

小米监控实践之路

陈帅

小米工程师

北 京

伦 敦

纽 约

旧金山

圣 保 罗

上 海

东 京

QCon

全球软件开发大会

[上海站]

主办方 **Geekbang** 极客邦科技 **InfoQ**

信息安全

人工智能

机器学习

黑产

互联网金融 (FinTech)

云计算

基础设施

团队管理

微服务

软件性能

硅谷

互联网架构

2017年10月17-19日
上海·宝华万豪酒店

——> 扫描二维码
开启软件开发新思路





Geekbang> | EGO EXTRA GEEKS' ORGANIZATION NETWORKS
极客邦科技

EGO会员招募季

EGO旨在组建全球最具影响力的技术领导者社交网络，联结杰出的技术领导者学习和成长。

2017年6月30-7月10

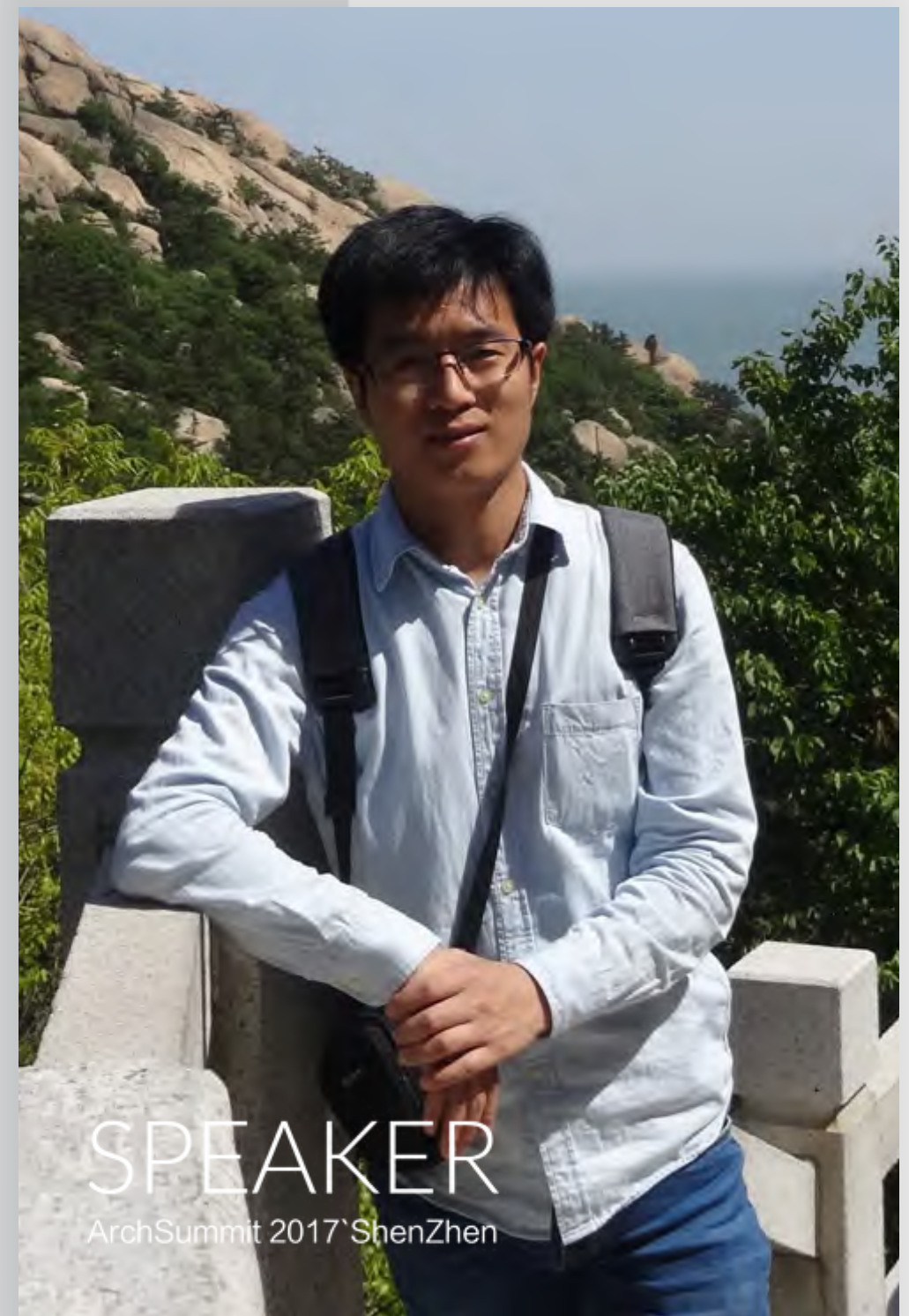


扫码报名

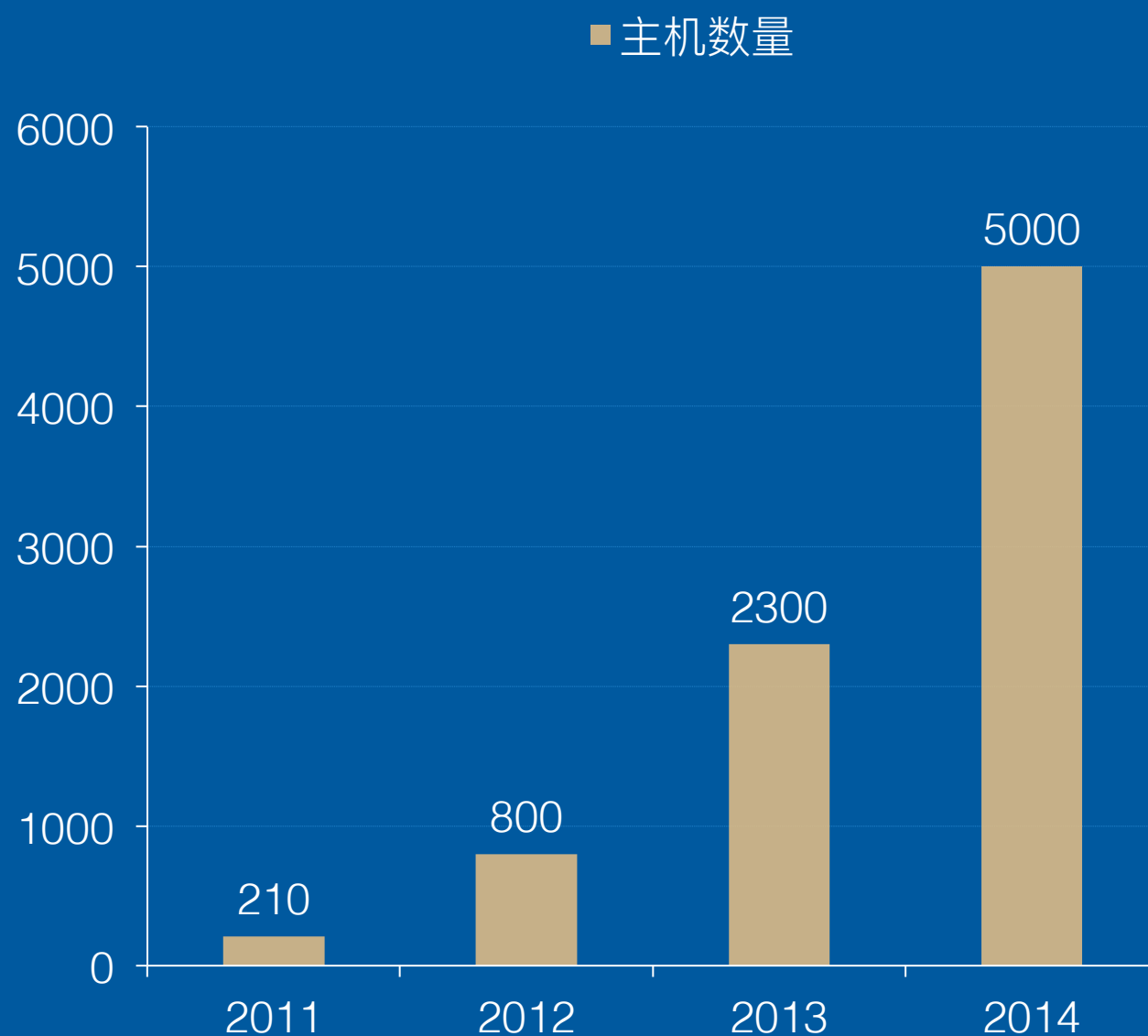
SPEAKER INTRODUCE

陈 帅 小米工程师

- 多年从事于监控系统、部署系统、缓存系统的研发工作，热衷于自动化运维理念；多次参与中国移动等竞标缓存系统的设计工作
- 现专注于小米监控的设计和研发，及开源版本Open-Falcon的维护工作，是主要开发者之一，对于监控系统的应用场景和实践有较深入的理解



曾经面临的困境



- ✕ 监控添加赶不上业务增长
- ✕ 监控扩容赶不上主机增长
- ✕ 监控规划赶不上机房变化
- ✕ 性能优化一次次到达瓶颈

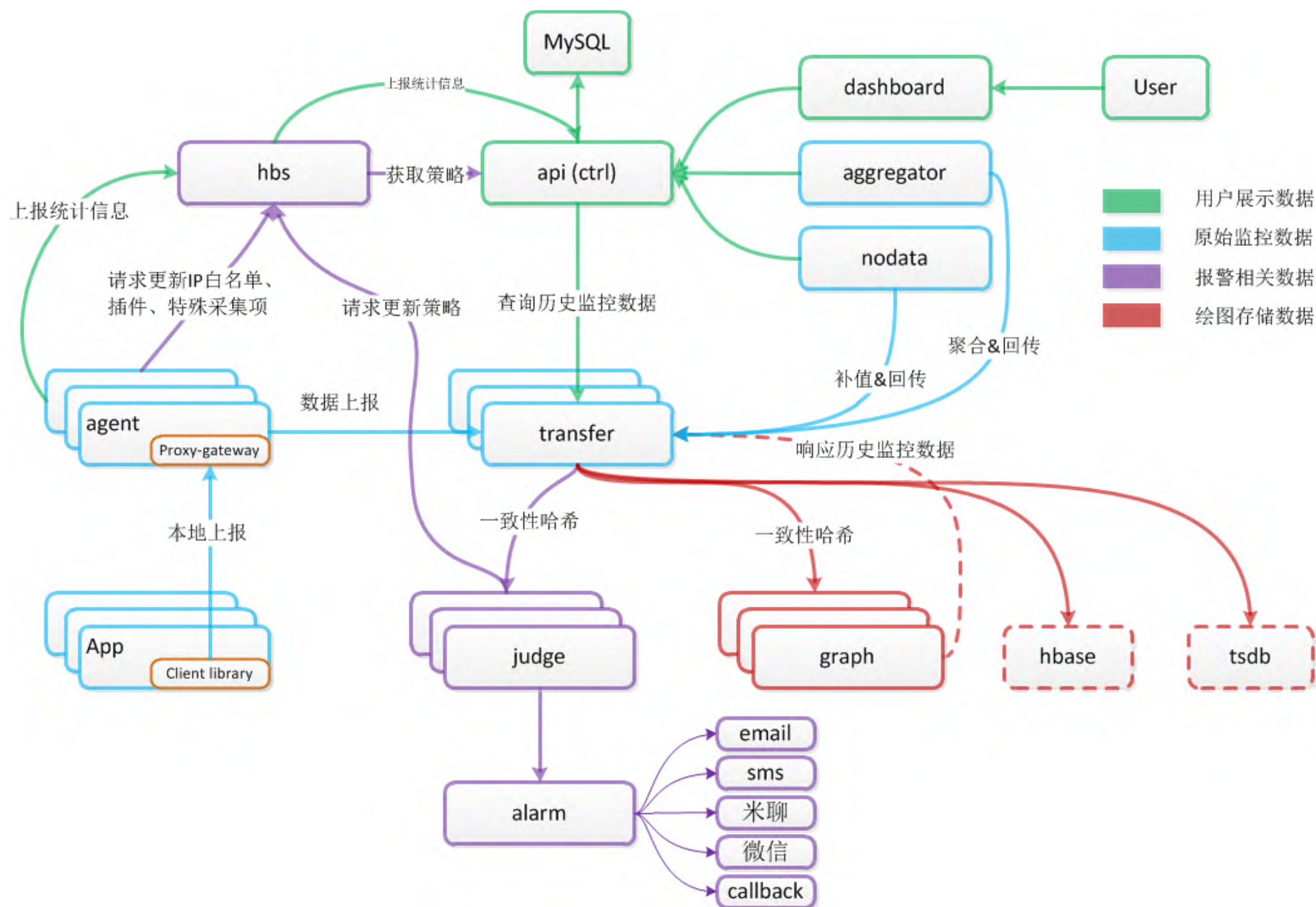
改变

全生命周期自动化

TABLE OF
CONTENTS 大纲

- **监控体系设计**
- 全周期自动化
- 体系持续演进

监控体系设计 - 监控架构图



监控体系设计 - 全push采集

■ 采集免配置



监控体系设计 - metric+tags

HostGroup : miui.apm.nginx

Strategy : df.bytes.free.percent all(#3)<80



df.bytes.free.percent/mount=/
df.bytes.free.percent/mount=/home
df.bytes.free.percent/mount=/data

监控体系设计 - 策略模板

- 继承
- 重载

The screenshot displays the 'Xbox' monitoring system interface. A modal window titled '模板基本信息' (Template Basic Information) is open, showing details for a template named 'tname'. The dialog includes a '模板中的策略列表' (List of strategies in the template) table with columns for id, strategy name, trigger conditions, maximum alarm count, alarm level, and validity period. The table lists two strategies: 'me1 / a=b' and 'me2 / c=d'. Below the table are navigation controls (previous, next, jump to page) and '取消' (Cancel) and '确定' (Confirm) buttons. The background interface shows a sidebar with a search tree of nodes (e.g., cop=xiaomi, owt=miui) and a main area with a list of templates (aaa, acx, tname, tname_test01, tname_test2) and their associated actions (clone, view binding, modify, delete).

模板基本信息

模板名称: tname 父模板名称: aaa_copyEBBC8D6F
报警接收组: build-alert

模板中的策略列表

id	策略基本信息(metric/tags [note])	触发条件	最大报警次数	报警级别	生效时间段
17240	me1 / a=b [无]	all(#3) >= 0	3	3	00:00 - 23:59
17241	me2 / c=d [无]	all(#3) >= 0	3	3	00:00 - 23:59

< 1 > 跳至 1 页

取消 确定

TABLE OF CONTENTS 大纲

- 监控体系设计
- 全周期自动化
- 体系持续演进

全周期自动化 - 服务树

JOBNAME

公司

部门

产品

集群

服务组

服务

任务组

任务

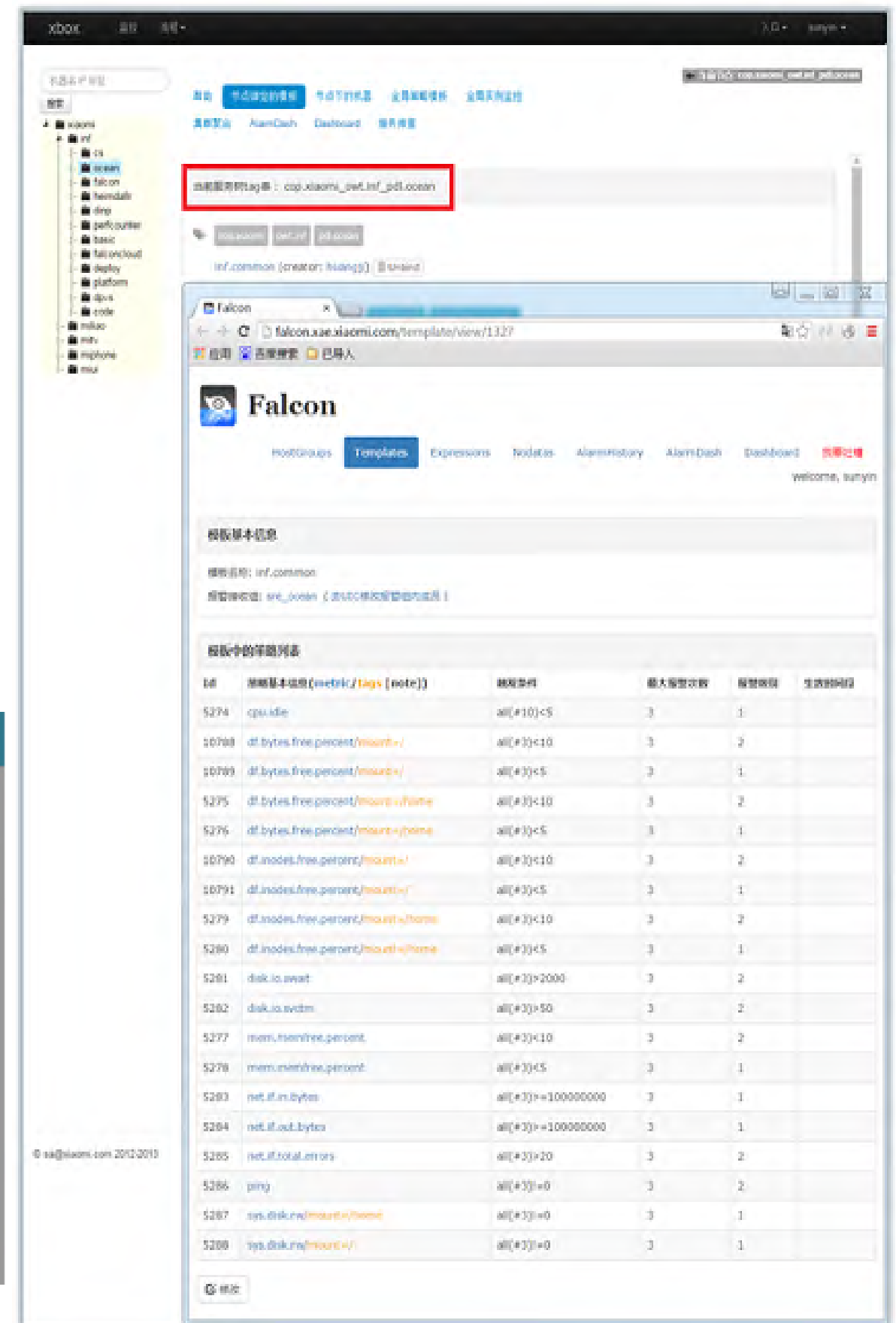
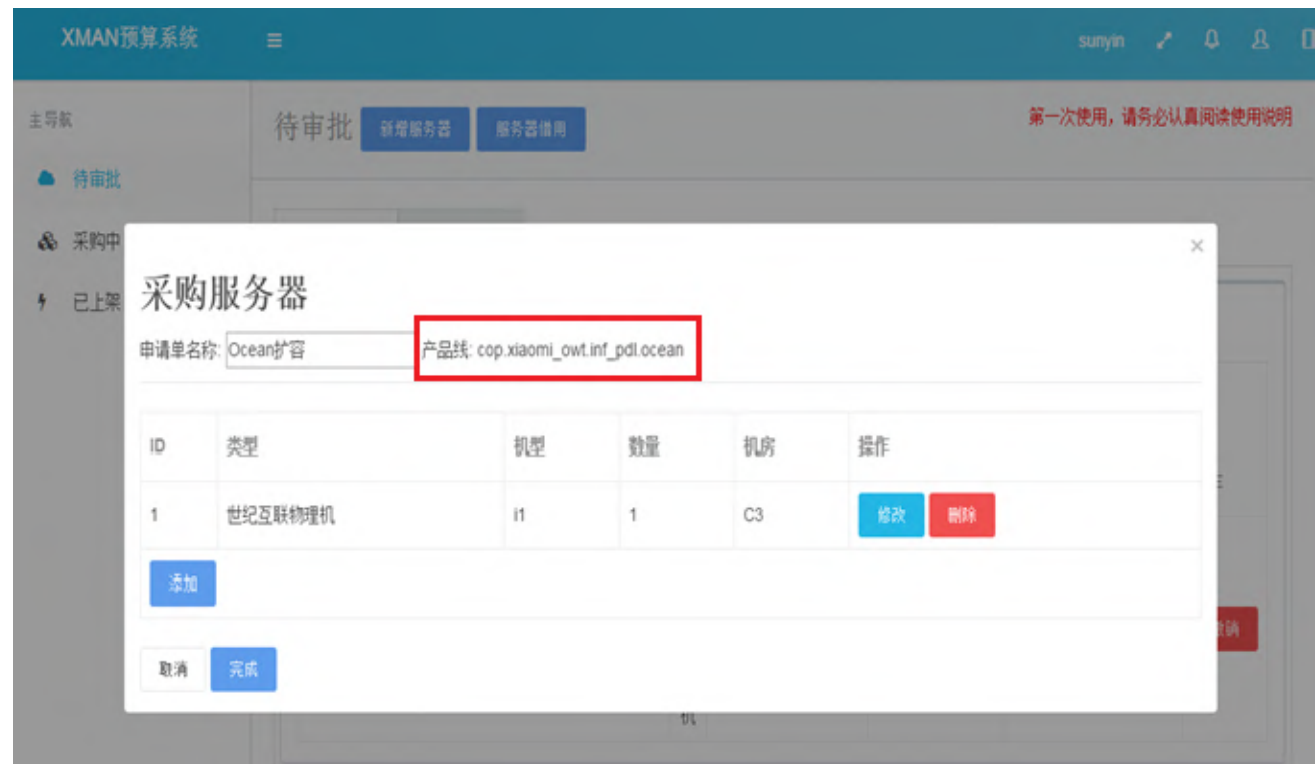
- 为主机打tag
- 8 tag定义唯一的服务

The screenshot displays the '服务树' (Service Tree) interface. On the left, a tree structure shows the hierarchy from 'xiaomi' down to 'mesos-dns' and 'c3'. The main panel shows the 'PATH: cop.xiaomi_owt.inf_pdl.ocean_servicegroup.ocean_service.mesos-dns_job.mesos-dns_cluster.c3'. Below this, there are buttons for actions like '添加到', '从删除', '改名', '故障', '重启', '初始化', '下线', '复制', '详情', '快照', and '回到老版'. A table lists hosts with their names, tag combinations, and actions. The tag combination for the first host is highlighted with a red box.

机器名	标签组合	操作
c3-com-mesosdns01.bj	cop.xiaomi_owt.inf_pdl.ocean_servicegroup.ocean_service.mesos-dns_job.mesos-dns_cluster.c3_idc.c3_loc.bj_status.online	改名 重启
c3-com-mesosdns02.bj	cop.xiaomi_owt.inf_pdl.ocean_servicegroup.ocean_service.mesos-dns_job.mesos-dns_cluster.c3_idc.c3_loc.bj_status.online	改名 重启

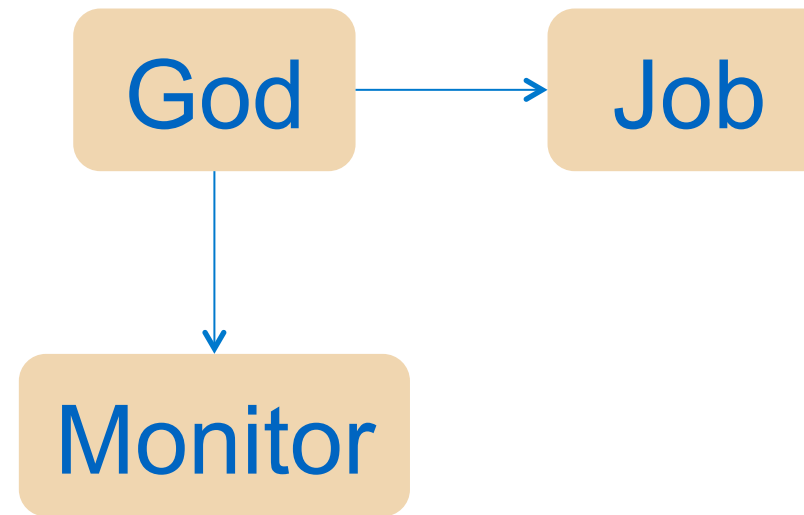
全周期自动化 - 主机交付

交付即监控



全周期自动化 - 产品发布

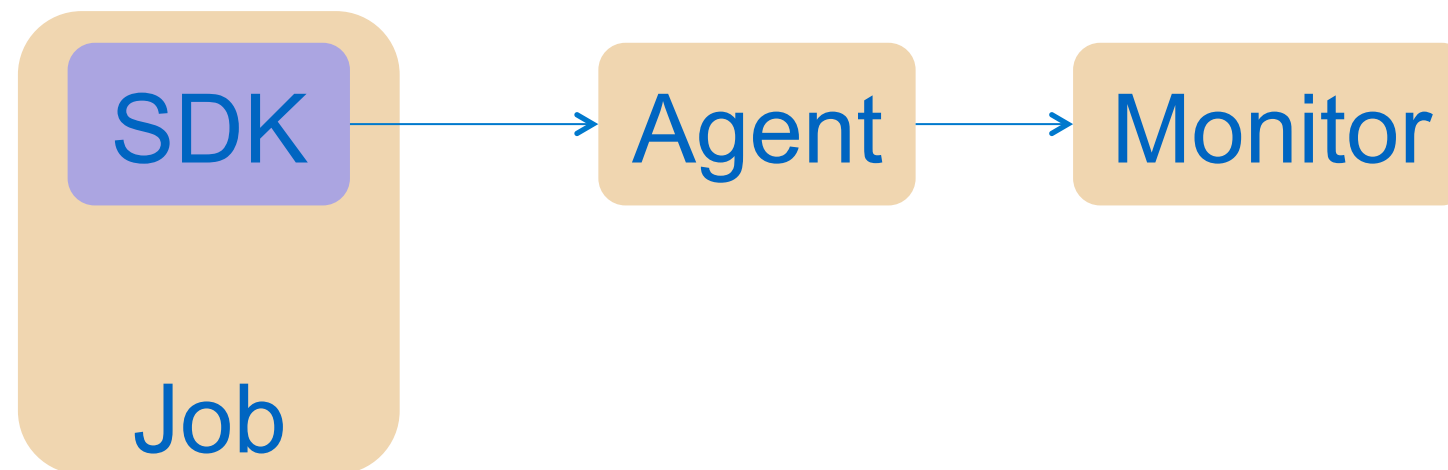
发布即监控



`common.servicerunning all(#3)!=1`

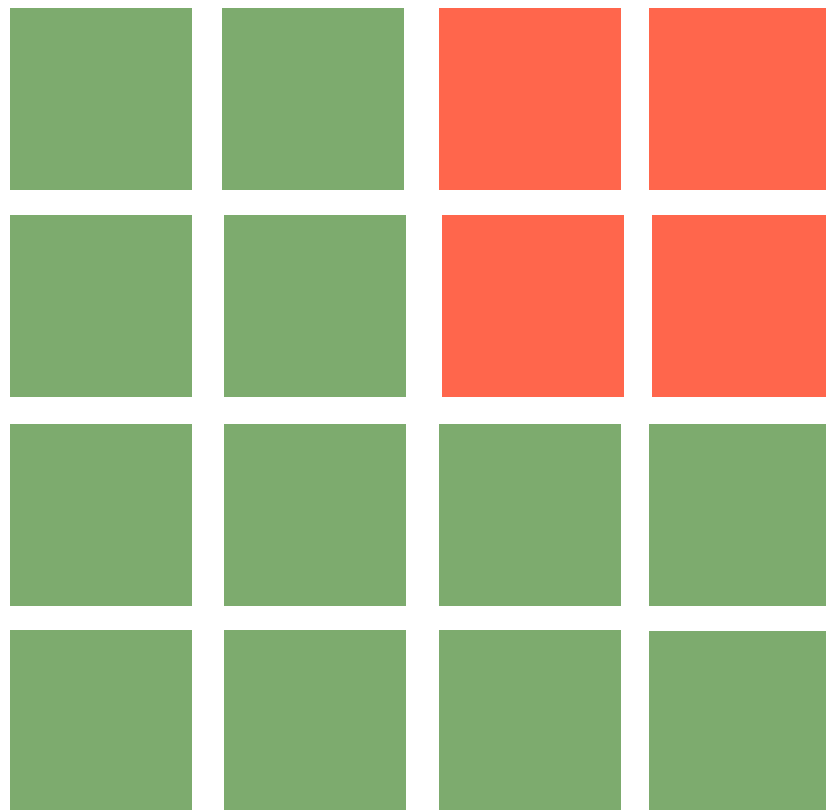
全周期自动化 - 服务启动

启动即监控



request-failrate all(#3)>5

全周期自动化 - 集群监控

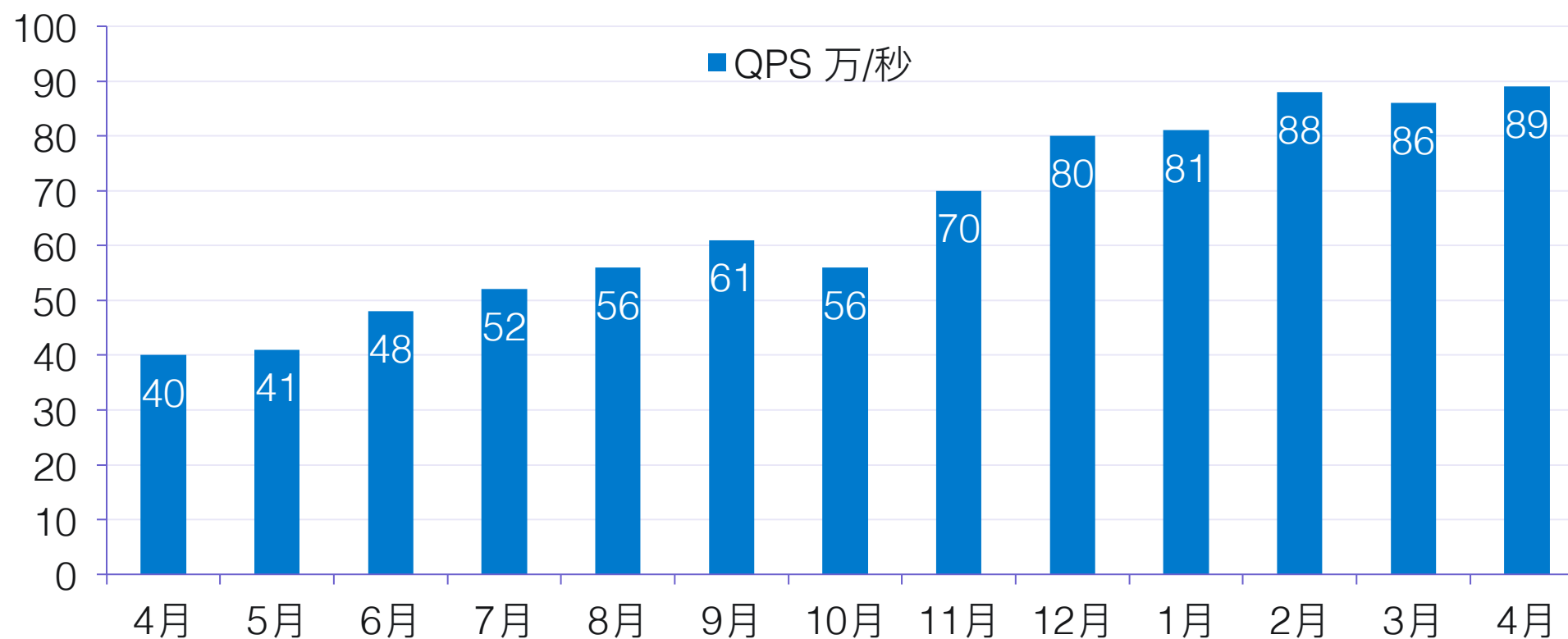


- 集群QPS总和
- 集群平均值
- 集群百分比

分子: $\$(cpu.busy) > 50$
分母: $\$ \#$

全周期自动化 - 成效

Endpoint: 40000+
Counter: 16000万+
SRE: 0.5人



	单位	秒	天	月	年
单个数据结构	183byte	157.07M	13.25T	397.58T	4.84P
单个监控数值	8byte	6.87M	579.36G	17.38T	211.47T

TABLE OF CONTENTS 大纲

- 监控体系设计
- 全周期自动化
- 体系持续演进

体系持续演进 - Nodatas

■ 填充默认值

modify nodata

name:

endpoint选择: ?

metric:

tags:

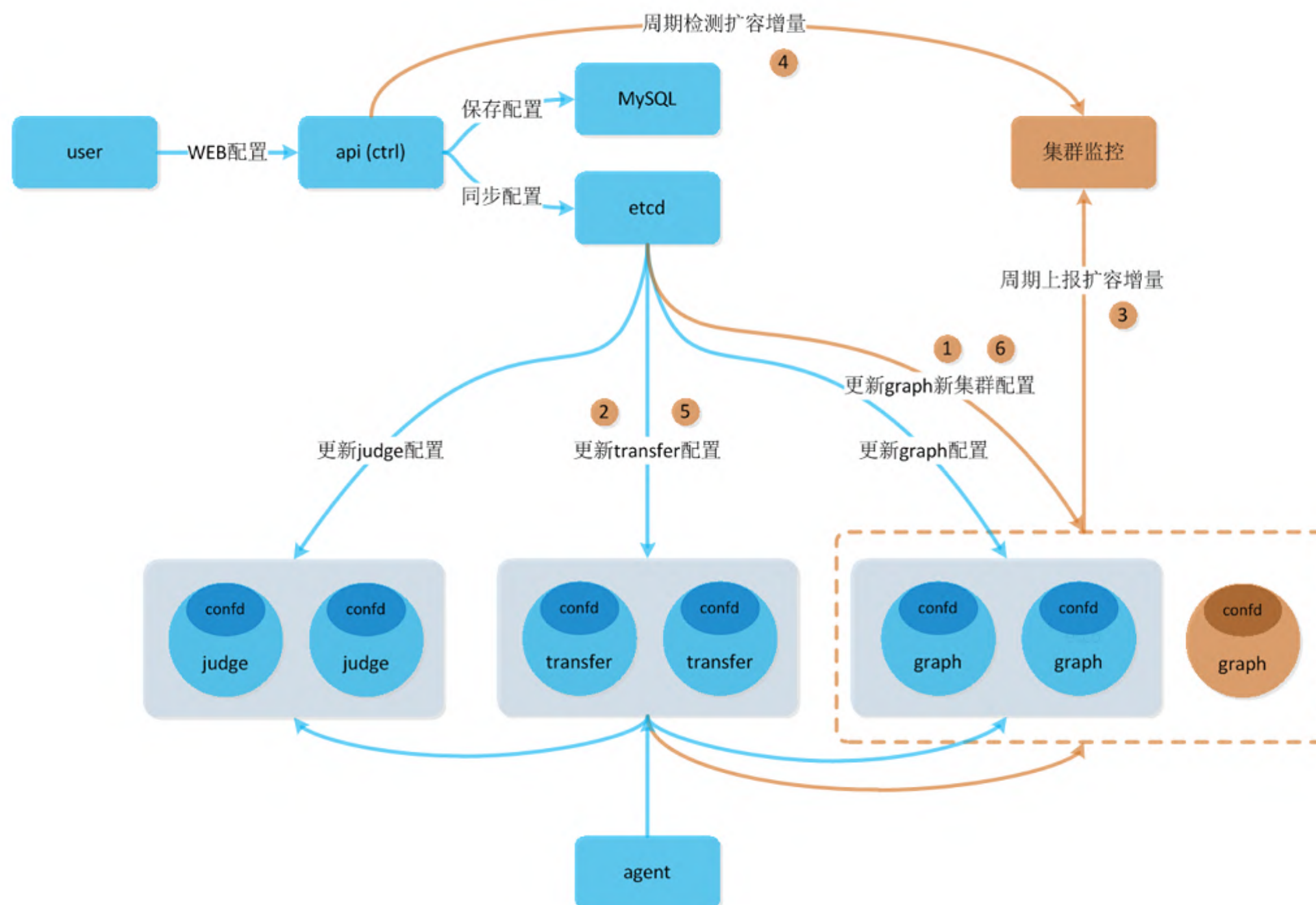
type:

周期(秒):

数据上报中断时, 补发如下值:

[? Nodata监控用户手册](#)

体系持续演进 - 自动配置扩容



体系持续演进 - 非连续性报警

■ 抖动型数据



体系持续演进 - 容器弹性环境

- 去除PING监控，改为JOB粒度实例数监控
- 容器退出时钩子清理告警事件
- 平滑dashboard曲线

体系持续演进 - 自动缩扩容

■ Falcon自动采集

CPU IDLE

MEM FREE

PROC QPS

PROC DELAY

■ Hook回调

Marathon API

■ 最短5s触发伸缩

动态调度部署系统

job.ocean-monitor-job_service.ocean-monitor-job_cluster.production-lg_pdl.ocean_owt.inf_cop.xiaomi

任务配置 运行时配置 自动伸缩 高级选项 取消 保存

伸缩指标 ?

CPU平均使用率

伸缩配置

实例数上限: 5 实例数下限: 1 实例增减步长: 1

是否通知: ☒ 短信通知组 (UIC): sre_ocean

触发条件 ?

连续次数

加实例: 3 > 减实例: 3

阈值 (单位: 百分比)

80 60

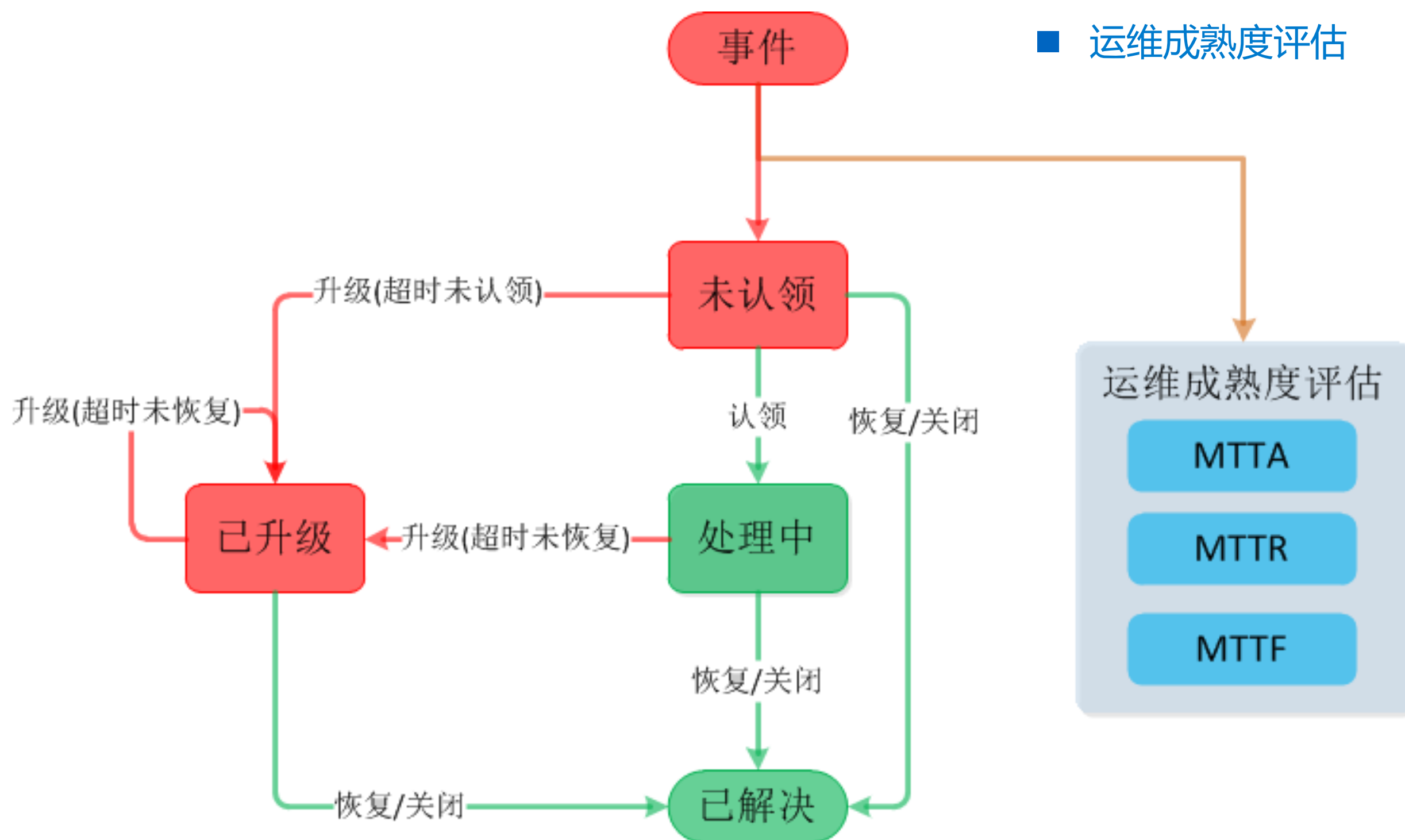
体系持续演进 - 网络监控

- 网络设备监控
- 南北、东西流量监控
- Pingmesh全网连通质量监控

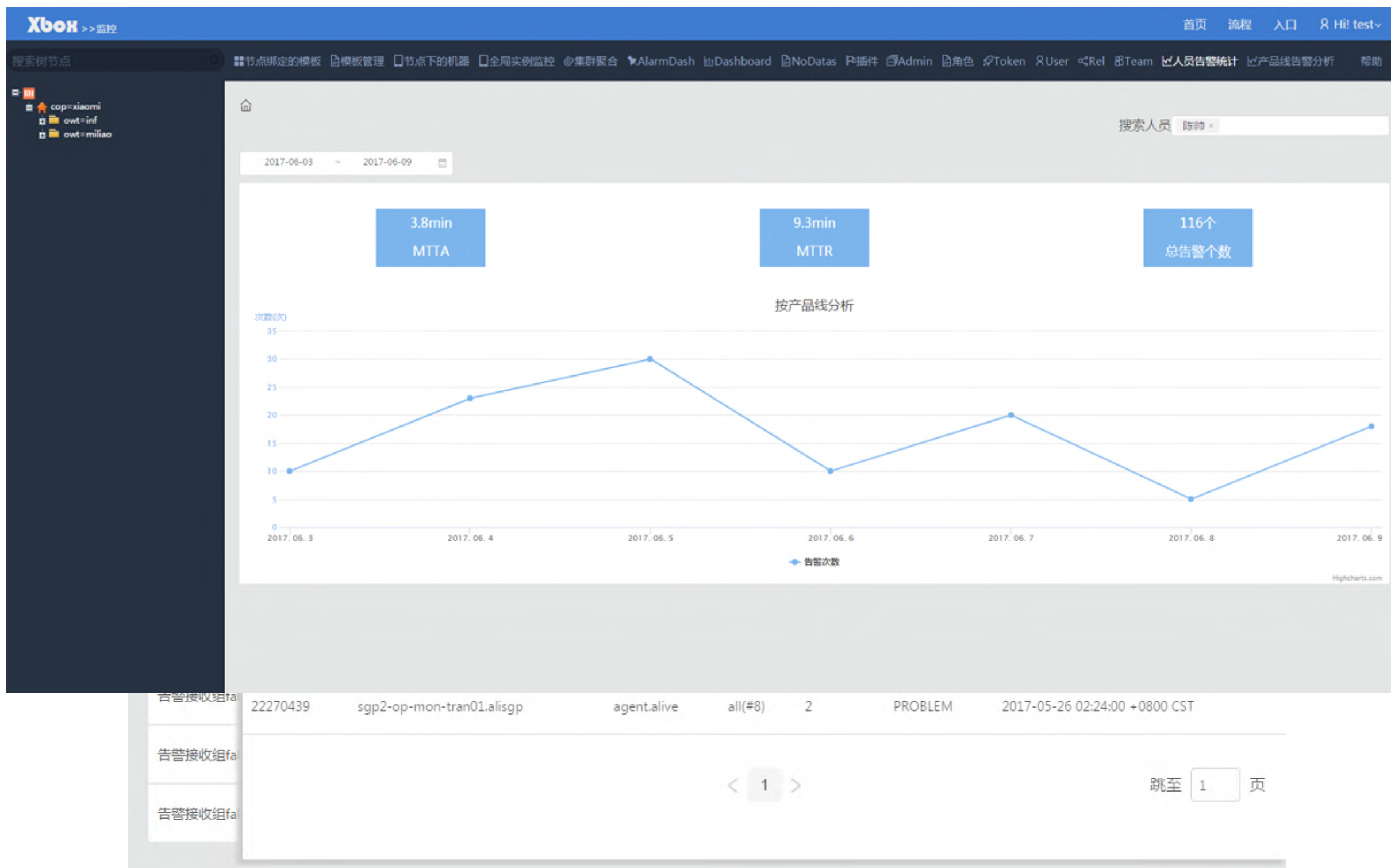


体系持续演进 - 告警管理

- 告警认领
- 告警升级
- 运维成熟度评估

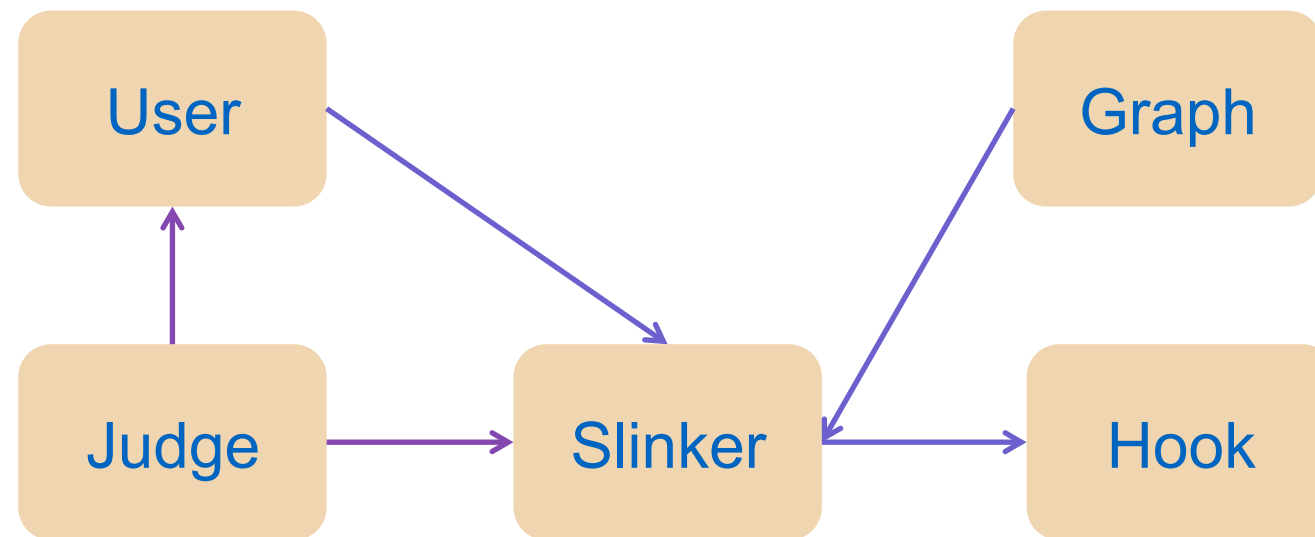


体系持续演进 - 告警管理



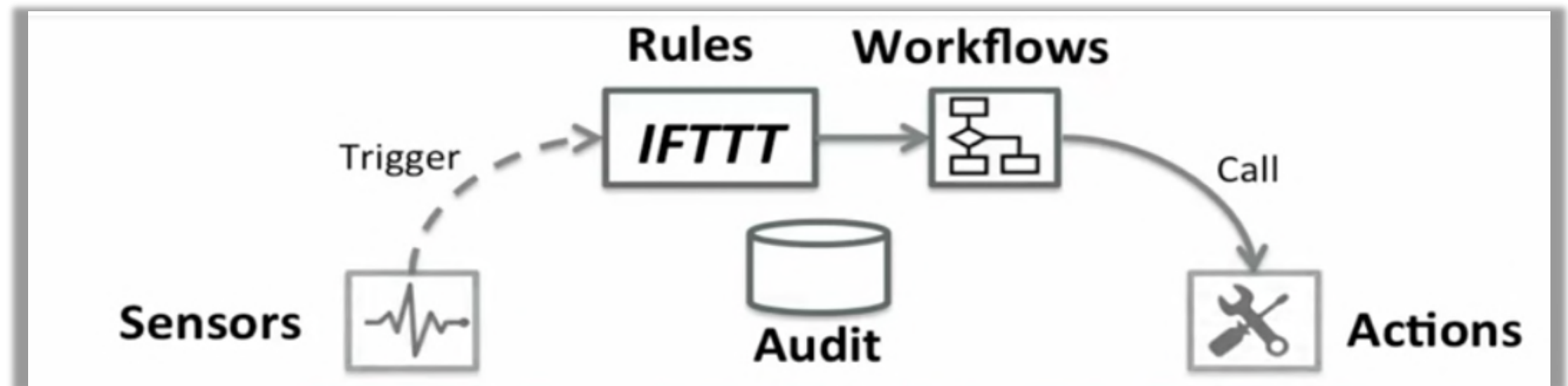
体系持续演进 - 智能监控

- 阈值推荐
- 告警内容丰富
- 故障精准定位（决策树、机器学习）



体系持续演进 - 告警驱动

stackstorm



- 接收响应报警事件
- Chat命令触发自动修复



- 故障设计
- 精准故障定位
- 服务治理框架
- 故障自愈

总结

- ✓ 基础设施
- ✓ 体系化
- ✓ 可扩展

THANKS!

让创新技术推动社会进步

HELP TO BUILD A BETTER SOCIETY WITH
INNOVATIVE TECHNOLOGIES

Geekbang >

极客邦科技

InfoQ

专注中高端技术人员的技术媒体



EGO EXTRA GEEKS' ORGANIZATION
NETWORKS

高端技术人员学习型社交平台



StuQ
斯达克学院

实践驱动的 IT 教育平台

