



Flink China

CODE A BETTER LIFE
一行代码 亿万生活

美团Flink实践与应用

刘迪珊 基础数据部 2018.08.11





刘迪珊
美团研发工程师

西安电子科技大学硕士毕业，2015年加入美团数据平台。致力于打造高效、易用的实时计算平台，探索不同场景下，实时应用的企业级解决方案及统一化服务。

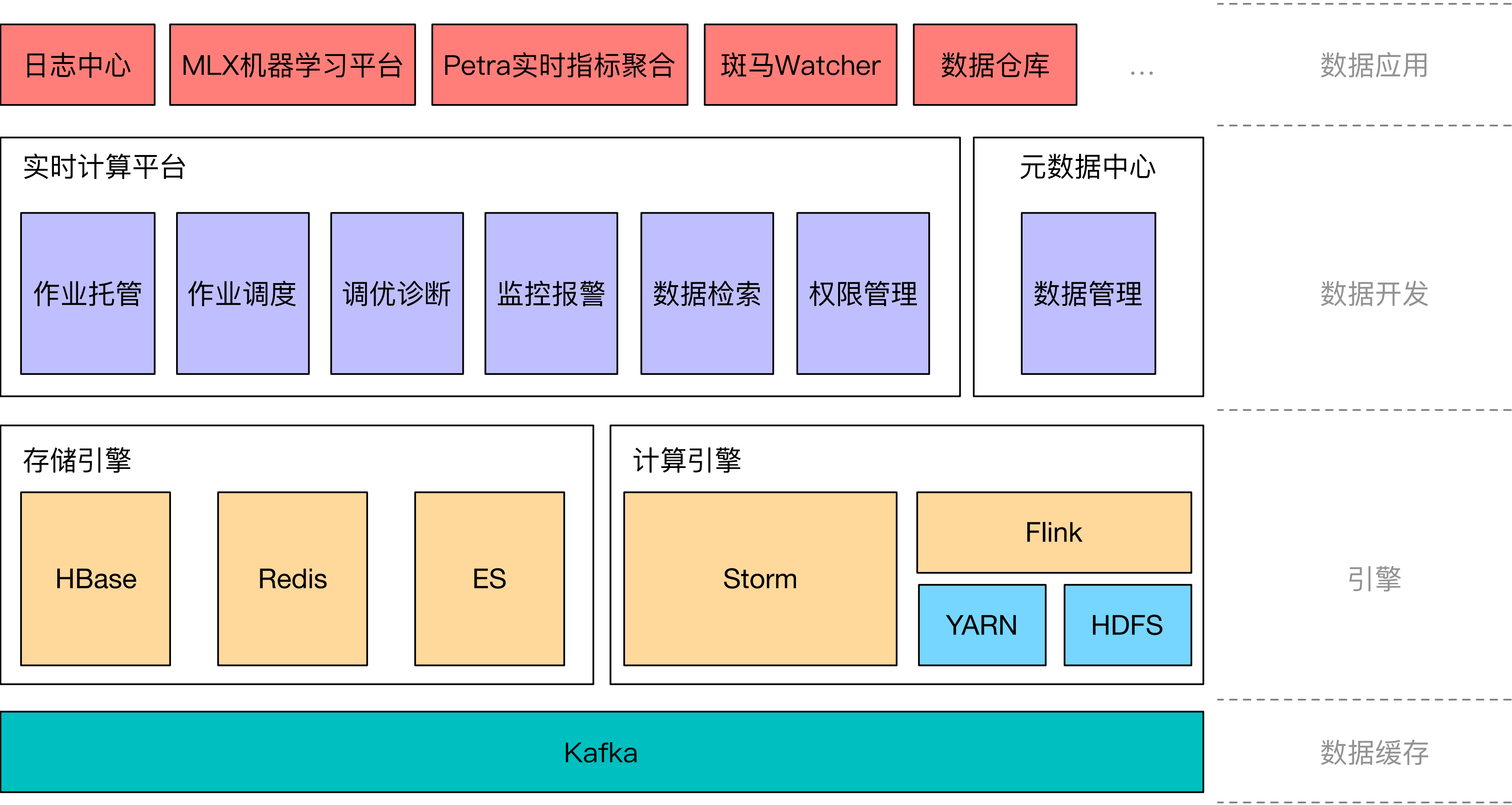
目录

1. 现状和背景
2. Flink在美团的实践
3. Flink在美团的应用
4. 未来展望

1 现状和背景

1 现状和背景

实时平台架构



1 现状和背景

实时平台现状

近万

作业量

千级

集群节点规模

万亿级

天级消息量

千万条/s

高峰期消息量

1 现状和背景

痛点和问题

- 实时计算**精确性**问题
- 流处理中的**状态管理**问题
- 实时计算**表义能力**的局限性
- 开发调试成本高

1 现状和背景

Flink探索关注点

- Exactly once计算能力
- 状态管理能力
- 窗口/Join/时间处理等等
- SQL/Table API

2 Flink在美团的实践

- 稳定性实践
- Flink平台化
- 生态建设

2 Flink在美团的实践

稳定性实践-资源隔离

- 资源隔离的考虑：分场景、按业务
 - 高峰期不同，运维时间不同；
 - 可靠性、延迟需求不同；
 - 应用场景，重要性不同
- 资源隔离的策略：
 - YARN打标签，节点物理隔离；
 - 离线DataNode与实时计算节点的隔离



按场景

- 离线生产
- 实时处理

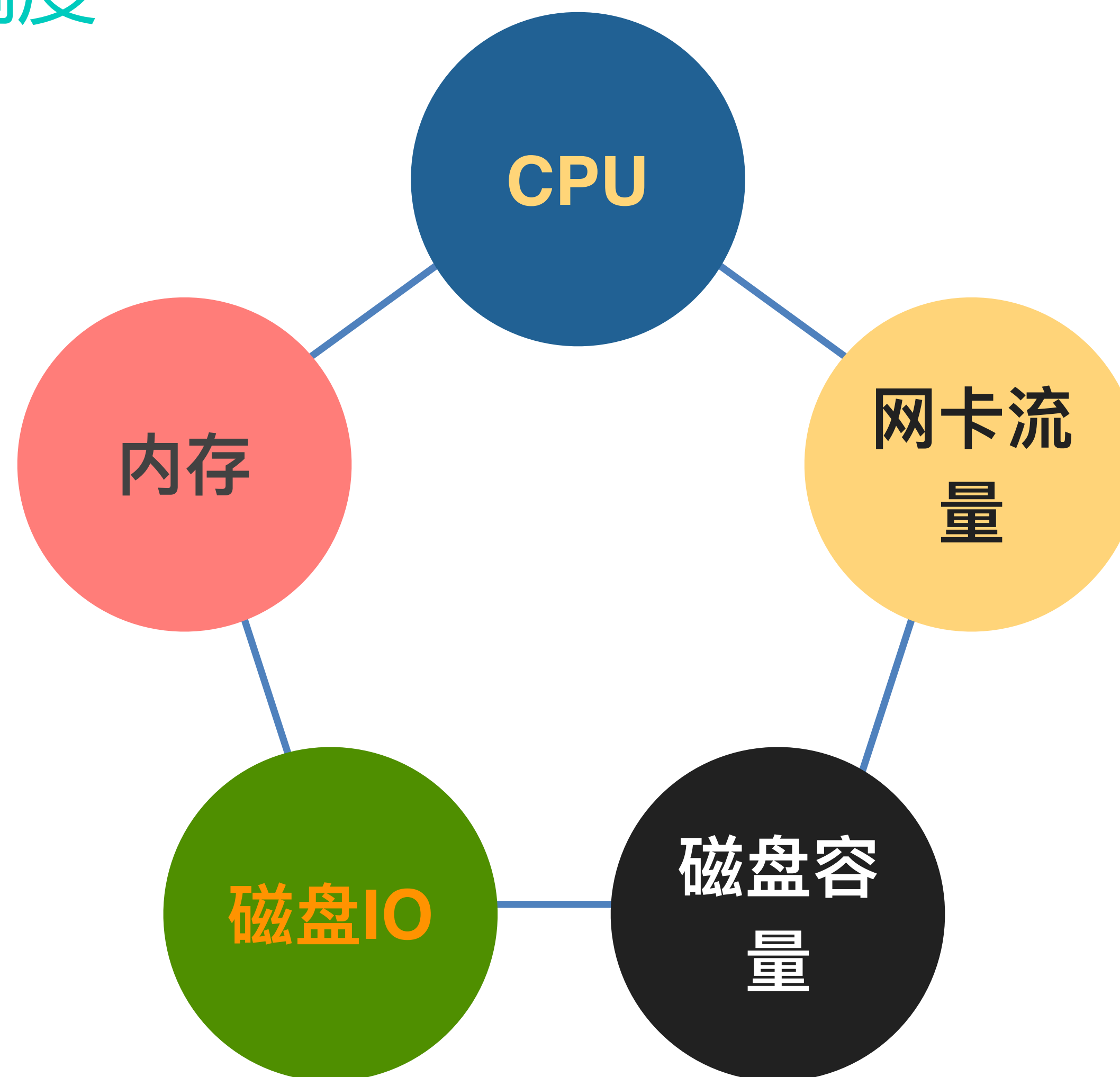


按业务

- 酒旅
- 广告
-

2 Flink在美团的实践

稳定性实践-智能调度



2 Flink在美团的实践

稳定性实践-故障容错

- 节点/网络故障
 - JobManager HA
 - 自动拉起
- 上下游容错
 - Flink Kafka 08异常重试
- 容灾
 - 多机房
 - 流热备

2 Flink在美团的实践

Flink平台化-作业管理

作业基本配置 [app_data_flinktest_biz]

基本配置

生产配置

调试配置

生产版本

Metrics

历史

依赖

延迟

TaskSet

FlinkDebug — liudishan

引擎

Flink

用户组

data

名称

flinktest_biz

Git仓库

报警方式

☒ 大家 ☐ 邮件 ☐ 短信

组织架构

负责人邮箱

备注信息

拓扑备注信息, 描述拓扑功能

修改基本配置

作业生产配置 [app_data_flinktest_biz]

基本配置

生产配置

调试配置

生产版本

Metrics

历史

依赖

延迟

FlinkFrame配置

主函数

并发度

1

TaskManager数量

1

Flink版本

1.4.0

从上一次结束Savepoint启动

☐

Flink arguments

选填

JobManager内存/MB

1024

TaskManager内存/MB

1024

JDK版本

jdk1.8.0_112

同步调试配置

保存生产配置

2 Flink在美团的实践

Flink平台化-监控报警

作业状态

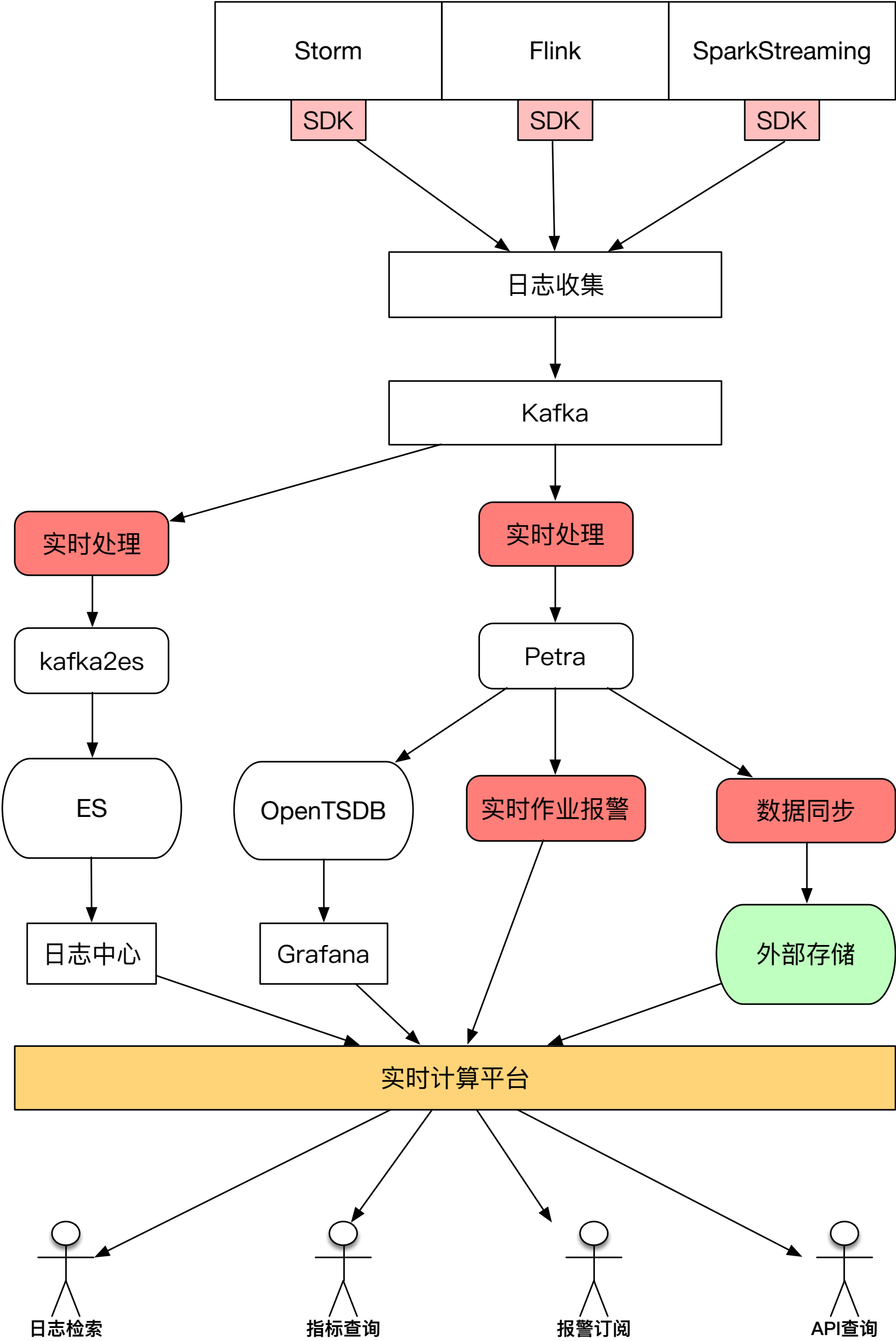
处理延迟

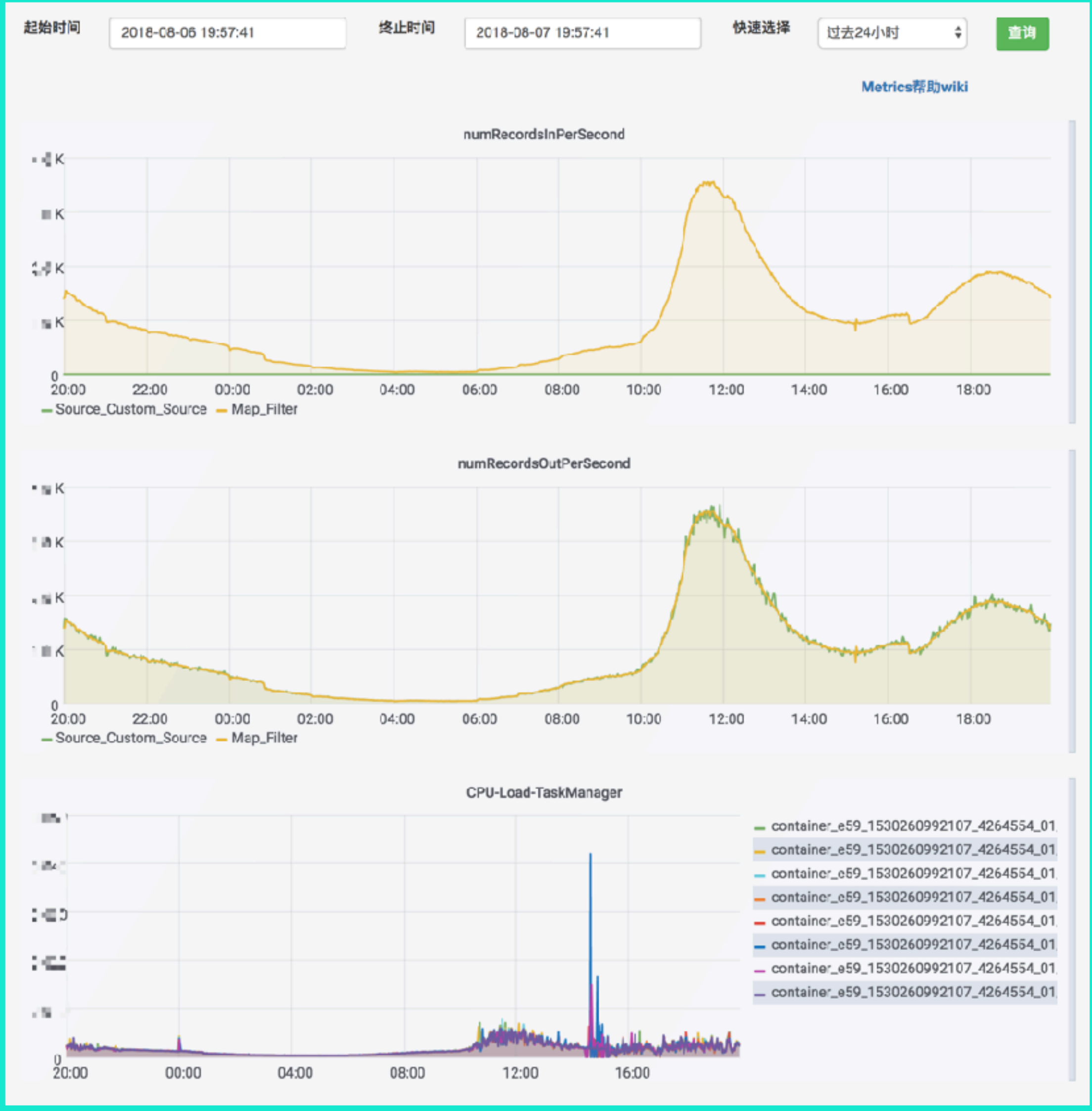
自定义
Metrics

2 Flink在美团的实践

Flink平台化-调优诊断

- 实时计算引擎提供**统一日志**和**Metrics**方案
- 为业务提供按**条件过滤**的**日志检索**
- 为业务提供自定义时间跨度的**指标查询**
- 基于日志和指标，为业务提供**可配置**的**报警**





➤ 首页

☰ 日志

🔍 搜索

⚠ 报警

📊 kibana2

📊 kibana5

👤 管理

❓ 文档

日志名

索引名

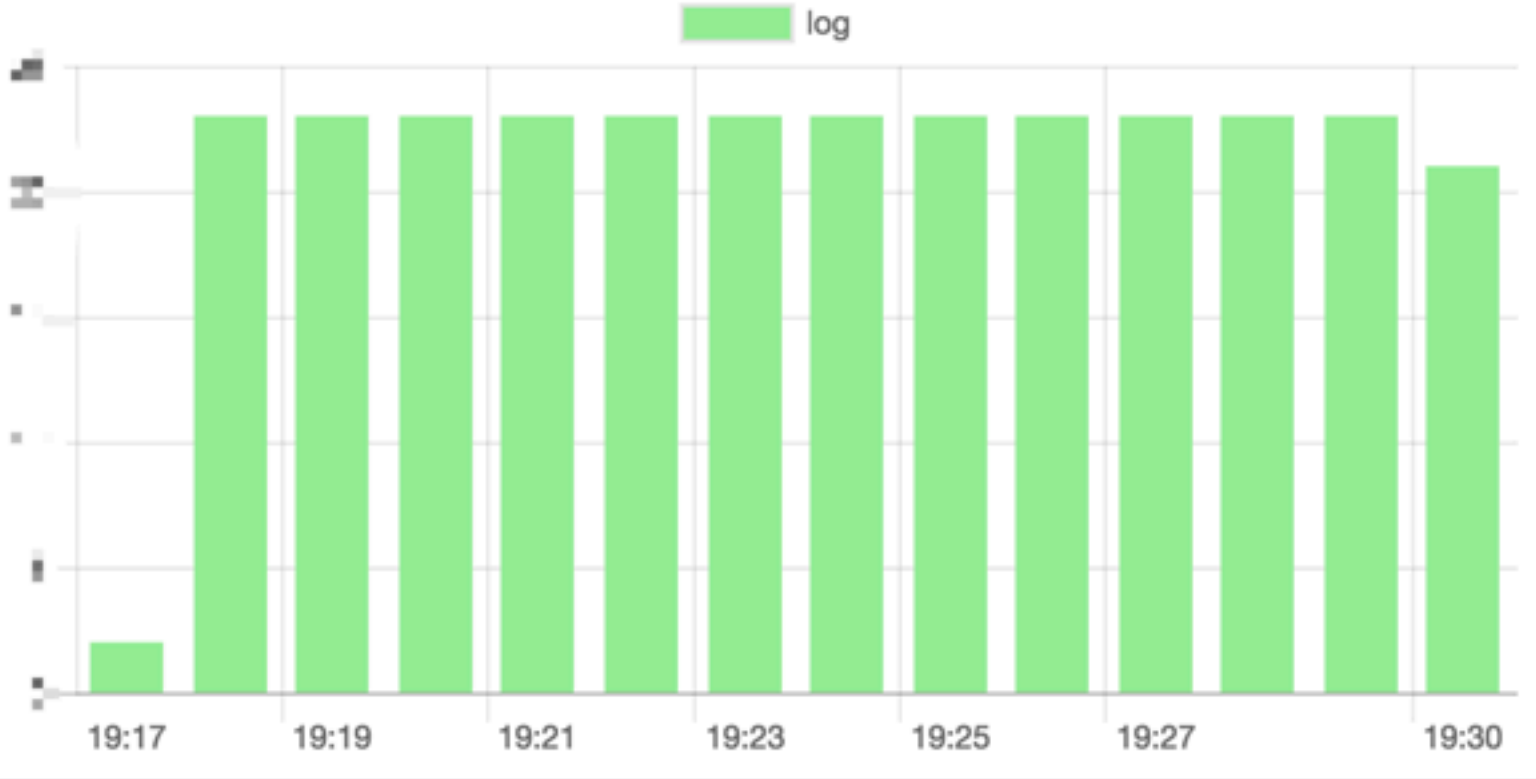
起始时间2018/08/06 19:17:50now

结束时间2018/08/06 19:30:50now

15m 30m 1h 3h 12h 1d 7d 14d

查询job_name AND apGO

log



raw

json

kv

☒ 全选

☐ 反选

☒ es_timestamp

☐ job_name

☒ applicationID

☒ message

☒ containerID

☐ mt_action

applicationID	containerID	es_timestamp	message	mt_datetime	mt_level	mt_logger_name	mt_thread
<input checked="" type="radio"/> application_1530780134899_563523	container_e59_1530780134899_563523_01_000003	2018/08/06 19:30:45 +0800		2018-08-06 19:30:45+0800	INFO	org.apache.flink.streaming.connectors.kafka.interna	Thread-43
<input checked="" type="radio"/> application_1530780134899_563523	container_e59_1530780134899_563523_01_000005	2018/08/06 19:30:44 +0800		2018-08-06 19:30:44+0800	INFO	org.apache.flink.streaming.connectors.kafka.interna	Thread-43
<input checked="" type="radio"/> application_1530780134899_563523	container_e59_1530780134899_563523_01_000004	2018/08/06 19:30:28 +0800		2018-08-06 19:30:28+0800	INFO	org.apache.flink.streaming.connectors.kafka.interna	Thread-43

2 Flink在美团的实践

生态建设

适配业务MQ

- 对接线上数据
- 一次同步，多次消费

Kafka封装

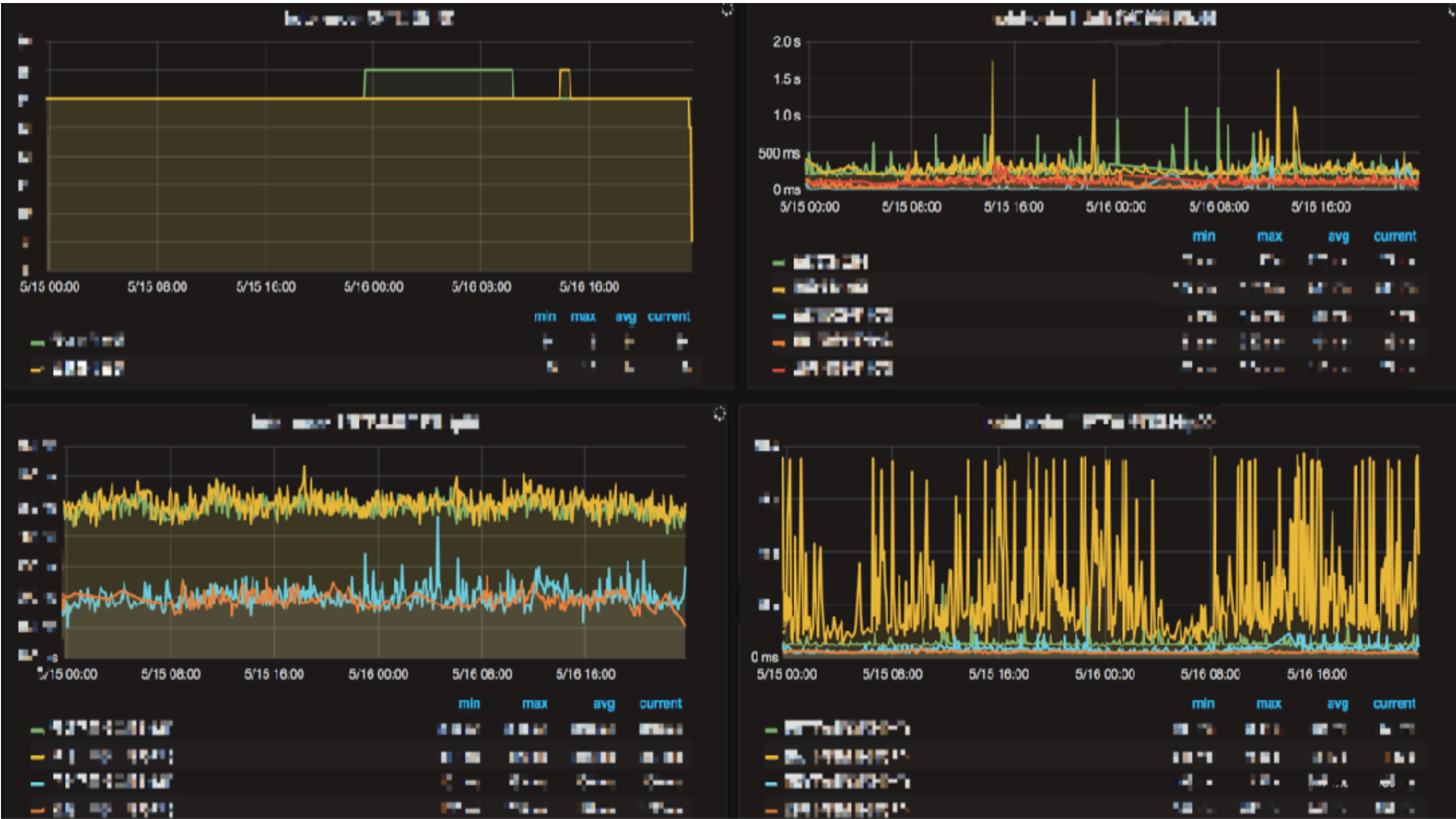
- 屏蔽地址、基础配置
- 权限管理
- 指标监控

3 Flink在美团的应用

3 Flink在美团的应用

Petra实时指标聚合

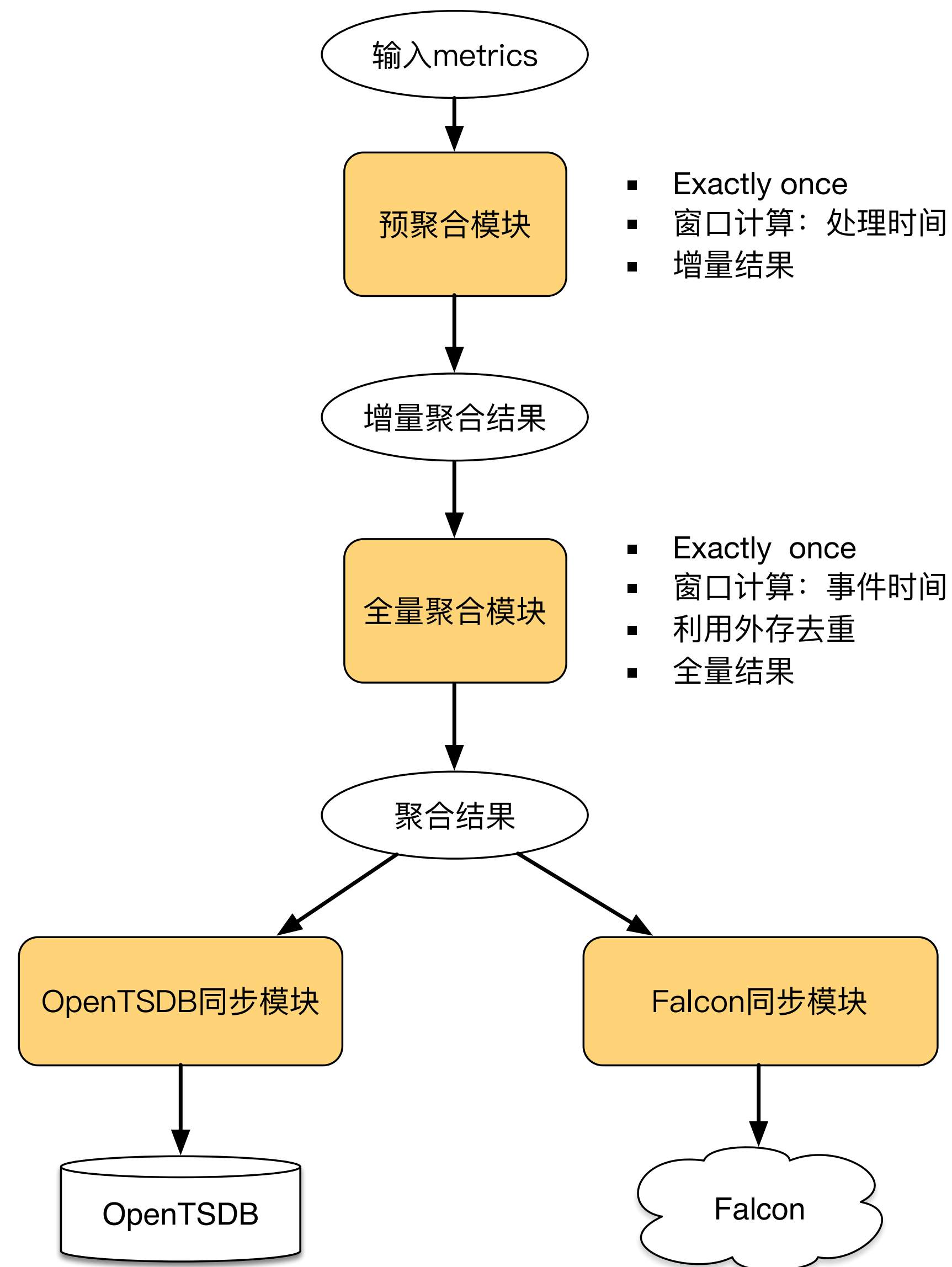
- 业务场景：
 - 基于业务时间（事件时间）
 - 多业务维度：如应用、通道、机房等
 - 复合指标计算，如：
 - 交易成功率=支付成功数/下单数
 - 低延迟：秒级结果输出



3 Flink在美团的应用

Petra实时指标聚合

- Exactly once的精确性保障
 - Flink checkpoint机制
- 维度计算中数据倾斜
 - 热点key散列
- 对晚到数据的容忍能力
 - 窗口的设置与资源的权衡



3 Flink在美团的应用

Petra实时指标聚合

应用名称

unionid

责任人



状态

上线

创建时间

2016-03-07 14:34:08

修改时间

2018-03-21 12:48:15

应用描述



文档链接



应用配置

输入配置

简单聚合配置

复杂聚合配置

Falcon同步配置

结果信息查看

线上数据追踪

聚合规则配置

✕ 删除配置

+ 新增配置

📄 批量导入

Search

↺

⌵

<input type="checkbox"/>	metric名称	聚合类型	聚合维度	聚合粒度	数据延迟时间	同步falcon	操作
<input type="checkbox"/>	count	counter	host, method, sub_name	60	Empty	False	

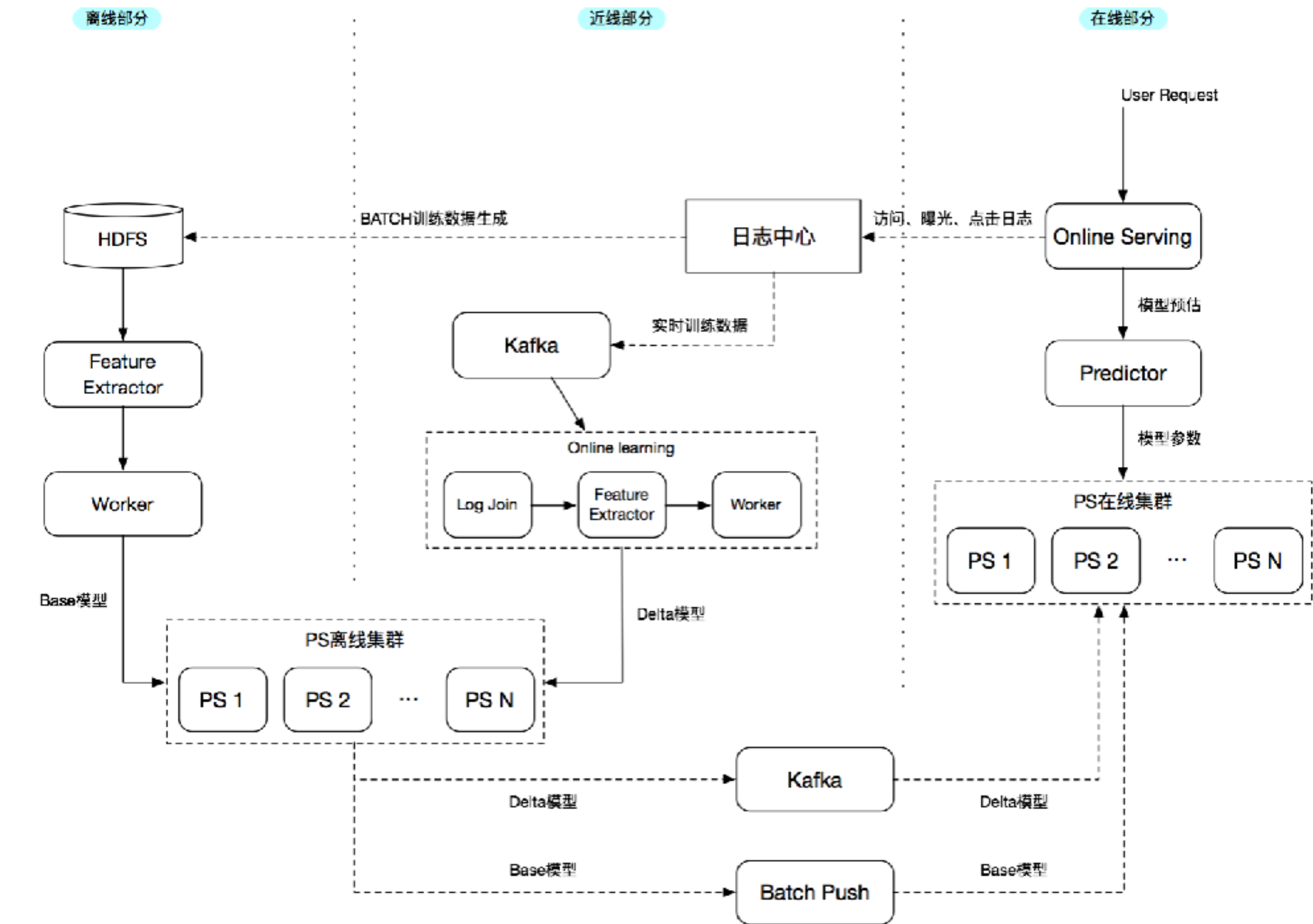
Showing 1 to 1 of 1 rows

数据过滤配置

流量细分配置

3 Flink在美团的应用

MLX机器学习平台



| 4 未来展望

- 状态管理
 - 统一管理
 - 大状态性能优化
- SQL
 - 开发效率提升
- 新场景探索

Q&A

CODE A BETTER LIFE

一行代码 亿万生活



更多技术干货
欢迎关注“美团技术团队”

招聘：后台开发岗位
邮箱：liudishan@meituan.com

