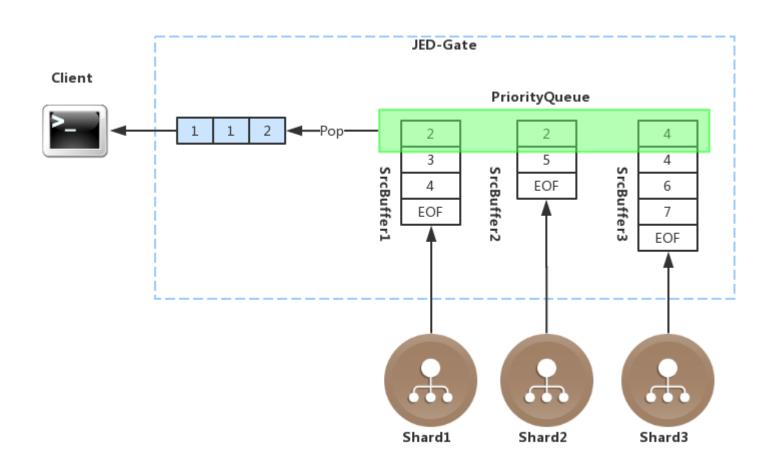
动态在线扩容

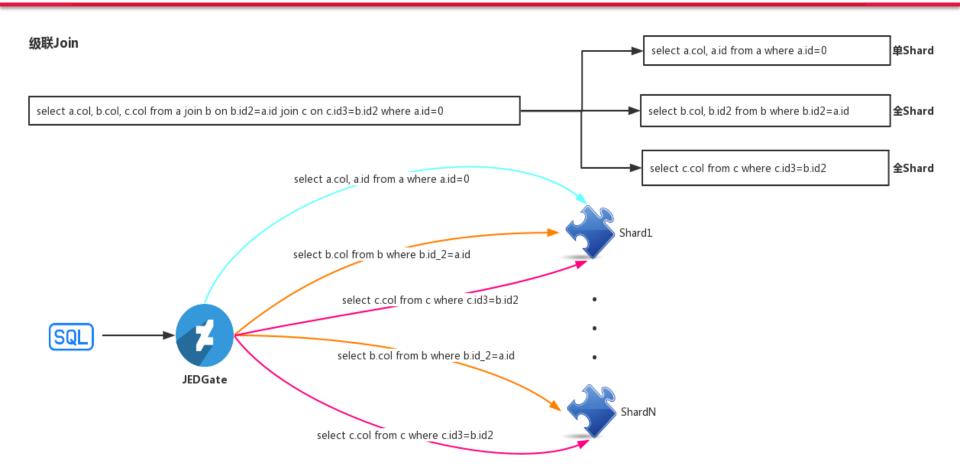
Step5:切Shard (1)保证Master正常写入的情况下,在保持复制关系的情况下,先切Slave Tablet至Target TargetShard1 (2) Source Master停写,等Target与Source binlog同步之后,停止过滤复制 JED-Tablet Master (3)将Target的状态改为Serving, Source状态改为NotServing,并停止Source服务 My5Q THE REAL PROPERTY OF SORTH BEATERS MySQL主从复制 SourceShard JED-Tablet JED-Tablet JED-Tablet JED-Worker Master My5Q MySQ Replica ReadOnly 0x00-0x80/Serving MySQL主从复制 TargetShard2 过滤复制[0x80-0xFF的增量数据] JED-Tablet JED-Tablet JED-Tablet Master My5Q My5Q Replica ReadOnly MySQL主从复制 0x00-0xFF/NotServing JED-Tablet JED-Tablet My5Q My5Q Replica ReadOnly

自动FailOver

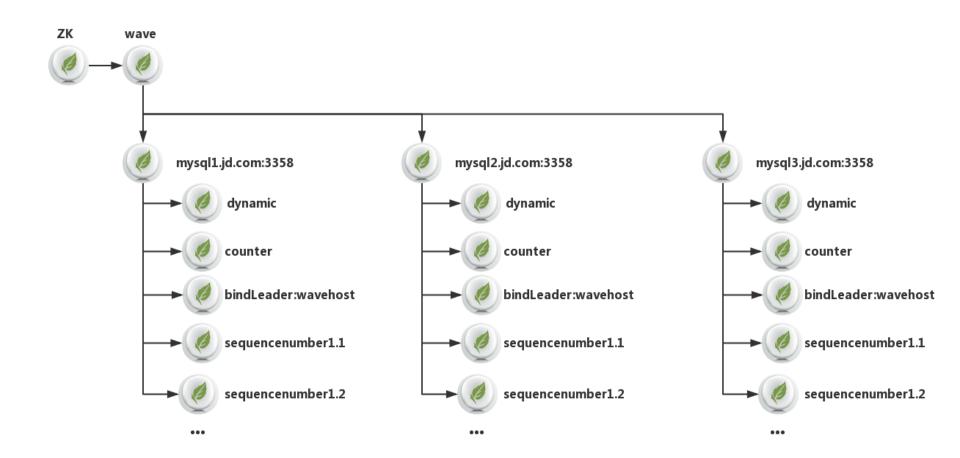


Streaming Process

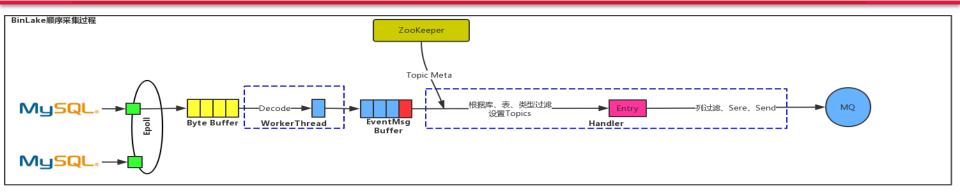


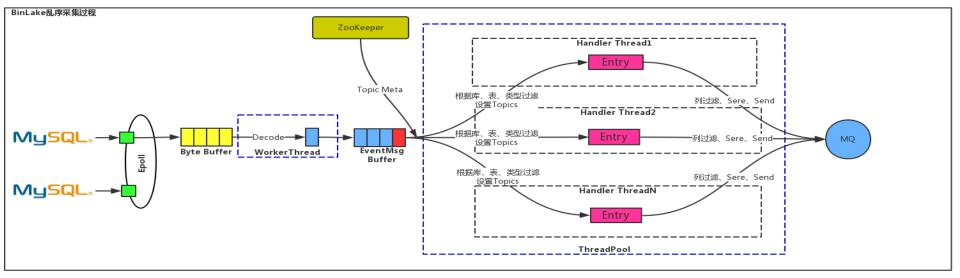


BinLake元数据架构图



集群化Binlog订阅

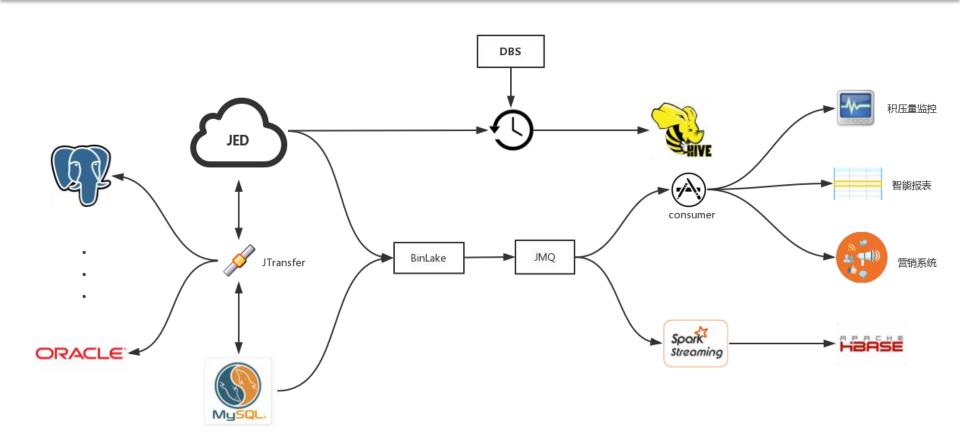


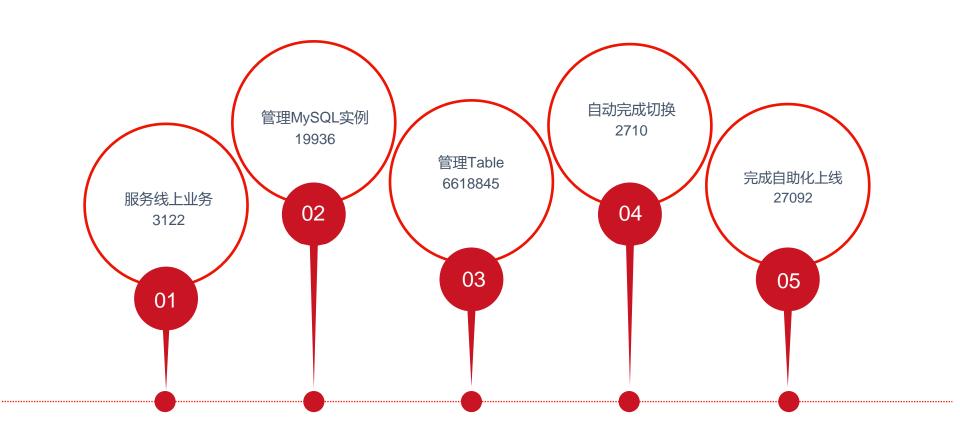


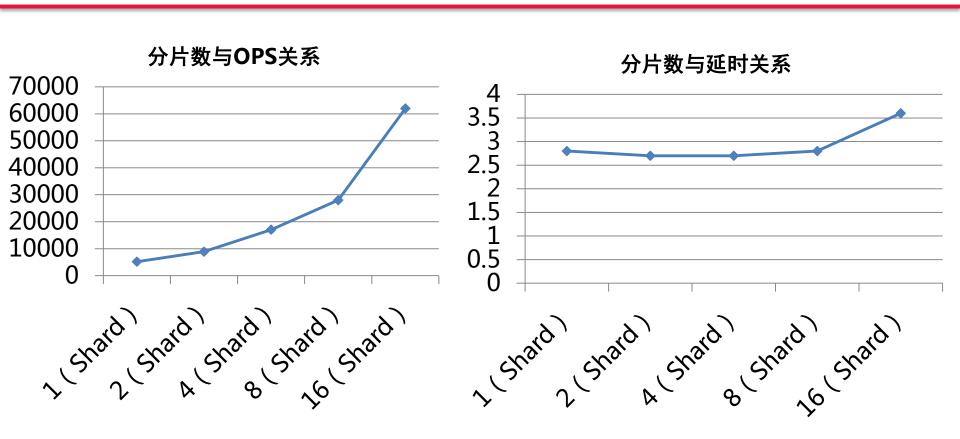


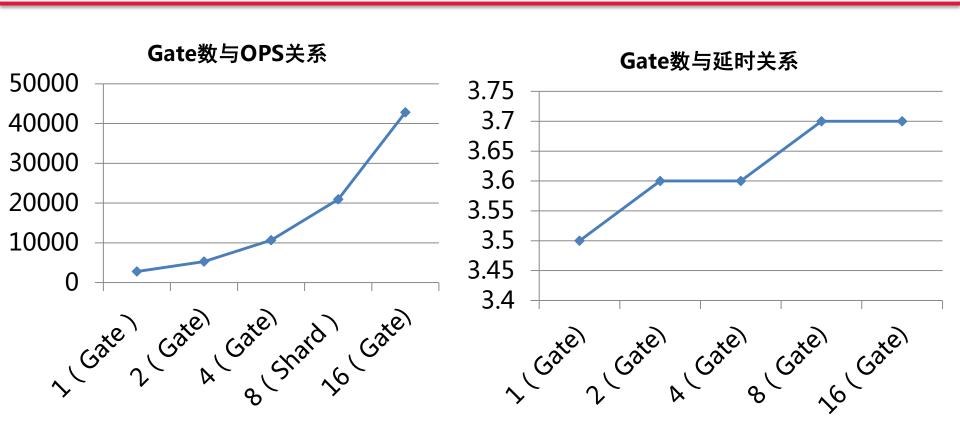
落地使用

弹性数据库在京东的集成使用方式、使用情况及性能 表现













团队简介

作为京东商城的一级部门与技术 基石,基础平台部专注于核心基础技术的自主研发与工程实施:数据中心 集群管理,电商中间件技术,分布式 存储与数据库系统,机器学习与知识 工程,商城整体架构提升。

基础平台持续运营多个数据中心、数万台服务器、一系列核心系统。京东商城总架构师刘海锋担任部门负责人。