

大数据在趣头条的演进

Kafka读写分离、Hadoop治理 机器学习平台

James Yu

趣头条大数据负责人



关注 QCon 公众号

收获国内外一线大厂实践 与技术大咖同行成长

✓ 演讲视频 ✓ 干货整理 ✓ 大咖采访 ✓ 行业趋势



自我介绍

虞沐, James。

- 2001 年安徽大学本科毕业后, 先后在上海 HP 和 SAP 工作。
- 2007 年赴美国硅谷工作, 先后在 eBay, 赛门铁克, 三星从事分布式系统开发。
- 2013 年开始从事大数据和公有云相关开发, 先后在百度美国, xAd, 和 Facebook, 从事技术管理和架构师工作。
- 2018 年底回到上海工作, 在趣头条担任大数据技术总监, 经历公司快速发展, 带领团队自研开发多个大数据和机器学习平台。

目录

- 趣头条简介
- Kafka读写分离
- Hadoop治理
- 机器学习平台

趣头条简介

- 2016年6月8日，趣头条1.0正式上线。
- 2018年3月16日，趣头条宣布完成超过2亿美元B轮融资，整体投后估值超过16亿美元。
- 2018年9月14日，趣头条正式挂牌纳斯达克交易所。

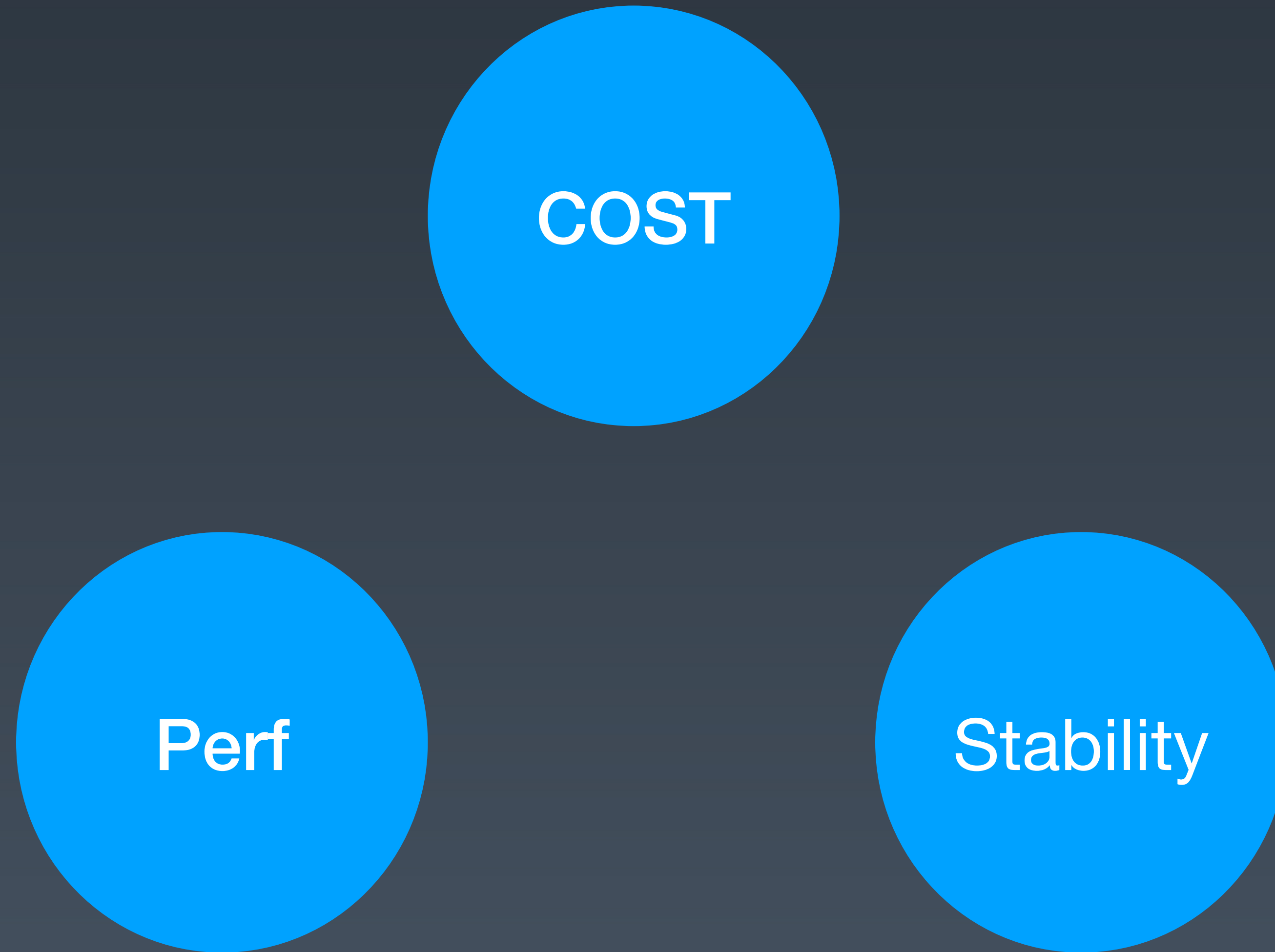
趣头条致力于打造一款新形式的资讯阅读软件，以平台、媒体和共赢的方式。以移动应用为载体进行内容创造、资讯阅读，提供更多有用、有趣、有益的内容给大家。

大数据部 简介

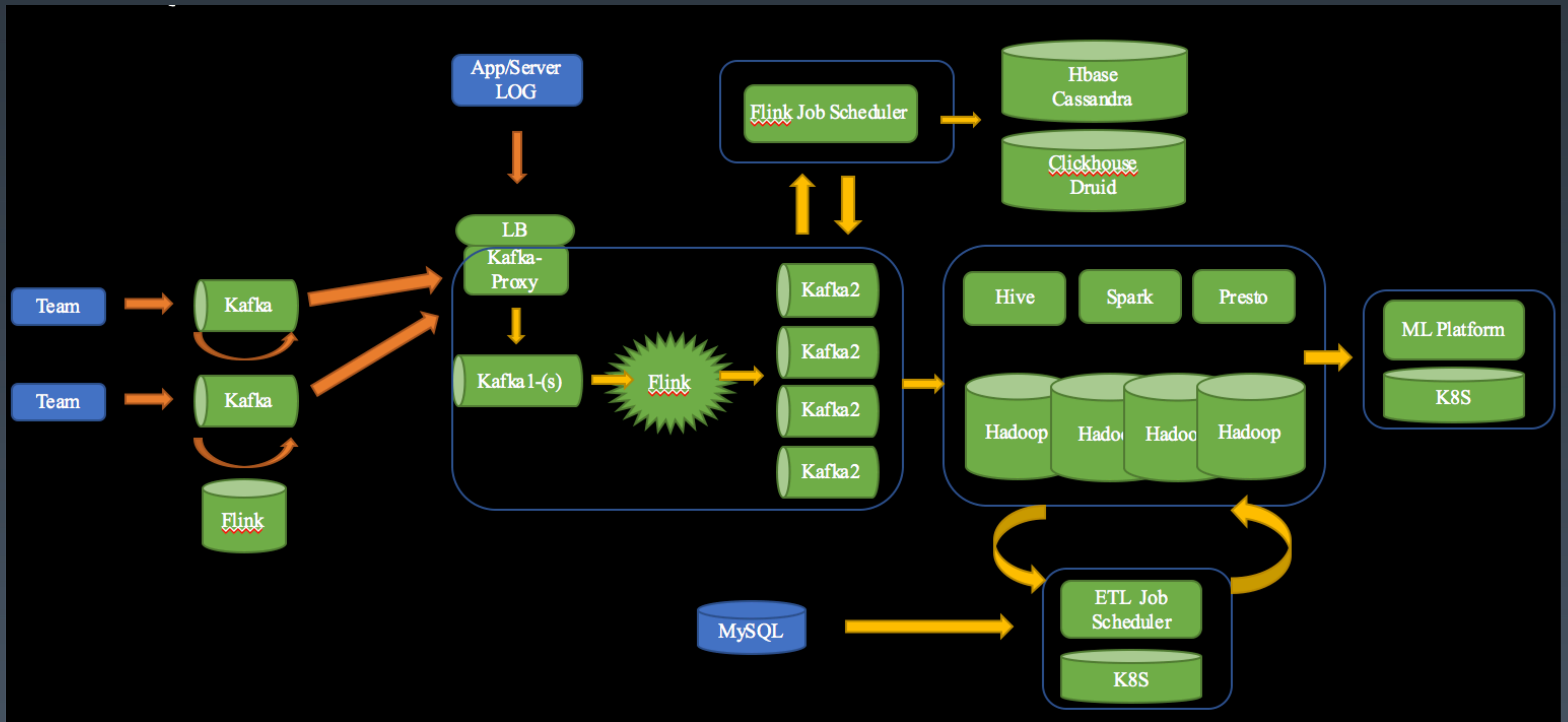
趣头条在2018-2019年经历了业务的高速发展，主App和其他创新App的日活增加了10倍以上。

相应的，大数据系统和平台也从最初的100台机器，增加到2000台以上，技术栈从单一的离线数据报表，发展到离线+实时+机器学习的完整系统。这个分享将从3个主要方面，阐述大数据系统的演进路线，和经验分享。

CPS

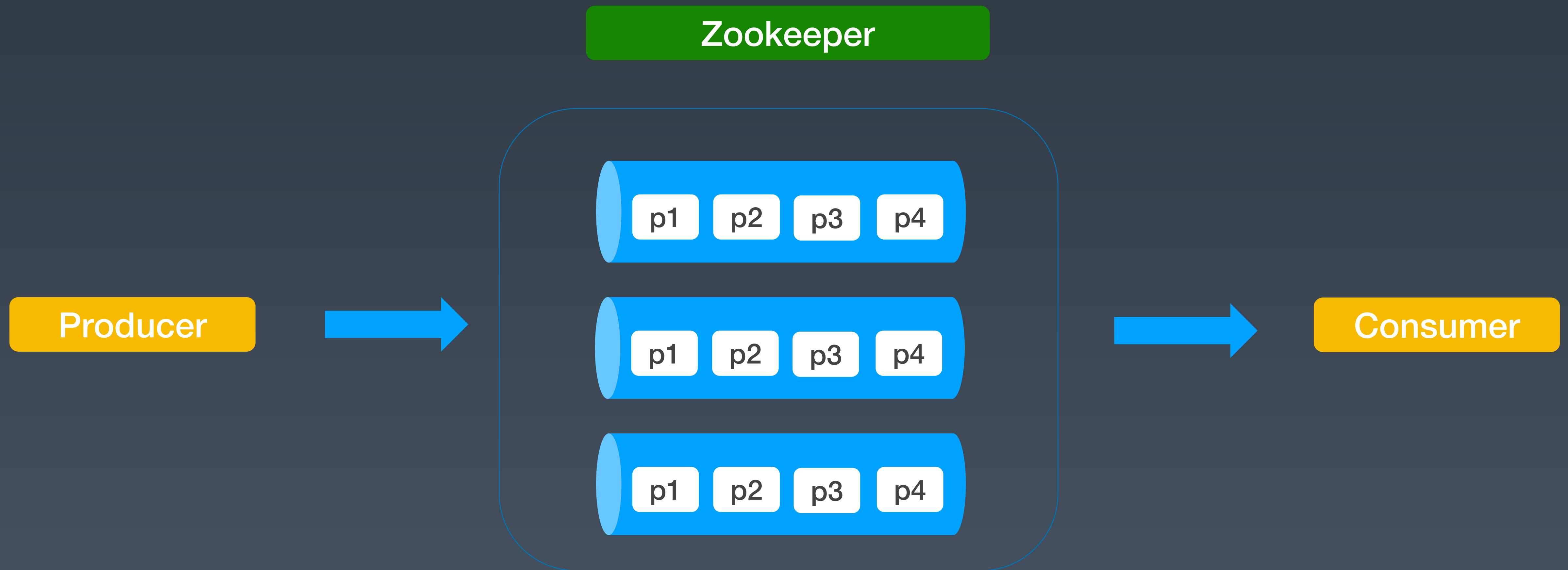


系统总图

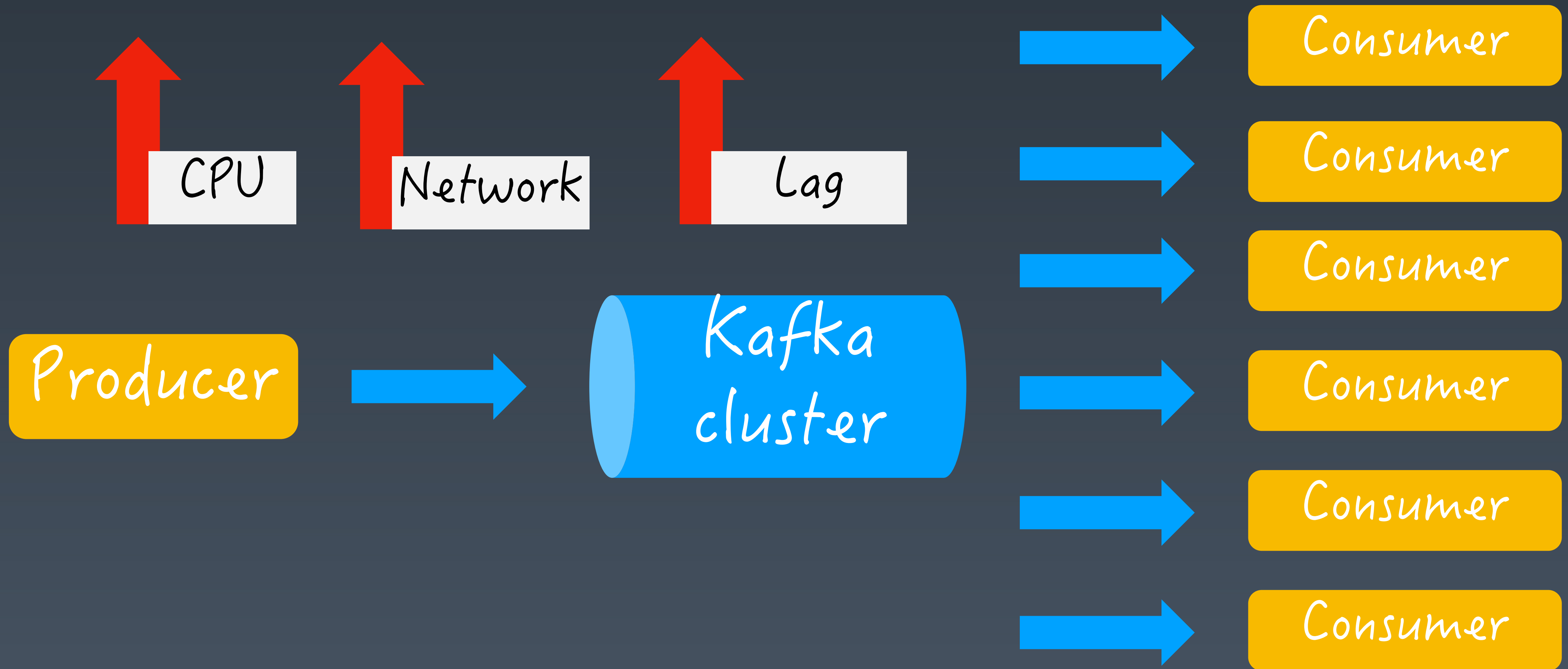


Kafka读写分离

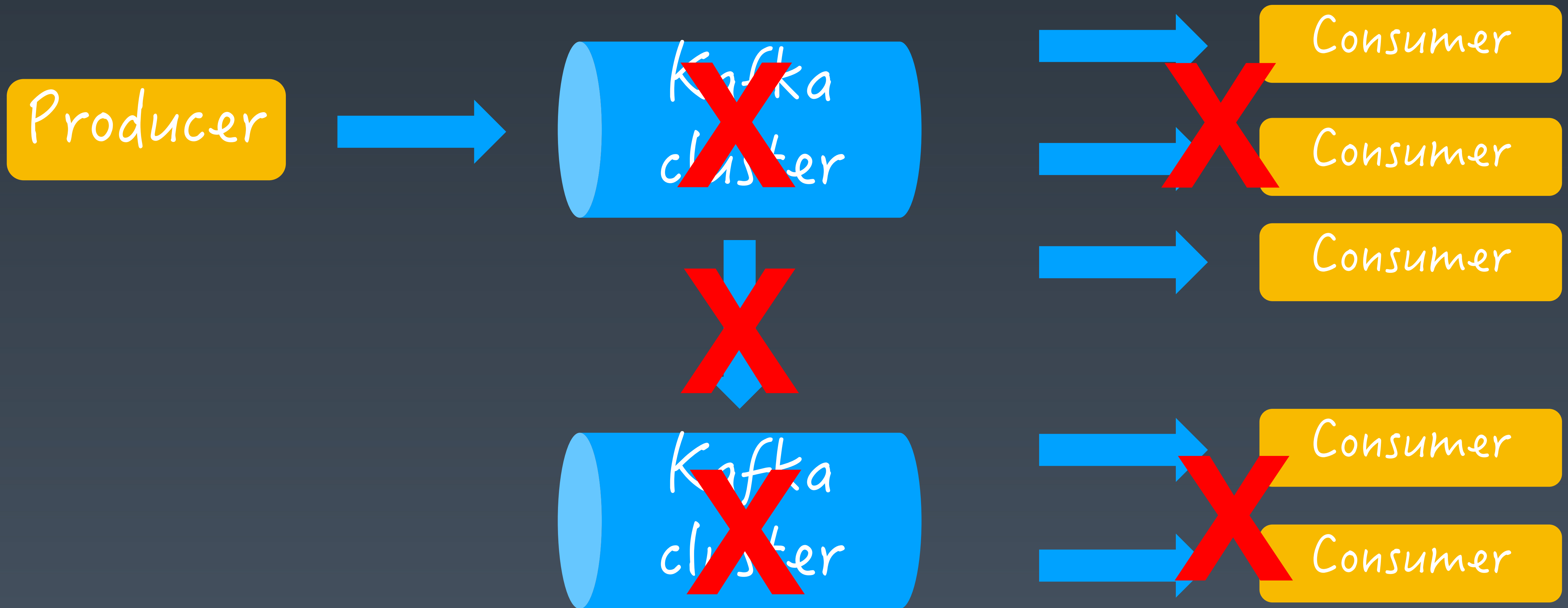
Kafka (Producer, Broker, Partition, ZK, Consumer)



Over-consumed topics

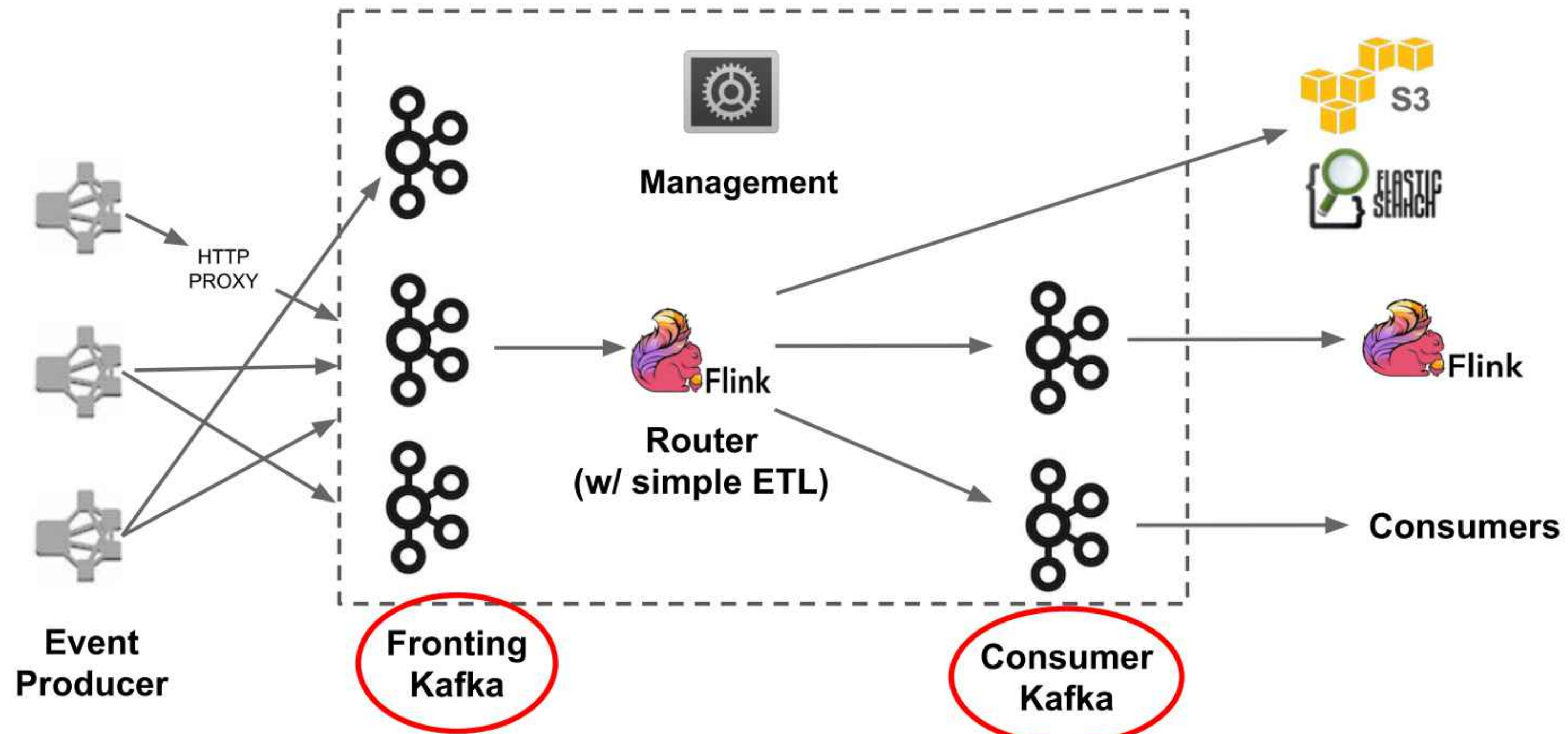


Over-consumed topics

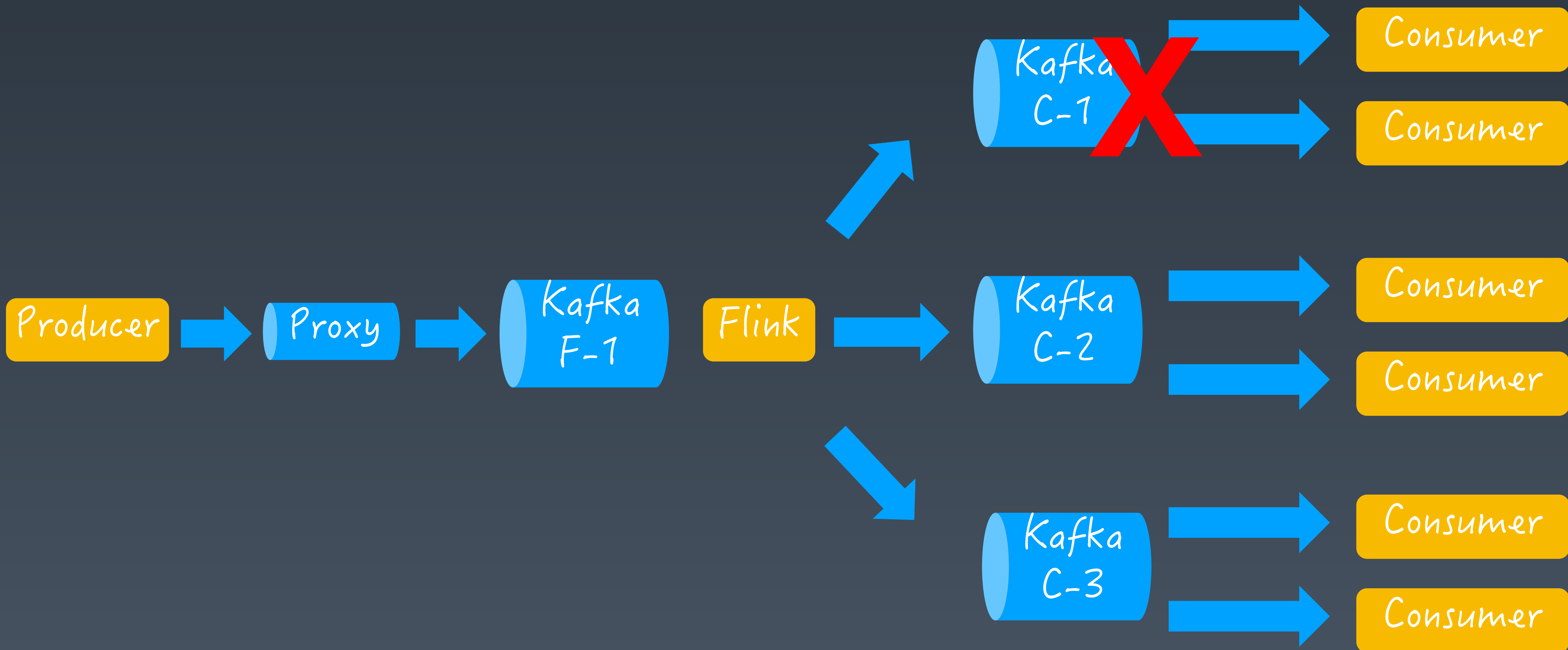


Kafka 读写分离 Netflix-KeyStone



Multi-Cluster Kafka Service At Netflix



Over-consumed topics



Kafka硬件选项

阿里云	旧	新
机型	ecs.d7ne.8xlarge	ecs.i2.8xlarge
CPU	32 	32
Mem	128G	128G
DISK	16 * 5.5TB HDD	4 * 1.8TB NVMe SSD 
Network	20Gb	6Gb

SSD

	旧	新
Retention	3 days	6-12 hours
Throughput	2GB	190MB

数据压缩

gzip

snappy

lz4



zstd



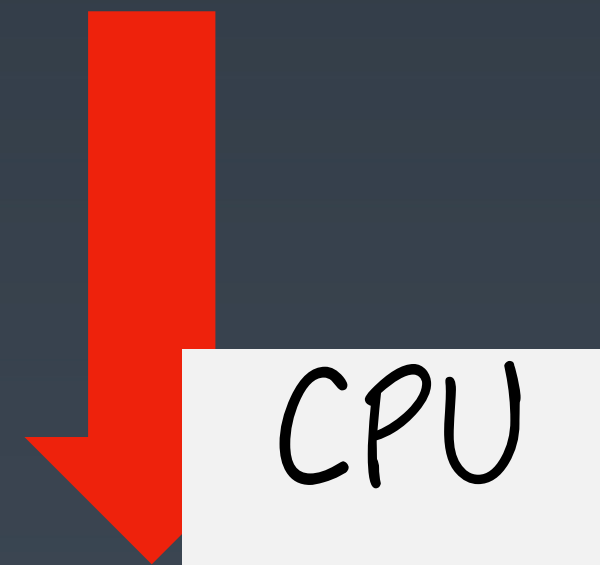
压缩比 40%



CPU 5-10%



性能提升



CPU



Network



Lag

系统监控

- 机器
CPU, 内存, 硬盘容量, 磁盘间平衡, 网卡出入流量
- 服务
消费延迟/积压, 数据流量速率, 数据压缩, consumer_group
注册报备

TODOs

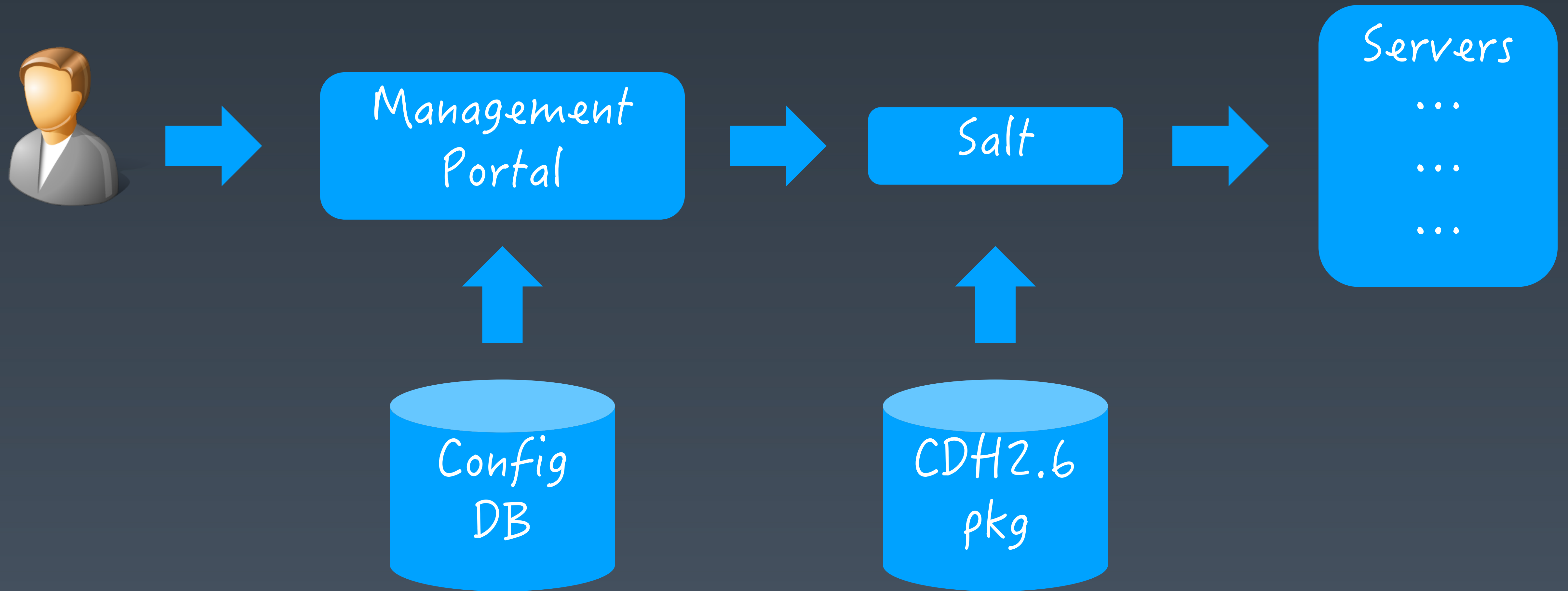
- 消费Consumer SDK封装, 自动更新集群配置, 系统迁移对用户透明。
- Partition Migration Tool开发, 避免大批量Partition迁移导致网络风暴和服务波动。
- 数据流量波动监控

Hadoop治理

存在的问题

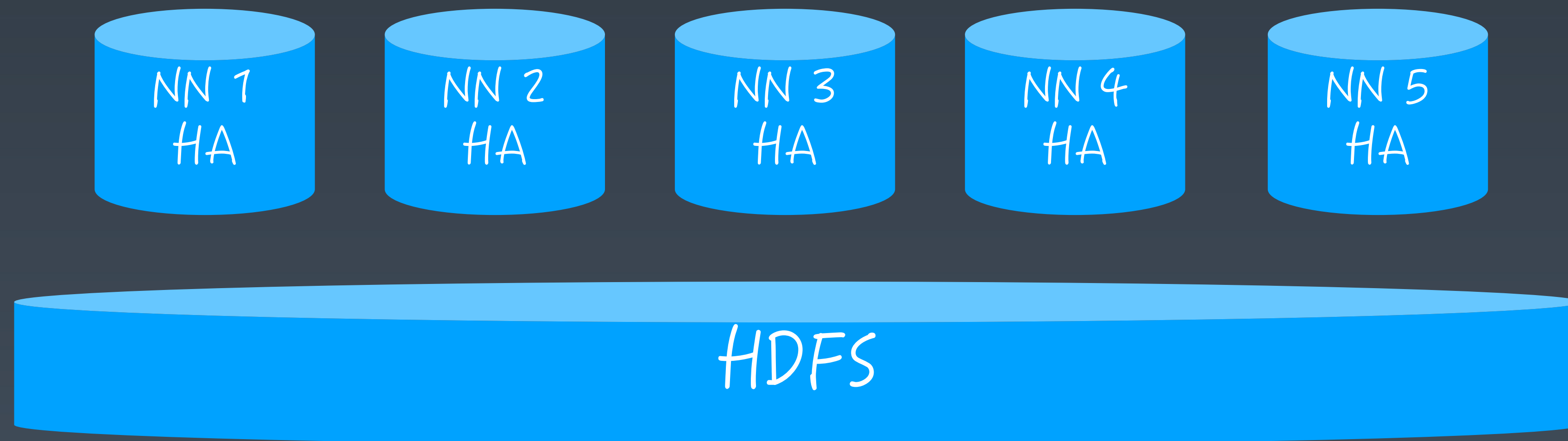
- HDFS单Namenode压力大, 响应慢, Callqueuelength高
- 队列分配混乱, 高优任务没有资源保障
- 系统监控报警不完善, 被动解决事故, 无法主动风险预警
- 计算引擎重依赖Hive, 任务效率低
- 云EMR无法支持大规模集群, 设计和功能上有缺陷
- 缺乏统一任务调度平台

自研集群管理平台

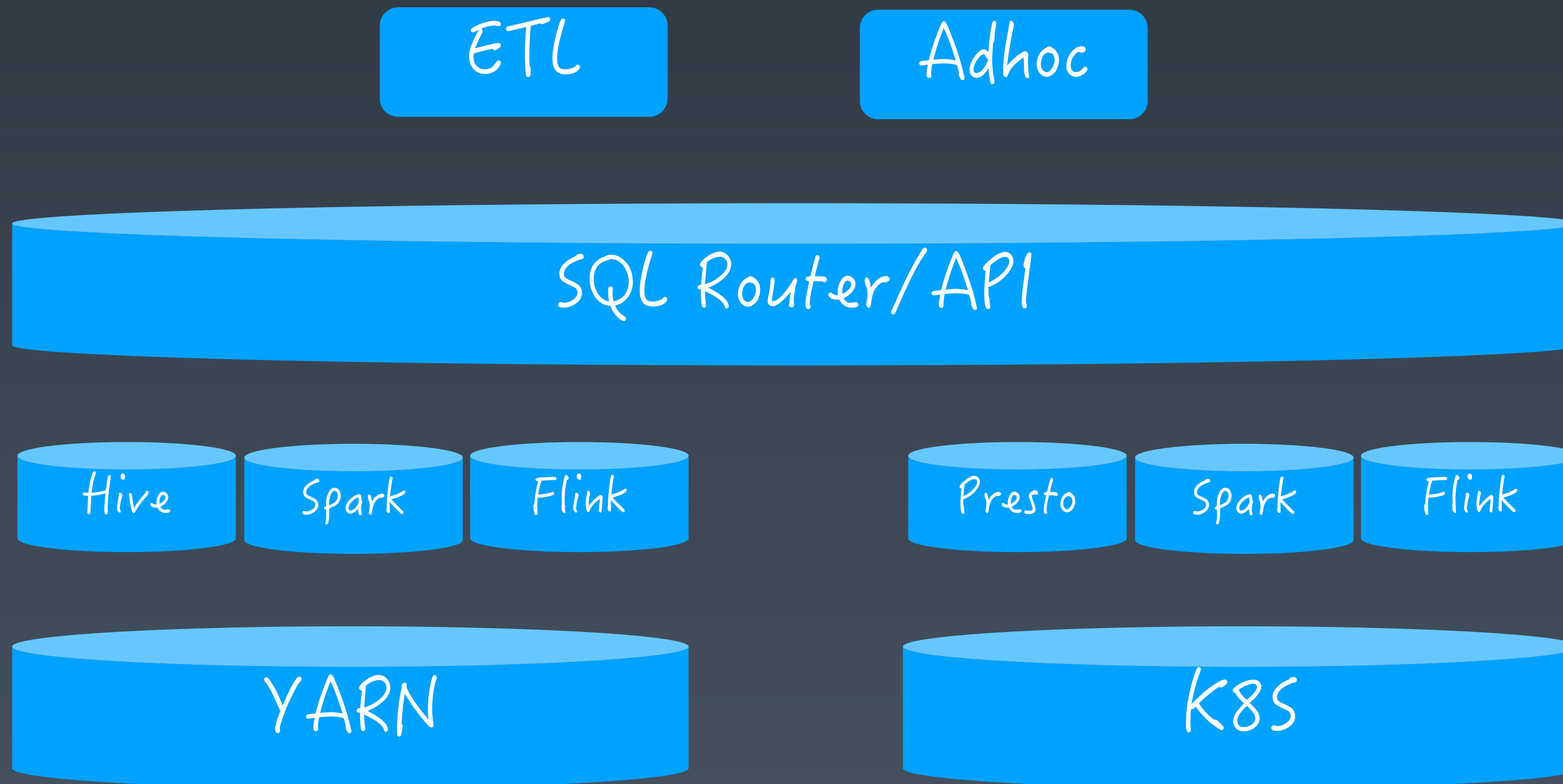


HDFS Federation HA

HDFS
Quota
Control



Resource Manager + Computation Engines



Master服务独立部署

NameNode

JournalNode

Resource Manager

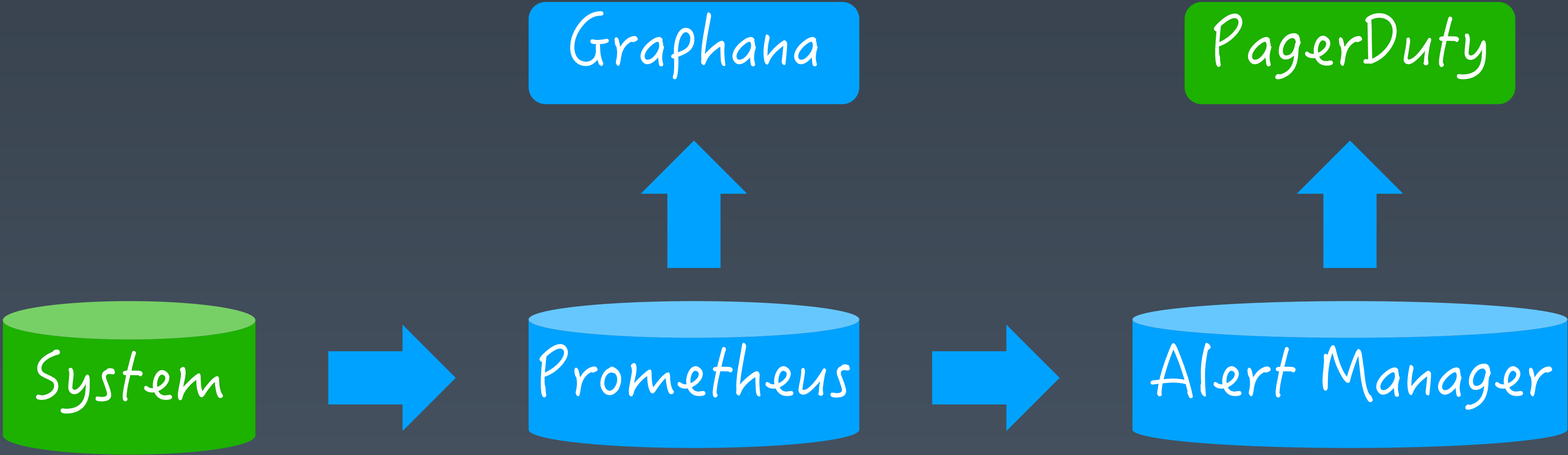
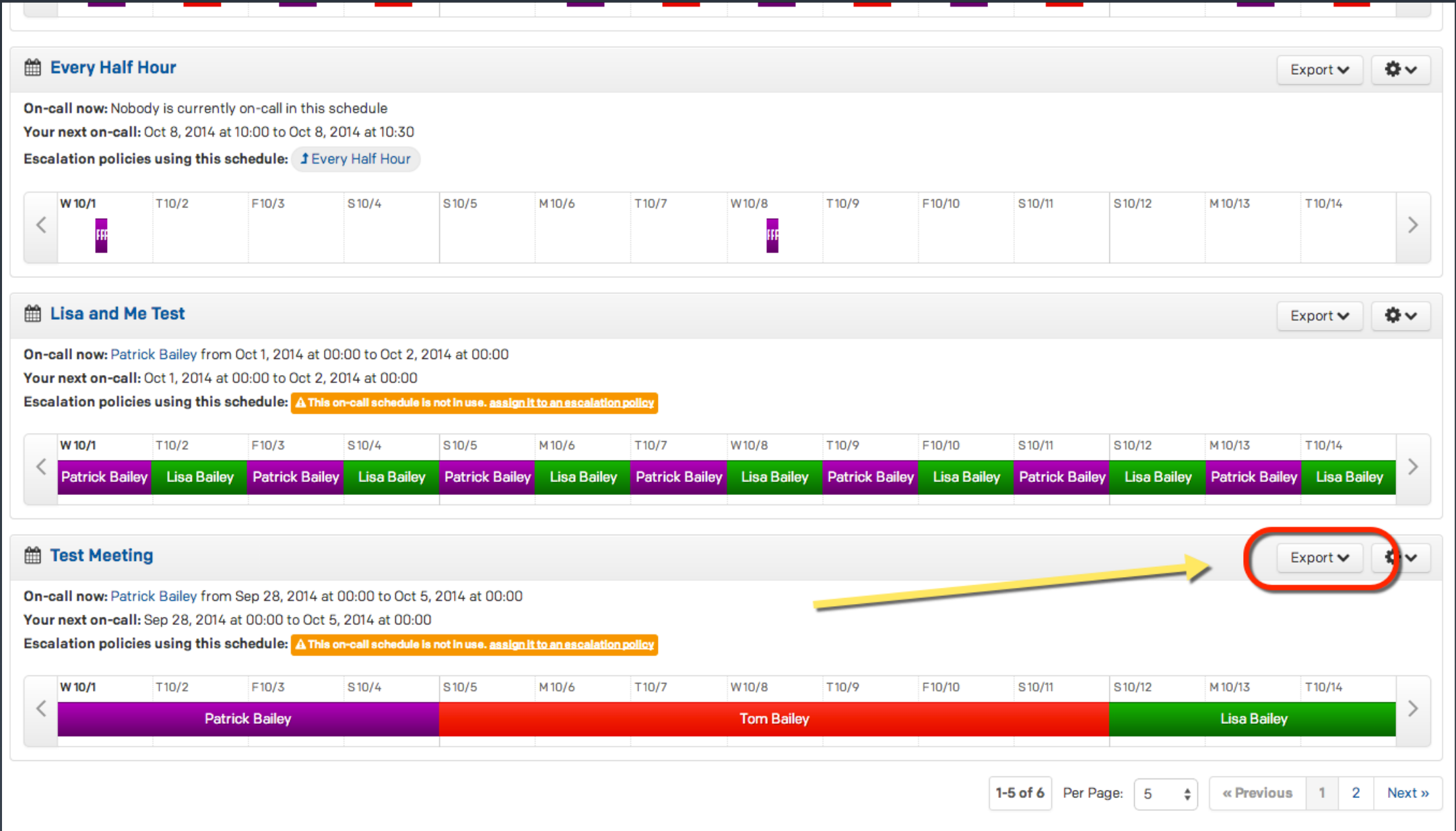
Job History

ZK

Hive Metastore

Hiveserver 2

监控报警



数据生命周期管理

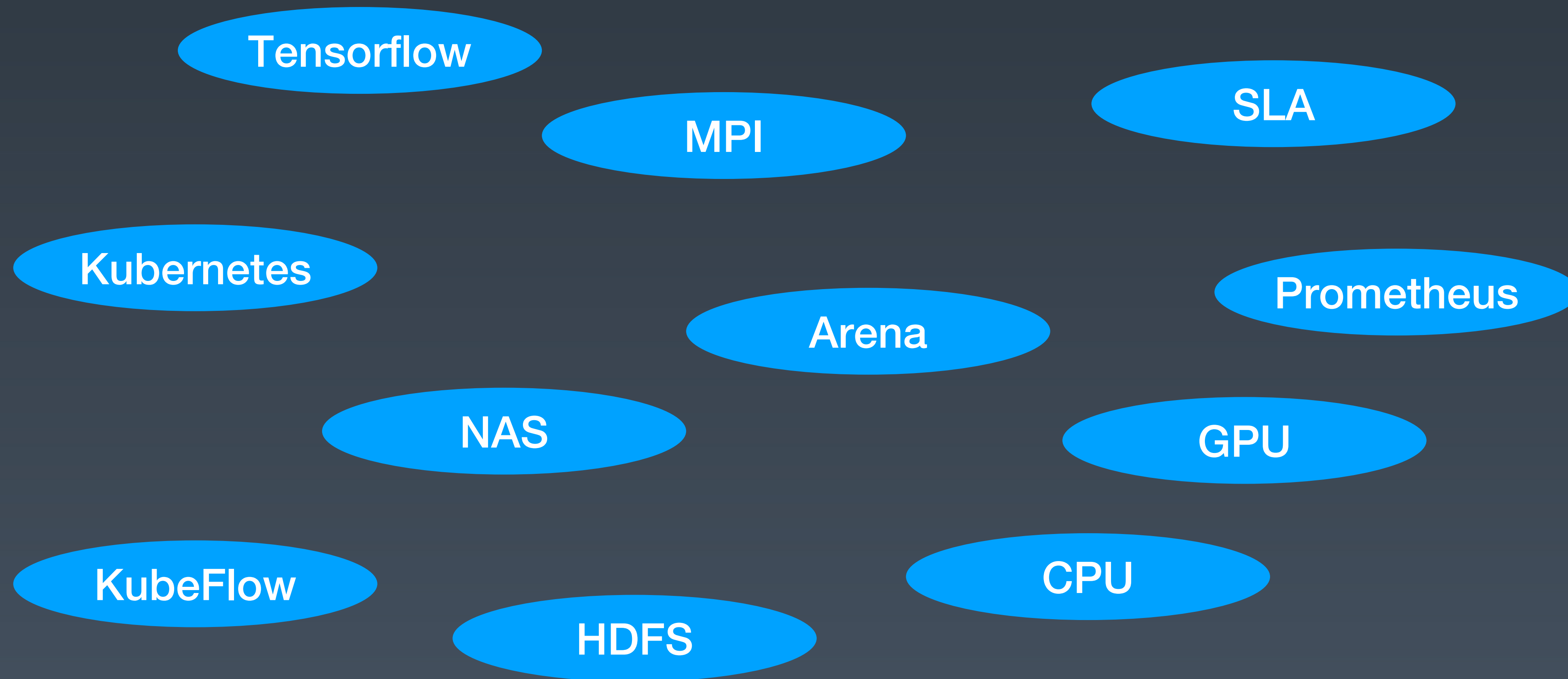
Table	热存储保留天数 HDFS	冷存储保存天数 OSS归档	成本
table_1	10		1TB/day 20K ¥/day
table_2	7	30	2TB/day 40K ¥/day
table_3	30	90	0.2TB/day 4K ¥/day
table_4	3		1TB/day 20K ¥/day
table_5	7	30	3TB/day 60K ¥/day

TODOs

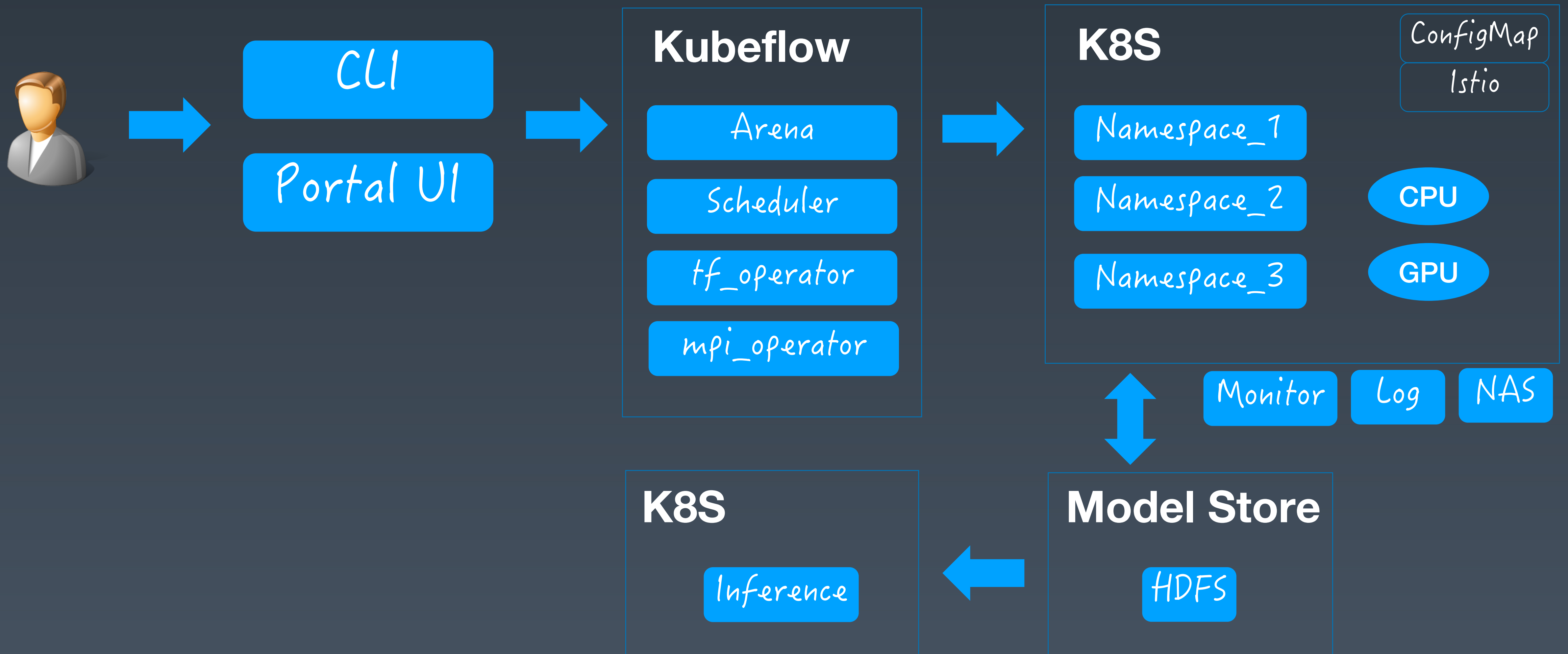
- Hadoop3 + Erasure Coding
- YARN队列资源抢占
- 云 + IDC 混部
- 存储计算分离
- Kerberos
- 成本统计：到目录，到任务级别

机器学习平台

组件s



System



解决的痛点

- 训练集群的资源利用率 (自研优化版scheduler)
- 训练任务提交和管理不便 (提供简易cli和UI, 提交, 查询任务状态。监控集群资源使用情况。)
- Ceph维护复杂 (替换成NFS/NAS)
- CPU深度训练效率低 (定制高内存GPU机器)
- 模型文件加载效率低 (迁移到HDFS存储, 模型版本管理平台)
- 任务排查困难 (丰富监控指标, 集中日志查询系统)

未来规划

- 云+IDC混部
- 存储与计算成本 细分治理
- 全链路数据监控报警
- 提高自动化运维程度
- 容器

InfoQ官网 全新改版上线

促进软件开发领域知识与创新的传播



关注InfoQ网站
第一时间浏览原创IT新闻资讯



免费下载迷你书
阅读一线开发者的技术干货

THANKS! | QCon th