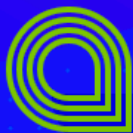




阿里巴巴-中间件 卫乐



Large-Scale Stream Processing in Alibaba



自我介绍

- 阿里巴巴技术专家
 - JStorm
 - 淘宝足迹
 - 推荐系统



团队介绍

- 团队职责
 - 应用开发
 - 技术支持
 - 集群运维管理
 - 平台/框架演进
 - 开源社区

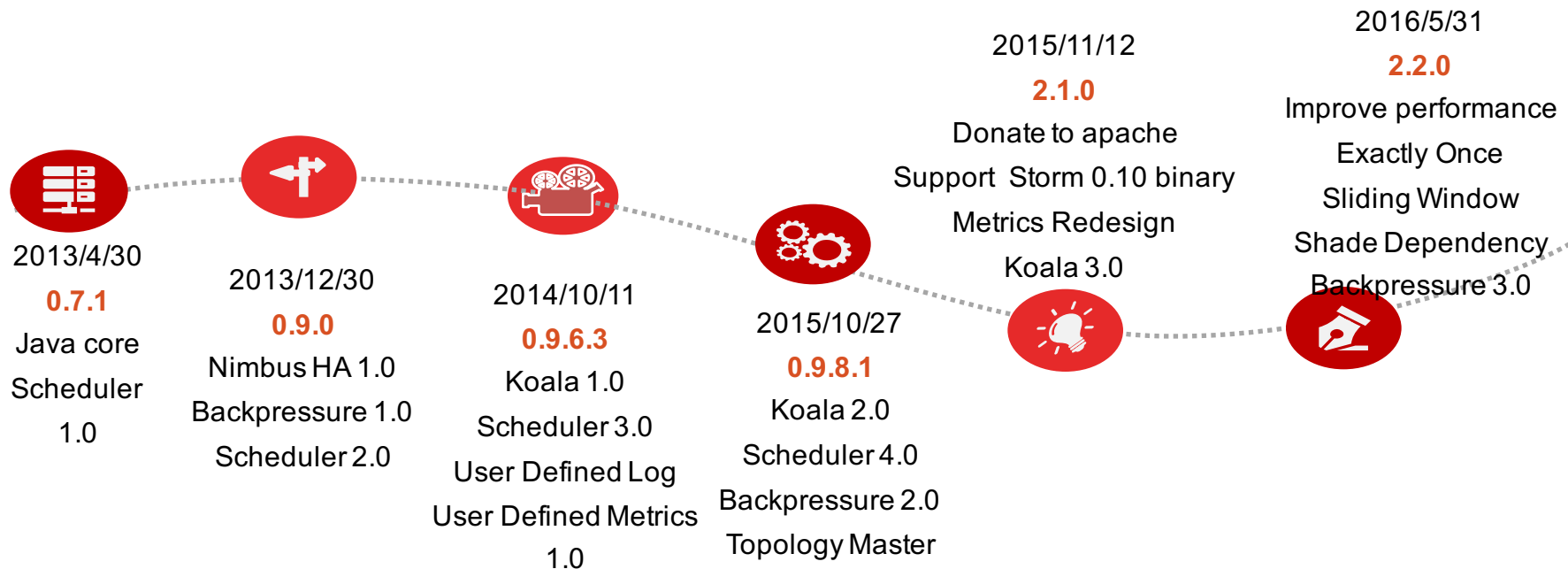
目录

- 发展历史
- 使用场景
- 架构
- JStorm增强
 - 性能
 - 稳定性
 - 新特性
- 开源社区
- JStorm Future

发展历史

JStorm Milestones

- 从2013-04-30起发布了超过25个版本
- 0.7.1/0.9.0/0.9.1/0.9.2/0.9.3/0.9.3.1/0.9.4/0.9.4.1/0.9.5/0.9.5.1/0.9.6/0.9.6.1/0.9.6.2/0.9.6.3/0.9.7/0.9.7.1/0.9.7.2/0.9.8/0.9.8.1/0.9.8.2/2.0.4/2.1.0/2.1.1/2.2.0/2.2.1/2.3.0**



使用场景

JStorm在阿里现状

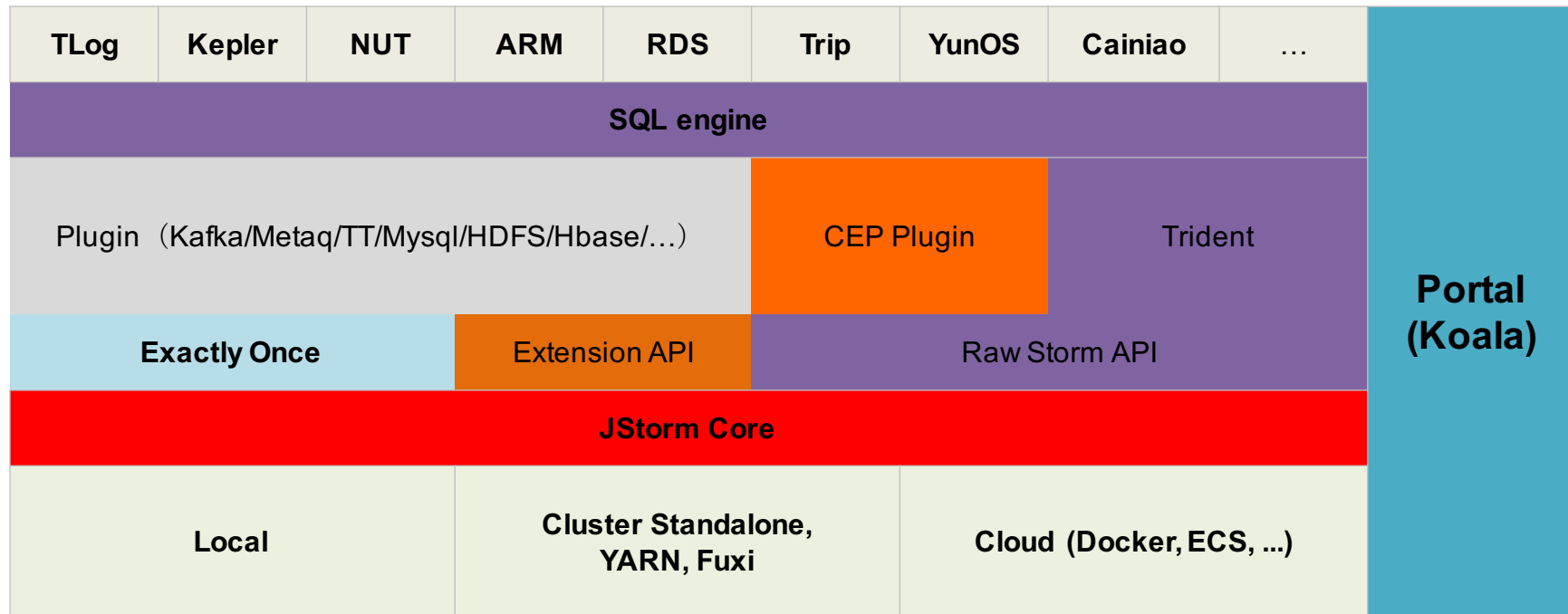
- 阿里集团、蚂蚁、优土、高德...
 - 4000+机器
 - Standalone
 - JStorm-on-YARN
 - JStorm-on-docker
 - 1500+应用, 2000+拓扑
 - 2PB+数据/天

应用场景

- 反欺诈：Nut/Velocity
- 审计：阿里妈妈广告/P4P，AMG
- 数据统计分析：Eagleeye，AE，BI
- 监控：Tlog，Rds-monitor，Oceanbase-monitor, Cainiao Radar, YunOSMonitor
- 数据同步：RDS-Log-Sync, Unify-Log
- 实时推荐：Alipay 1315, Alipay Hyperloop, TPP
- 应用调度：AE邮件实时分析，网销宝

架构

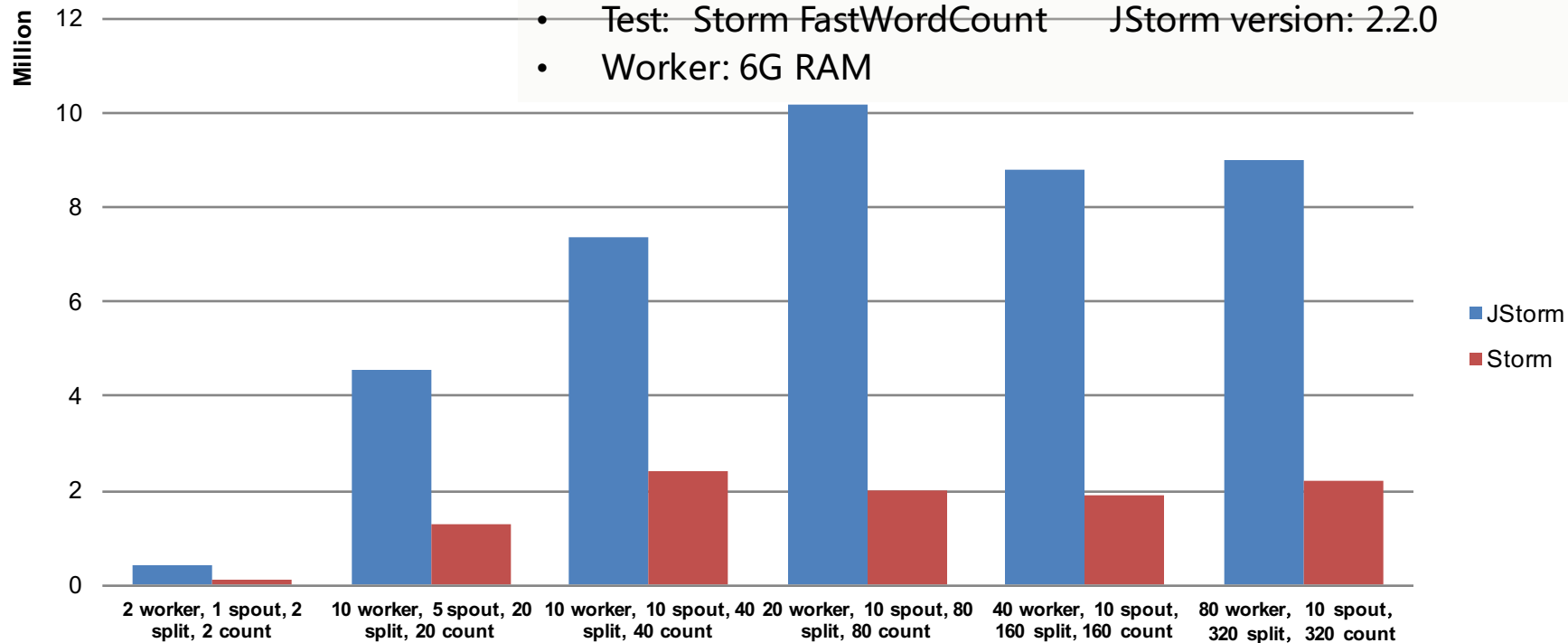
JStorm架构



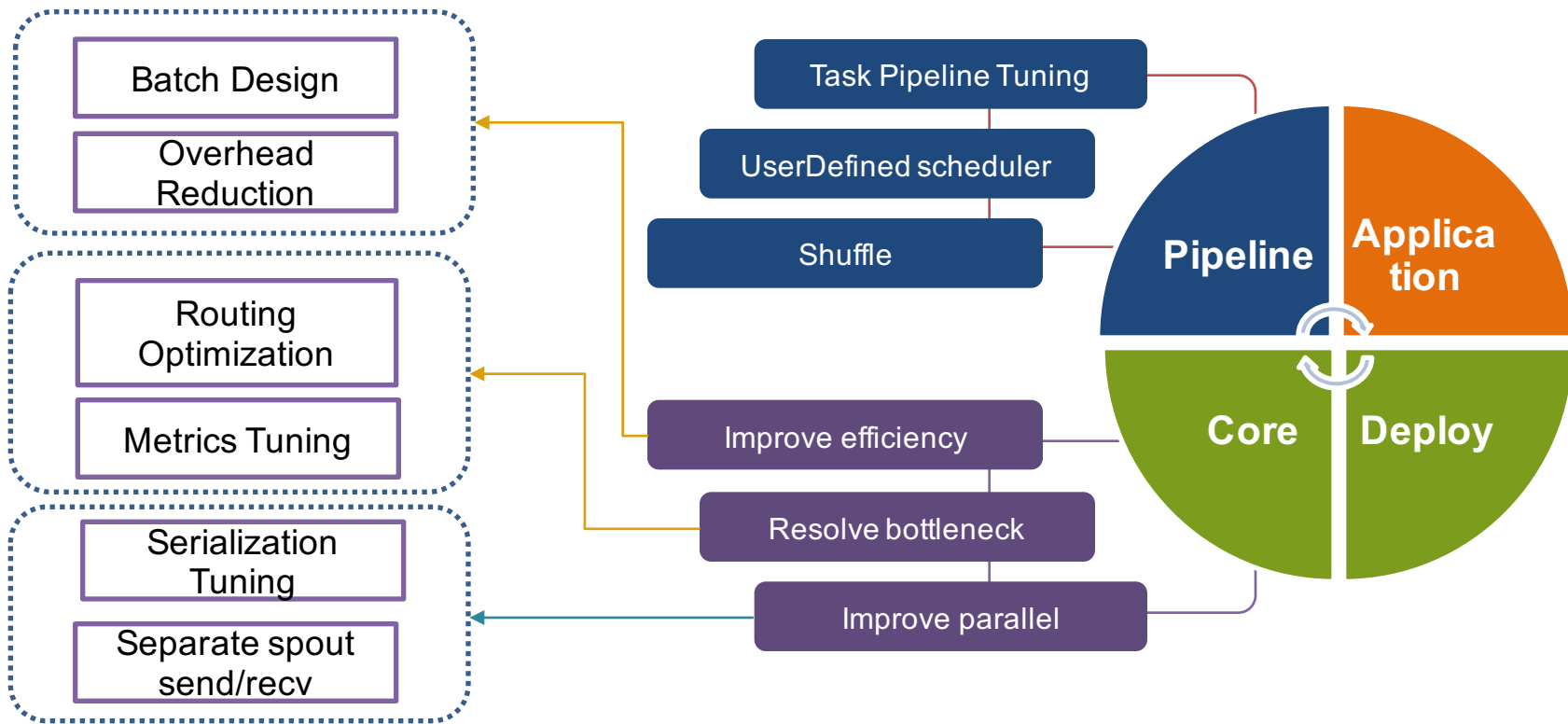
JStorm增强

性能

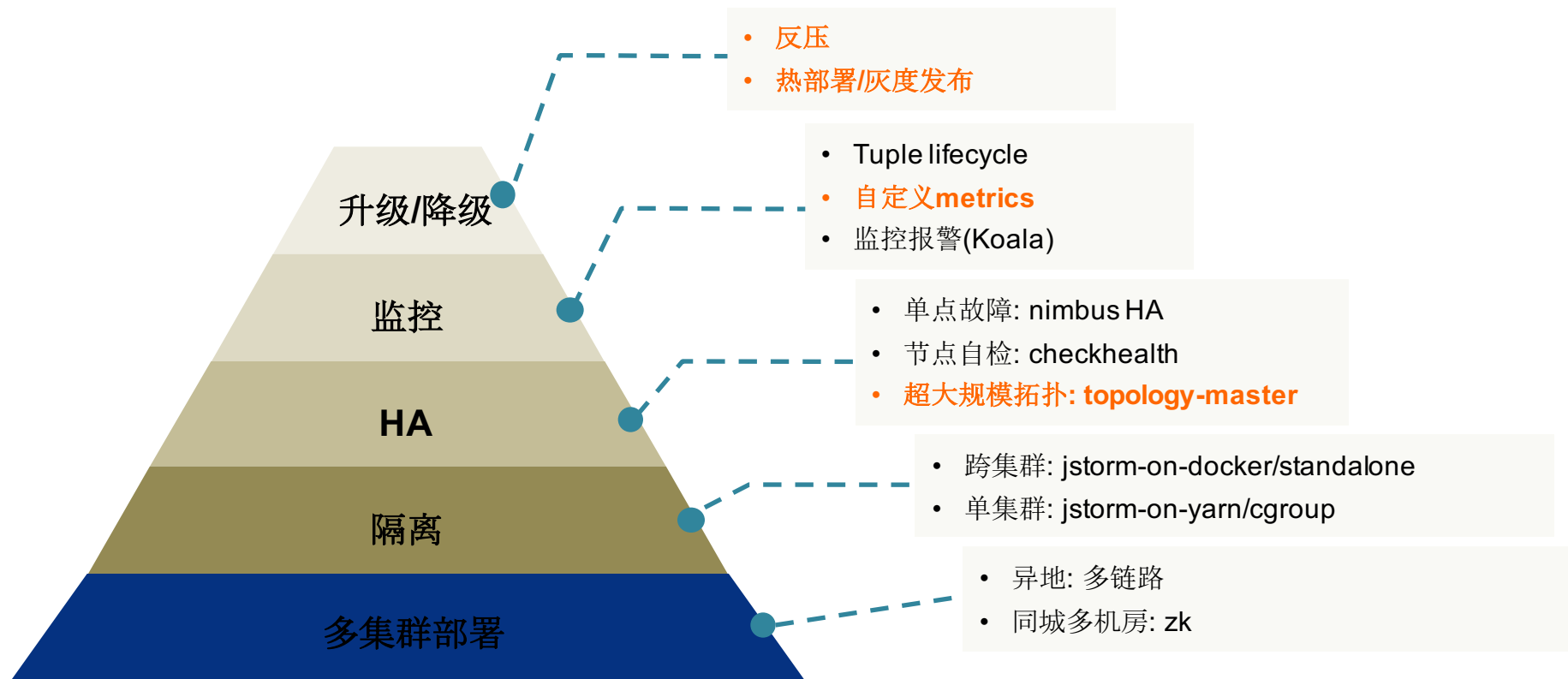
- HW: 32 CPU core/128G
 - Test: Storm FastWordCount
 - Worker: 6G RAM
- Storm version: 1.0.2
JStorm version: 2.2.0



How do we optimize?

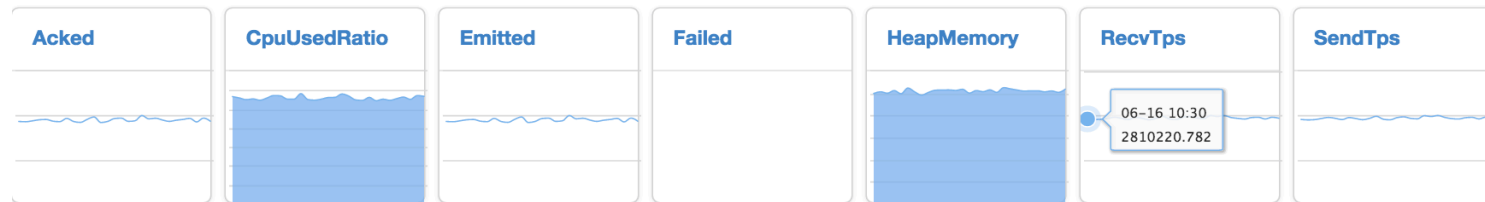


稳定性

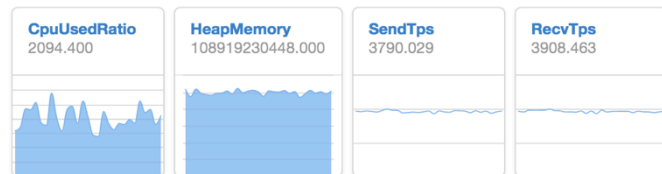


监控系统

Cluster Metrics



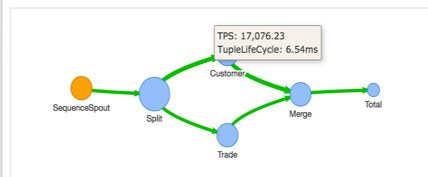
Topology Metrics

[指标含义](#)

Emitted 拓扑操作

显示历史 | 日志搜索 | 监控报警 | 修改信息 | 日志级别 | 结束拓扑 | 重启拓扑 | 一键优化 | 其它操作...

拓扑结构图

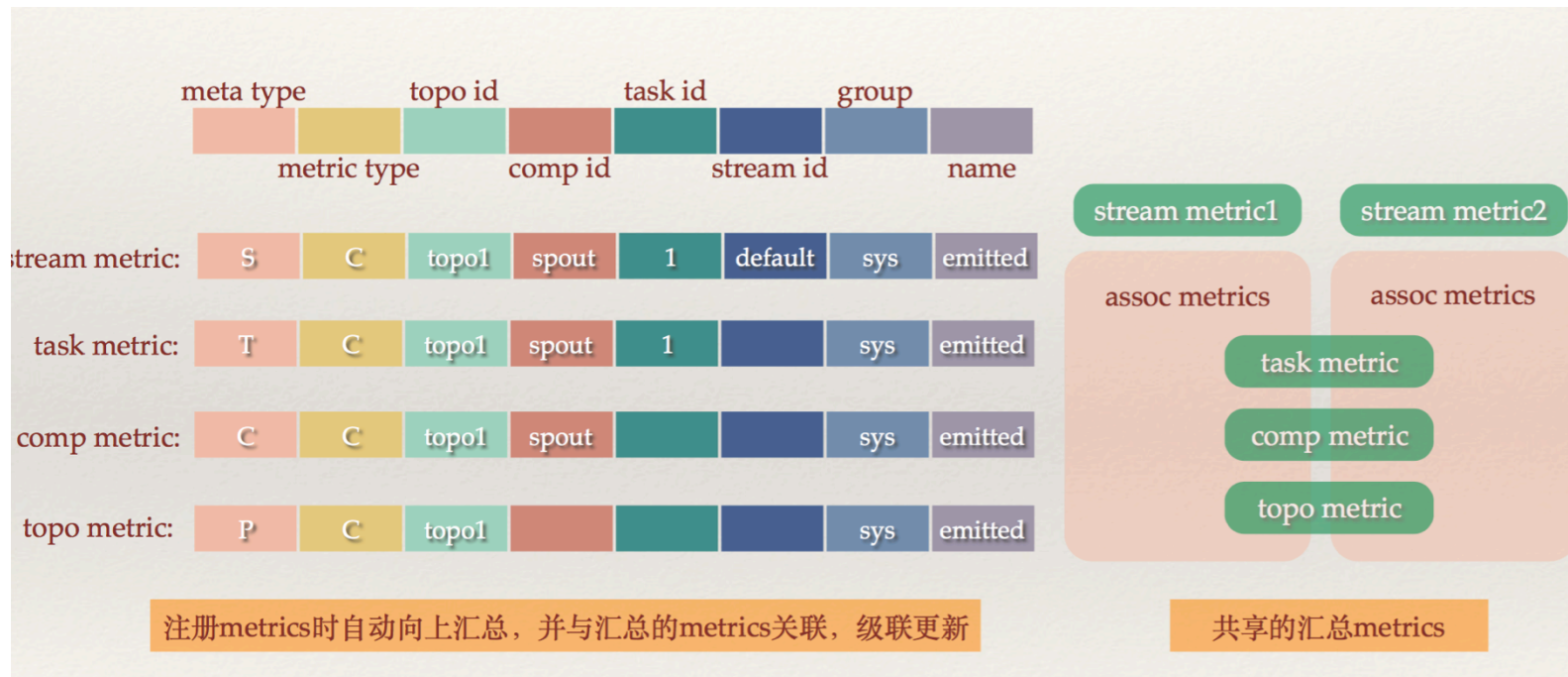


Customer -> Merge

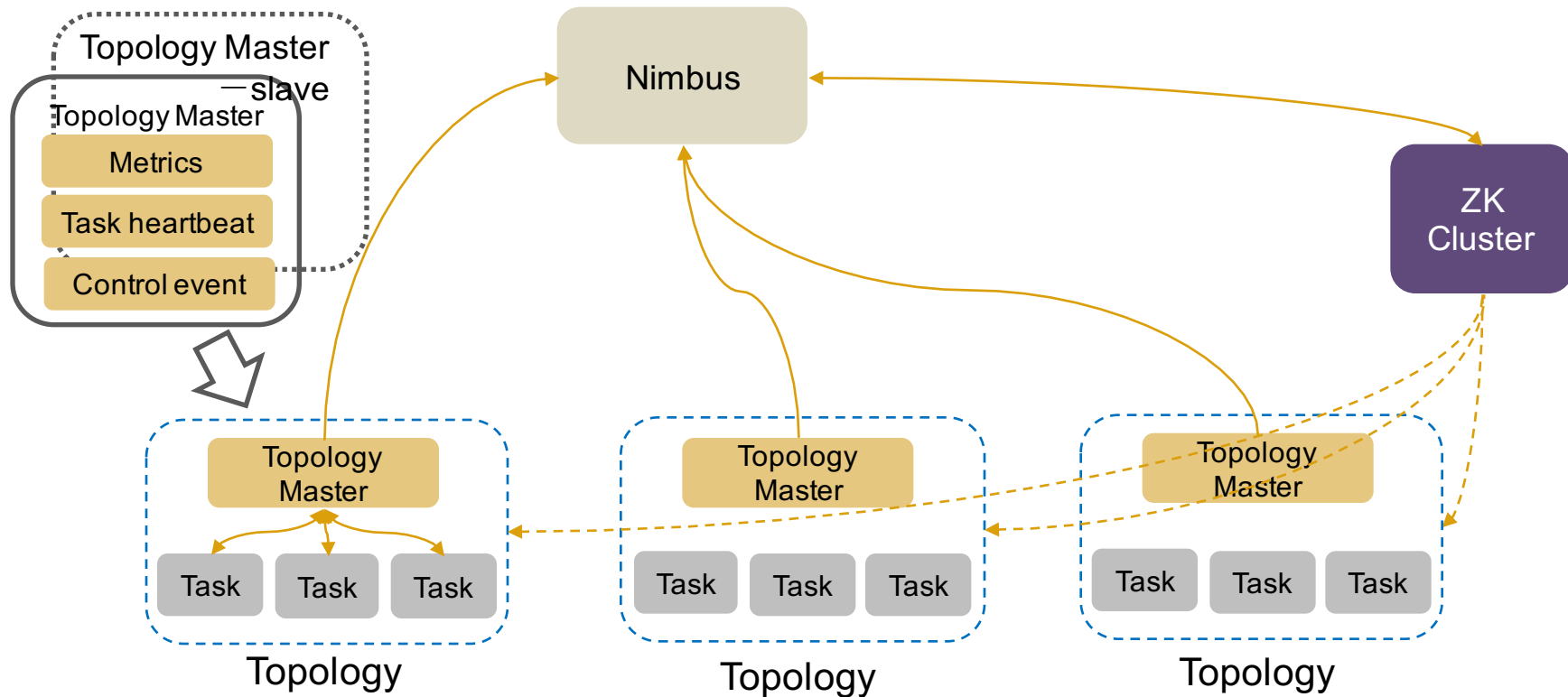
Window	TPS	TupleLifeCycle(ms)
1 min	17,076.43	7.83
10 min	16,978.8	6.75
2 hour	8,447.99	7.38
1 day	17,294.04	6.43

ackner has been removed for clarity, the stream color represent tuple life cycle time.

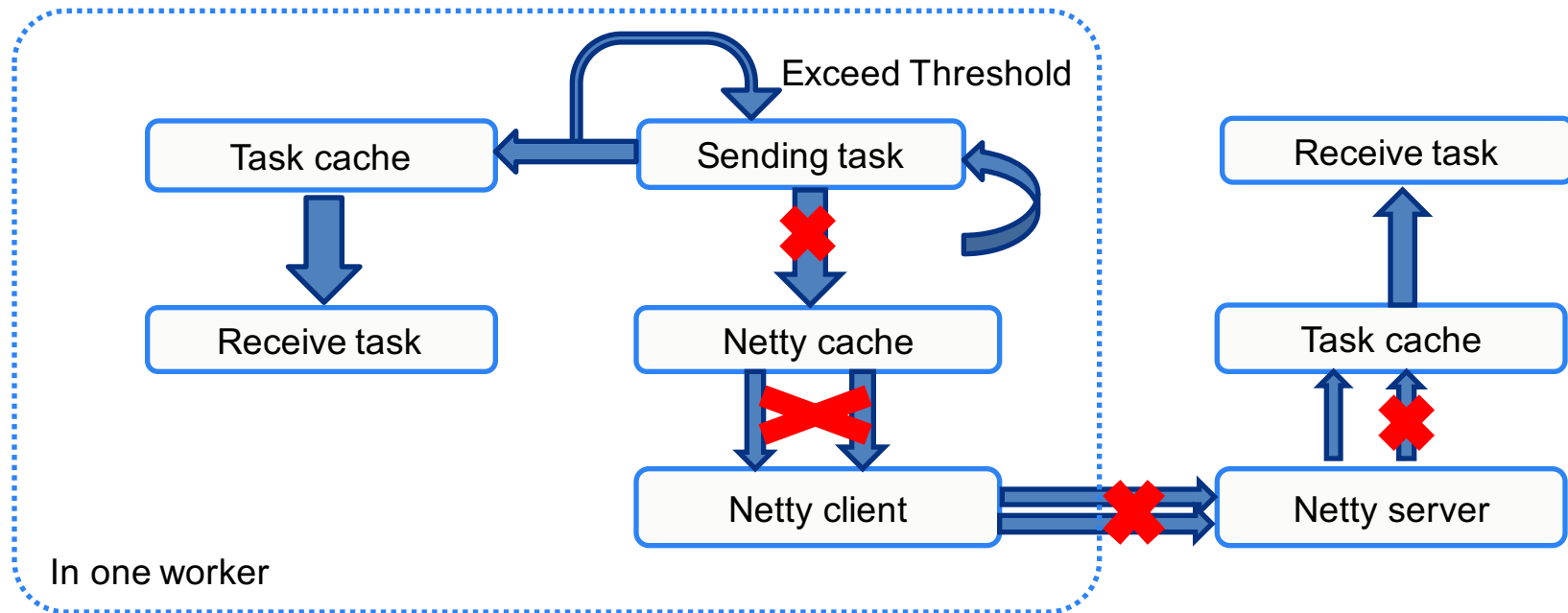
自定义监控



大规模部署

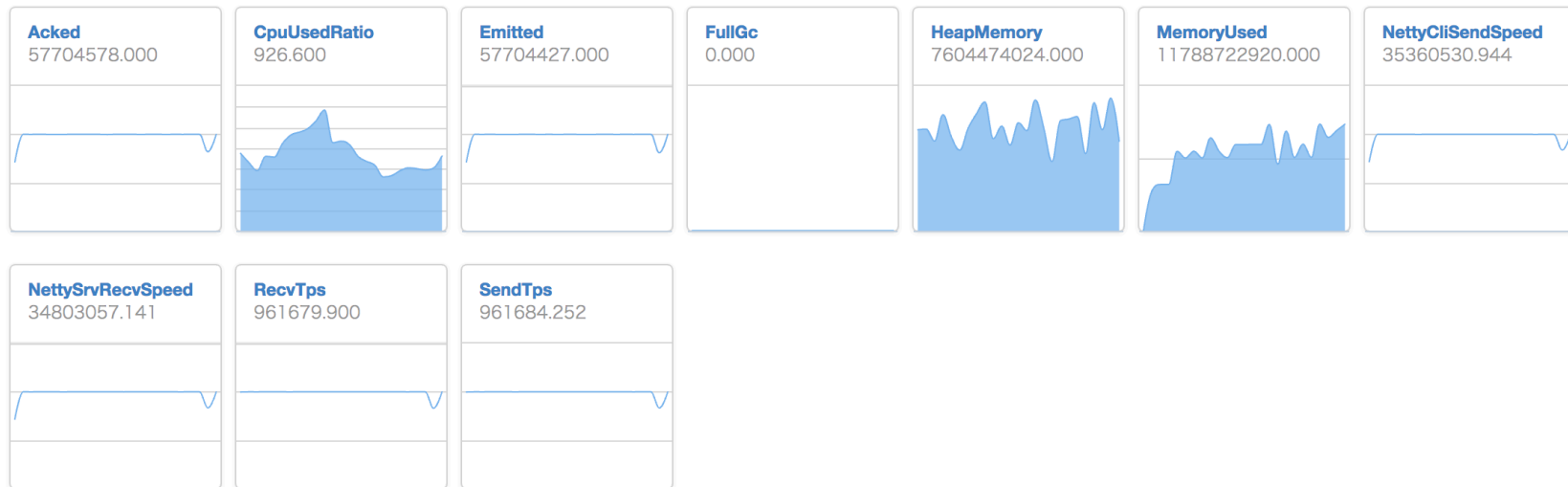


反压



灰度发布&热升级

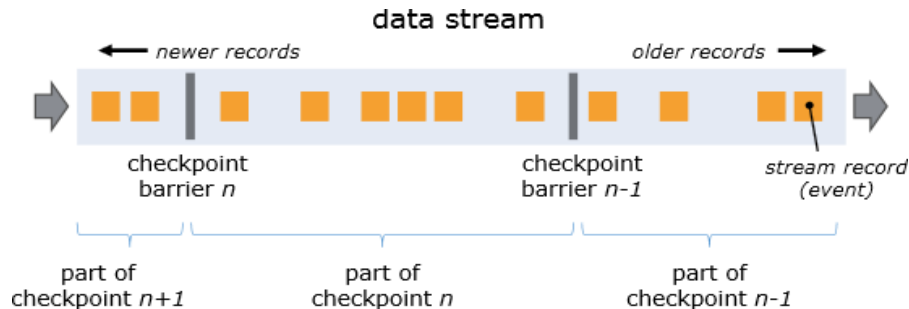
Topology Metrics

[指标含义](#)

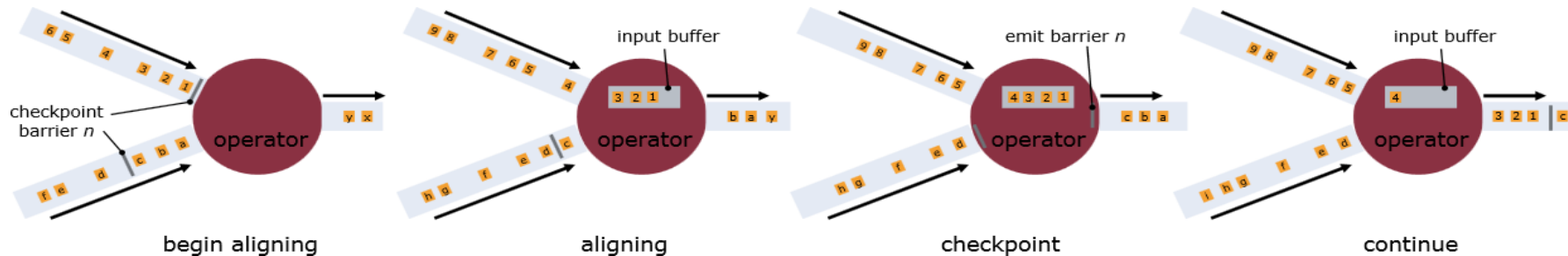
Exactly-Once

Chandy-Lamport Algorithm

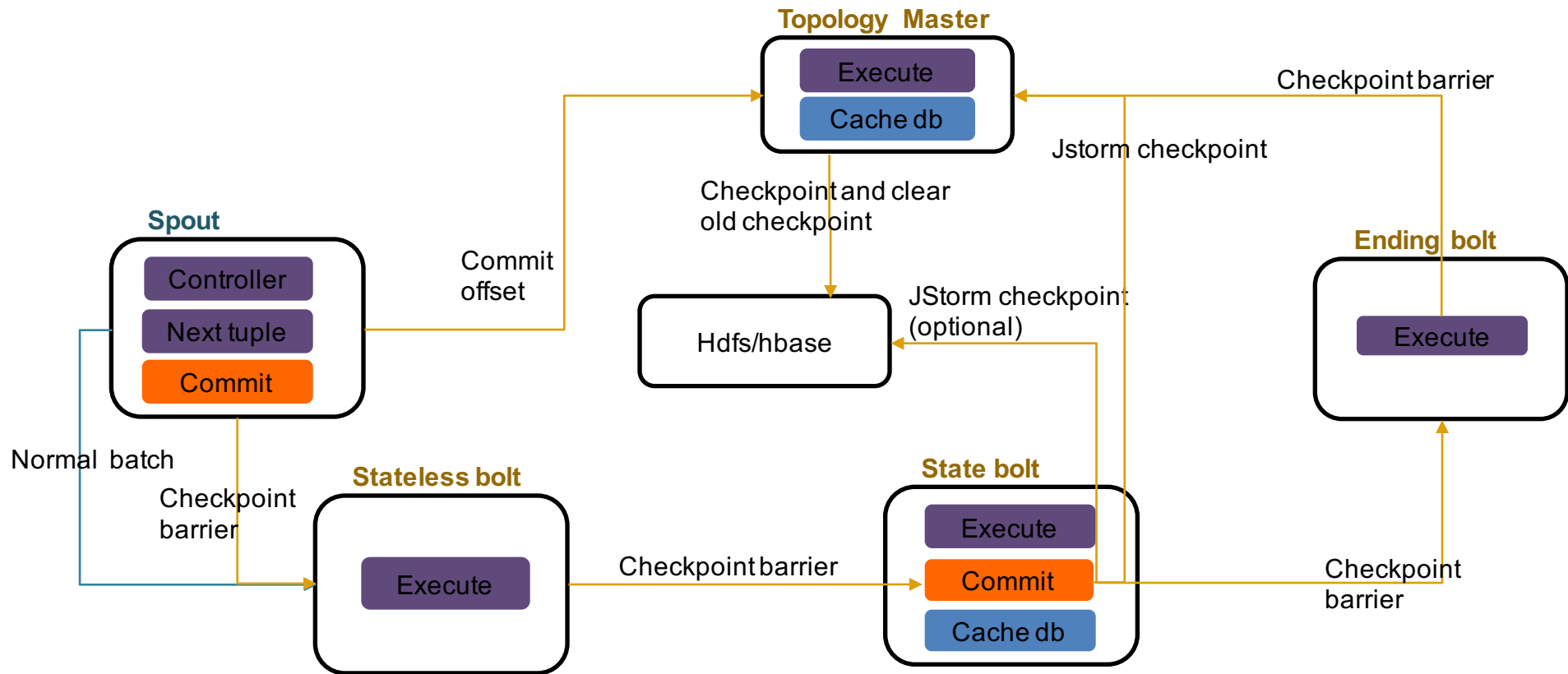
Barrier



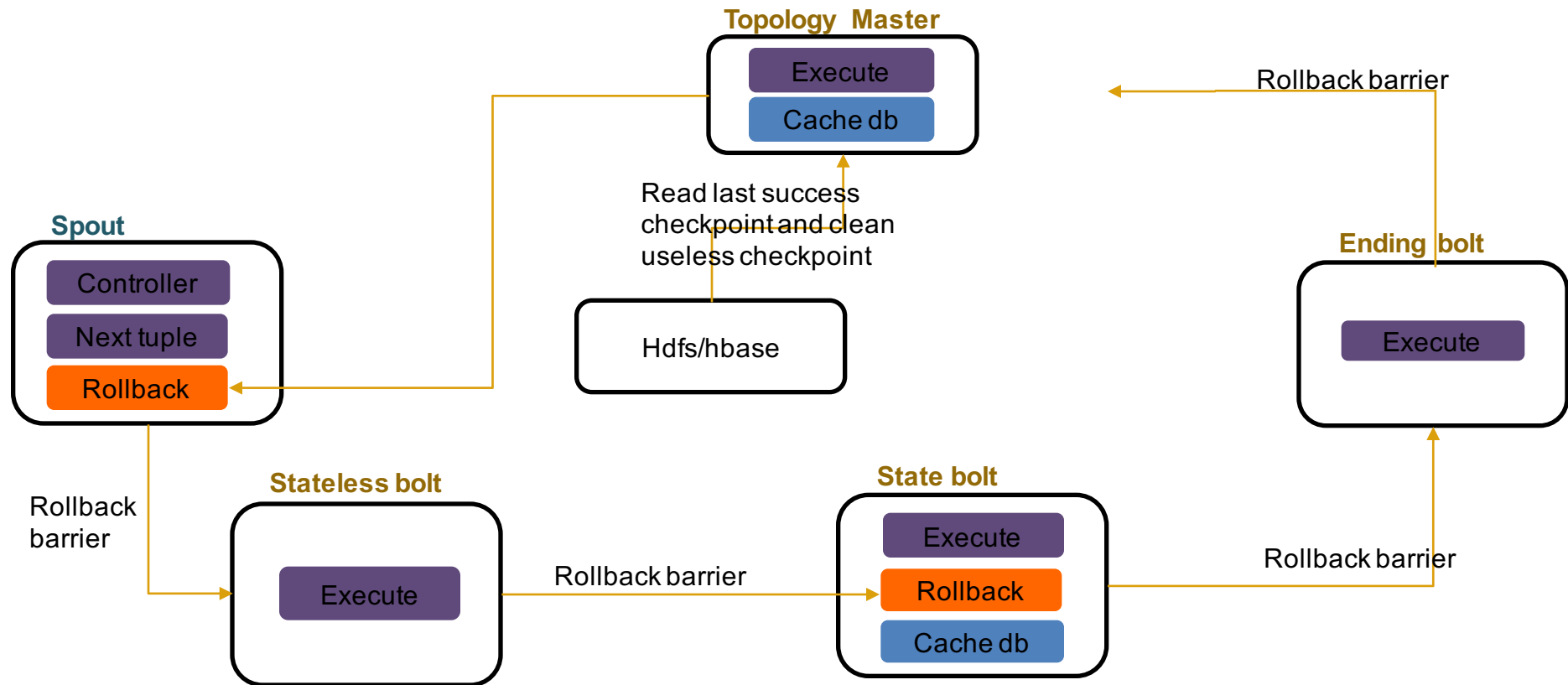
Stream Aligning



Exactly-Once



Exactly-Once



其他feature

- 自定义调度
- 动态调整日志级别
- 自定义日志
- 用户自定义metrics
- 优雅shutdown
- 动态扩容、更新配置
- ...

新的窗口机制

- 重构的窗口机制，支持：
 - Processing time
 - Ingestion time
 - Event time
 - Watermark (late element)
 - Window early fire

开源社区

开源社区现状

50+公司



JStorm Future

- Apache Beam JStorm runner
- JStorm SQL
- 编程框架：易于开发、调试，性能最优化



THANKS

