



GOPS2018
Shenzhen

GOPS

全球运维大会 2018

2018.4.13-4.14

中国·广东·深圳·南山区 圣淘沙大酒店（翡翠店）





GOPS2018
Shenzhen

千亿交易背后的0故障发布

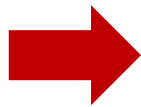
陆叶平(少荃) 阿里巴巴研发效能

自我介绍

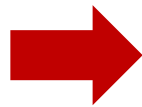


GOPS2018
Shenzhen

业务



基础核心系统



运维中台

- 主要从事发布相关的运维中台建设
- 重点关注发布的效率和稳定性方面
- SP2P文件分发系统蜻蜓的创始人之一



GOPS2018
Shenzhen

目录

- ➔ **1** 线上发布之痛
- 2** 我们的解决方案
- 3** 典型的发布案例
- 4** 无人值守发布实现



GOPS2018
Shenzhen

线上发布之痛-变更故障

昨晚一对情侣朋友拉我出去吃宵夜，我说不去不想看到他们虐狗，然后他们说还有四个朋友不用怕，我就放心去了，结果我跟三对情侣一起吃宵夜。



理想和现实的差异

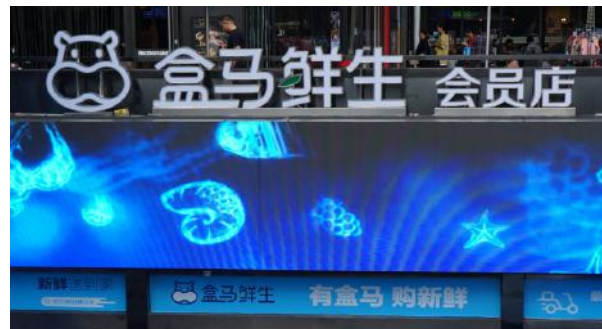
骑士资本4.4亿美元的教训



线上发布之痛-业务挑战



GOPS2018
Shenzhen





GOPS2018
Shenzhen

线上发布之痛-如何避免



1. 发布前
 - 单元测试、集成测试
2. 发布中
 - 预发、灰度、分批发布
3. 发布后
 - GOC故障预警



GOPS2018
Shenzhen

目录

1 线上发布之痛

➔ 2 我们的解决方案

3 典型的发布案例

4 无人值守发布实现



GOPS2018
Shenzhen

我们的解决方案



“人工”智能的发布保障

- 盯监控
- 盯发布单
- 盯机器
- 盯GOC等各种屏幕

不足之处

- 成本高
- 人不可靠



GOPS2018
Shenzhen

我们的解决方案

无人值守发布是什么？

是一个智能化变更故障检测和异常推荐系统

能力：检测发布风险，拦截异常发布

价值：减少发布投入成本，提升发布稳定性，减少故障



GOPS2018
Shenzhen

智能化发布路线图



什么是无人值守发布



GOPS2018
Shenzhen

部署单

应用:

发布方式: 分批发布

发布进度: 完成 1 | 总数 2

部署内容: 版本12345678

发布结果: 部署中

分批策略: 按机房均分

发布批次: 当前批次 1 | 分 2 批

计划发布时间: 2018-03-01 00:00:00

发布单状态: **暂停**

暂停方式: 分批发布第一批暂停

部署单提交人: 少奎

部署说明: test

主机失败率: 0.00%

纯发布耗时: 133s

提交时间: 2018-03-01 00:00:00

发布系统指标分析: [点击查看](#)

监控数据分析



分析时间:-

无人值守发布检测到 1 个异常,发布被拦截,如确认没有问题请点击[恢复],如确实有问题,请评估是否回滚

分析指标	是否异常	异常项	提示	操作
基础监控指标	未发现异常		未接入, [参考手册] 接入查看更多数据	查看详情
业务监控指标	是	订单创建_金丝雀测试	请确认是否由发布引起	查看详情
异常日志	未发现异常			查看详情



什么是无人值守发布

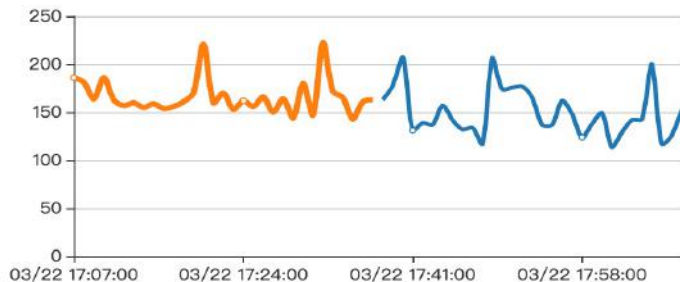


实时数据对比

发布前后对比 昨天今天对比 日志详情

appExceptionFluctuation

○ 发布前 ○ 发布后



异常检测报告

基础监控指标 :未接入

业务监控指标

⚠️ P0 [03-01 18:03 订单创建_金丝雀测试]最近3分钟求和与上3分钟求和的环比下跌: 1.22% > 1%

⚠️ P0 [03-01 18:03 订单创建_金丝雀测试]最近1分钟求和值为: 10000 < 10000001.00

⚠️ P0 [03-01 18:03 订单创建_金丝雀测试]最近6分钟求和与上6分钟求和的环比下跌: 1.29% > 1%

已接入的业务智能基线监控指标未发现异常

应用日志

🕒 应用纬度发布后异常数量 正常

波动图

🕒 已发布机器发布后异常Top排行 波动正常

波动图



GOPS2018
Shenzhen

什么是无人值守发布

应用名:

🔍 可接入查询

一.基础监控接入

❌ 基础监控未接入, 请参考[\[无人值守发布接入指南\]](#)接入基础监控指标

二.业务监控接入

1.业务监控项

✅ 业务监控已经接入, 分析自动开启

2.开启智能基线业务的监控项

❌ 没有可以接入的业务指标, 请参考[\[无人值守发布接入指南\]](#)接入业务监控指标

三.日志接入

❌ 请输入日志路径包含文件名,换行分隔.如此处已填写路径,请确认路径是否正确,缺失路径请换行追加,没有问题请点击提交则下次发布不再提示接入



GOPS2018
Shenzhen

目录

1 线上发布之痛

2 我们的解决方案

➔ 3 典型的发布案例

4 无人值守发布实现



典型的发布案例

总览



应用: ██████████
 是否异常: **是**
 异常评分: -
 分析区间: 2018-01-19 16:50到2018-01-19 17:28

异常检测报告

基础监控指标

业务监控指标

已接入的业务监控指标未发现异常
 已接入的业务智能基线监控指标未发现异常

应用日志

✔ 应用纬度发布后异常数量 增长百分之399.53

波动图

⊗ 拦截时间:2018/01/19 17:18:26至2018/01/19 17:28:26涨幅top排行的异常

ⓘ 异常新增IllegalStateException getKeyByKeyRef(KeyServiceImpl.java:100) 出现
 23743次

✔ 已发布机器发布后异常Top排行 波动正常

波动图

实时数据对比

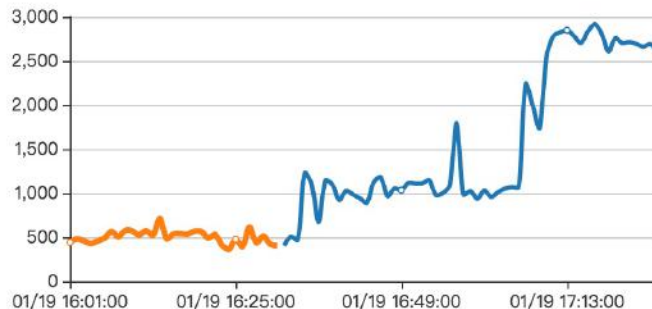
发布前后对比

昨天今天对比

日志详情

appExceptionFluctuation

○ 发布前 ○ 发布后



发布单状态: 已结束

发布单结果: **发布失败**

发布批次: 当前批次 2 | 分 6 批

发布进度: 完成 10000 | 总数 10000
 主机失败率: 67.1%



典型的发布案例

152人未读

a3+无人值守 分分钟接入完成，但是从结果来看，核心应用已经有几十个bug发现

147人未读

库存最近接入无人值守后，效果立竿见影

已读

hi, tripmt刚刚的回滚 是因为发布时检测出的这个日志异常吗?



已读

嗯 是的



GOPS2018
Shenzhen

怎么说清楚价值

准确率

- 有效拦截/总拦截 = 60%

召回率

- 有效拦截/(有效拦截+漏报) = 90%



GOPS2018
Shenzhen

目录

1 线上发布之痛

2 我们的解决方案

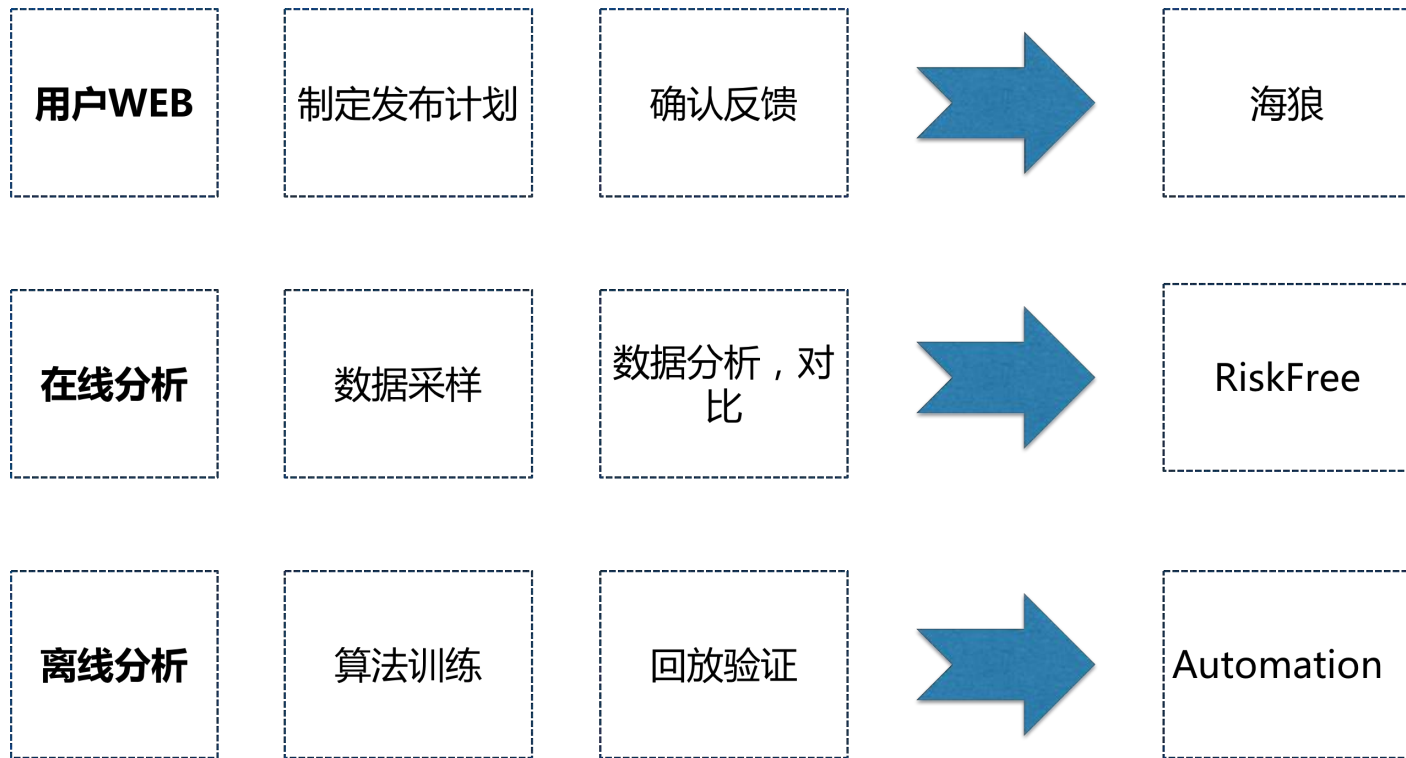
3 典型的发布案例

➔ 4 无人值守发布实现



GOPS2018
Shenzhen

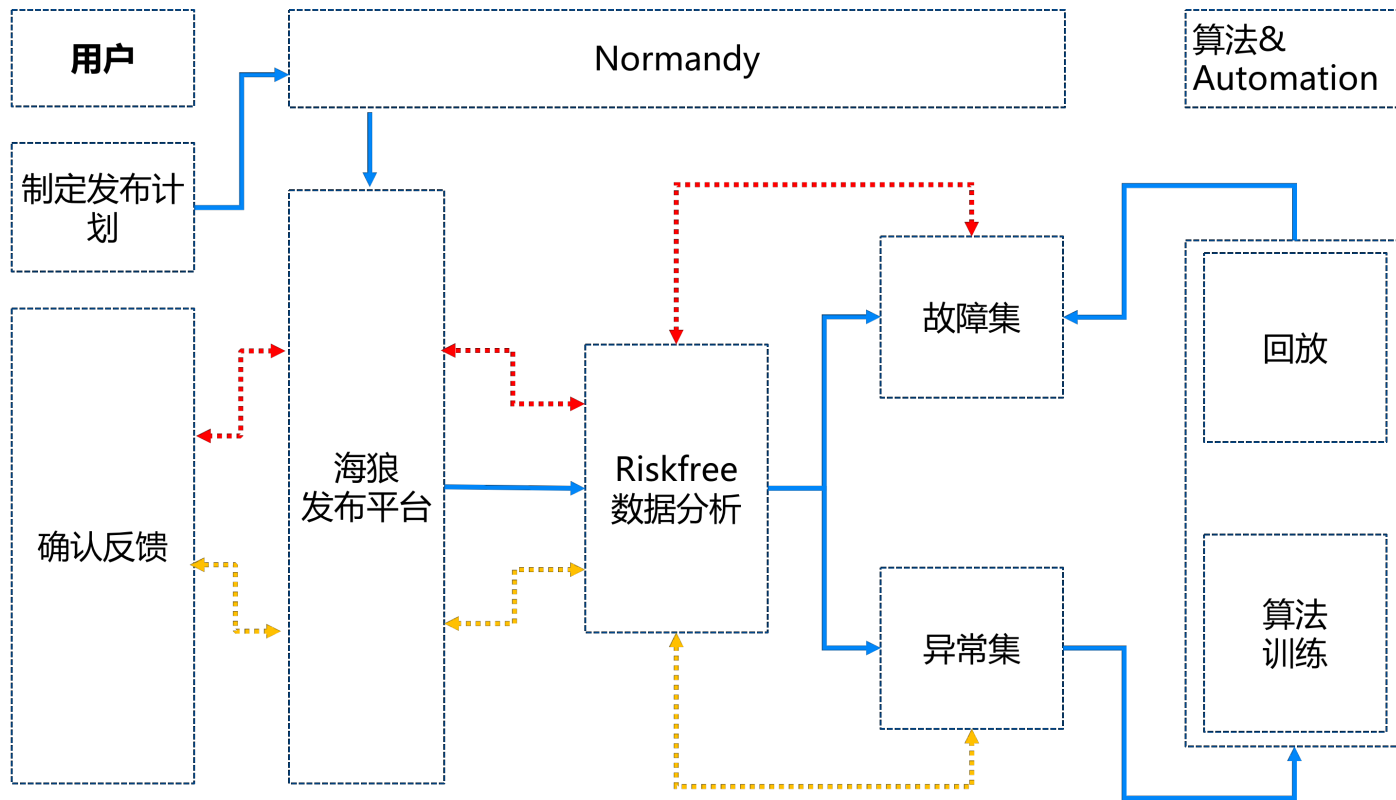
产品架构分层



业务处理流程



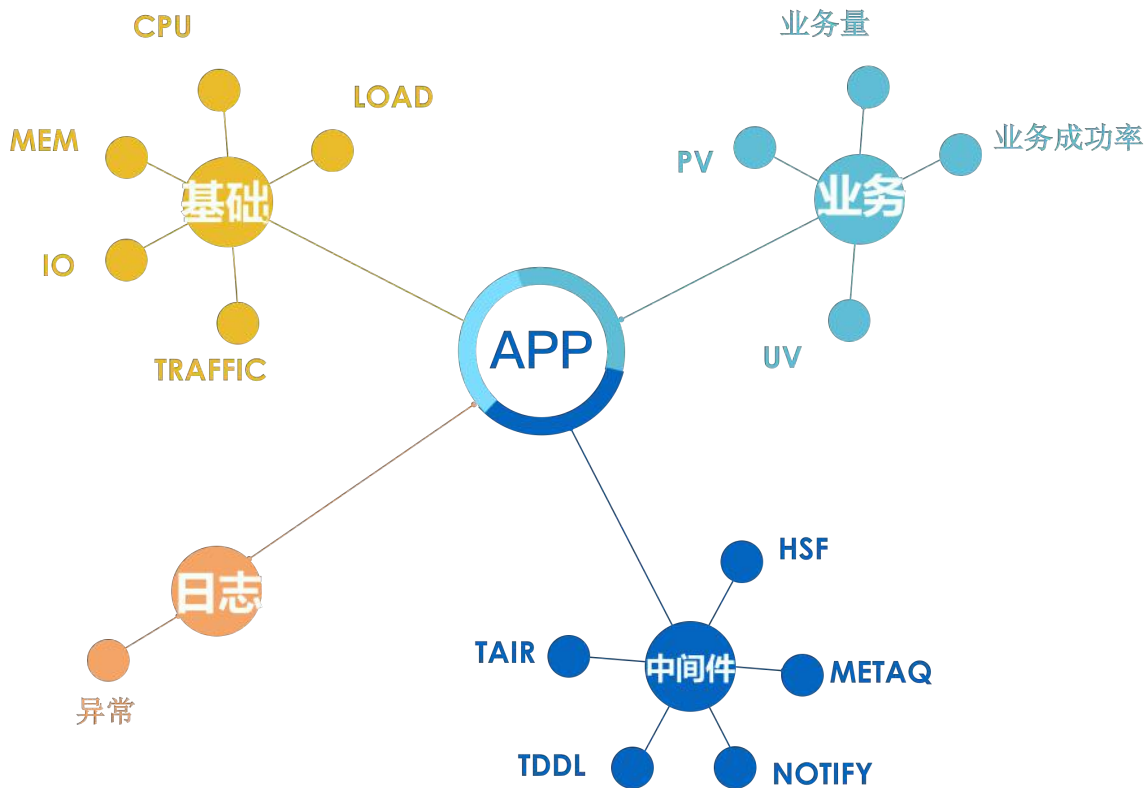
GOPS2018
Shenzhen





GOPS2018
Shenzhen

无人值守发布1.0实现-应用健康度指标





GOPS2018
Shenzhen

无人值守发布1.0实现

特征选取

基础监控

日志

业务监控

中间件指标

异常识别方式

简单规则

复杂算法



GOPS2018
Shenzhen

数据异常识别的挑战



如何在发布过程中快速的采集数据



如何排除各种干扰数据的影响



如何在各种具有不同特性的指标中选择合适的检测方法

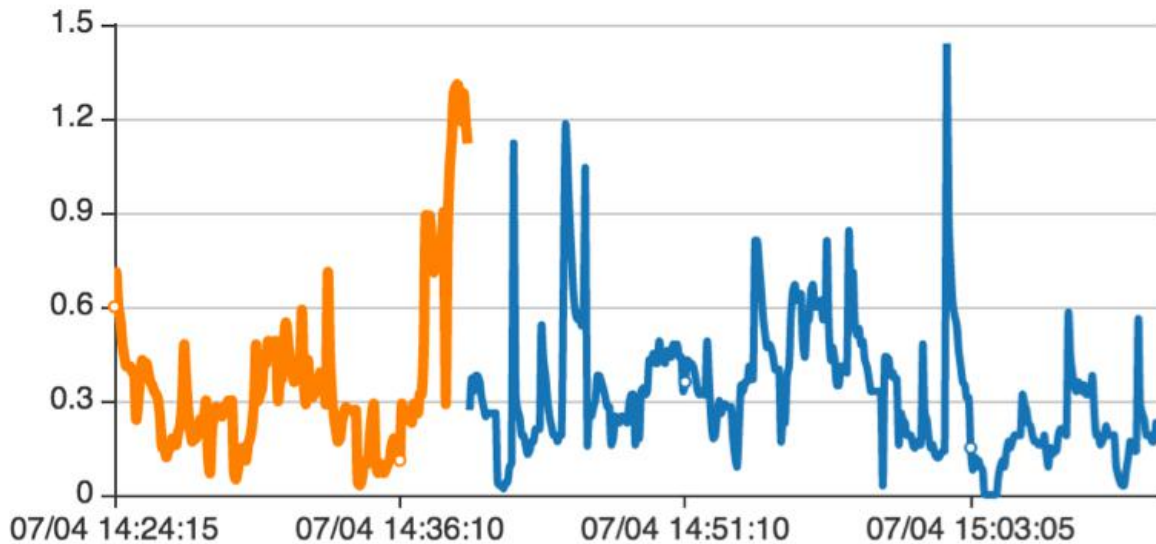
基础监控指标



GOPS2018
Shenzhen

middleware.hsf.provider.qps

—○— 发布前 —○— 发布后





数据采集



- **采集的Metric**
- 系统指标：load,cpu,mem
- 基础服务指标：hsf,notify,metaq,tair,tddl等的rt,qps,success_rate指标
- tomcat: http.failure.count,http.request.error_count ...
- jvm:thread.deadlock.count, gc.concurrentmarksweep.count ...



- **采集哪些机器数据**
- 发布机器
- 参照机器



- **采集时间区间**
- 发布时间点前5分钟，取7天
- 发布完后成当前到前5分钟数据

数据预处理



GOPS2018
Shenzhen





GOPS2018
Shenzhen

我们选择的方法

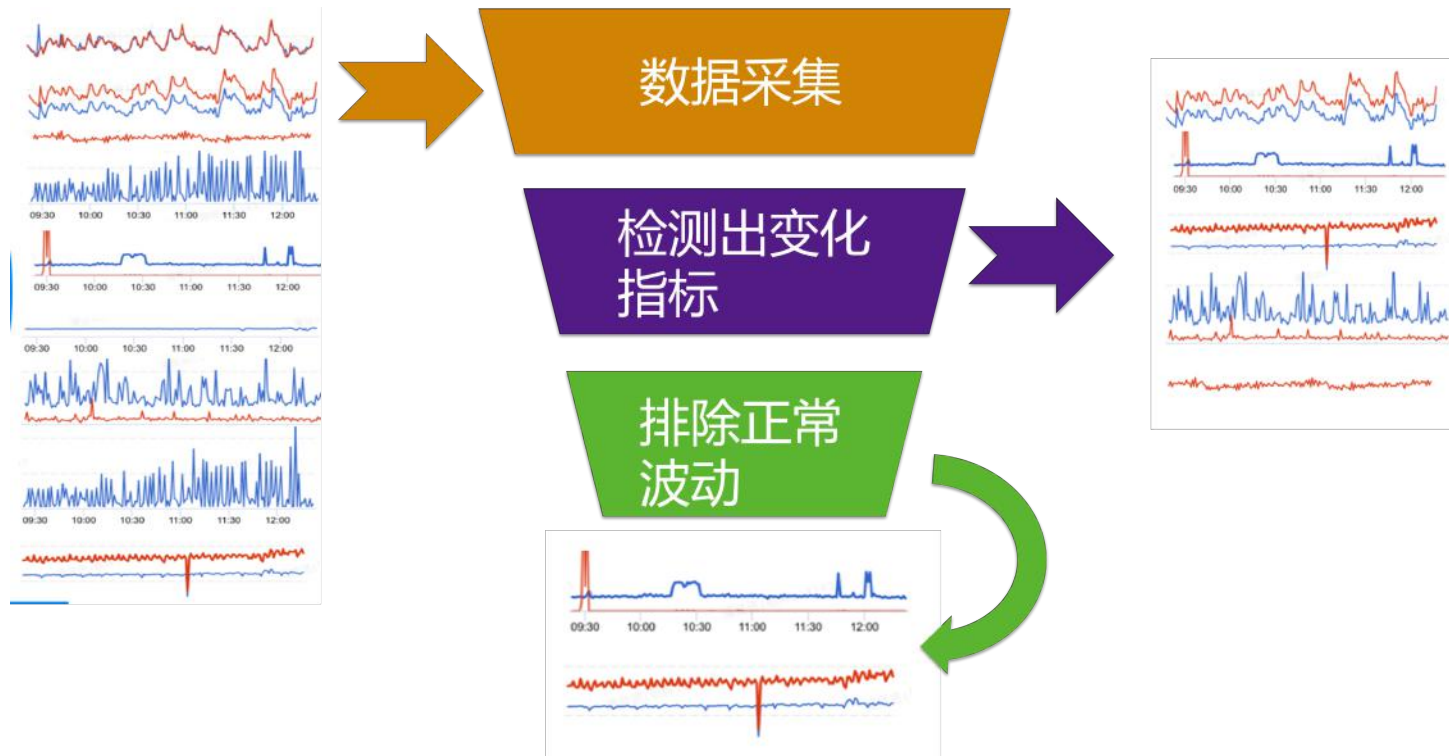
改进型的FUNNEL检测模型

- 按照不同指标特性选择不同方法
- 检测速度快，延迟低
- 计算资源消耗少
- 同时支持很多指标一起分析



GOPS2018
Shenzhen

改进型FUNNEL检测模型





GOPS2018
Shenzhen

如何保障有效的准确率和召回率

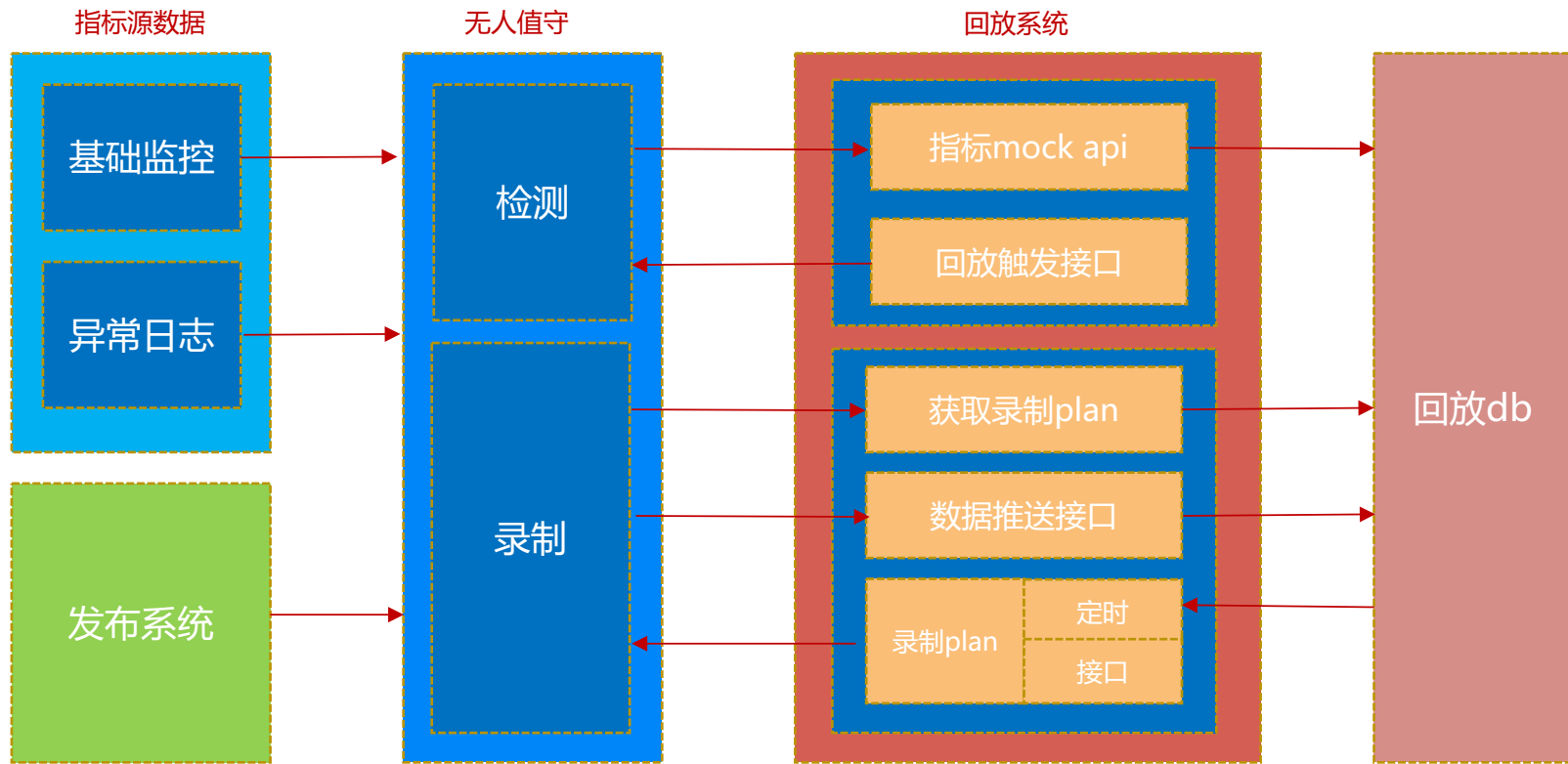
不断的分析误报和漏报数据，微调算法

带来的新的问题：怎么保证算法调整的质量？

答案：无人值守故障回放



无人值守故障回放





GOPS2018
Shenzhen

算法是否解决了我们的问题

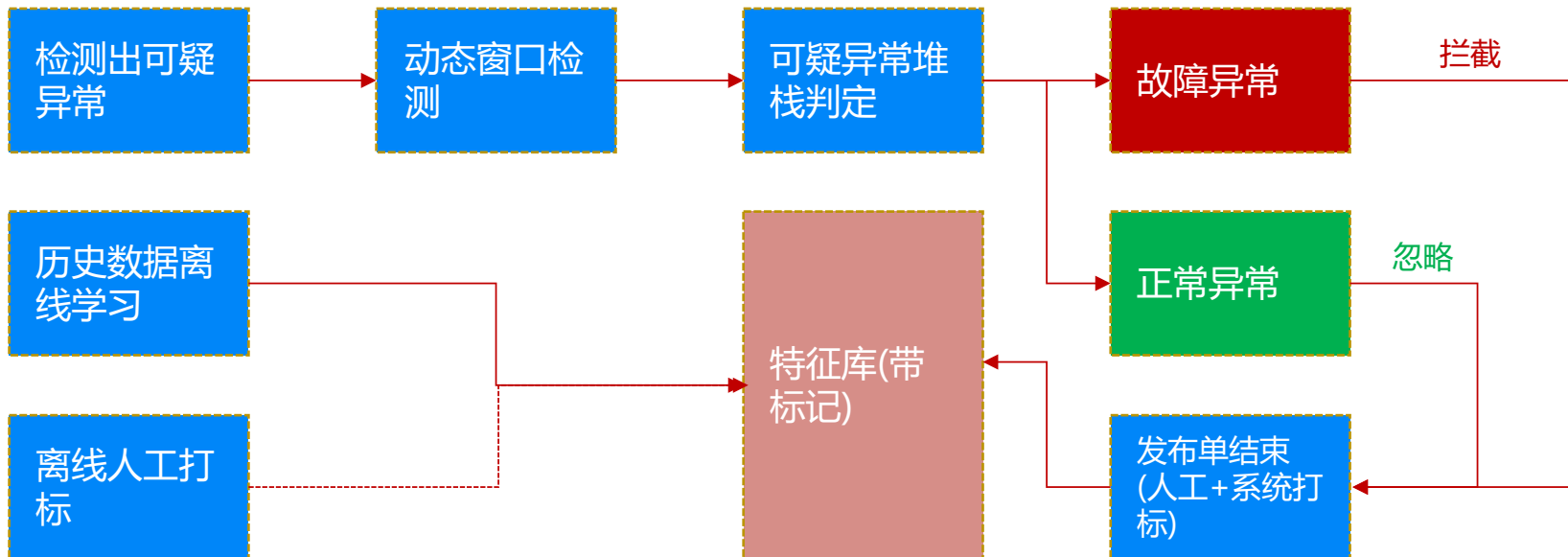
- 需要不断的分析检测数据并调整算法
- 没有办法明确识别出来是否会造成故障
- 之前已经误报的后续发布的时候继续在重复误报

改进方案：

基于机器学习来提升检测准确率



日志检测(机器学习改进)



机器学习要素



GOPS2018
Shenzhen

学习用的样本数据

机器学习的算法

学习结果和实时检测的结合



GOPS2018
Shenzhen

学习用的样本数据

多

VS

少

正常

VS

异常

行为

VS

反馈

关键

VS

不靠谱

拦截指标反馈



GOPS2018
Shenzhen

分析时间:-

分析指 是否

4 个异常

无人值守...
如确认没有...
请评估是否...

强制恢复

无人值守检测到以下指标有异常请确认:

指标类:	异常项	是否应该据此暂停发布: ?	操作
基础监控指标	未发现异常		查看详情
业务监控指标	[03-01 19:56 订单创建_金丝雀测试] 最近1分钟求和值为: 200 < 10000001.00	<input type="button" value="是"/> <input type="button" value="否"/>	查看详情
	[03-01 19:55 订单创建_金丝雀测试] 最近1分钟求和值为: 100 < 10000001.00	<input type="button" value="是"/> <input type="button" value="否"/>	
异常日志	未发现异常		查看详情

有问题请联系[xx] 本次发布不再暂停 ?

部署详情

全部暂停 恢复 切换

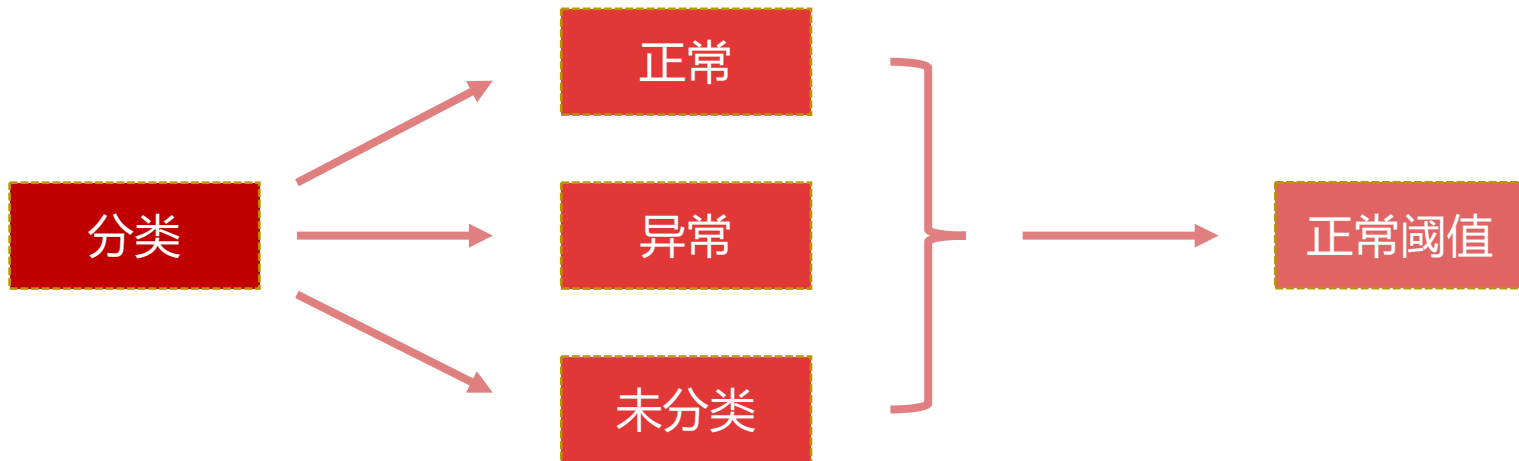
进行中 未开始 已完成

请输入IP

机器学习的算法



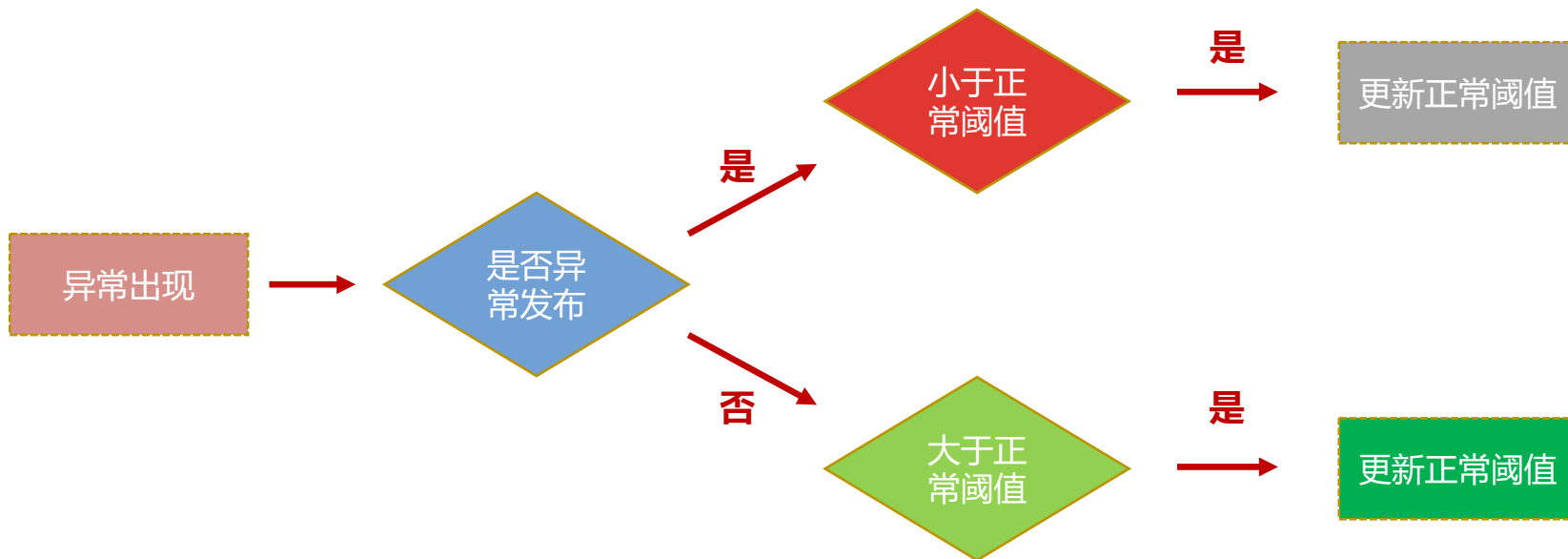
GOPS2018
Shenzhen



机器学习的算法



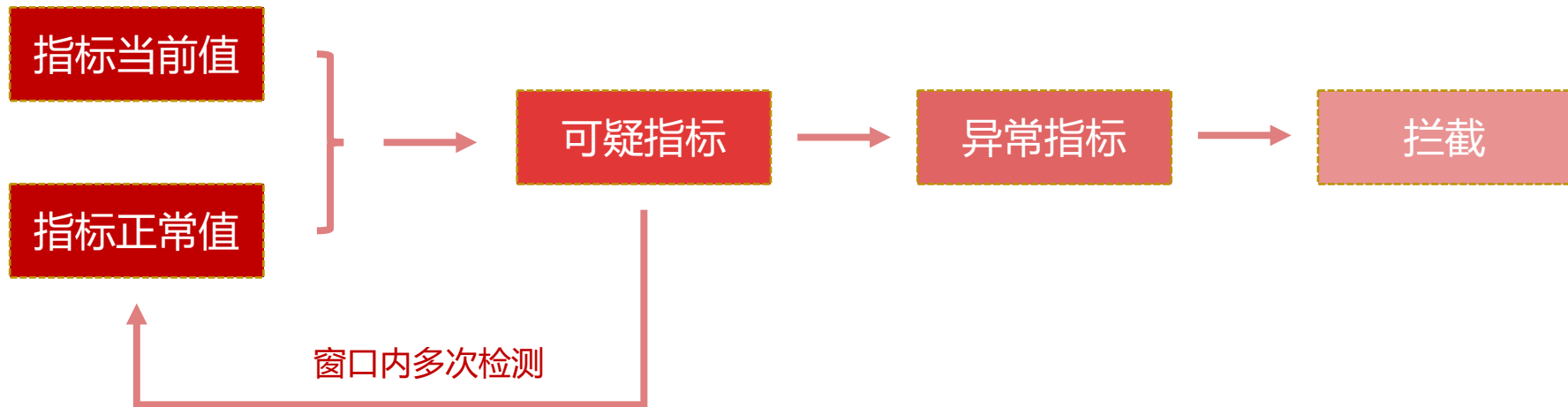
GOPS2018
Shenzhen



学以致用



GOPS2018
Shenzhen





GOPS2018
Shenzhen

Q & A



钉钉



微信



团队公众号



GOPS2018
Shenzhen



Thanks

高效运维社区
开放运维联盟

荣誉出品



GOPS2018
Shenzhen

想第一时间看到高效运维社区的
最新动态吗？

