

Apache Pulsar

社区演进和规划

翟佳

上海谕流科技有限公司



Apache Pulsar 社区演进

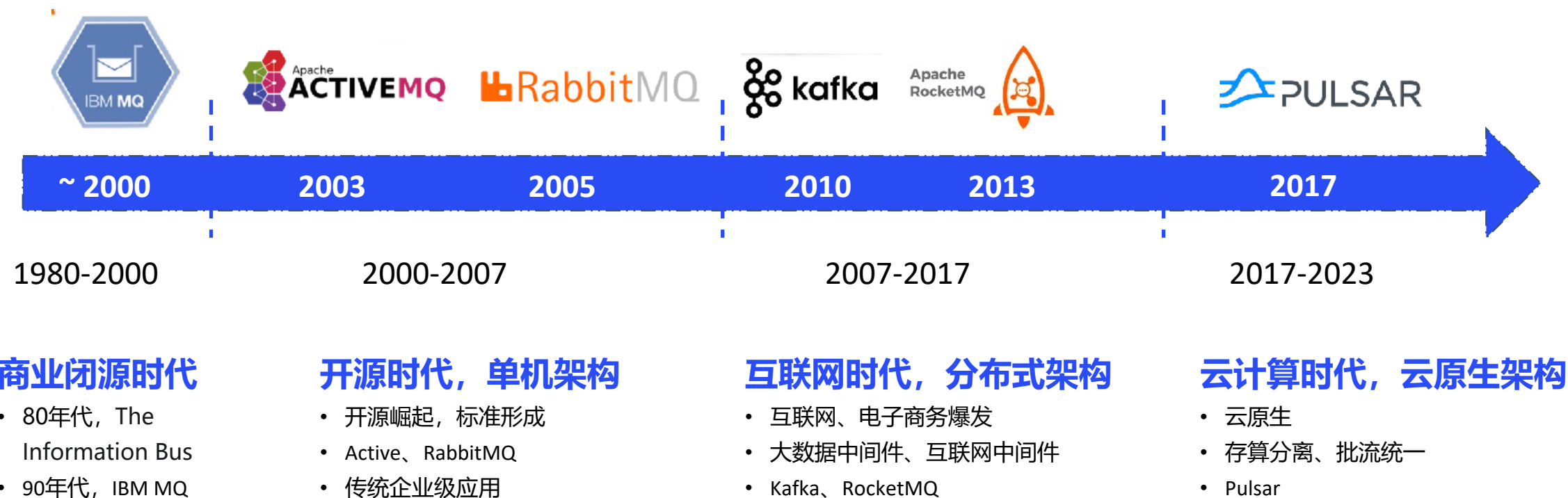
Apache Pulsar 功能和场景优势

Apache Pulsar 社区近况

Apache Pulsar

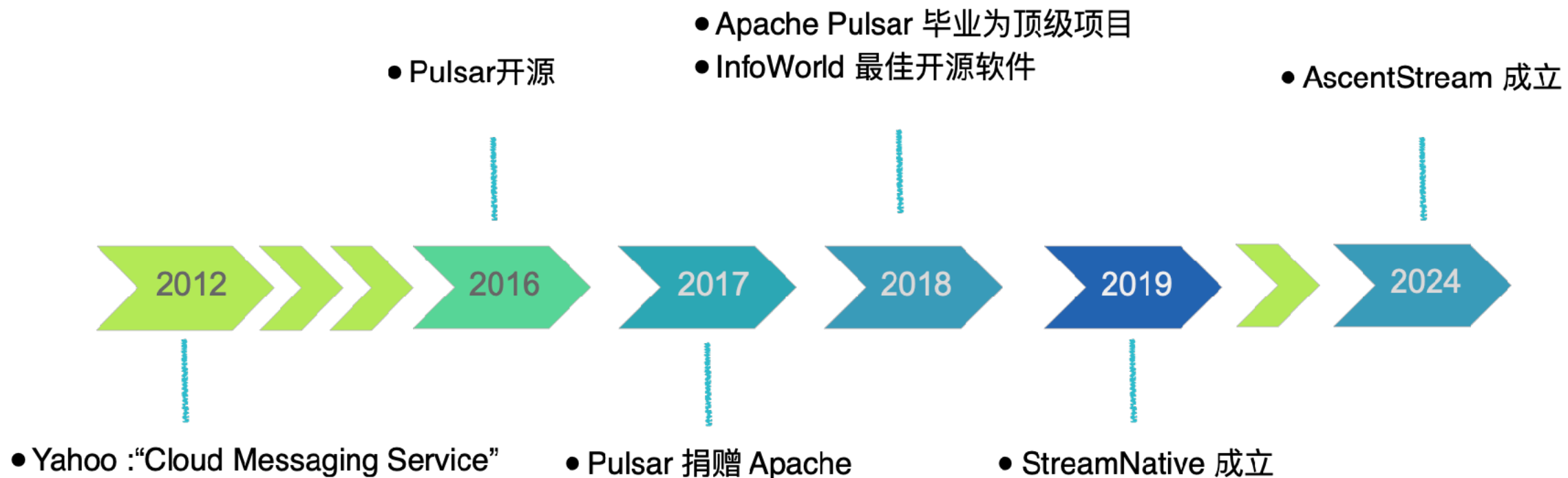
社区演进

消息队列发展历程



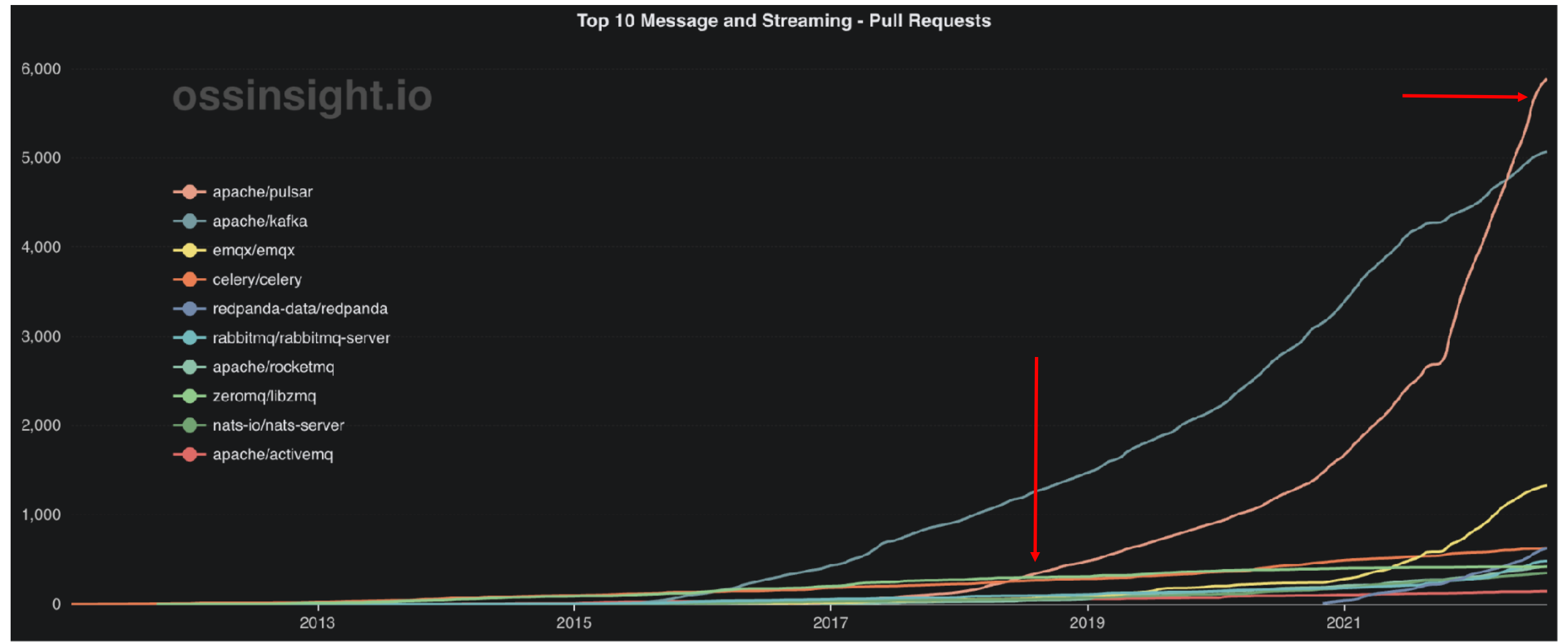
- 从需求上看: 消息 -> 流 -> 消息和流融合
- 从架构上看: 单机 -> 分布式 -> 云原生

Apache Pulsar 发展历程



Apache Pulsar: 深度使用于中国头部企业





Unified API

Cloud-Native
Architecture



Multi Tenancy

Geo-replication



Data Offloaders
(Tiered Storage)



Client Libraries



为数据处理提供批流融合的存储层



Processing Engines



Data Offloaders (Tiered Storage)



Client Libraries





Pulsar Functions
(Lightweight Stream Processing)



Processing Engines



Data Offloaders
(Tiered Storage)



Client Libraries



Pulsar Connectors: 无代码化的数据集成



Pulsar Functions (Lightweight Stream Processing)



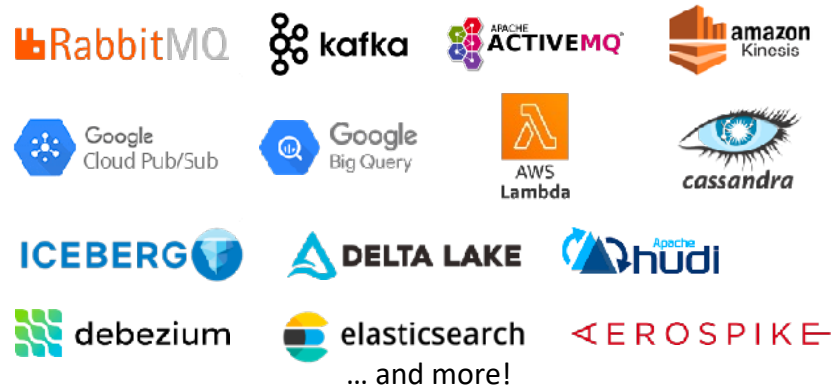
Processing Engines



Data Offloaders (Tiered Storage)



Connectors (Sources & Sinks)



Client Libraries



协议插件：云原生助力平台构建运维

Protocol Handlers



Pulsar Functions (Lightweight Stream Processing)



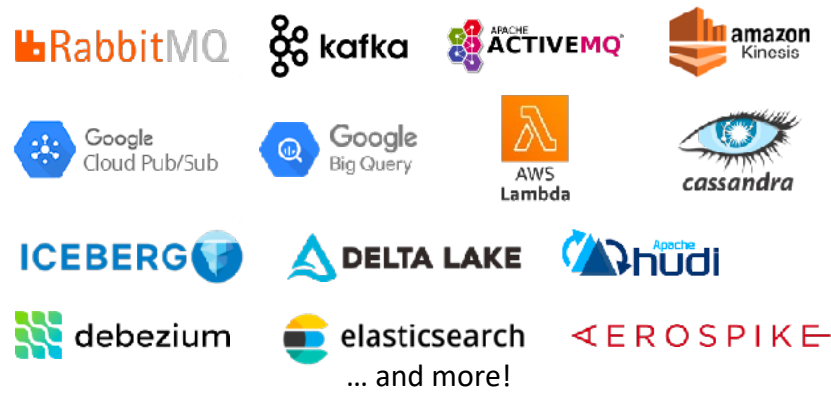
Processing Engines



Data Offloaders (Tiered Storage)



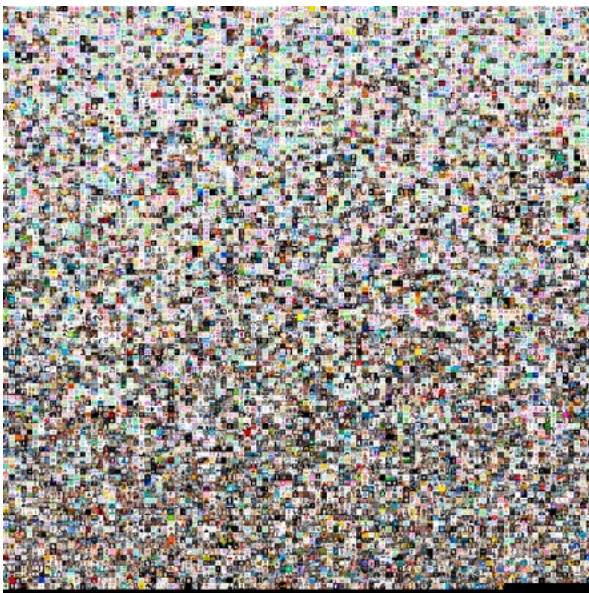
Connectors (Sources & Sinks)



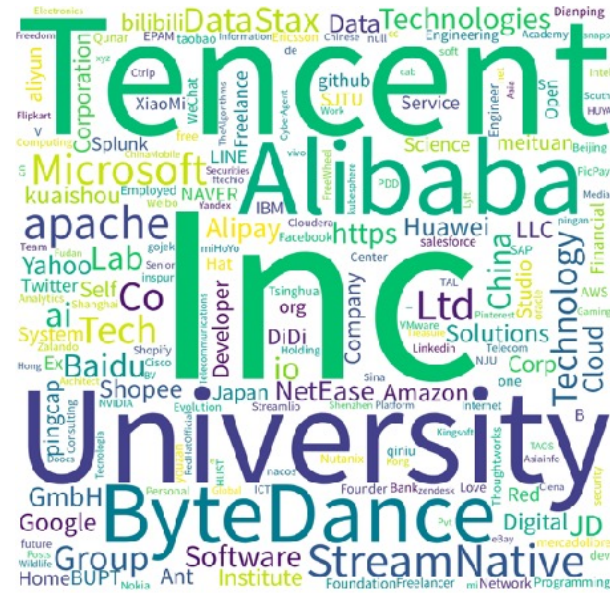
Client Libraries



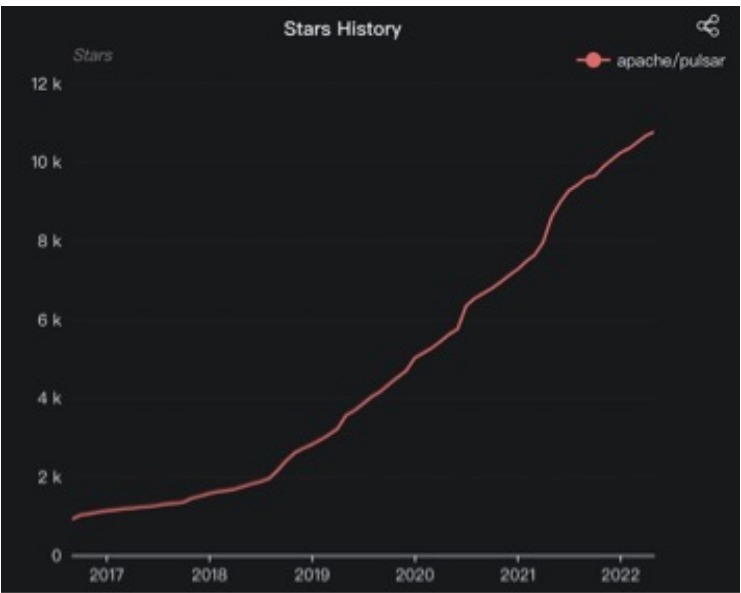
Apache Pulsar: 社区状态



- 600+ contributors



- 最活跃的MQ社区



- 10K+ Github Stars

上海谙流科技有限公司 (AscentStream)

- 致力于打造**金融级云原生**消息流平台，助力企业挖掘实时数据价值。
- 由 Apache Pulsar 与 Apache BookKeeper **团队**成员组建。
- 专注中国商业化和社区，**信创**。



Apache Pulsar

MQ场景优势

多业务部门取代原有各种MQ

巨大的数据量

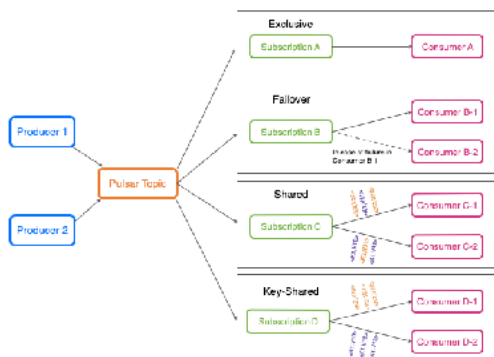
2012 yahoo 7亿用户、60个国家

广告计费等关键一致性场景

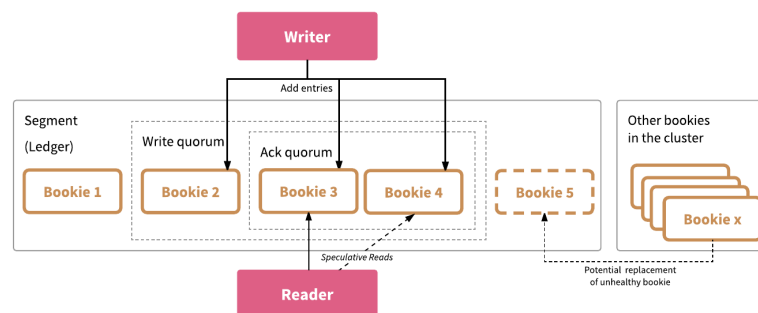
跨全球各大洲的数据可用性



应用数量增长：满足MQ灵活统一需求



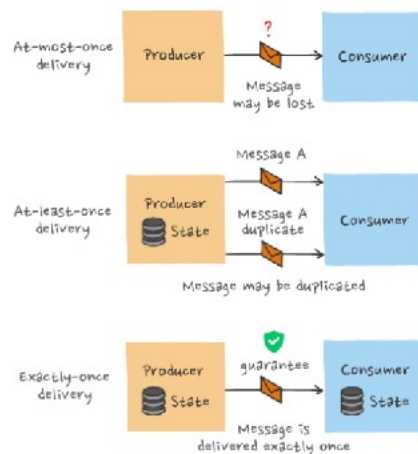
Flexible Subscribe modes



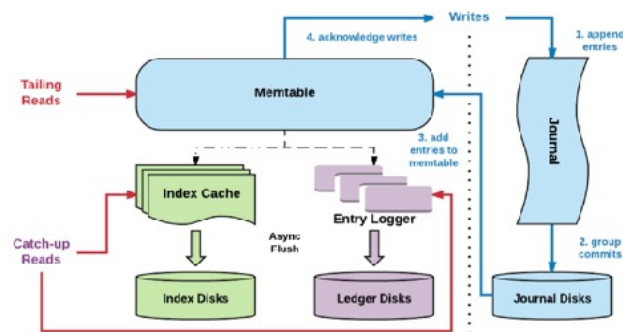
Flexible replicas



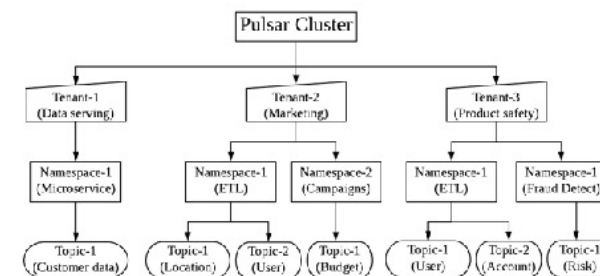
Multi Region



Flexible Semantics

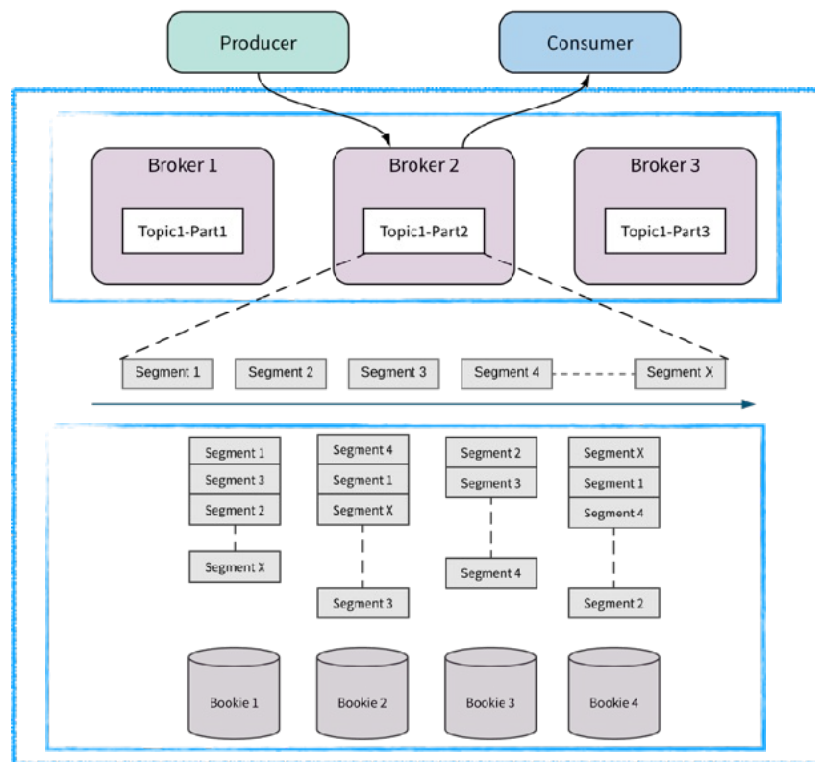


Flexible persistence

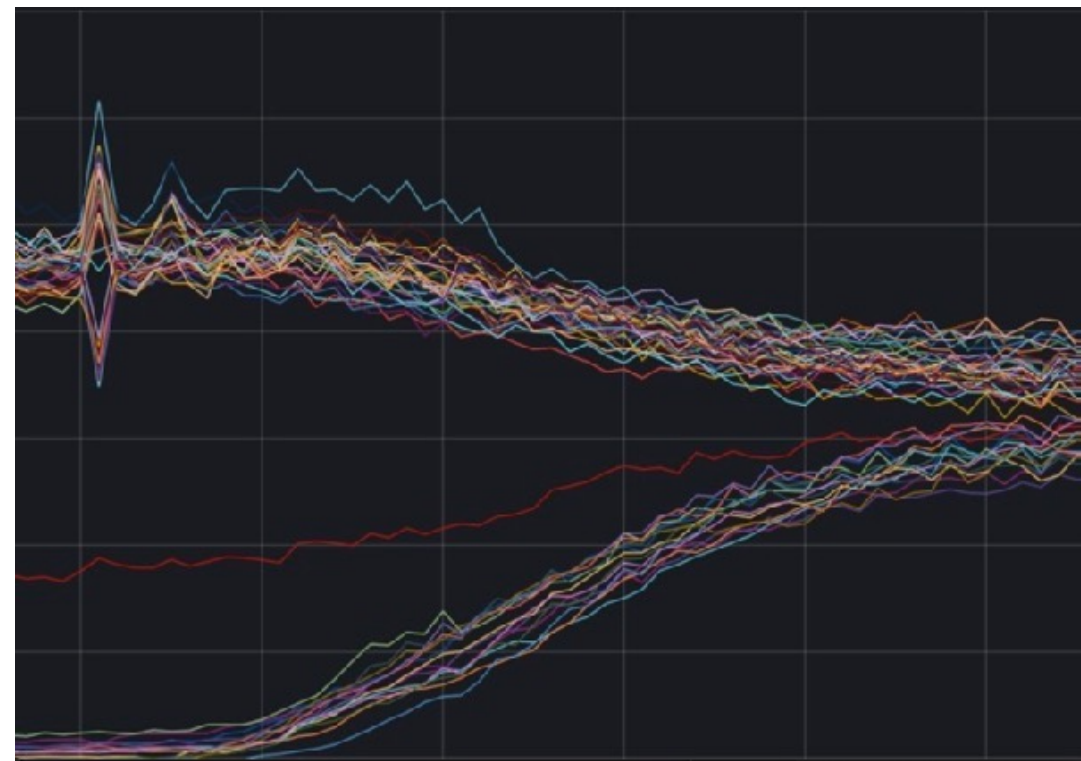


Multi tenant

单个应用数据规模增长：满足MQ扩展和弹性需求



分布式 — 存算分离

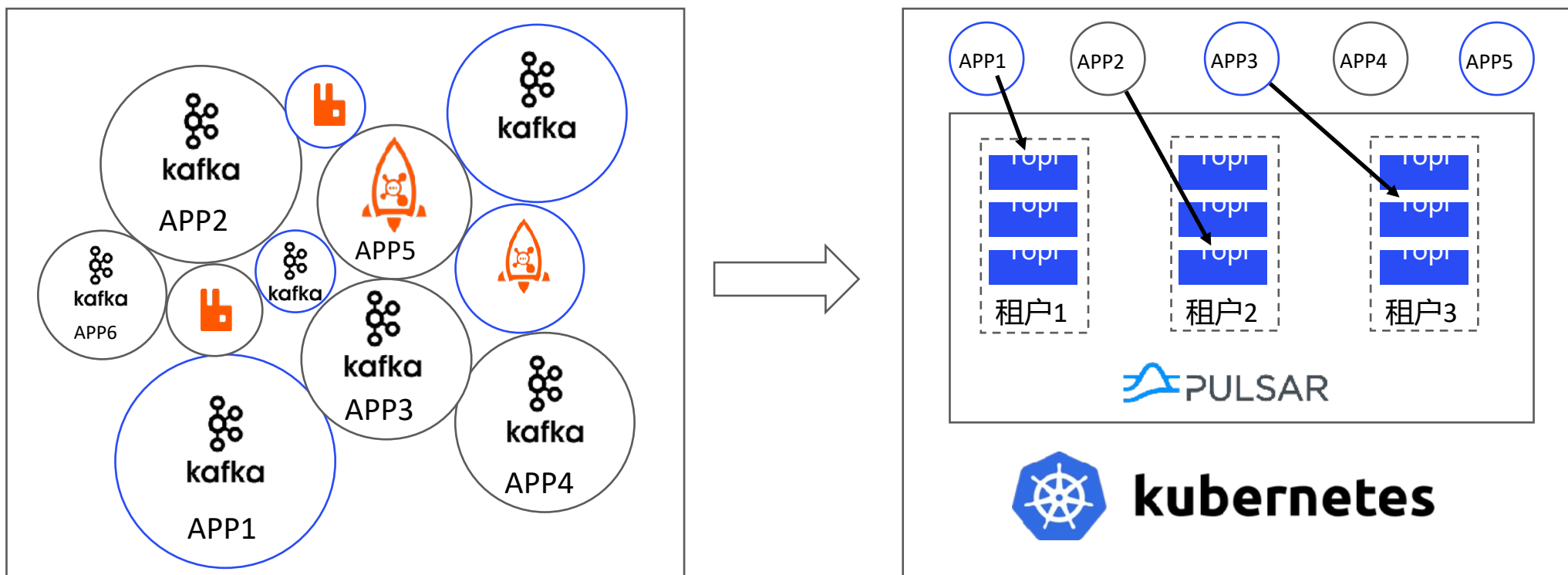


实时的弹性

MQ场景的例子

案例

某客户前期使用MQ过程中，由各业务方独立建设，系统中并存了近百套集群，机器利用率低，运维复杂，开发运维成本居高。

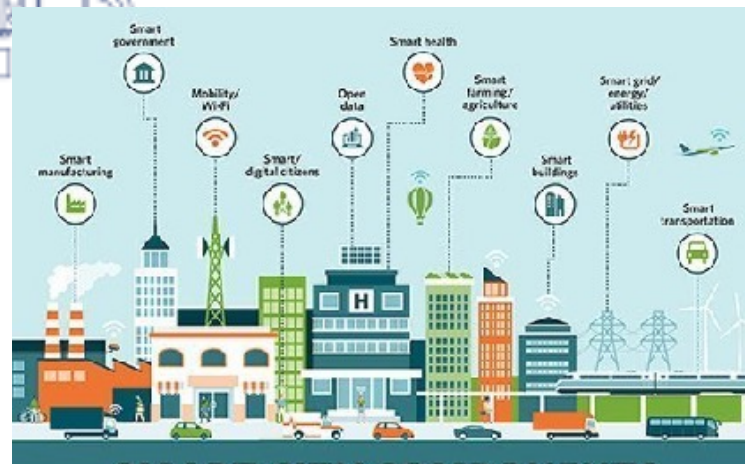


利用**Pulsar多租户**和**云原生弹性扩缩容**的特性，统一了不同技术栈，将近百套集群缩减为6套集群，大大降低了人员维护成本和硬件成本，提升了业务协同效率

Apache Pulsar

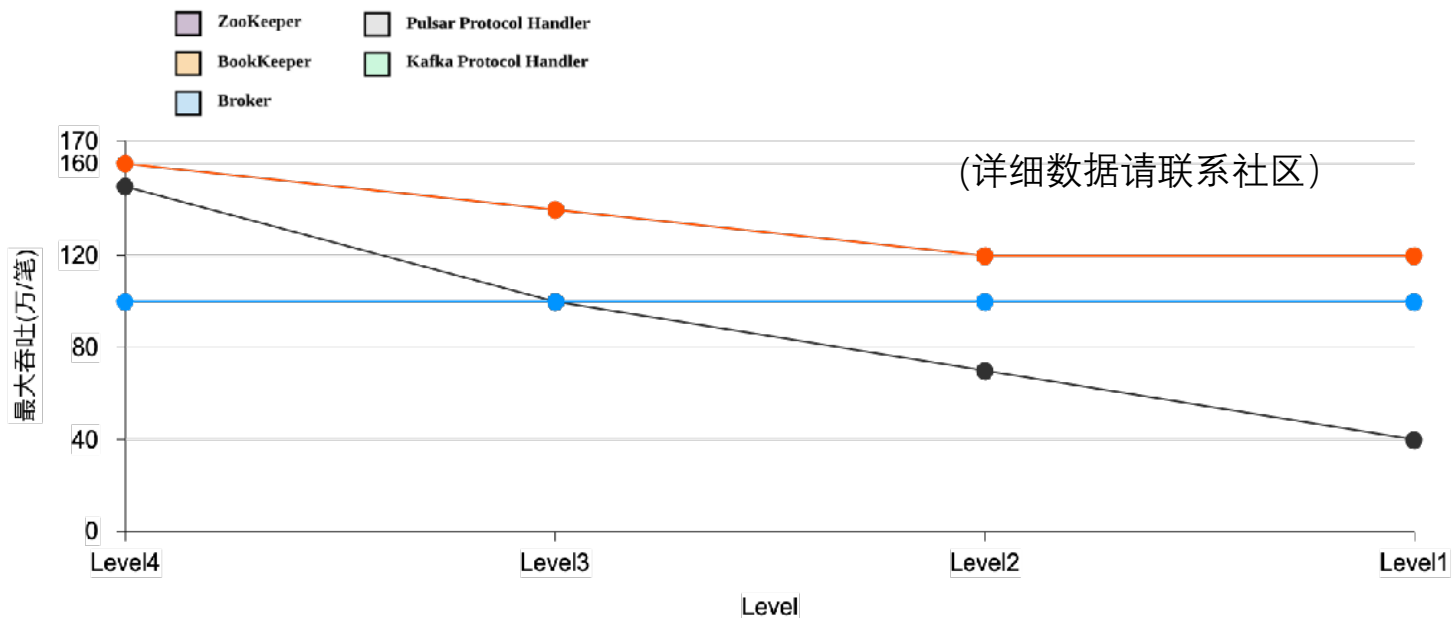
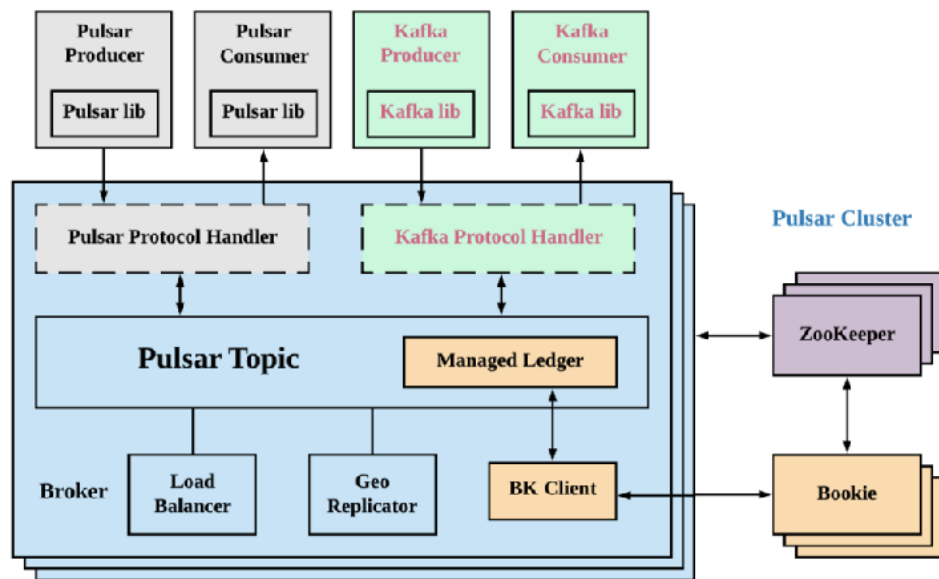
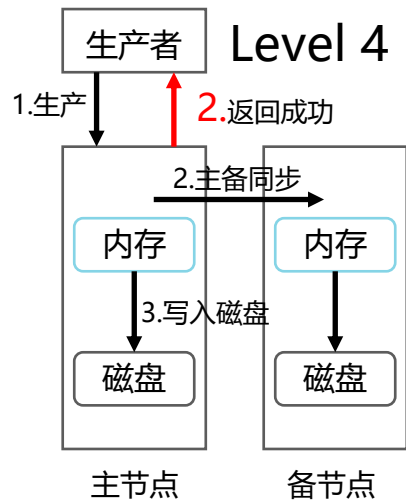
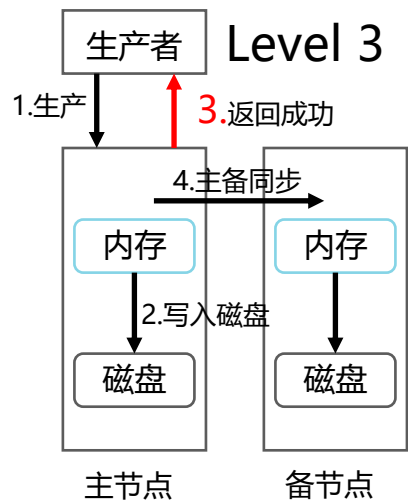
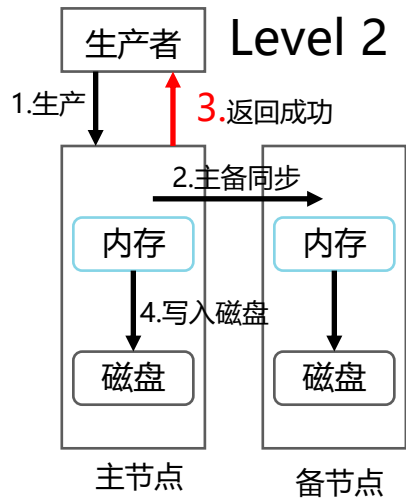
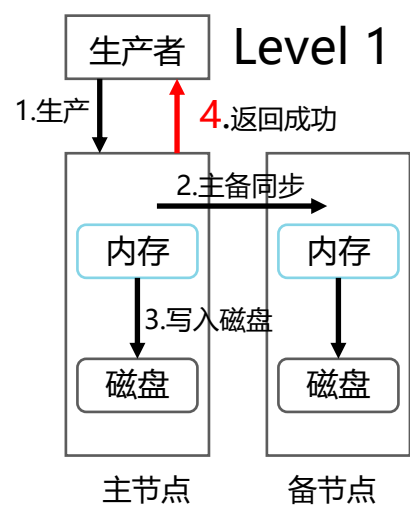
Stream 场景优势

Stream 场景下的新需求



- High throughput
- 1M Topics
- Scale
- Low latency
- Long storage

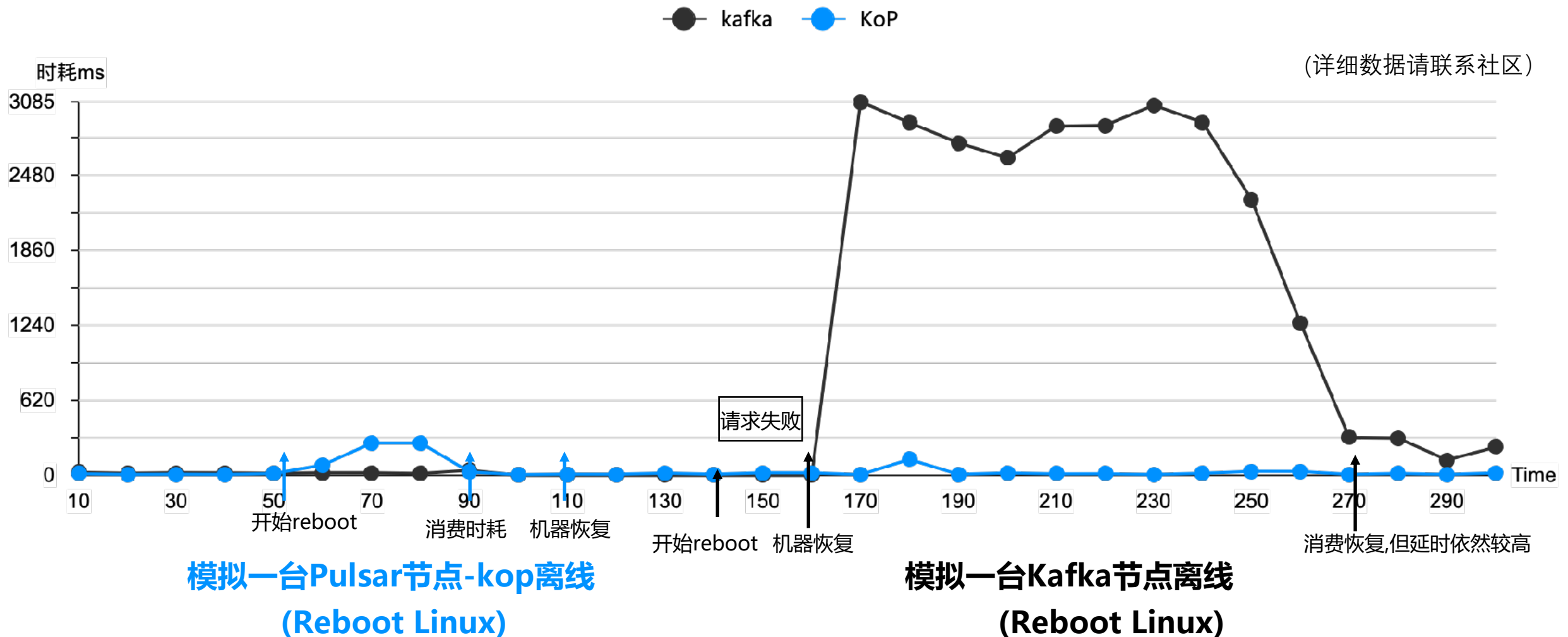
Kafka 生态无缝切换 — KoP (Kafka-on-Pulsar)



便捷运维：故障自愈 提高可用性

验证实验：

- 1) 3台Server机器组成Kafka/Kop集群， Topic配置为2副本存储。
- 2) 通过openmessage向集群10Topic共80Partition进行消息生产和消费，写入流量10W笔/秒，消息大小1KB。
- 3) 过程中模拟reboot直接重启Linux模拟机器异常



Apache Pulsar

云原生的优势

Easier to use & scale



Unified Messaging Model



Scale elastic



High availability

Lower Cost



Multi-tenancy



Tiered storage



Low & Stable Latency

Multi & hybrid env



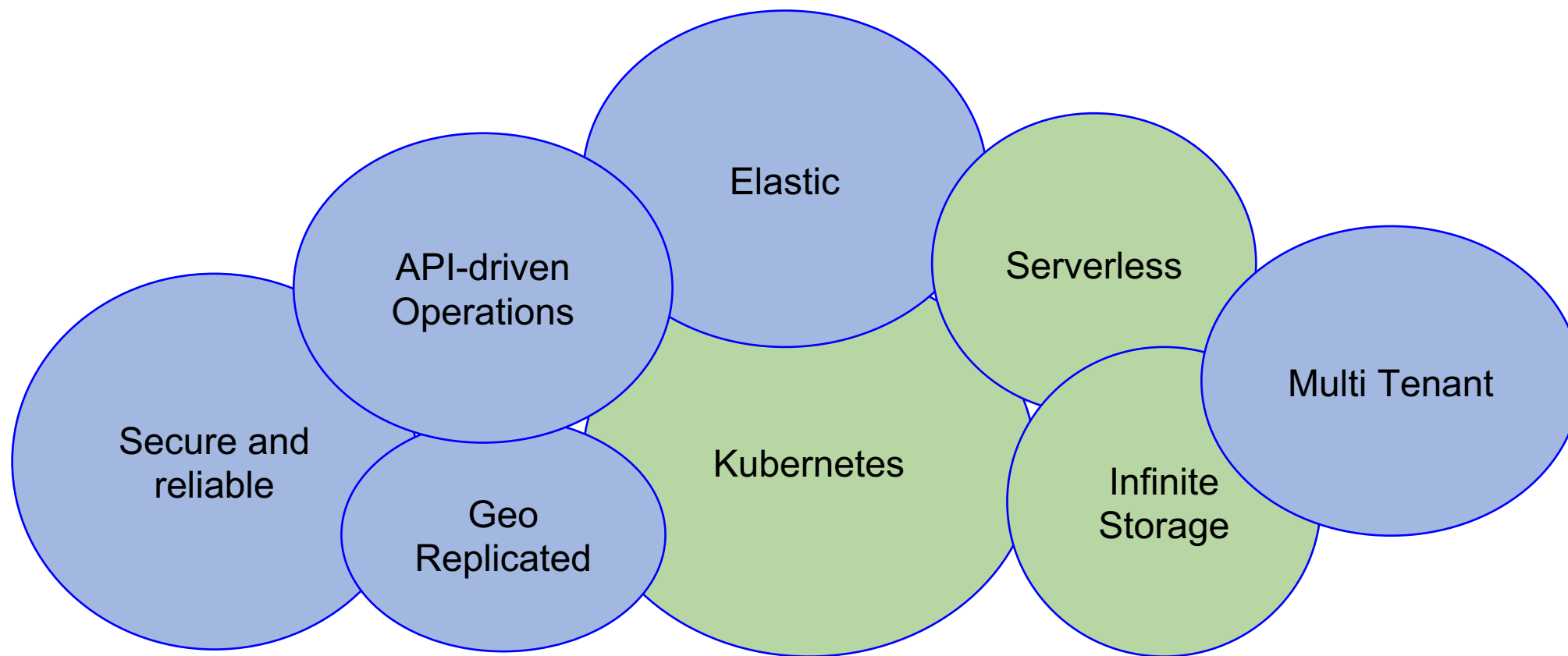
Container based



Geo-replication



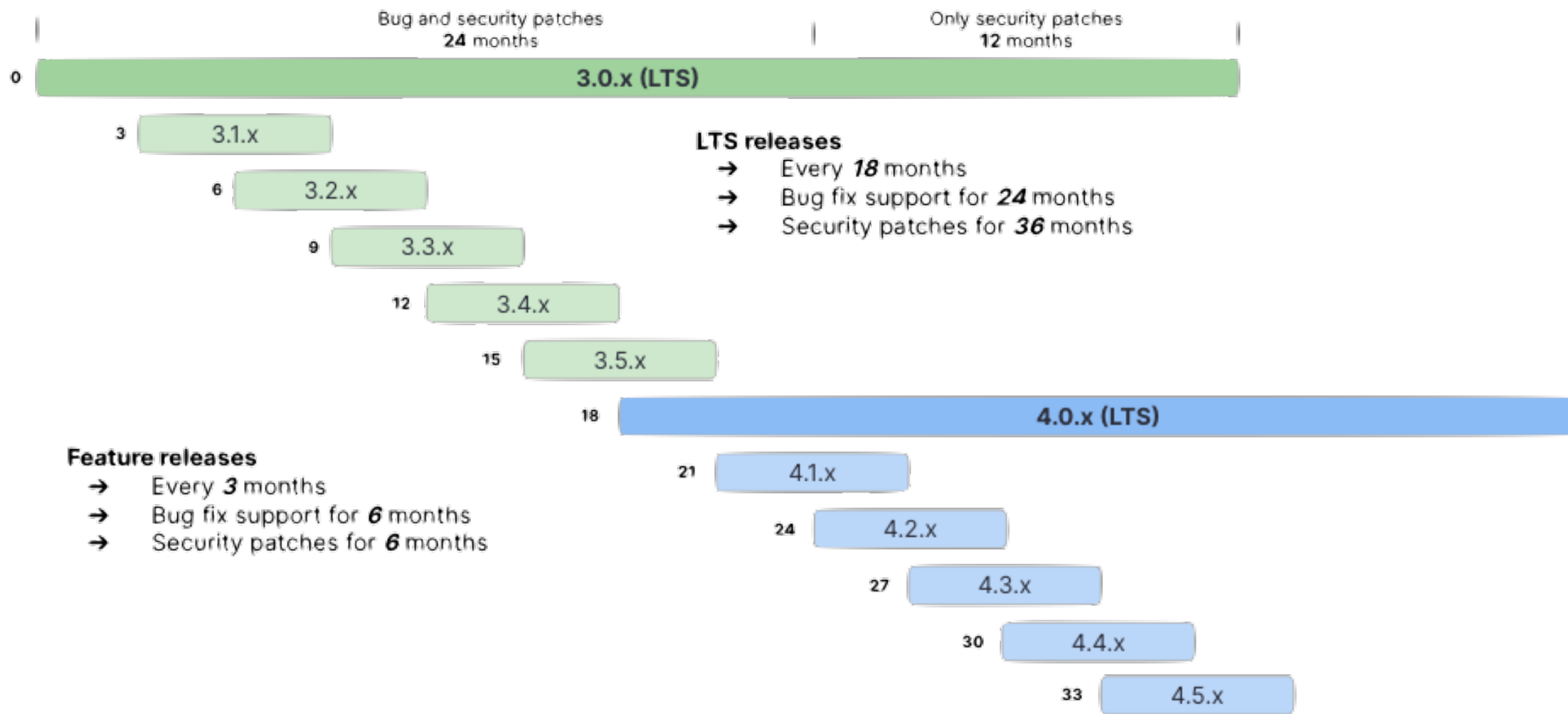
Standard security API

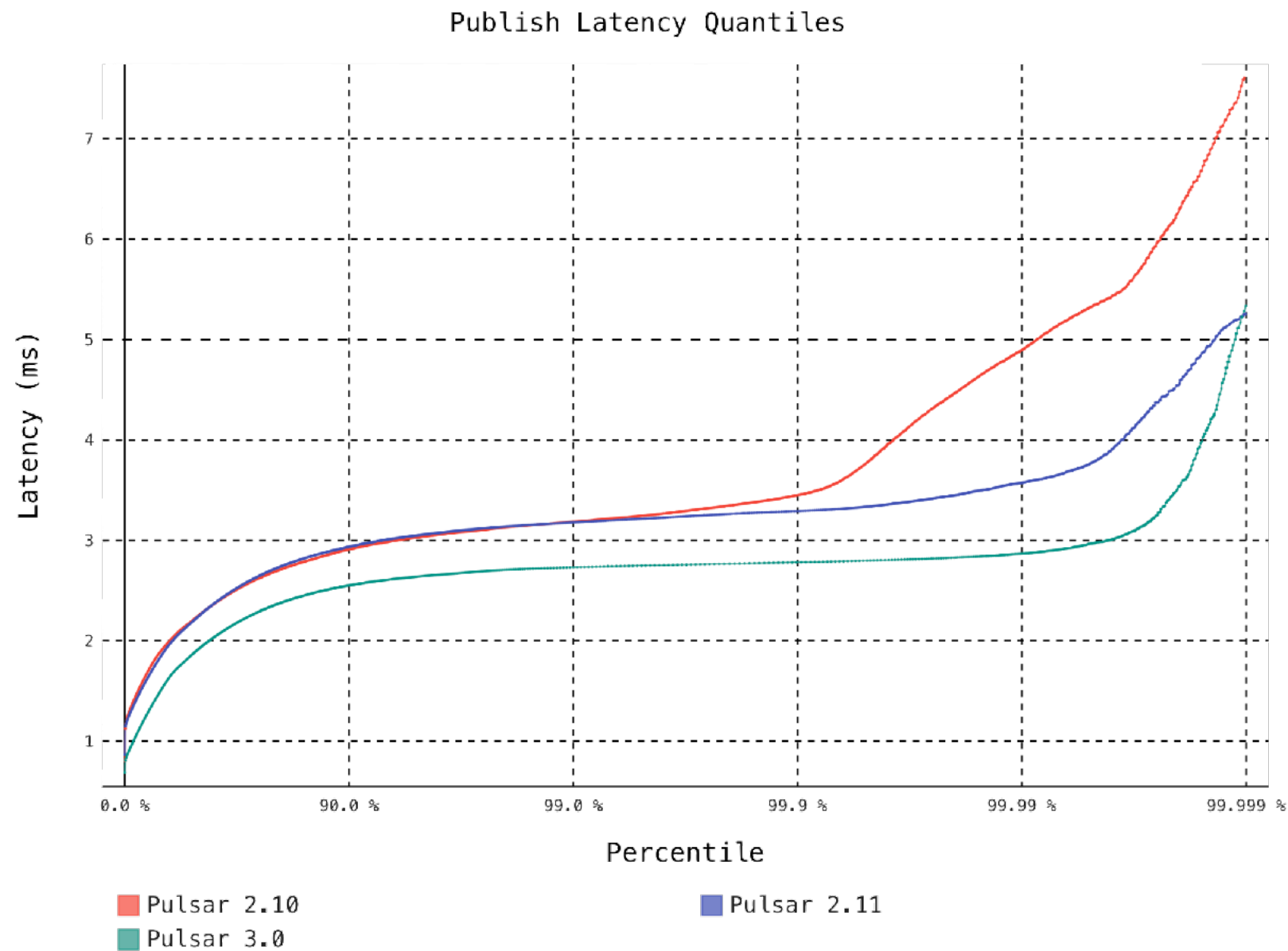


Apache Pulsar

近況

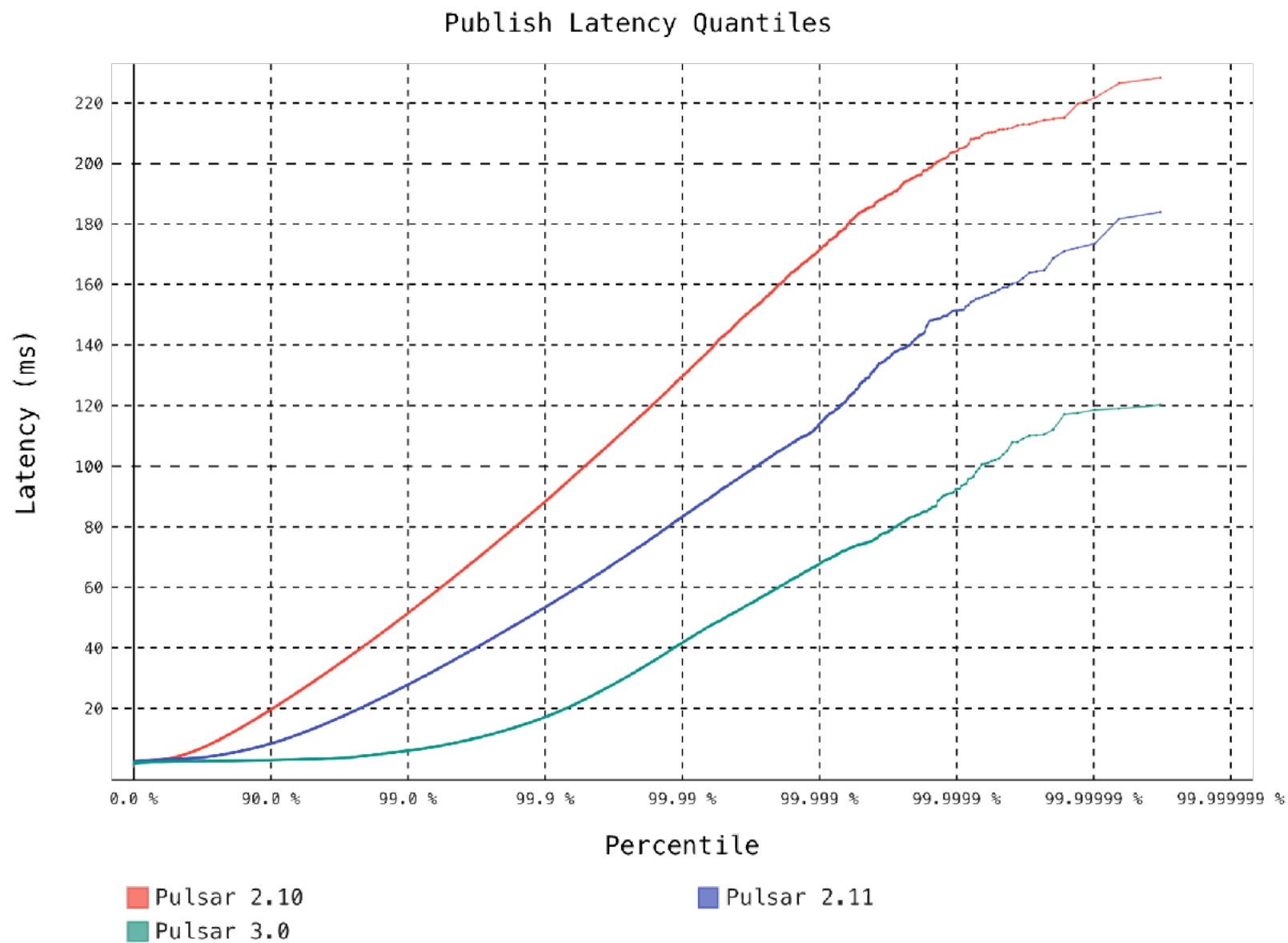
长期支持版本 — LTS Support





3 nodes cluster

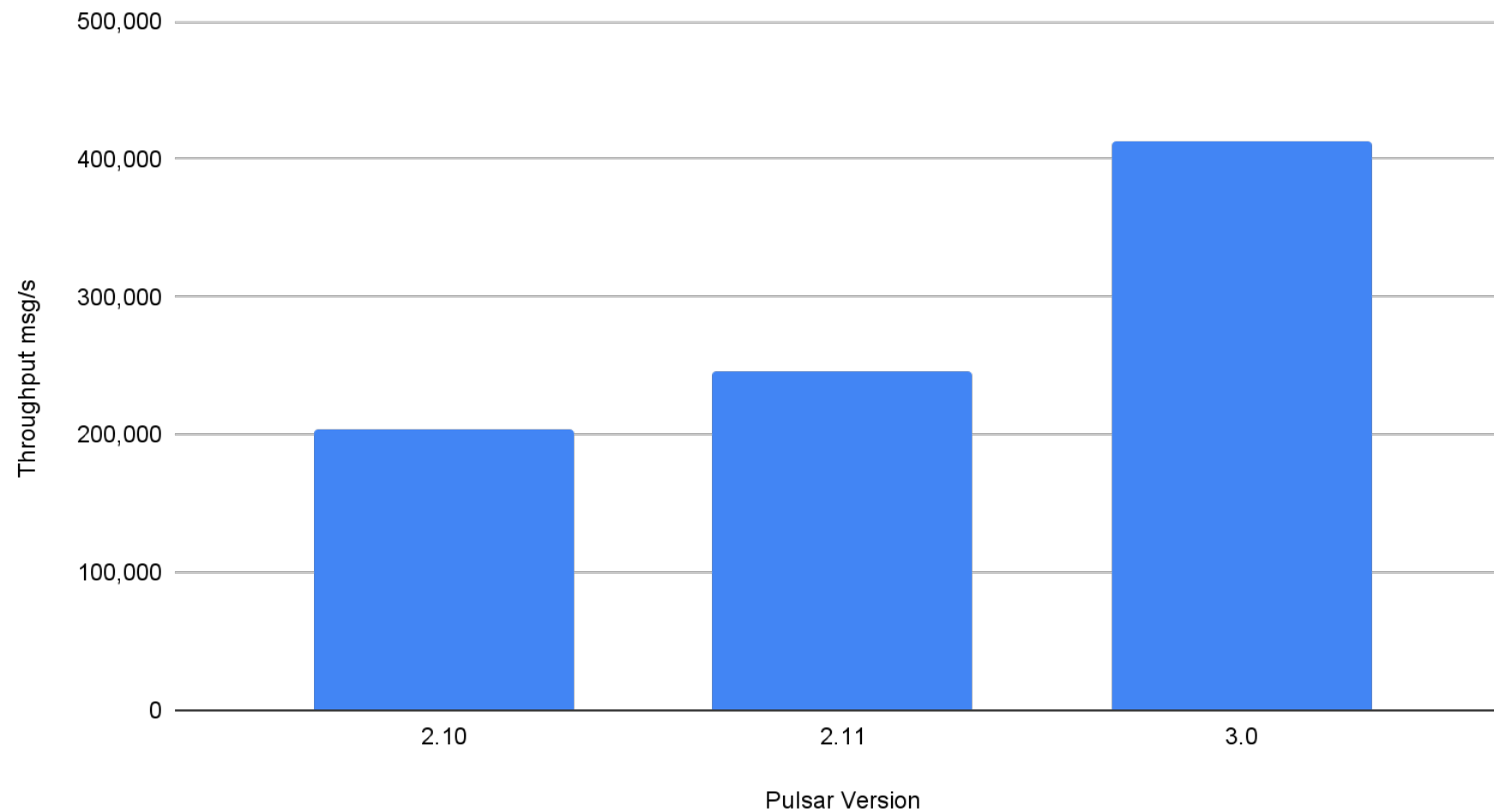
1 topic
100 MB/s



3 nodes cluster

10K topics
100 MB/s

Throughput over 10k topics - 3 nodes cluster



PIP-192 - New Load Balancer

PIP-195 - Scalable delayed delivery message

PIP-264 - Enhanced OpenTelemetry-based metrics

◦ ◦ ◦ ◦ ◦ ◦



Thanks