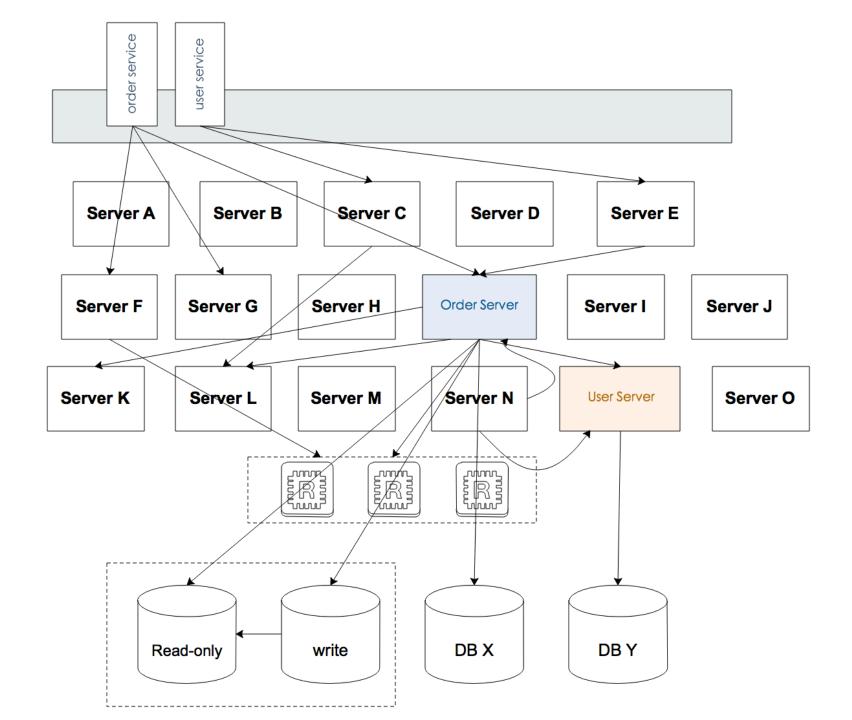
新牛计划

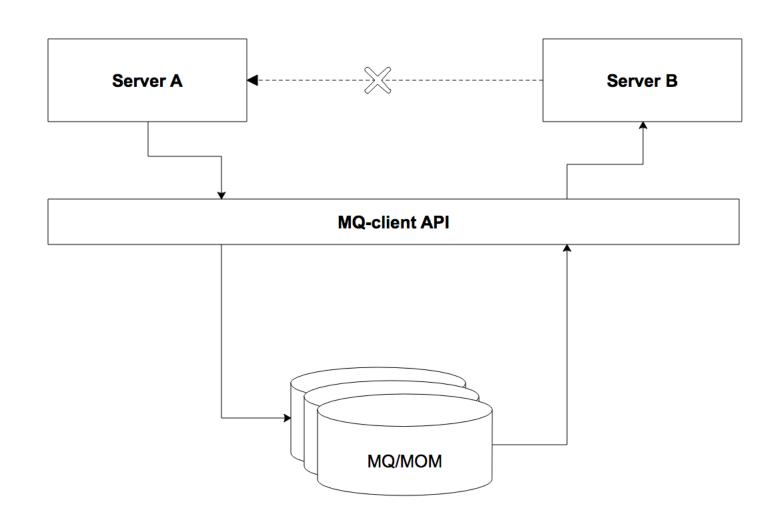
消息中间件

概要介绍

服务依赖调用



消息中间件



业务述求

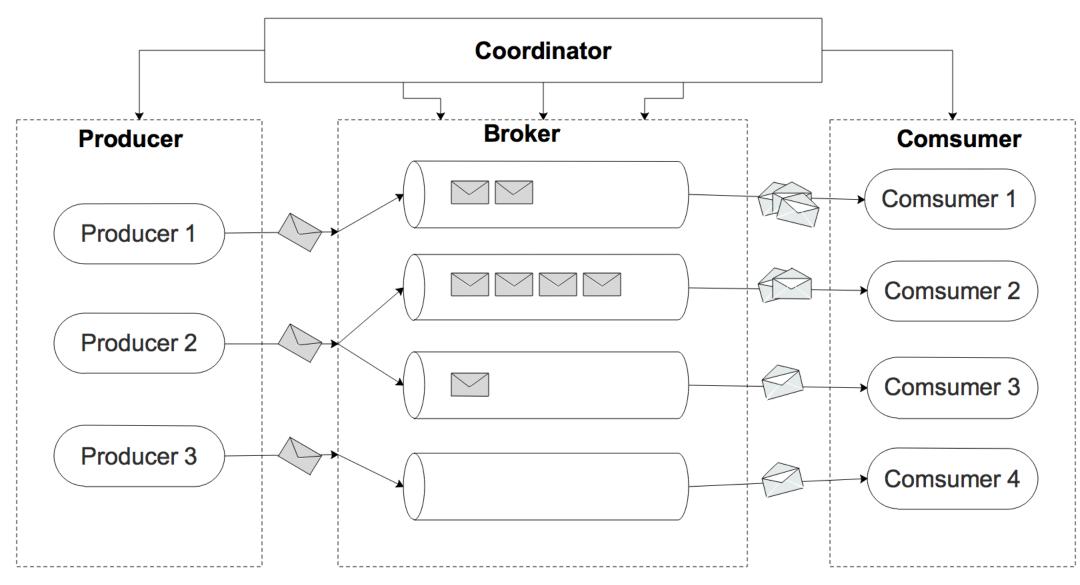
- 磁盘不够用,没用就删了
- 处理完了再通知我
- 一堆定时任务, 烦!
- 操作事件没有调用我呀
- 数据在好多服务器上,怎么办
- 每次都得轮询数据库,才知道数据更新了
- XXX的服务真烂, 老是把我的服务拖挂了
- 又要秒杀了,心累,服务别挂了
- 能不能按顺序通知我呀
- 数据放内存的,重启之后丢了

作用

- 解耦
- 削峰
- 高效性
- 伸缩性
- 可靠性
- 可恢复
- 异步

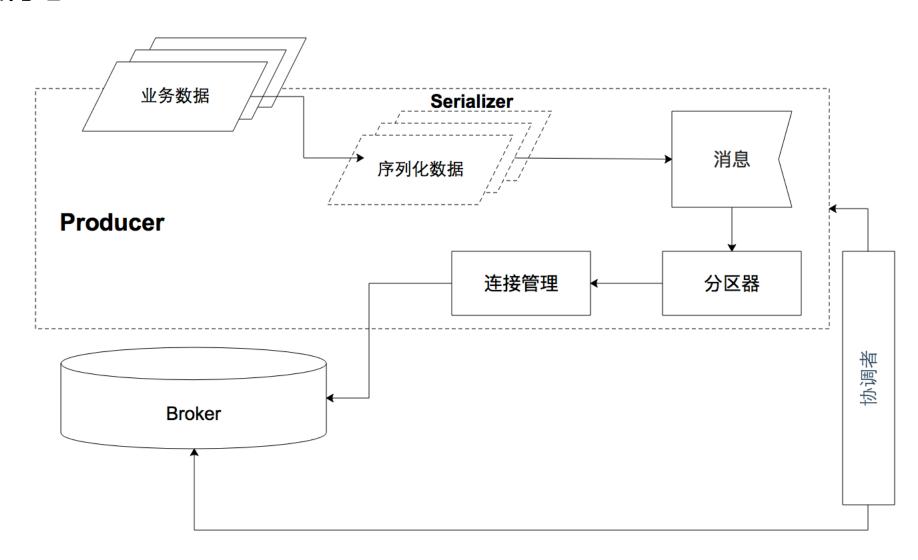
- 缓冲
- 顺序性
- 持久化

结构



Producer

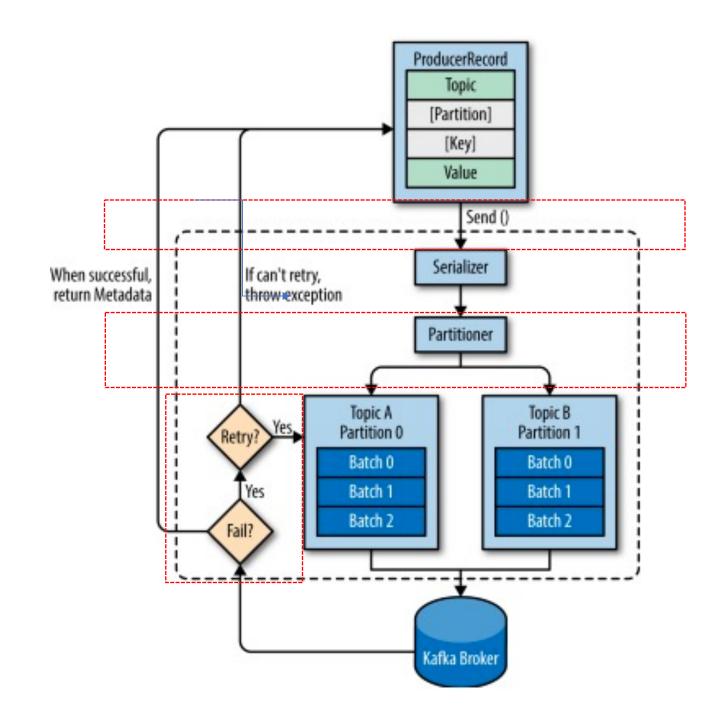
发布消息

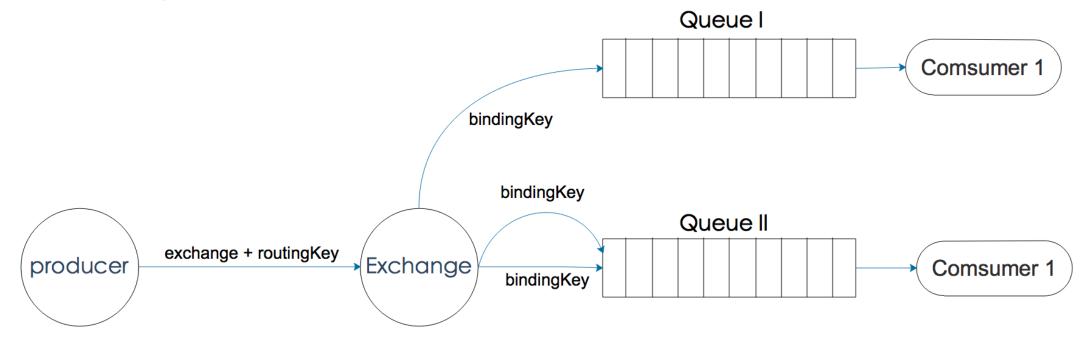


序列化

- Protobuf
- Json
- Thrift
- Avro

- 发送消息
 - 同步
 - 异步
 - 批量
- 路由/负载均衡
 - 默认 Round Robin
 - 自定义分区策略?
 - 修改分区数量?
- 超时/重试
 - 提交策略



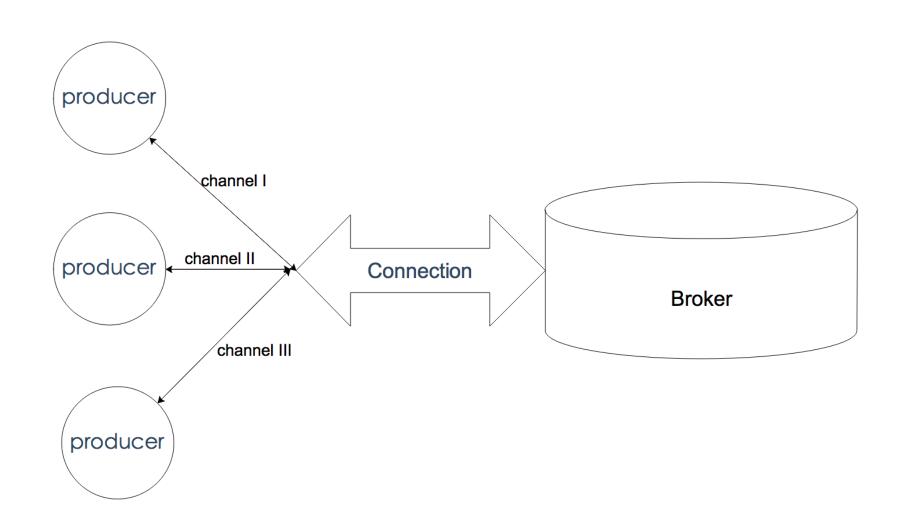


Exchange

bindingKey

routingKey

● 连接复用



Consumer

反序列化

• • •

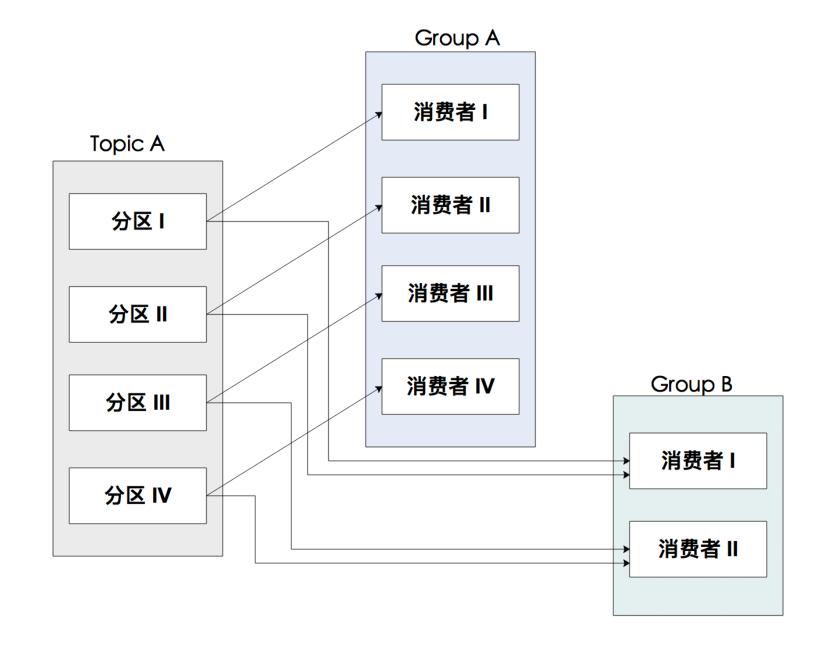
Pull or Push?

● 优势

● 劣势

- pull
 - 轮询
 - 长轮询?长连接?

- Partition vs consumer
- 均衡策略
- 再均衡



Offset

● 提交

● 存储: zk? Broker?

rabbitMQ

- Push or Pull?
 - 均衡策略
 - 过载保护

复习

消息传输保障

ACID

BASE

CAP

- At most once
- At least once
- Exactly once

两个讨论

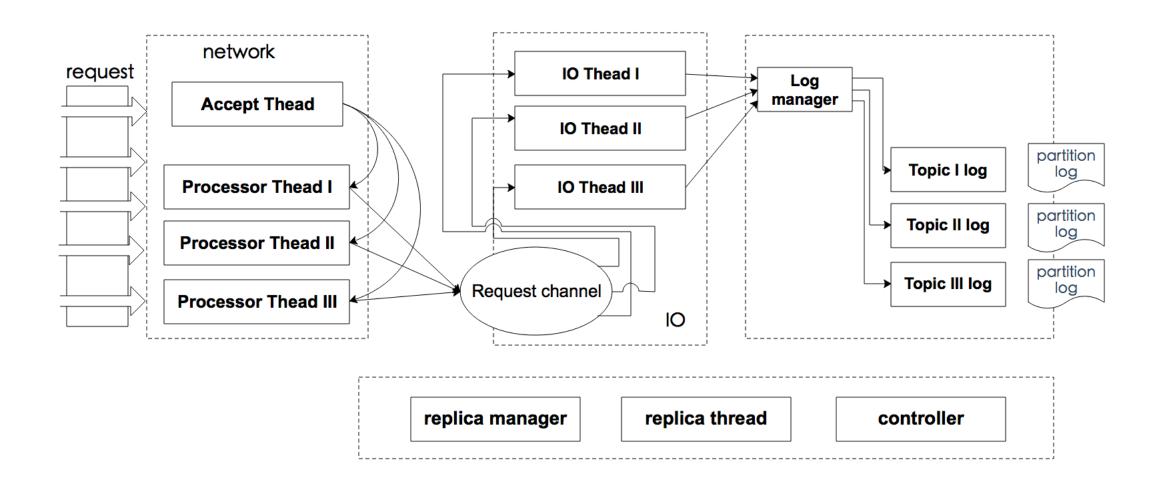
● 有序队列

● 事务

Broker

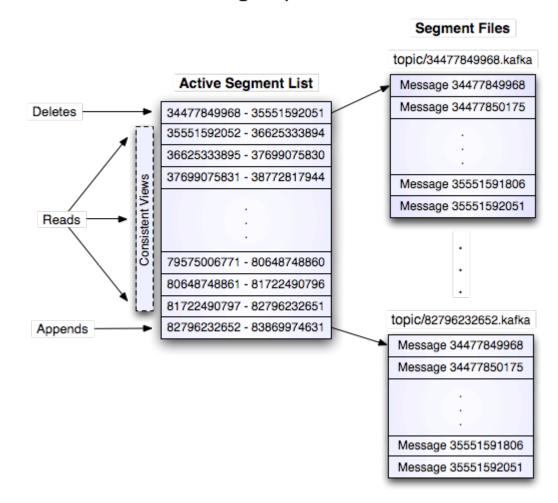
核心,核心,核心

几乎全是它



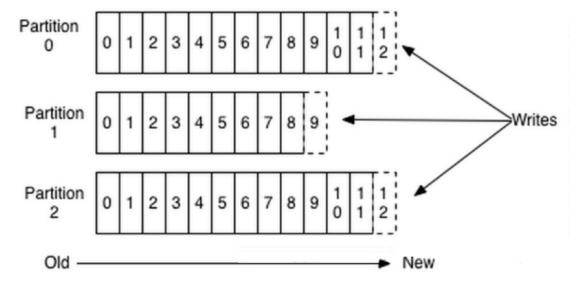
● 存储

Kafka Log Implementation



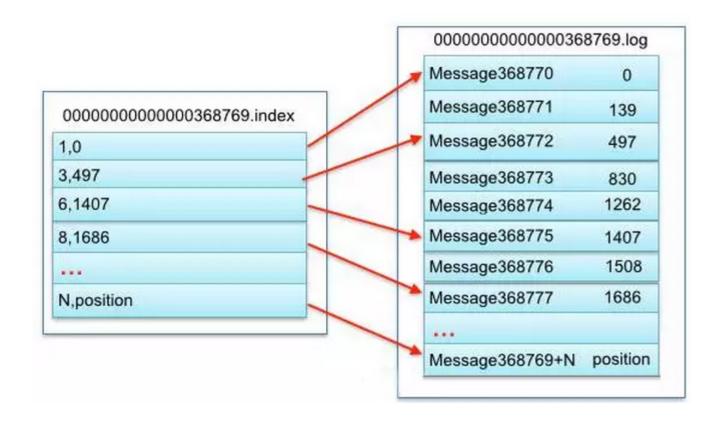
● 存储





00000000000000000000.index 000000000000000000000.log 000000000000000368769.index 000000000000000368769.log 00000000000000737337.index 000000000000000737337.log 000000000000001105814.index 000000000000001105814.log

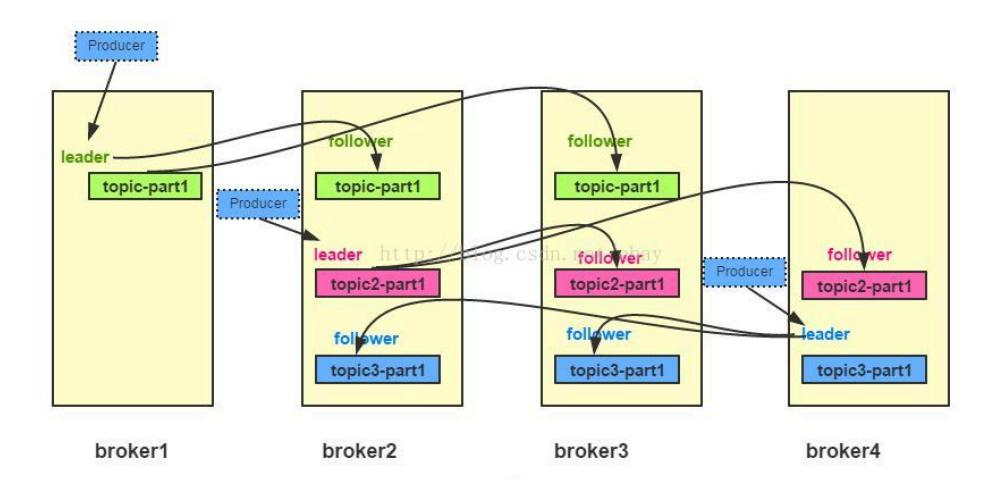
● 存储

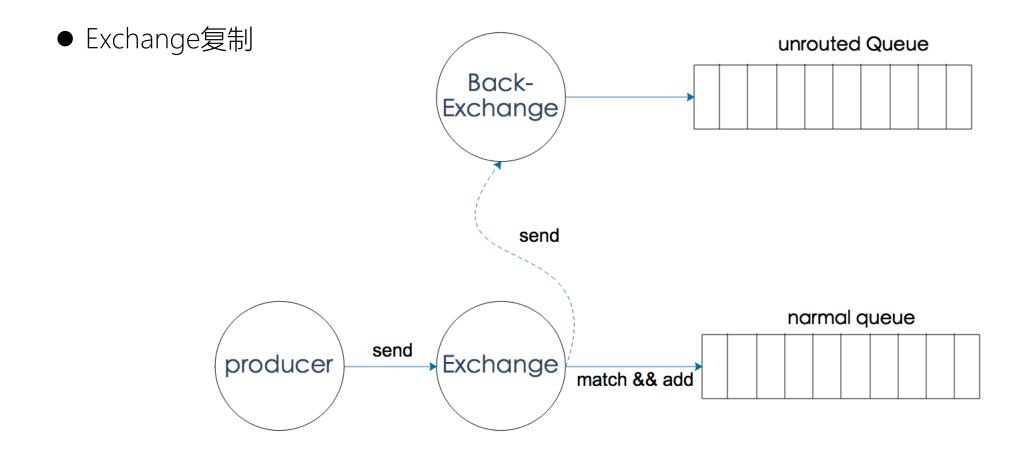


● 查询

● 复制

• ISR





• TTL

- 全局过期(队列维度)
- 单条过期

如何删除?

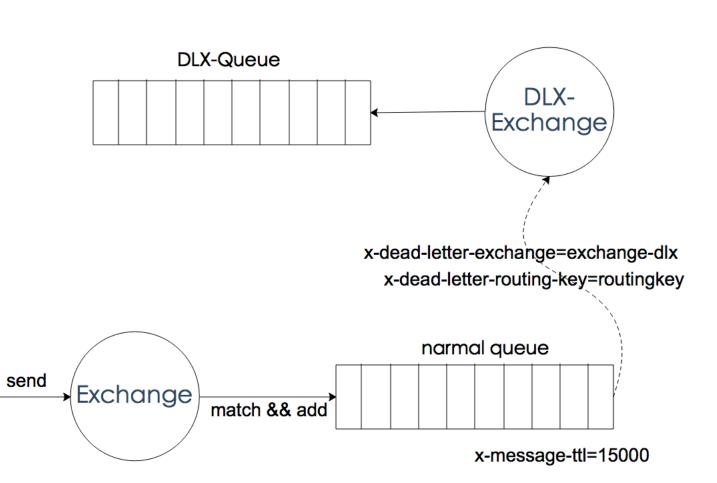
● 死信队列 (DLX)

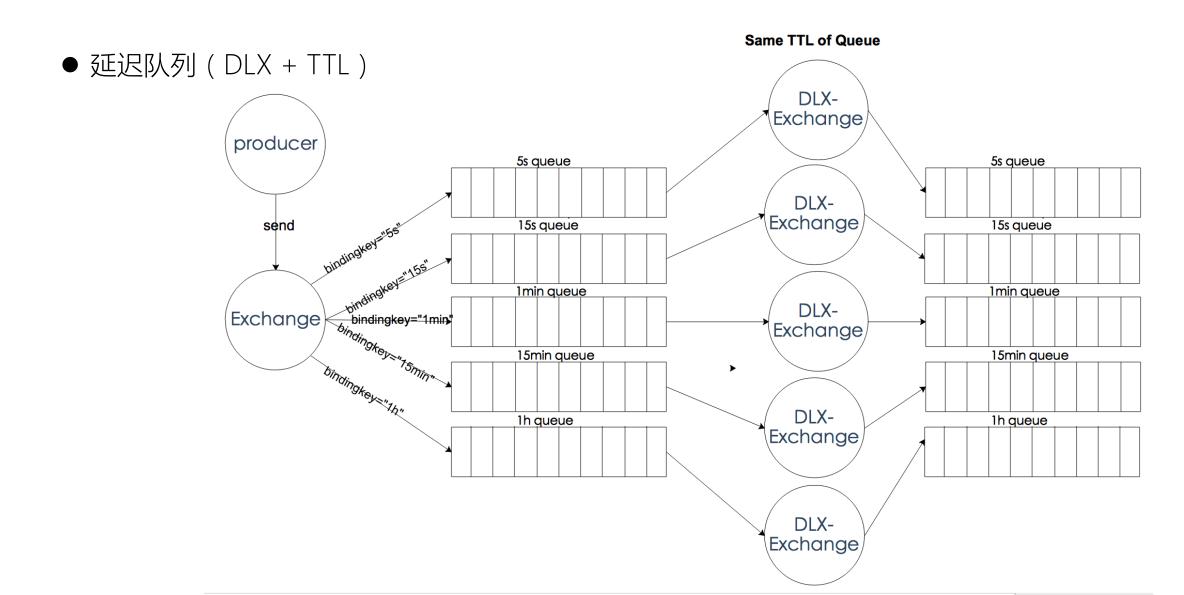
● 消息被拒

• TTL

● 队列超长

producer





● 优先级队列

● 持久化

Exchange

Queue

Message

如何选择??影响?

持久化了消息就不丢了吗?

● 存储机制

RAM

□ 消息内容

□ 消息索引

DISK

• 都在内存

• 内容在磁盘,索引在内存

• 内容在磁盘,索引在内存和磁盘均有

• 均在磁盘

● 惰性队列

存文件系统

内存 - 10 十

非持久化模式,会丢吗?



END