



GOPS2016  
Beijing

# 全球运维大会

2016  
DevOps 2.0: 重塑运维价值



北京站

会议时间：12月16日 - 12月17日

会议地点：北京国际会议中心

主办单位：



# 建立基于Kafka的企业级数据传输平台

方伟 eBay中国研发中心 架构师

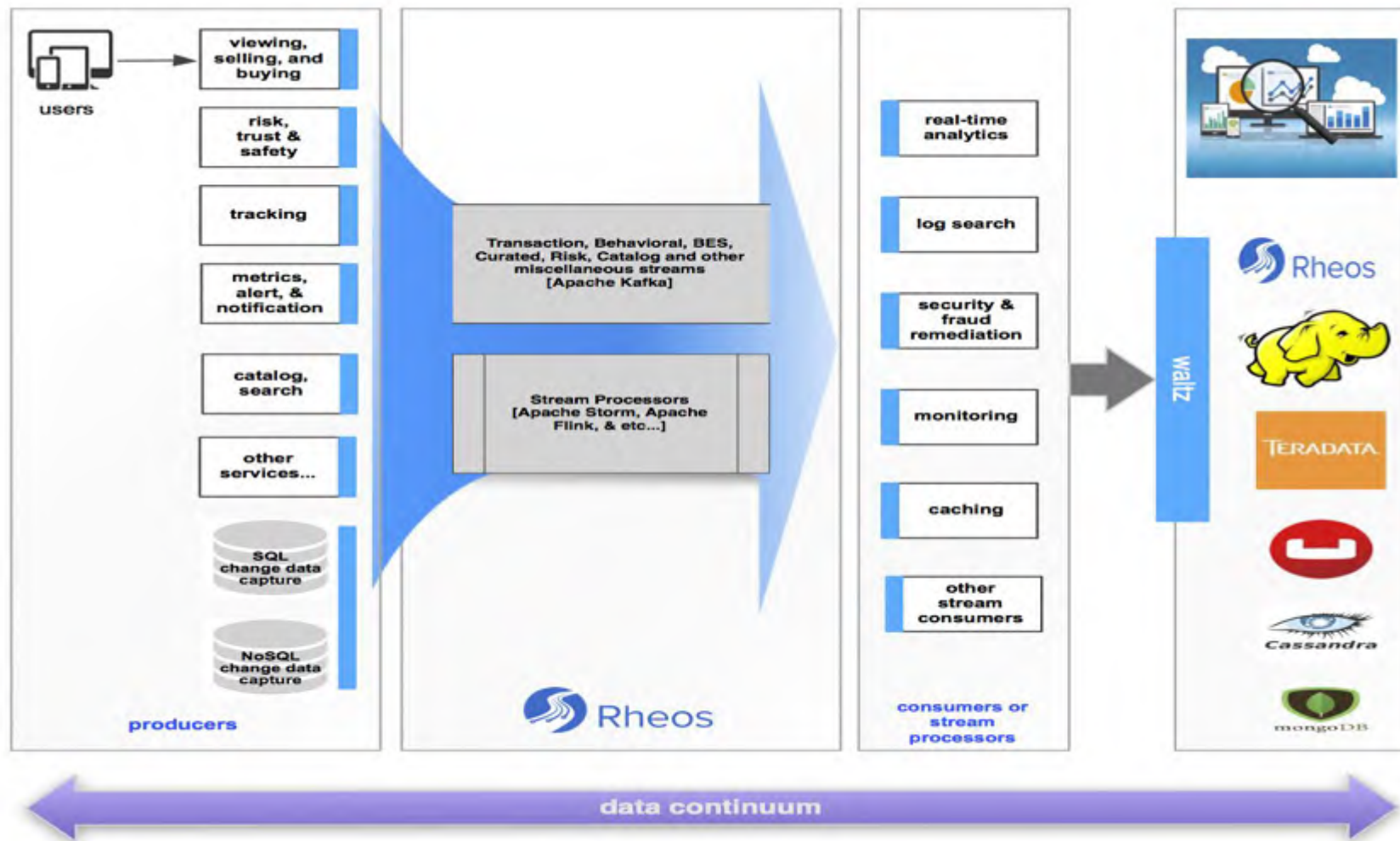


# 目录

- ➔ 1 eBay数据传输平台概述
- 2 平台核心服务
- 3 系统监控与自动化
- 4 Kafka性能优化

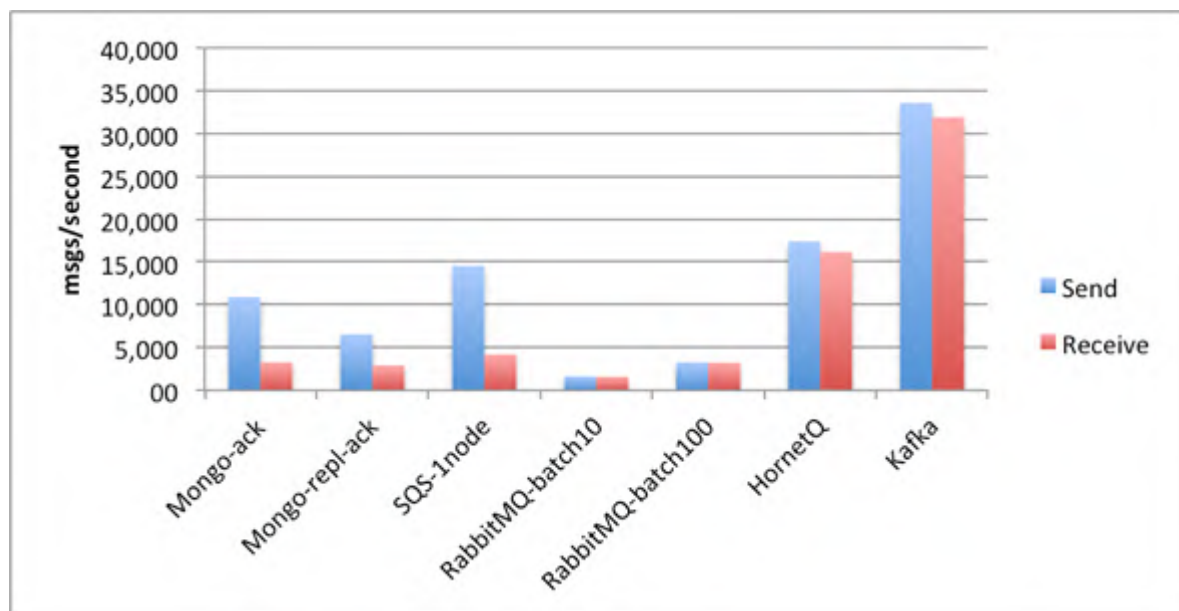


# eBay数据传输平台定位



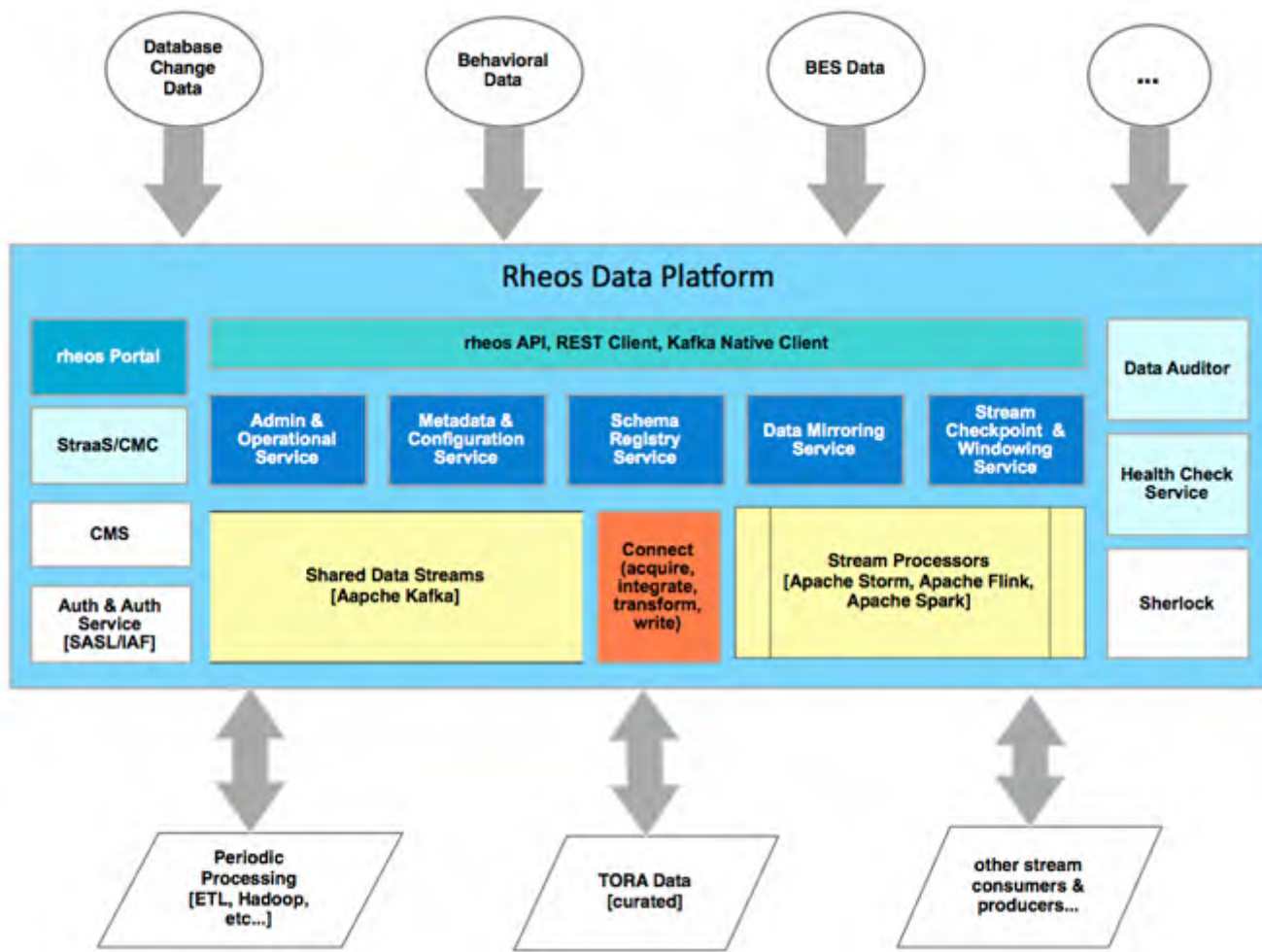
# 为什么要用Kafka

- 高吞吐率
- 高性能
- 多重订阅
- 消息持久性
- 良好的伸缩性
- 高可用性





# 平台架构设计



# 平台运行状况

- 30+ Kafka集群
- 800+ Kafka broker节点 (虚拟机)
- 1200+ Topic
- 25000+ Partition
- 1000亿+ 消息/天



用户行为



数据库事务



业务事件



应用共享



数据分析



系统监控



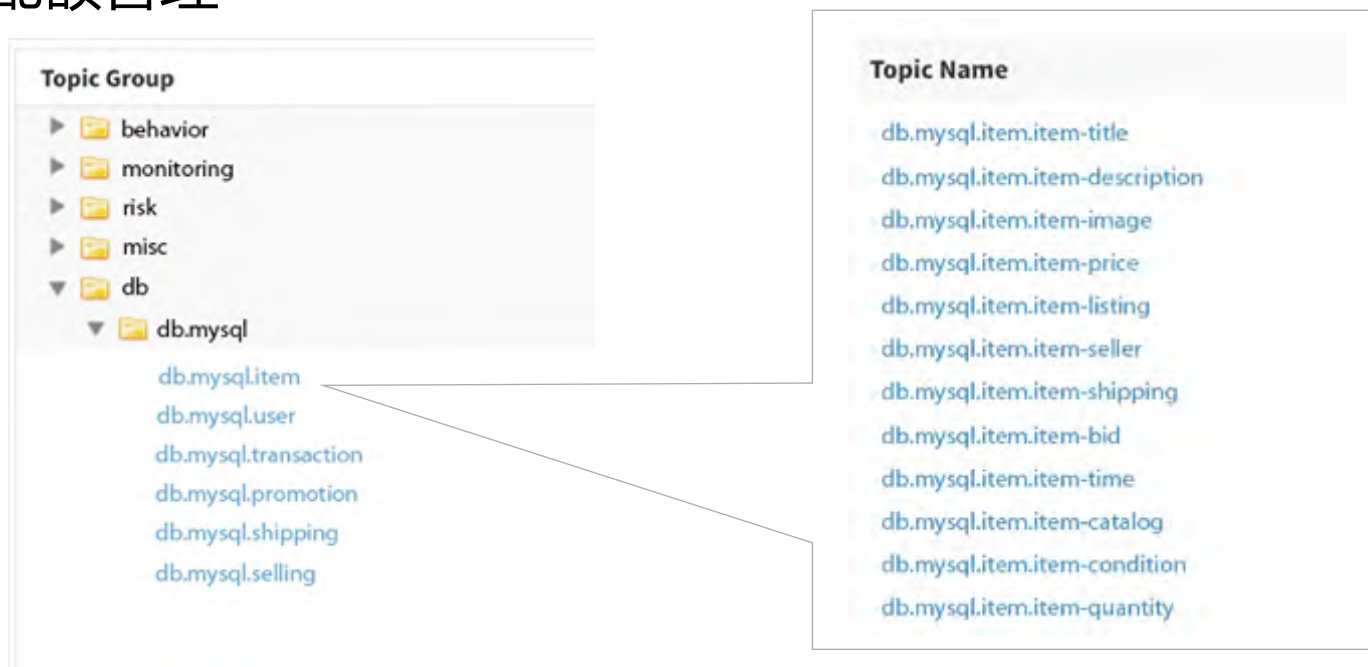
风险控制



结构化数据

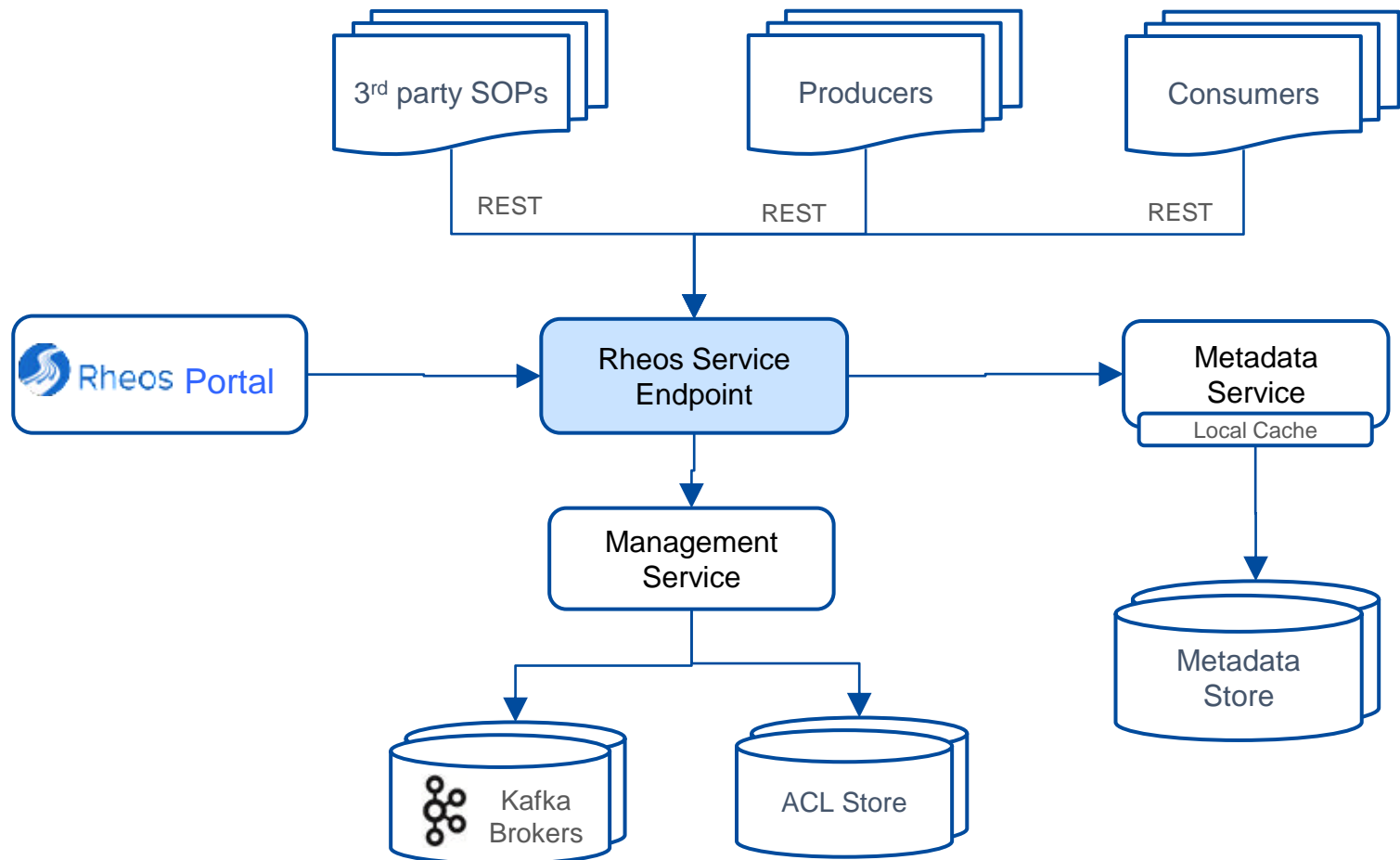
# 元数据服务的目的

- Topic组
- 统一名称空间
- 便于授权管理
- 便于配额管理



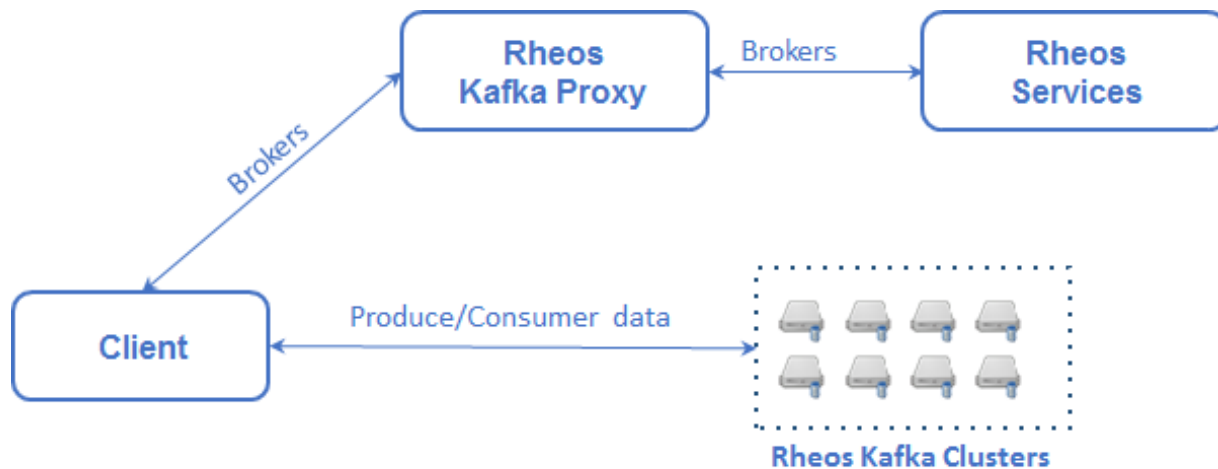


# 元数据服务

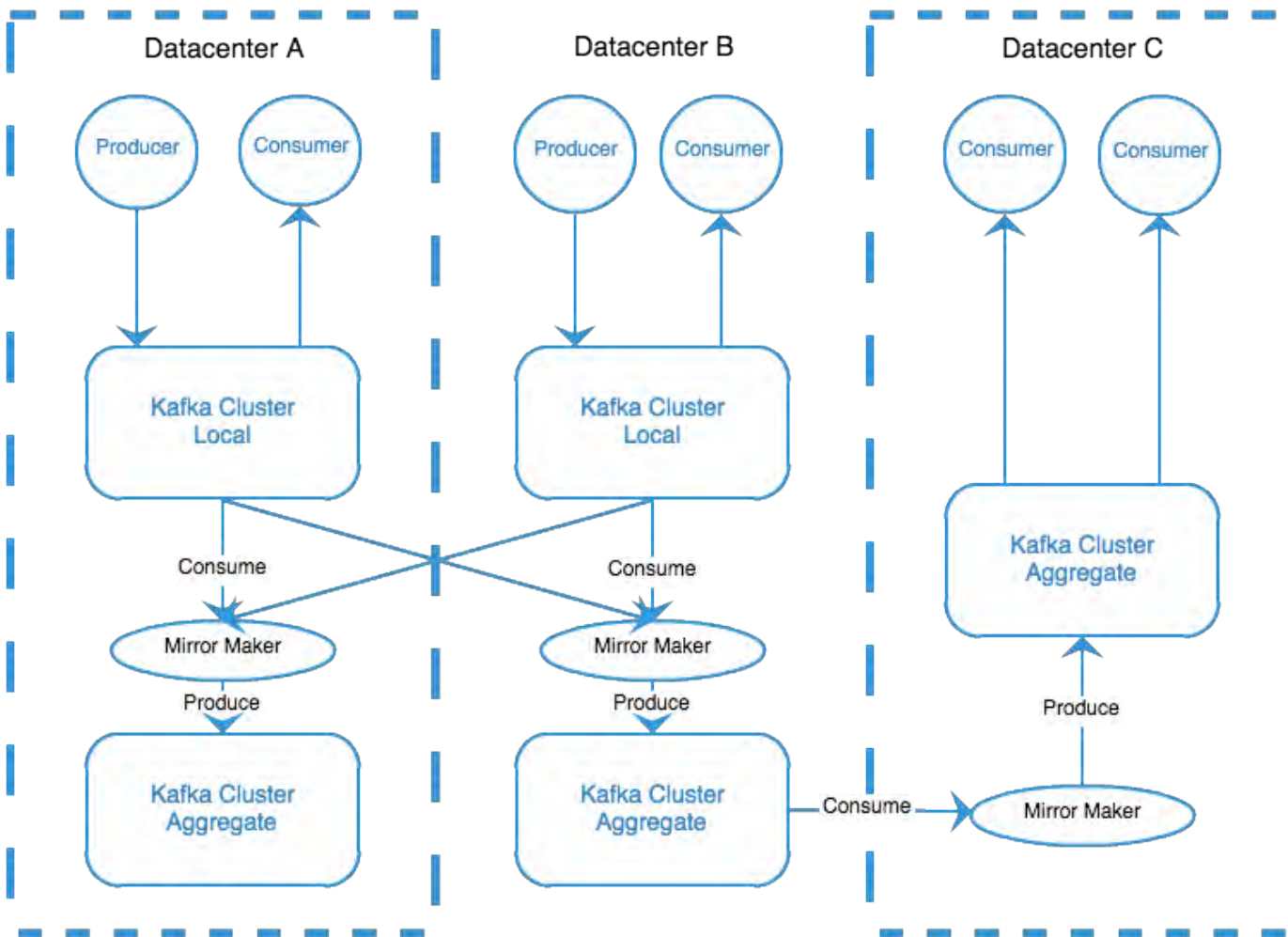


# Kafka代理

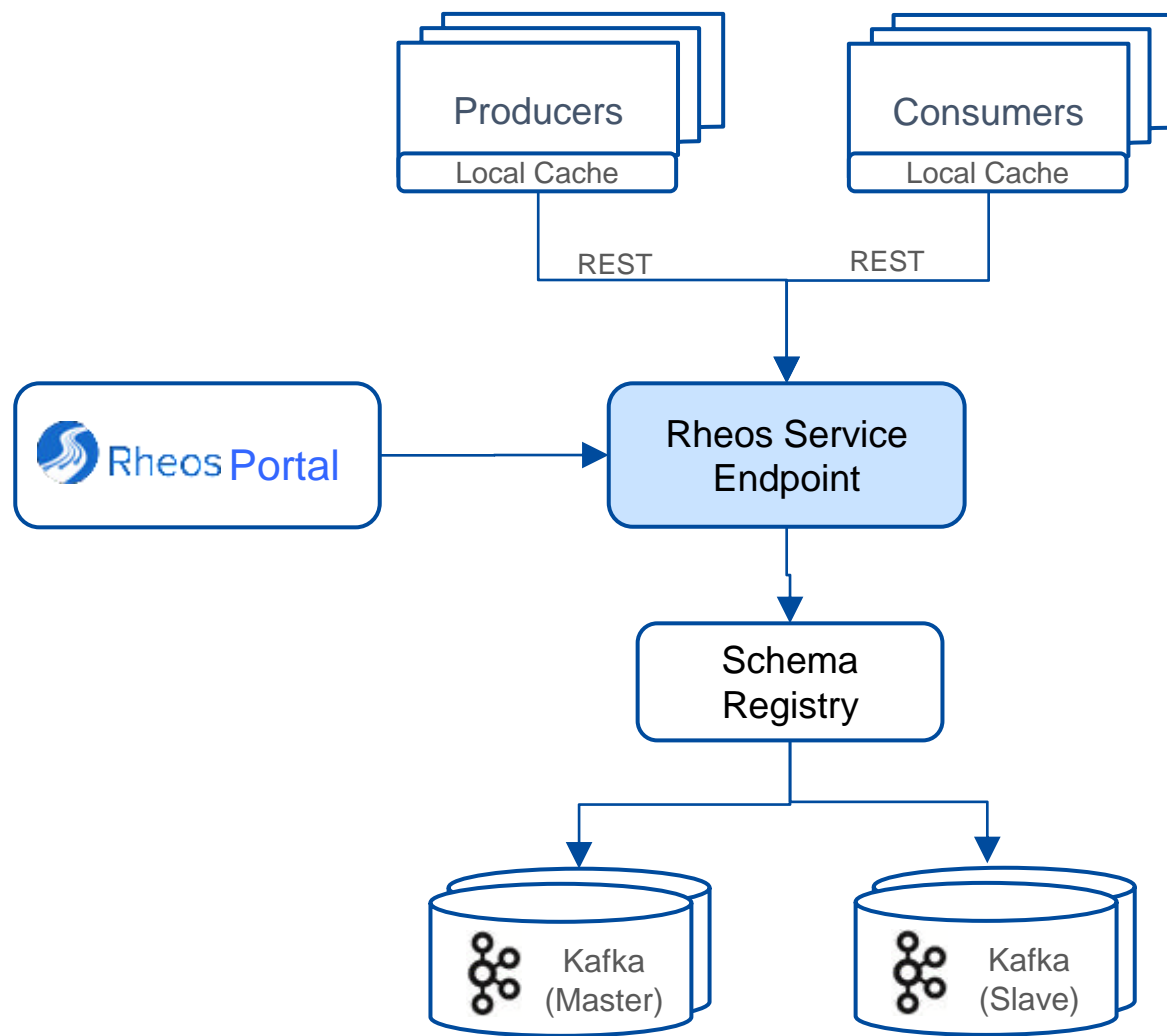
- 为客户端提供统一访问入口
- 与元数据服务整合
- 便于容灾备援



# Tier-Aggregation 模式和数据镜像服务

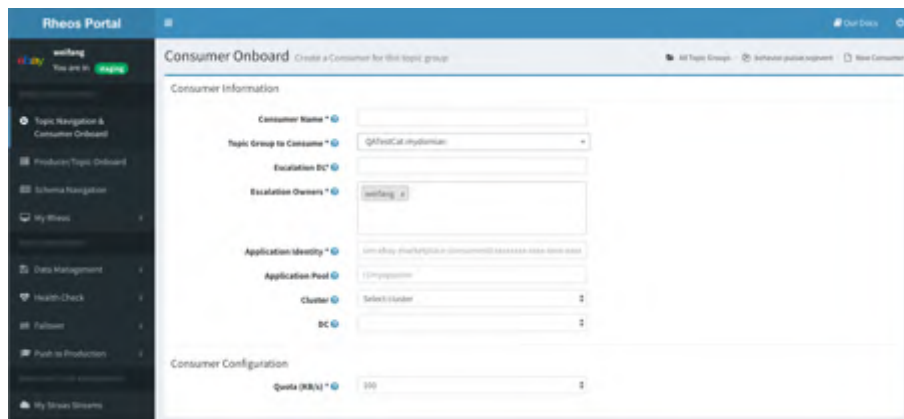


# Schema注册服务

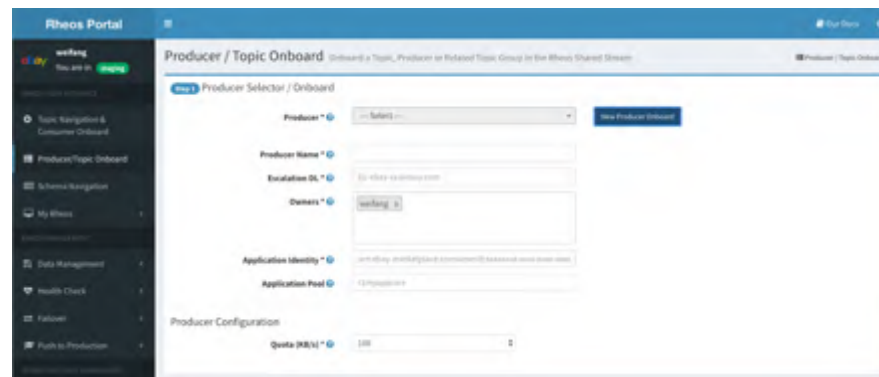


# 用户自服务

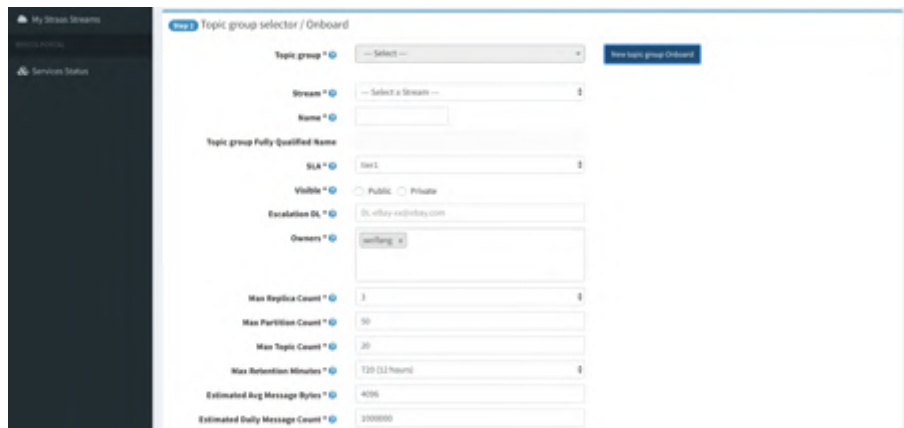
- Consumer 注册



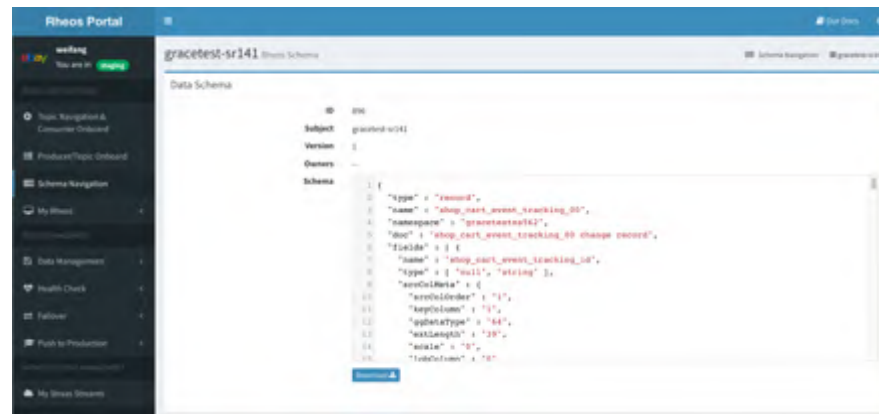
- Producer注册



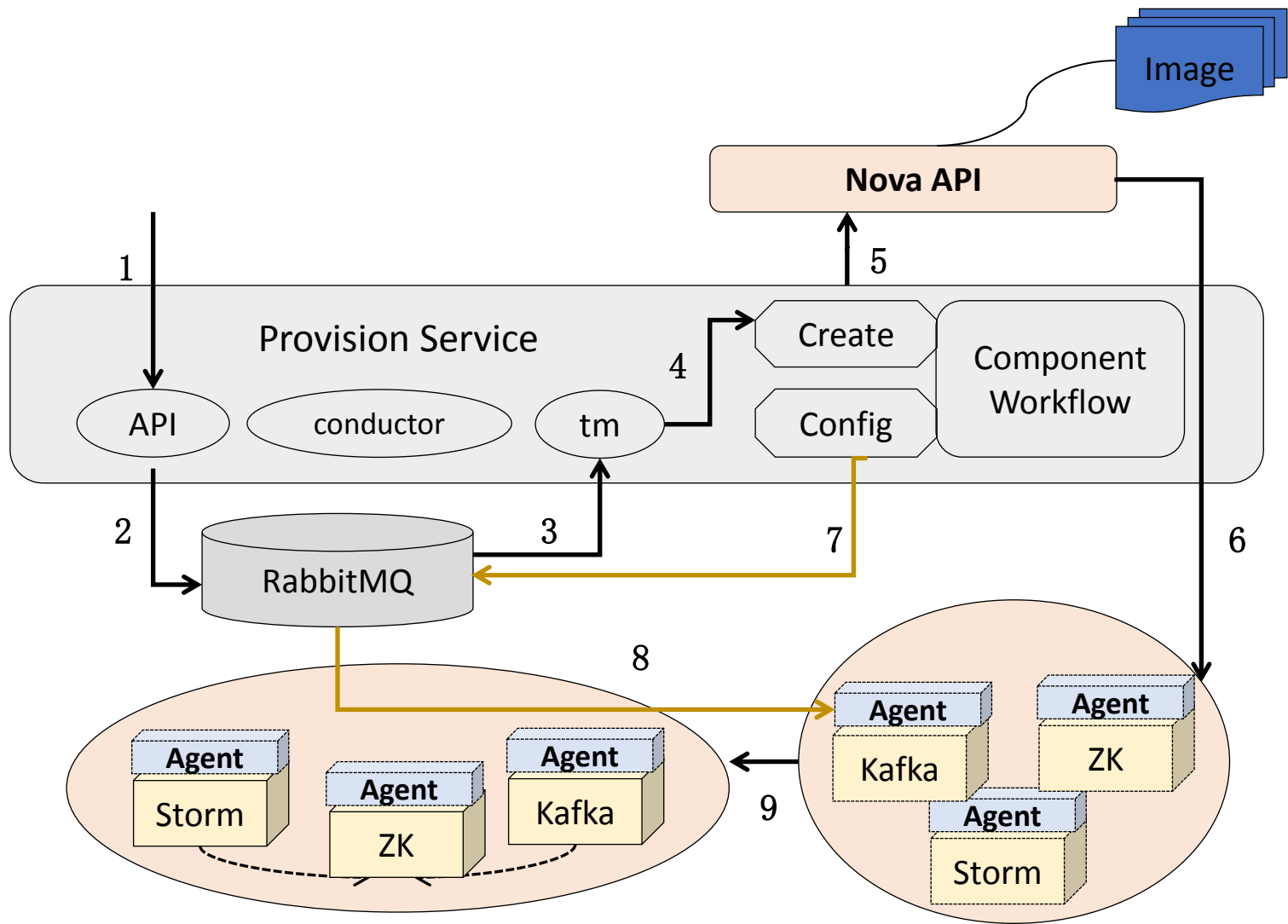
- TopicGroup/Topic 注册



- Schema 管理



# 集群部署和管理



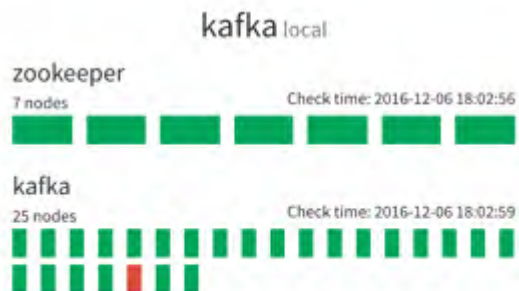


# 集群节点监控

Datacenter1



Datacenter2

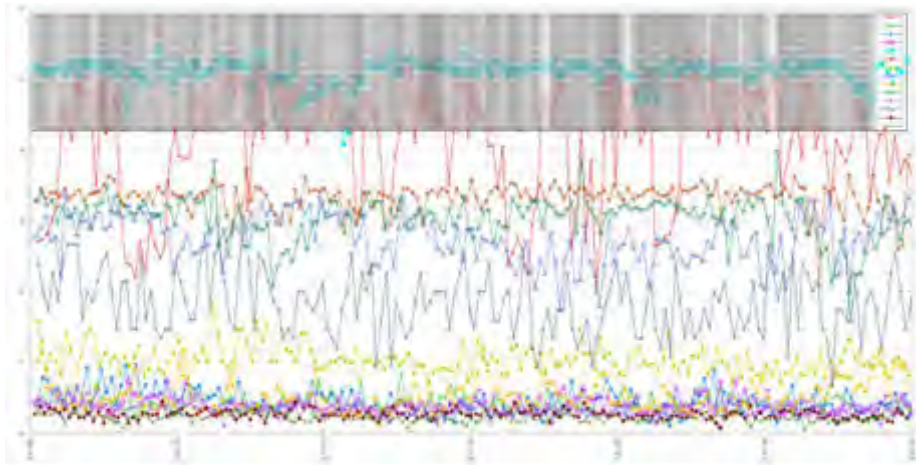


Datacenter3



# Kafka状态监控

- 系统状态监控



- Consumer状态监控



- Producer状态监控



- 端到端延迟监控

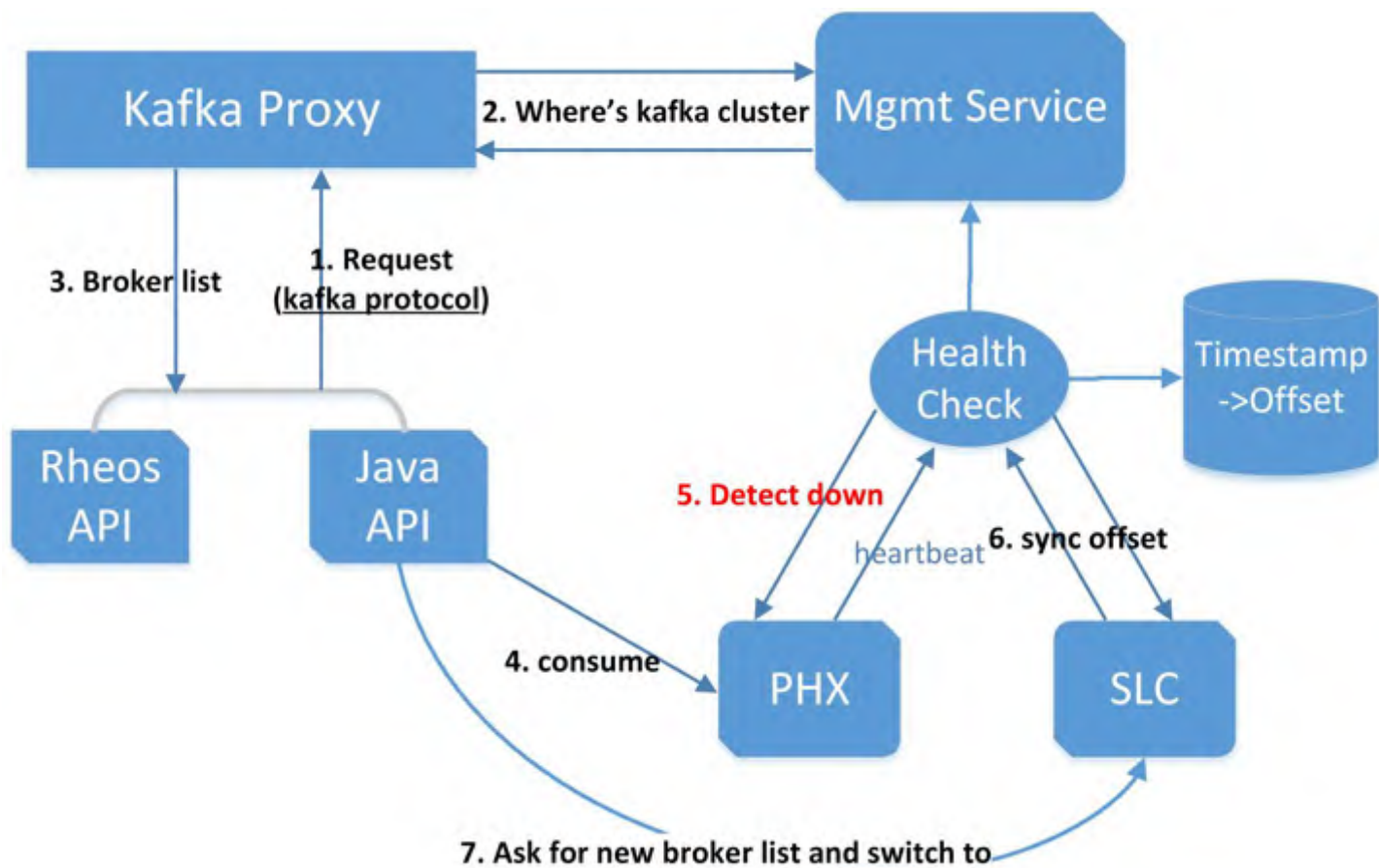


# 故障节点监测与处理

- 坏节点日常维护
- 慢节点监测
  - 系统性能参数 (CPU, 内存, 磁盘IO)
  - 异常日志捕捉与分析
  - footprint Topic测试
- 慢节点处理
  - 重启/停掉Kafka进程
  - 重启节点
  - 替换节点
  - Partition reassignment



# Offset索引和自动备援



# Kafka性能优化

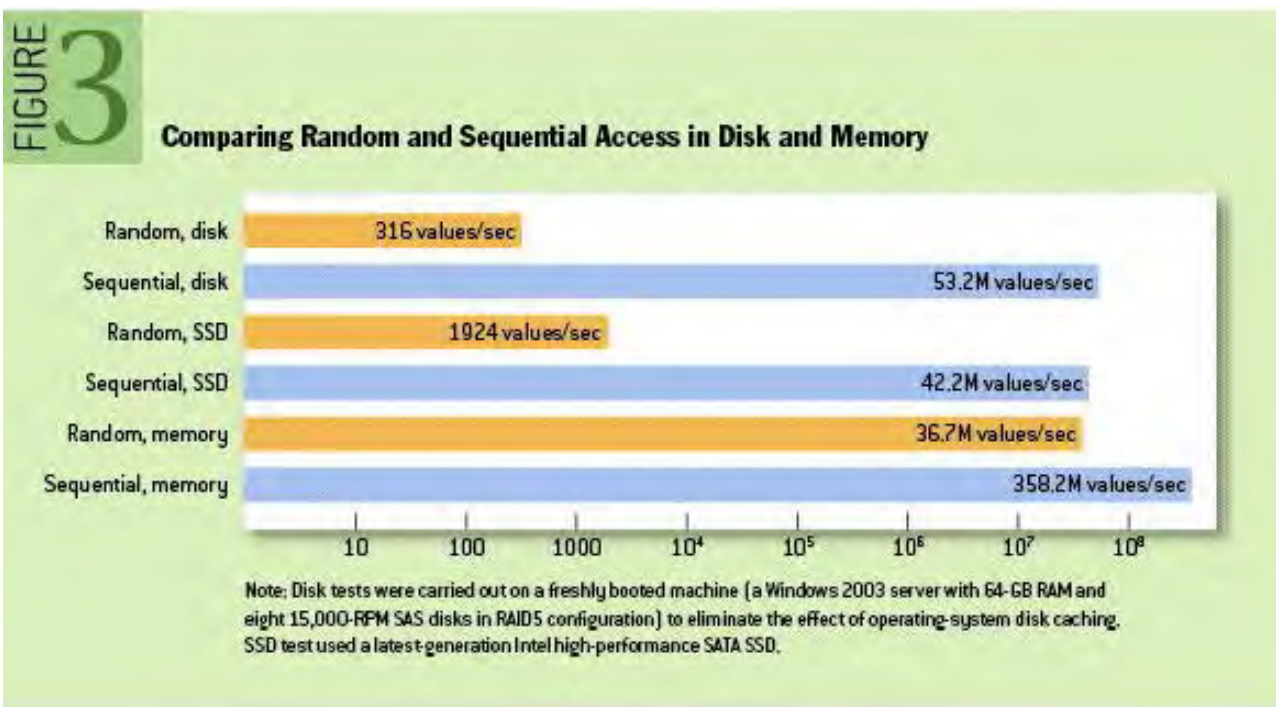
- Kafka为什么高性能
  - 磁盘顺序读写
  - 以page cache为中心的设计
  - Zero copy





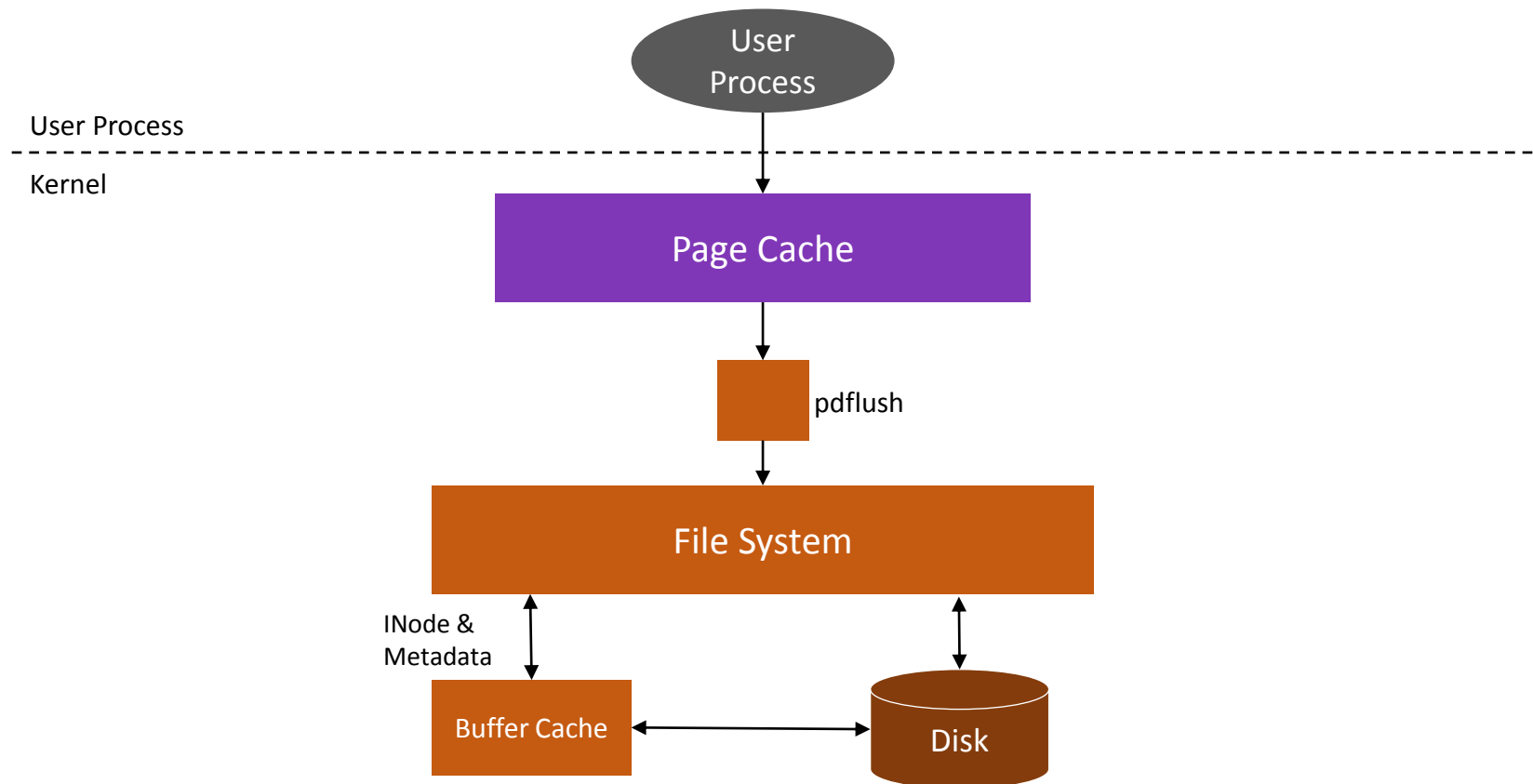
# Kafka性能优化

- 保证磁盘的顺序读写





# Kafka性能优化 – Page Cache

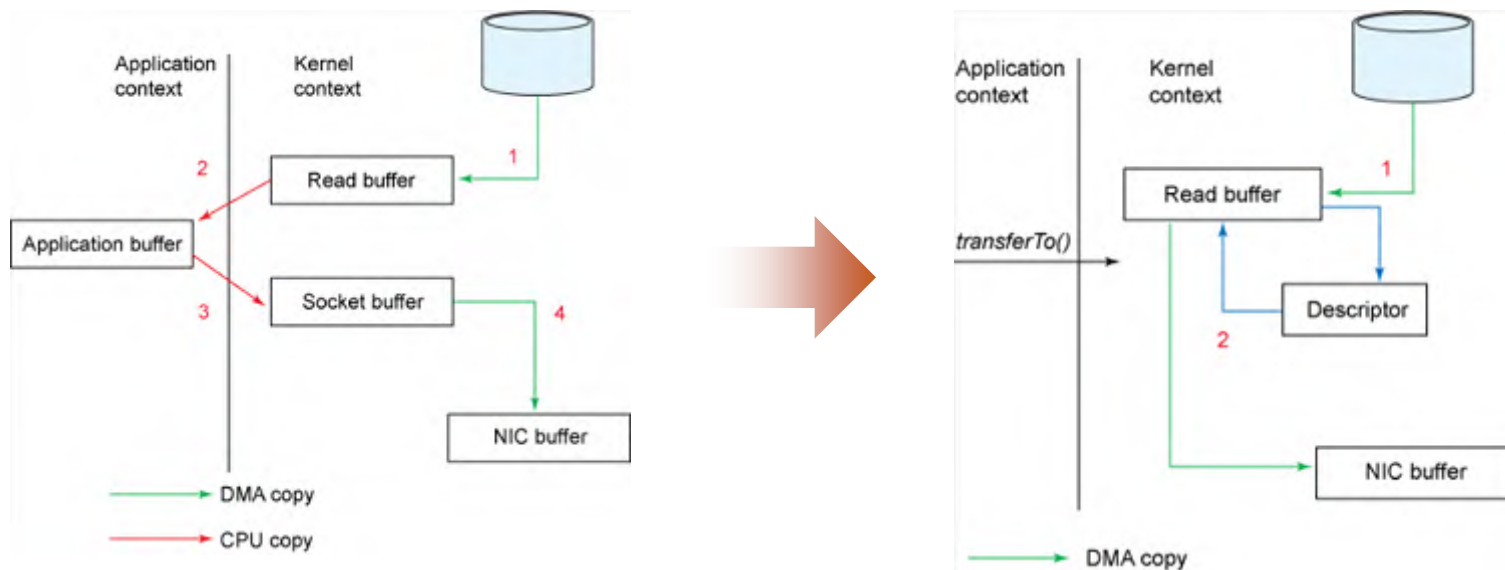


# Page Cache相关系统参数

- 小心Swap
  - set /proc/sys/vm/swappiness to 0
  - VM上的NUMA问题
- Linux page cache相关设置
  - vm.dirty\_background\_ratio (default 10)
  - vm.dirty\_background\_bytes (default 0)
  - vm.dirty\_ratio (default 20)
  - vm.dirty\_bytes (default 0)
  - vm.dirty\_writeback\_centisecs (default 500)
  - vm.dirty\_expire\_centisecs (default 3000)



# Kafka性能优化 – Zero Copy



- 小心Zero copy 失效

- 系统升级，数据格式改变 (Eg. 0.9.0 -> 0.10.0)
- SSL/TLS

# Kafka性能优化 - 其他参数调整

- File descriptor 设置
  - 增加/etc/security/limits.conf
  - 推荐 > 100000
- 增加max socket buffer size
- 权衡unclean.leader.election.enable
- 增加fetch线程数量num.replica.fetchers
- 处理leader选举auto.leader.rebalance.enable

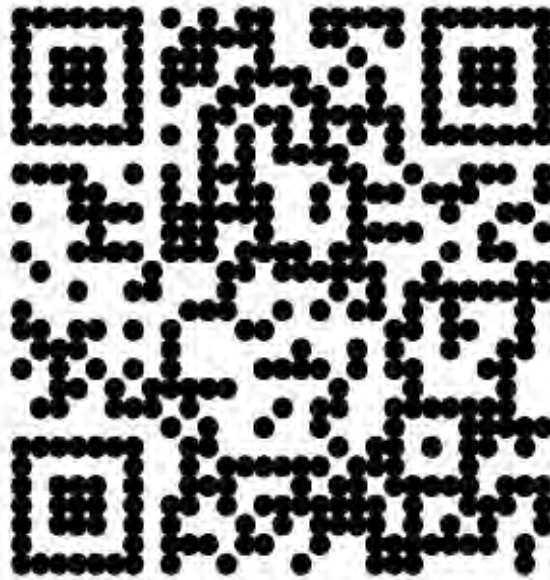


# DevOpsDays 即将首次登陆中国



DevOps 之父 Patrick Debois 与您相约

DevOpsDays 北京站 2017年3月18日



门票早鸟价仅限前100名，请从速哟

<http://2017-beijing.devopsdayschina.org/>





想第一时间看到  
高效运维社区公众号  
的好文章吗？

请打开高效运维社区公众号，点击右上角小人，如右侧所示设置就好







# Thanks

高效运维社区  
开放运维联盟

荣誉出品

