

# 大数据实战

## 测试工具领域应对海量数据的解决方案

孔祥云

2017/12/16



# 目录

CONTENTS

1

稳定性监控系统产生的背景和目标

2

难点分析

3

技术方案

4

智能接口平台数据推荐产生的背景和目标

5

整体方案

# 我是谁？



《京东系统质量保障技术实战》作者之一，10年测试领域从业经验，擅长功能测试、自动化测试、测试工具开发等。

在京东任职期间参与开发EAT测试框架、Mock接口测试平台、自动部署系统、商城质量门户等；主导应用大数据技术的接口稳定性监控平台、接口变更探测工具等。



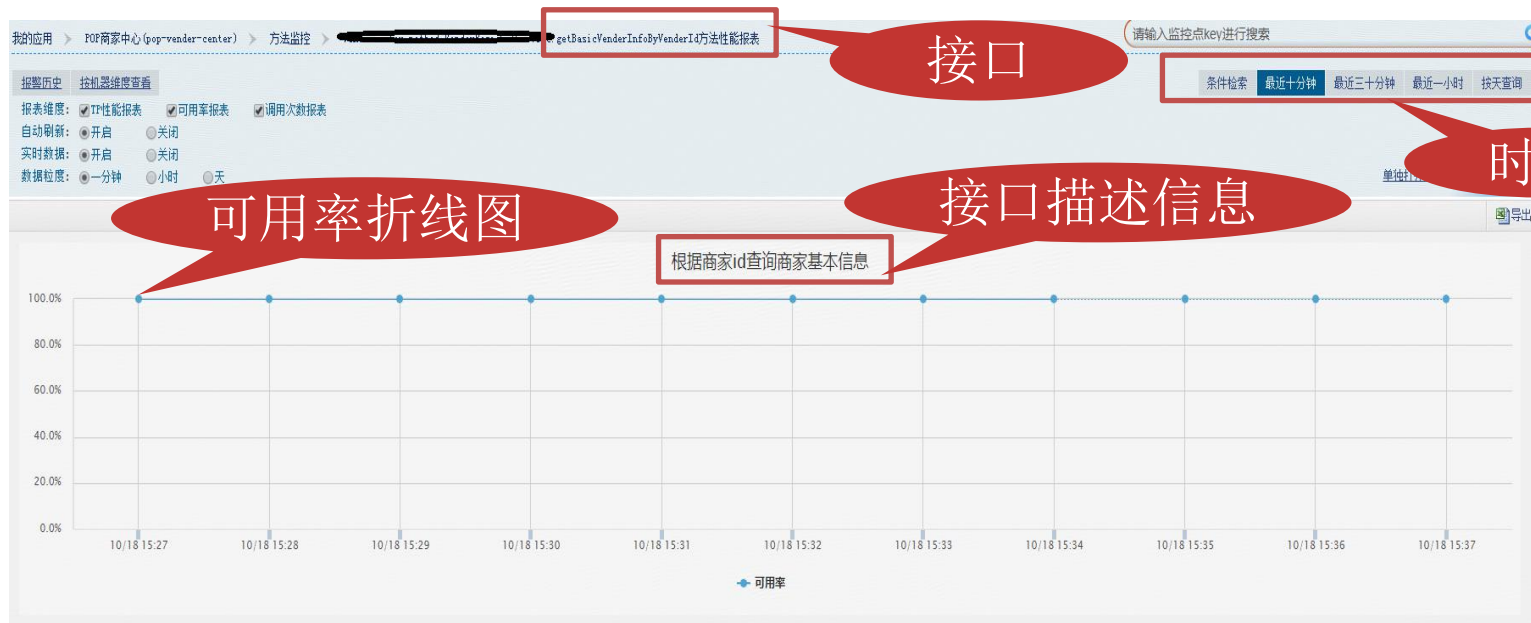
# 稳定性监控系统产生的背景和目标

接口

监控

稳定性

# 稳定性监控系统产生的背景和目标



- 接口的稳定性数据采集后，虽然直观，但是不能对整个应用有好的表达
- 一般查询只提供了单个接口的方式，不能查询较长时间、多接口的稳定性数据
- 不是基于业务场景的，非开发人员无法映射到具体的业务场景上

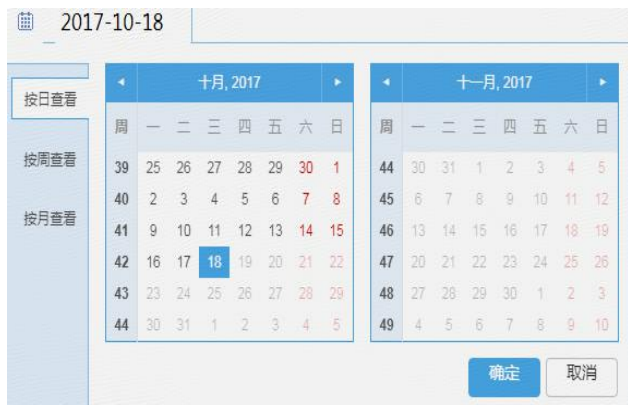
# 稳定性监控系统产生的背景和目标

- 直观展示多维度稳定性趋势

多时间跨度

多业务场景

多层次



服务名称
获取单个商品的广告词
设置单个商品的广告词
通过类目和属性类型，获取类目的属性列表

商品中心系统	>
商品中心系统	
商品中心API	
商品中心RPC	

# 系统难点分析

## ●实时性

- 每秒稳定性数据采集频次：

$$100（应用）* 100（接口） * 5（分钟） / 60（秒） = 1000$$

- 及时反映稳定性变化



采集频率5分钟



变化频率10分钟

# 系统难点分析

## ●海量数据

➤ 数据量大：

单个应用：

$$100 \text{ (接口)} * 24 \text{ (小时)} * 60 \text{ (分钟)} / 5 \text{ (分钟)} \\ = 28800$$

$$28800 \text{ (条)} * 365 \text{ (天)} = 10,512,000$$

➤ 维度多：





# 系统难点分析

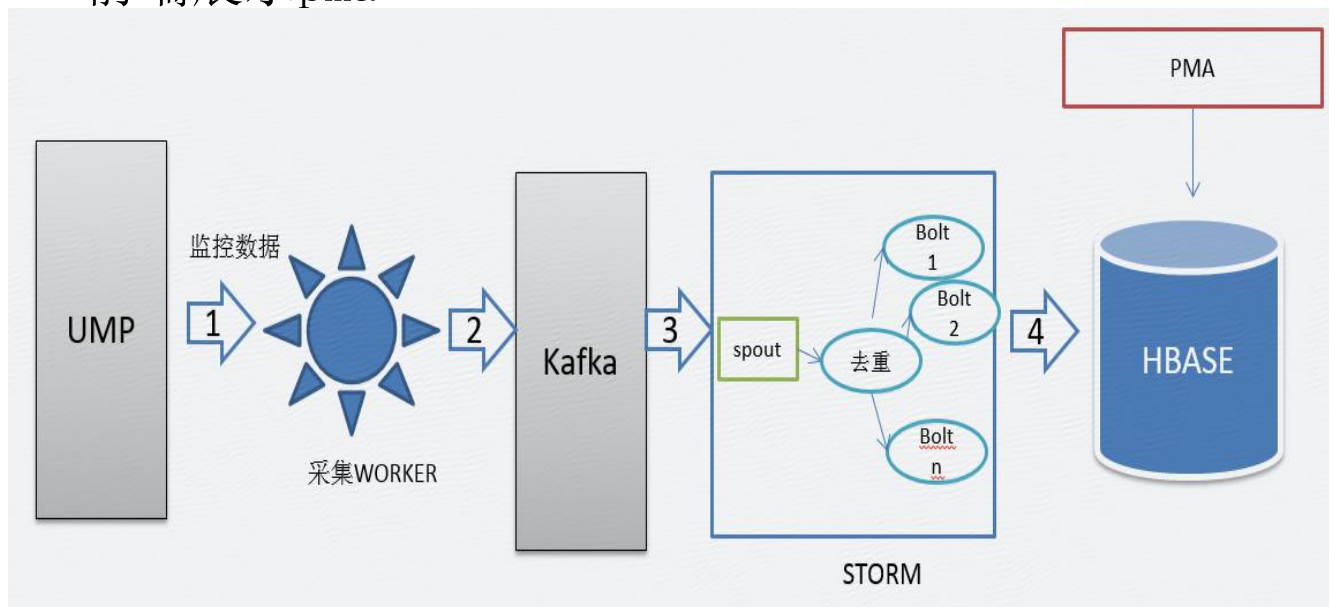
## ● 数据可视化展现

- 稳定性折线图上突出展示异常点
- 折线图时间跨度可任意选择

# 技术方案

## ● 涉及到的技术模块

- ✓ 采集worker
- ✓ 实时计算storm
- ✓ 前端展示pma

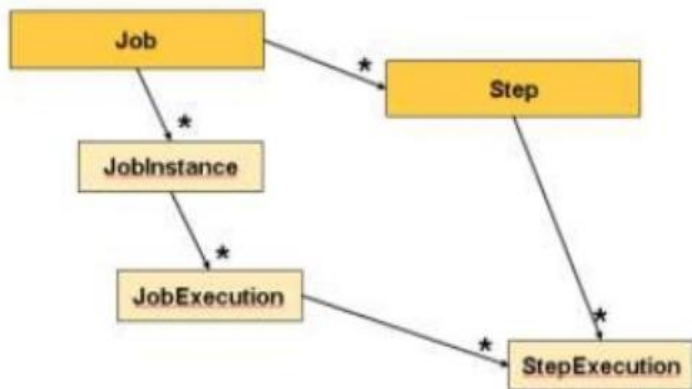


# 碰到的问题及解决方案

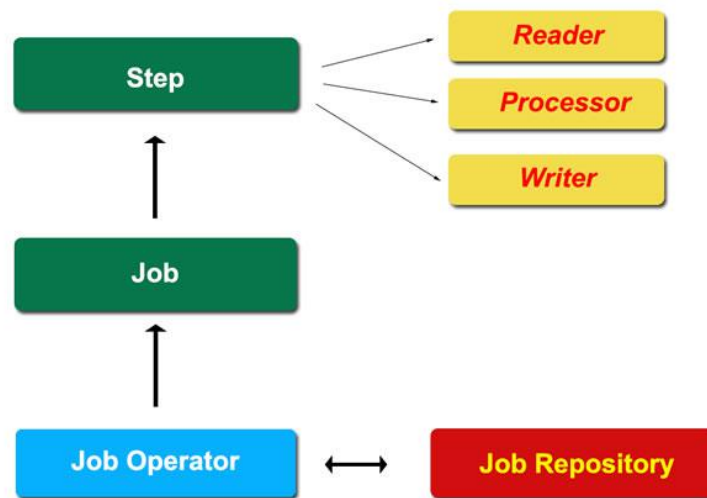
## ● 采集worker的及时性



并行

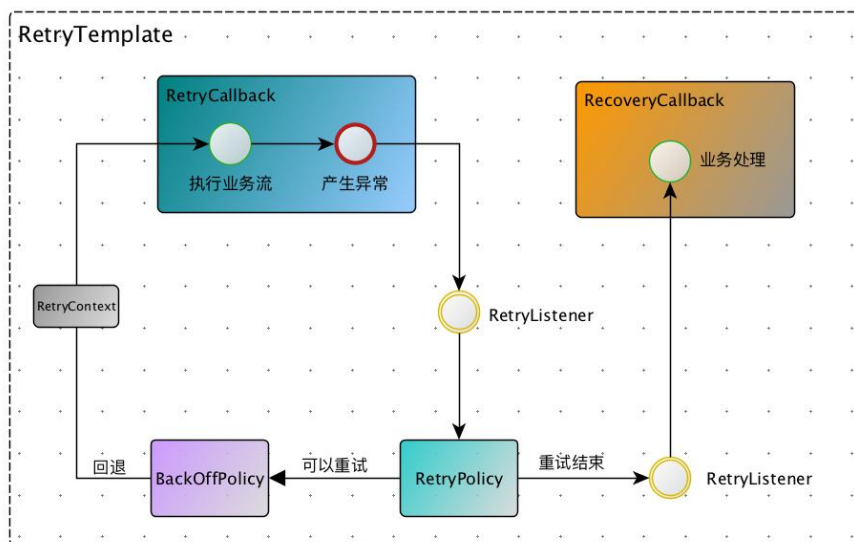


解耦



# 碰到的问题及解决方案

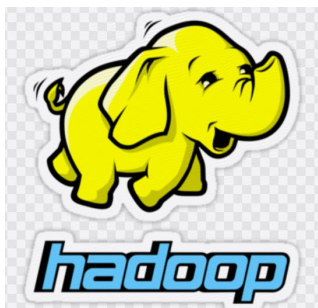
- 采集worker的失败重采机制



# 碰到的问题及解决方案

- 及时展示变化数据的大数据方案

离线计算



准实时计算



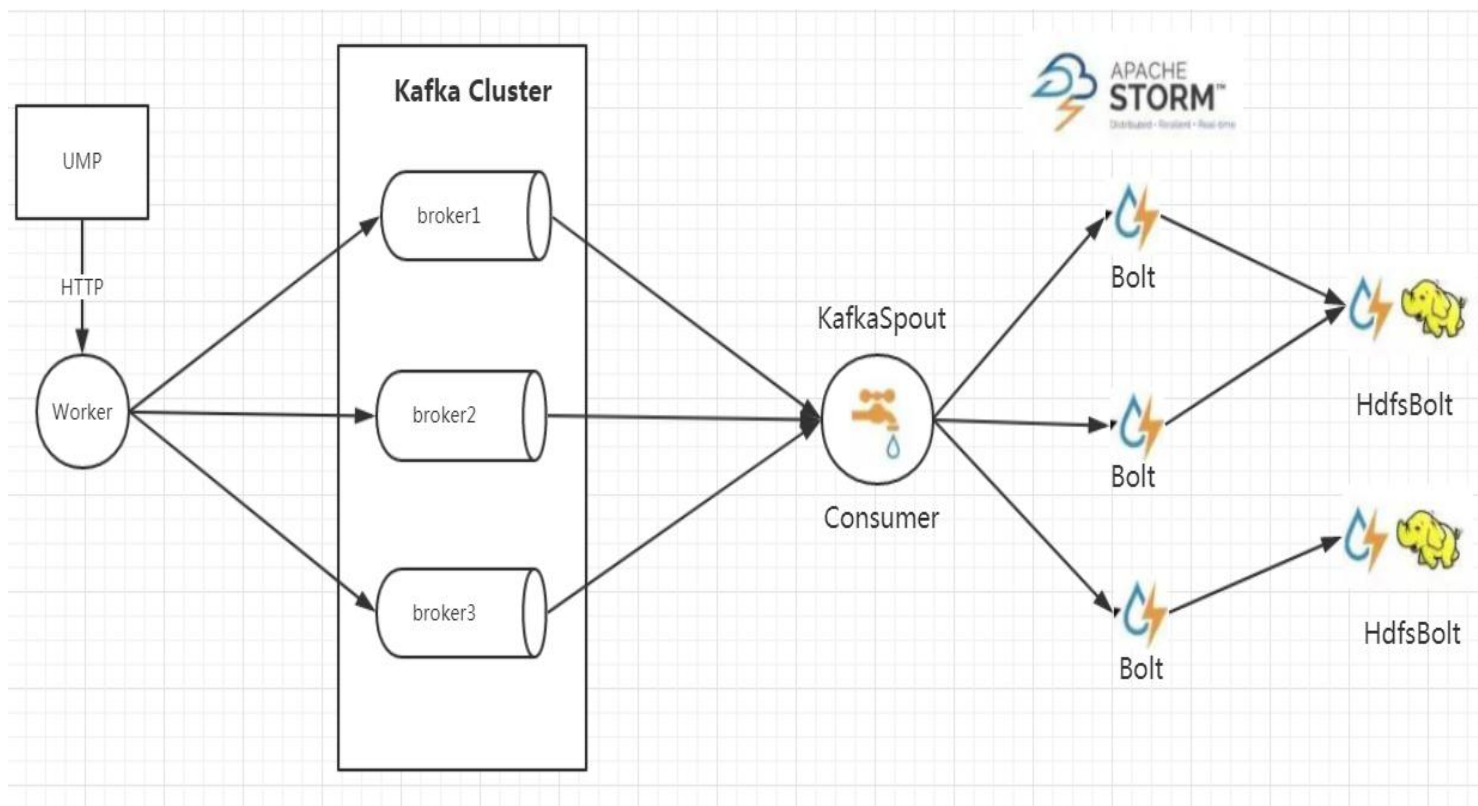
实时计算



# 碰到的问题及解决方案

## ● Storm 特性

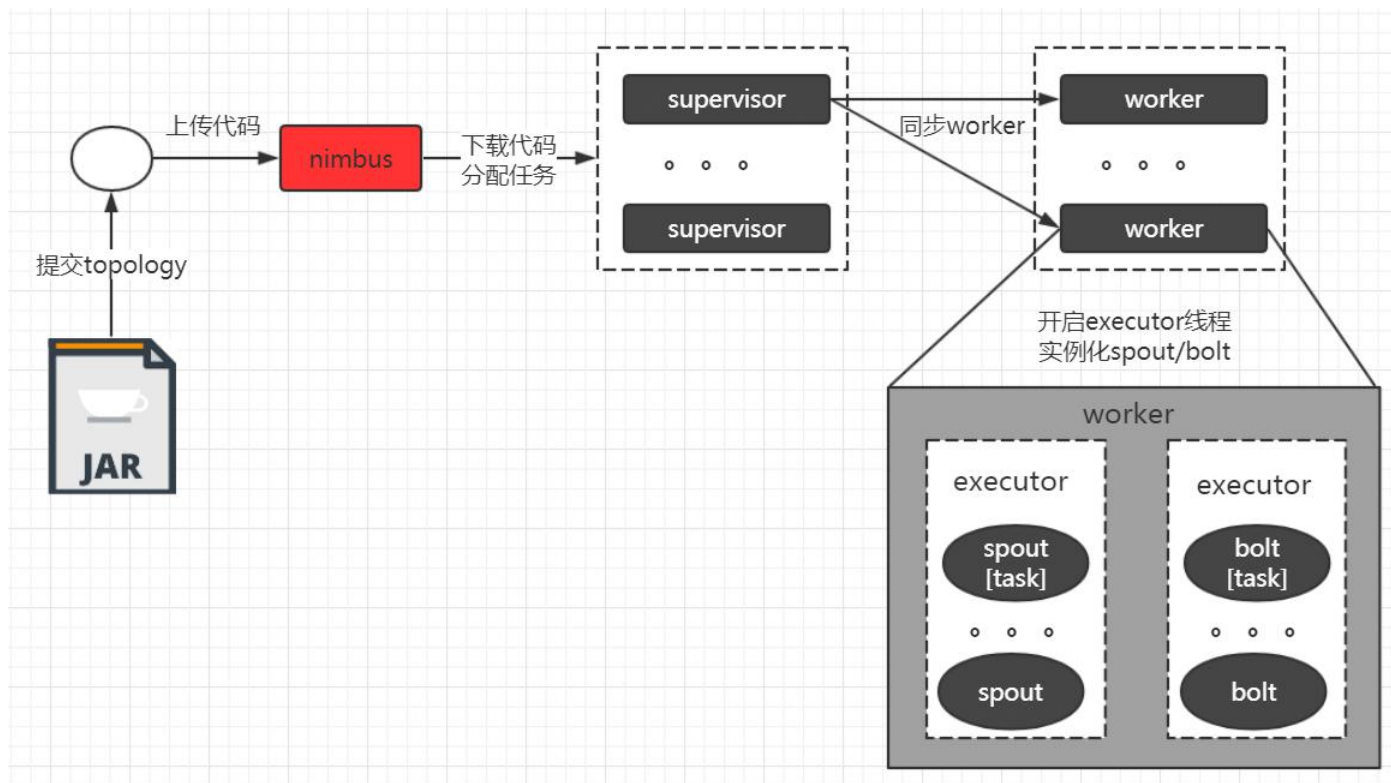
### ➤ 实时性



# 碰到的问题及解决方案

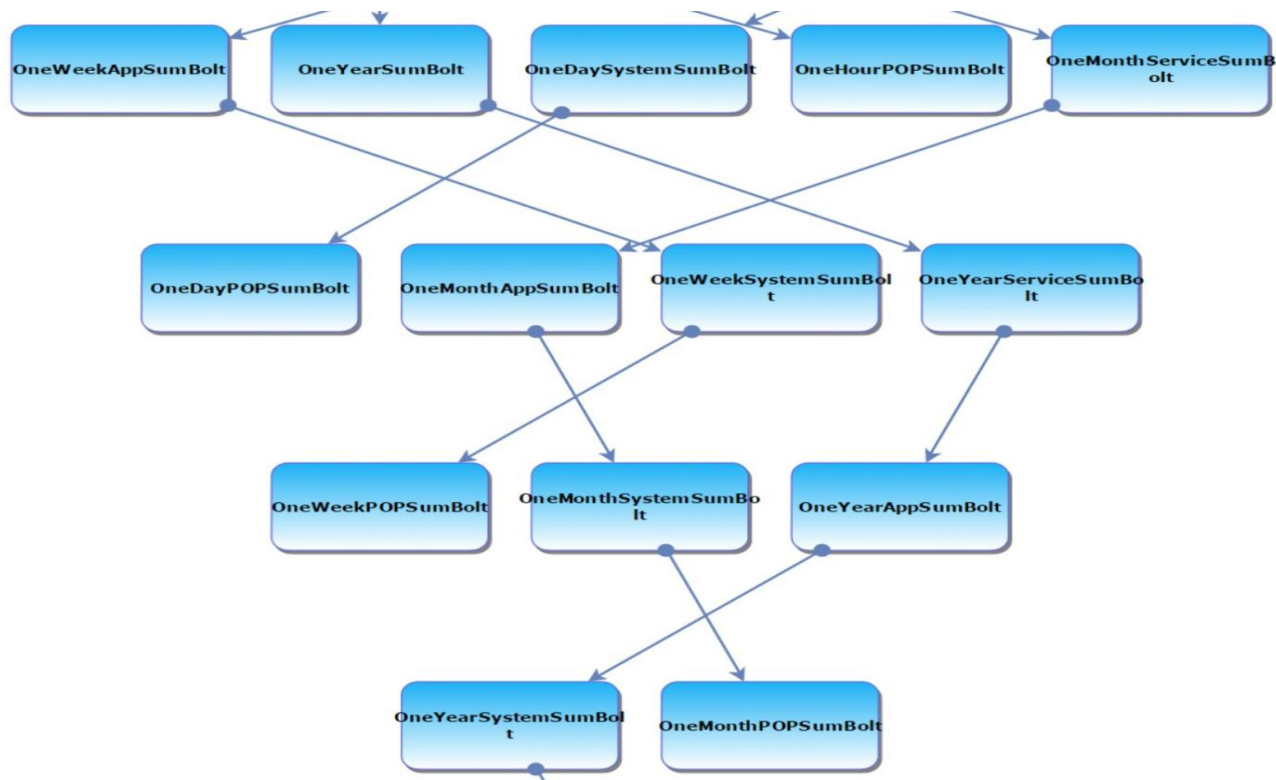
## ● Storm 特性

### ➤ 分布式



# 碰到的问题及解决方案

## ➤ Storm实际部署的拓扑图（部分）

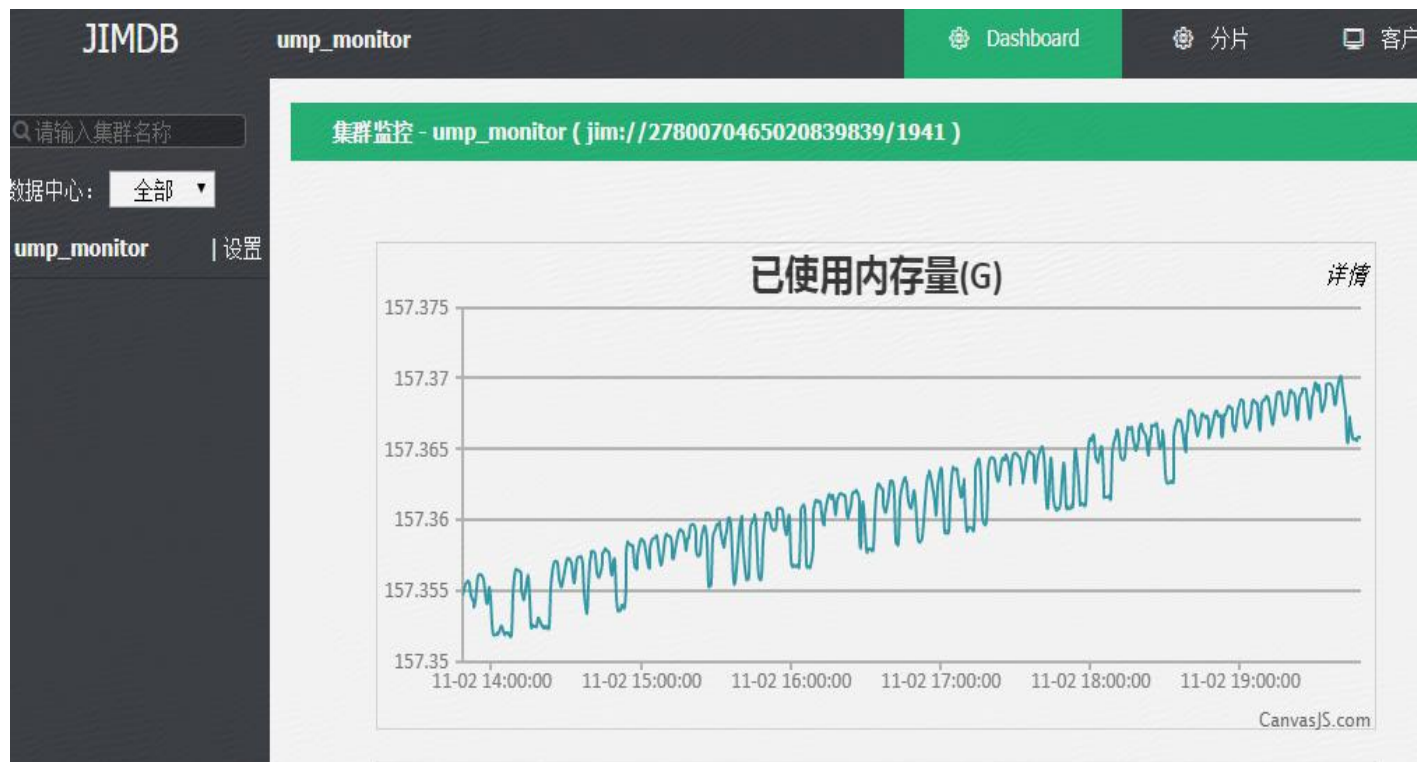




# 碰到的问题及解决方案

## ● 实时计算多维度计算的数据存储策略

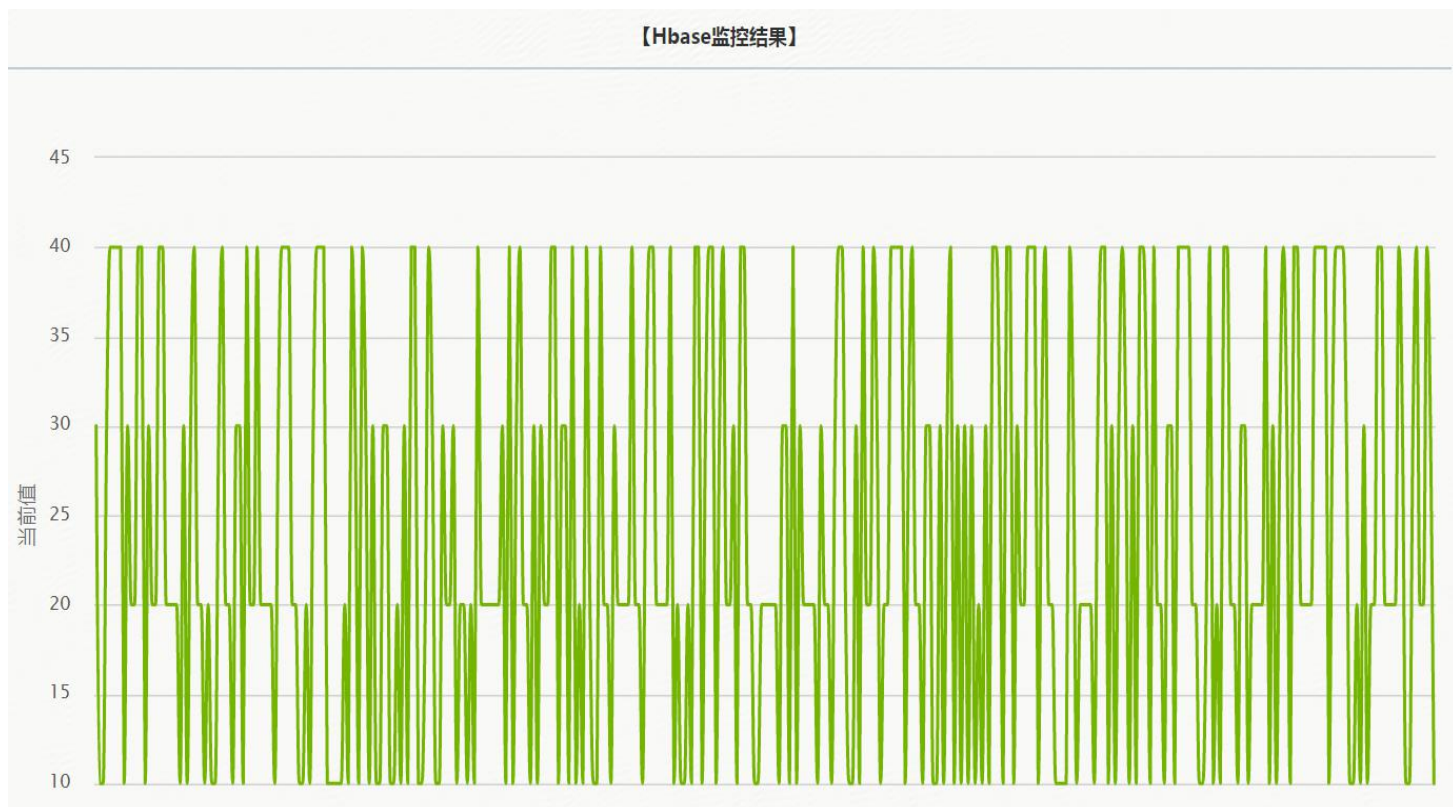
- 利用JimDB缓存中间数据进行计算
  - ✓ 汇总 + 计算



# 碰到的问题及解决方案

## ● 实时计算多维度计算的数据存储策略

- 利用HBase进行落盘存储
  - ✓ 落盘 + 去重

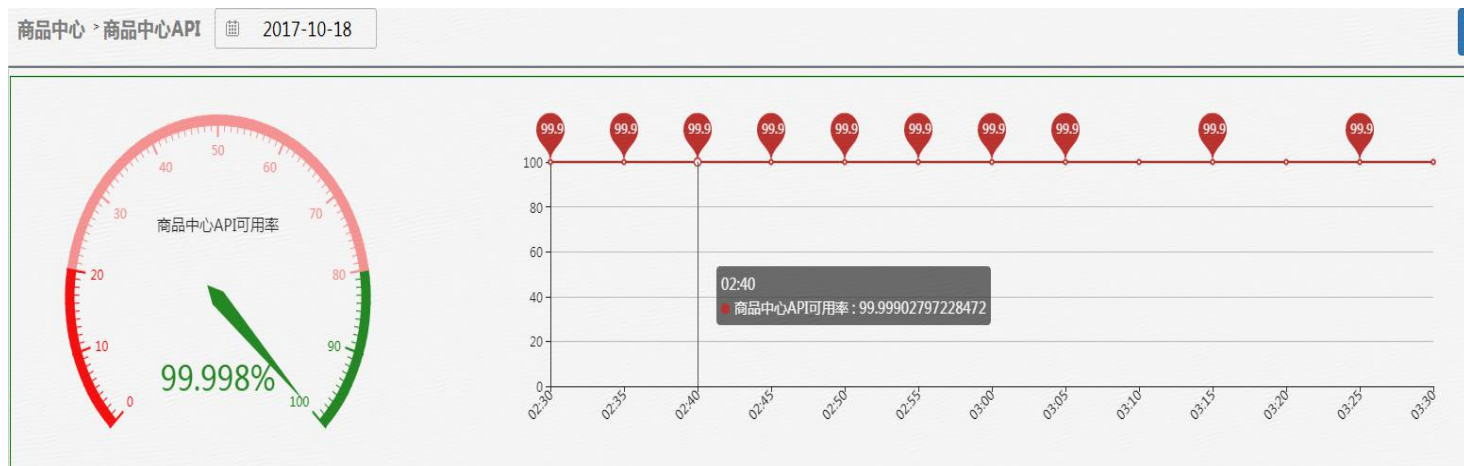


# 碰到的问题及解决方案

## ● 前端展示图表的方案

### ➤ ECharts的折线图

- ✓ ECharts拥有丰富的图表：折线图、柱状图、散点图、饼图、K线图等
- ✓ 时间跨度选择——通过设置dataZoom设置数据窗口的范围



# 效果展示

pop 监控平台

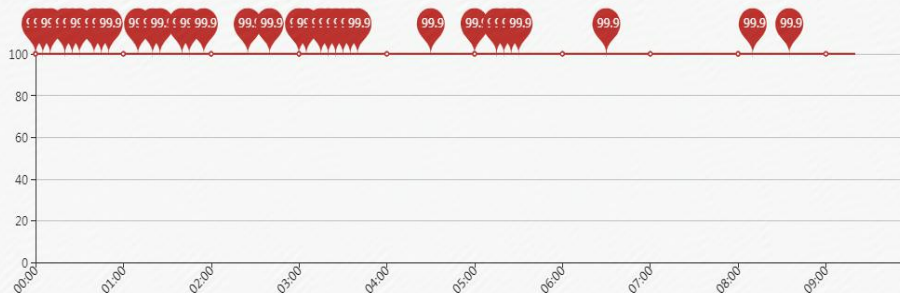
欢迎您, kongxiangyun [退出]

商家后台 京麦 数据管家 JSHOP 系统稳定性

- 整体监控概况
- 商品中心系统
  - 商品中心系统
  - 商品中心API
  - 商品中心RPC
- 订单中心系统
- 商家中心系统
- 营销中心系统
- 结算中心系统
- 商品中心系统配置
- 订单中心系统配置
- 结算中心系统配置
- 商家中心系统配置
- 营销中心系统配置

快报 实时 100.0% 今天 99.999% 本周 99.999% 本月 99.999%

商品中心 > 商品中心API 2017-11-03



商品中心API各服务可用率趋势图

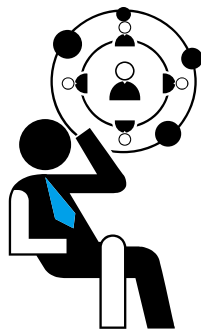


# 智能接口平台数据推荐产生的背景和目标

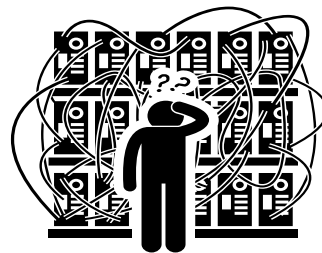
## ● 背景



数据准备



回归用例



用例选择

# 智能接口平台数据推荐产生的背景和目标

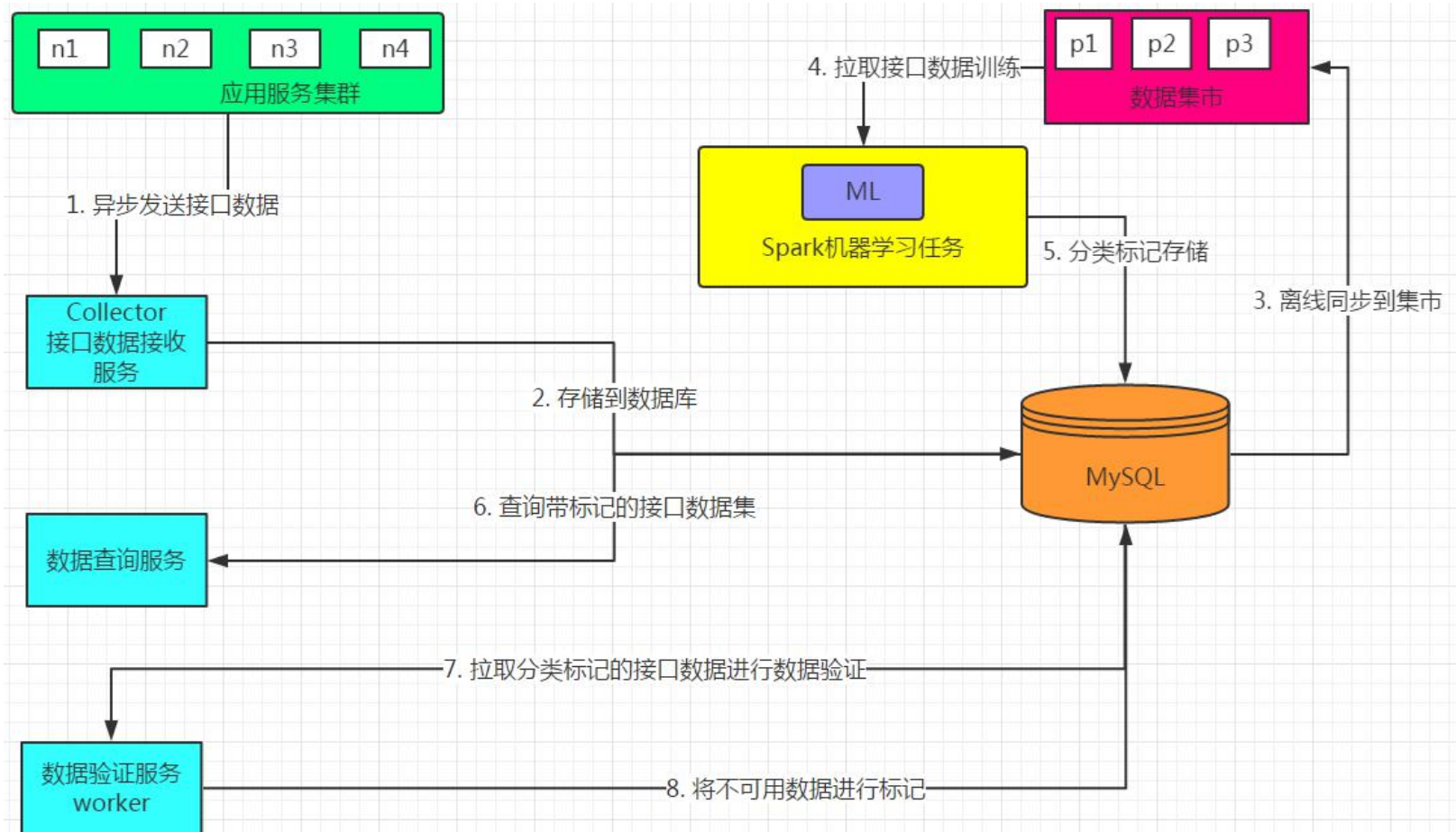
- 目标

真实数据



分类

# 整体方案



# 未来规划

数据分类  
优选数据集

1

2

3

精准测试  
接口覆盖率

无人值守测试  
Code Diff



# Q&A

# 谢谢

