

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



全球软件开发大会

[上海站]

主办方

Geekbang > InfoQ
极客邦科技

2017年10月17-19日
上海·宝华万豪酒店

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

→ 扫描二维码
开启软件开发新思路





Geekbang | EGO EXTRA GEEKS' ORGANIZATION
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

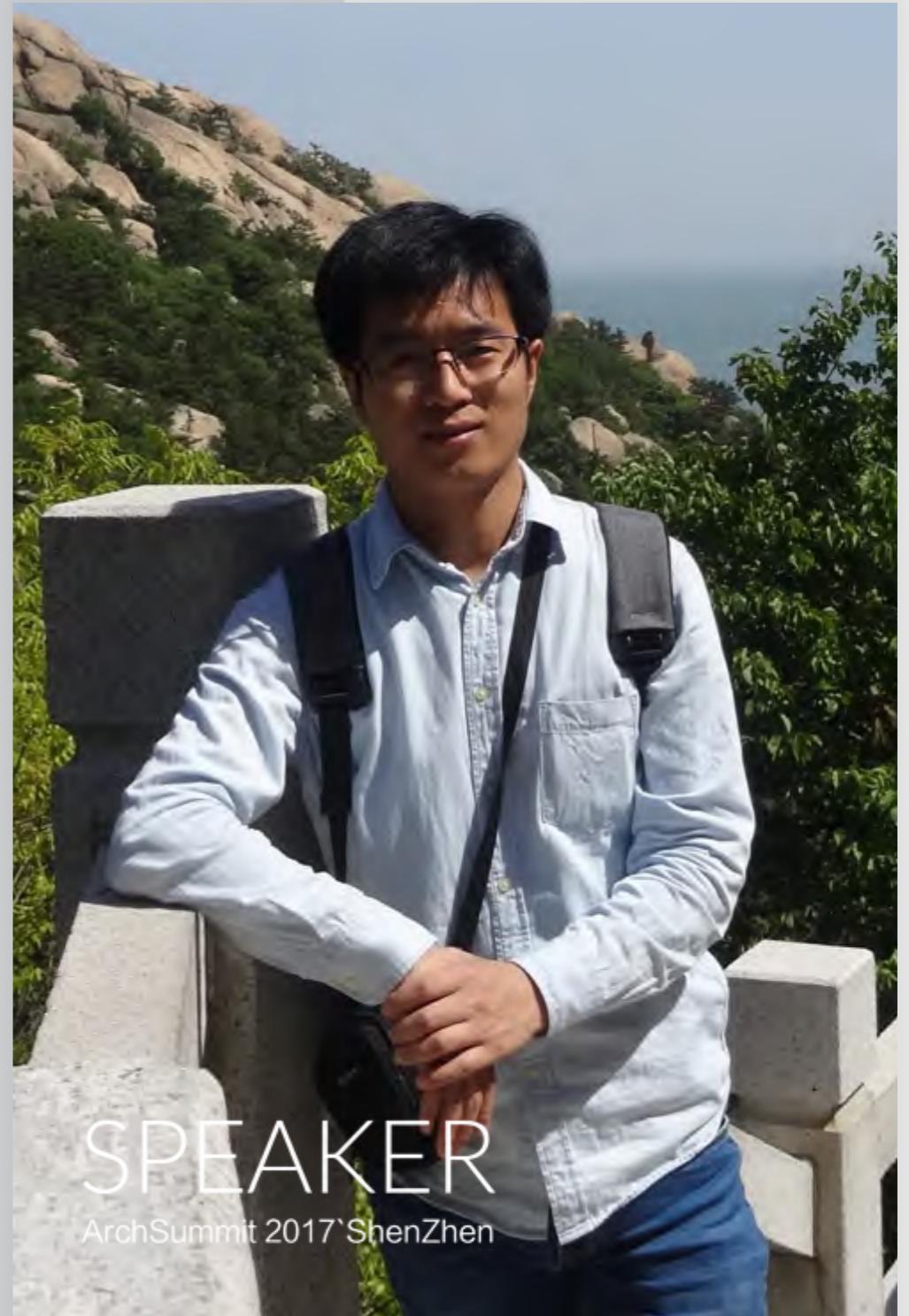


扫 码 报 名

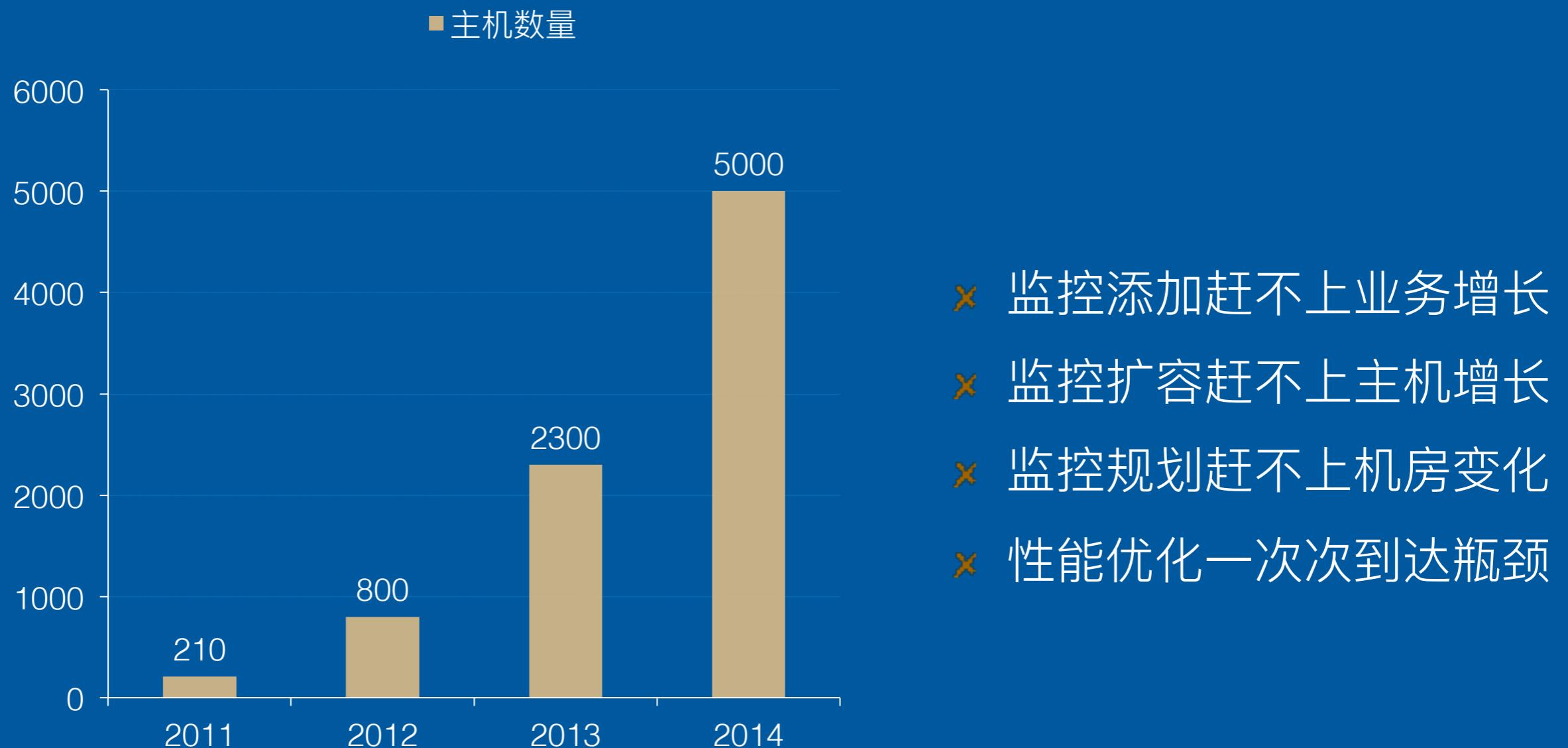
SPEAKER
INTRODUCE

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



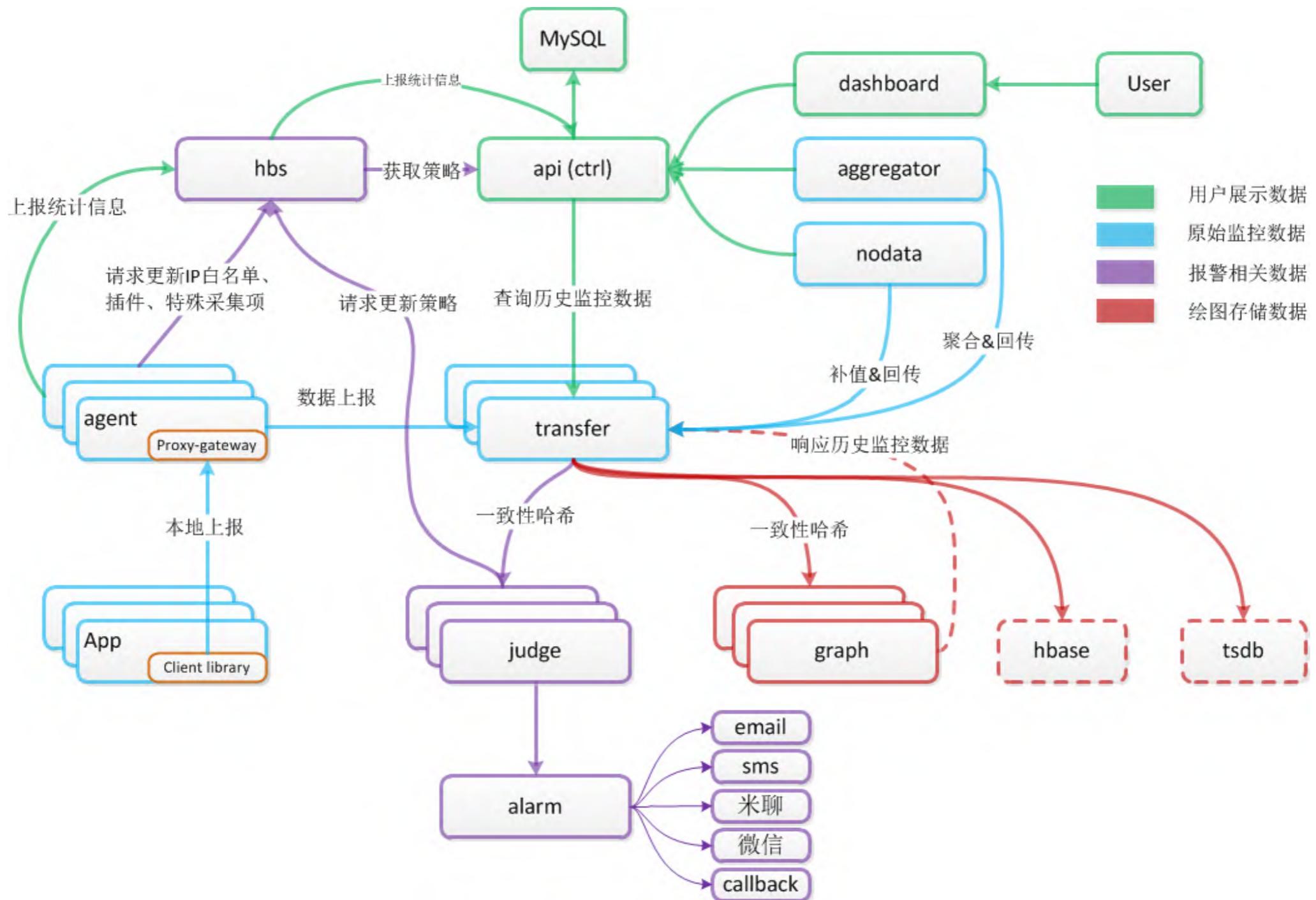
改变

全生命周期自动化

TABLE OF
CONTENTS 大纲

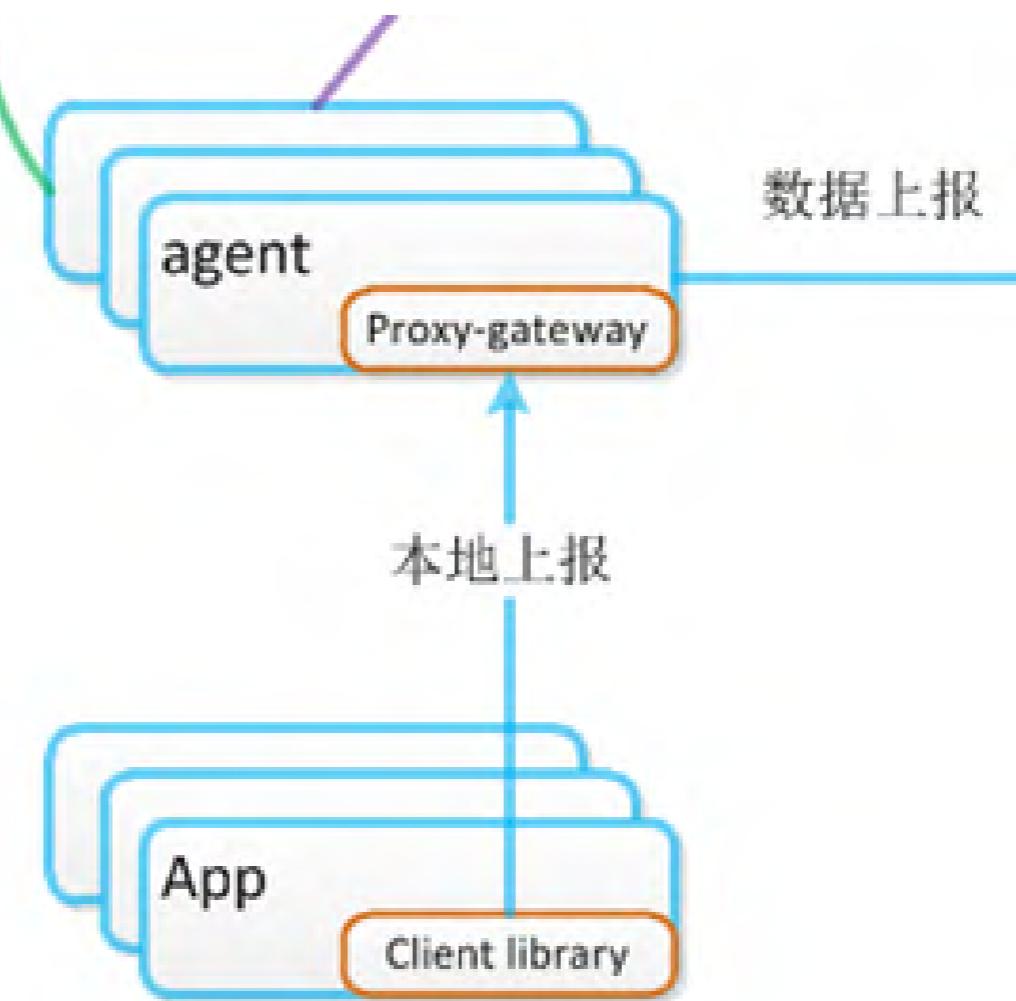
- 监控体系设计
- 全周期自动化
- 体系持续演进

监控体系设计 - 监控架构图



监控体系设计 - 全push采集

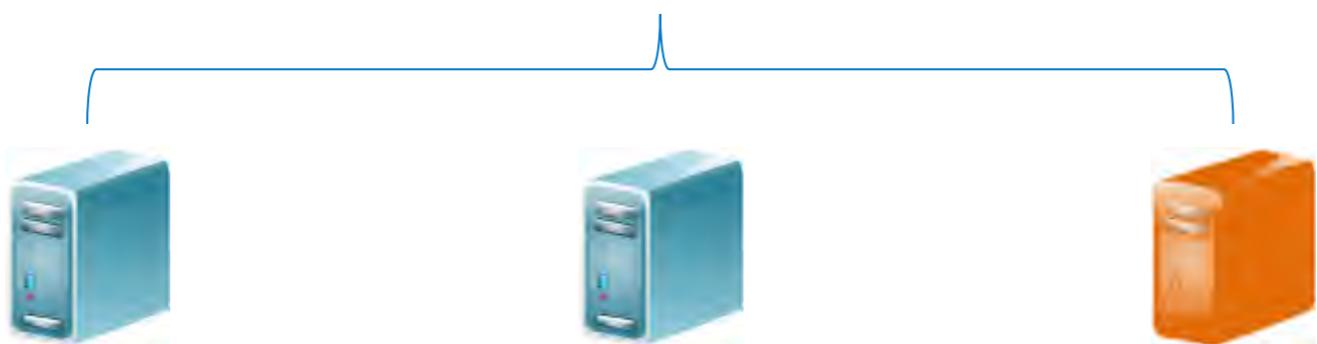
■ 采集免配置



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 shows the Xbos monitoring system interface. At the top, there's a navigation bar with links like 首页 (Home), 流程 (Process), 入口 (Entry), and a user profile. Below the navigation is a search bar labeled '搜索树节点' (Search tree nodes) and a toolbar with icons for '节点绑定的模板' (Node-bound template), '模板管理' (Template Management), '全局实例监控' (Global instance monitoring), '集群聚合' (Cluster aggregation), 'AlarmDash' (AlarmDash), 'Dashboard', 'NoDatas', 'Plugins', 'Admin', and 'Roles'.

The main area displays a tree view of templates on the left, with items like 'cop=xiaomi', 'owt=miui', 'owt=duokan', etc. A modal window titled '模板基本信息' (Template Basic Information) is open, showing details for a template named 'tname'. It states: 模板名称 : tname 父模板名称 : aaa_copyEBBC8D6F 报警接收组 : build-alert. Below this, a table titled '模板中的策略列表' (List of strategies in the template) shows two entries:

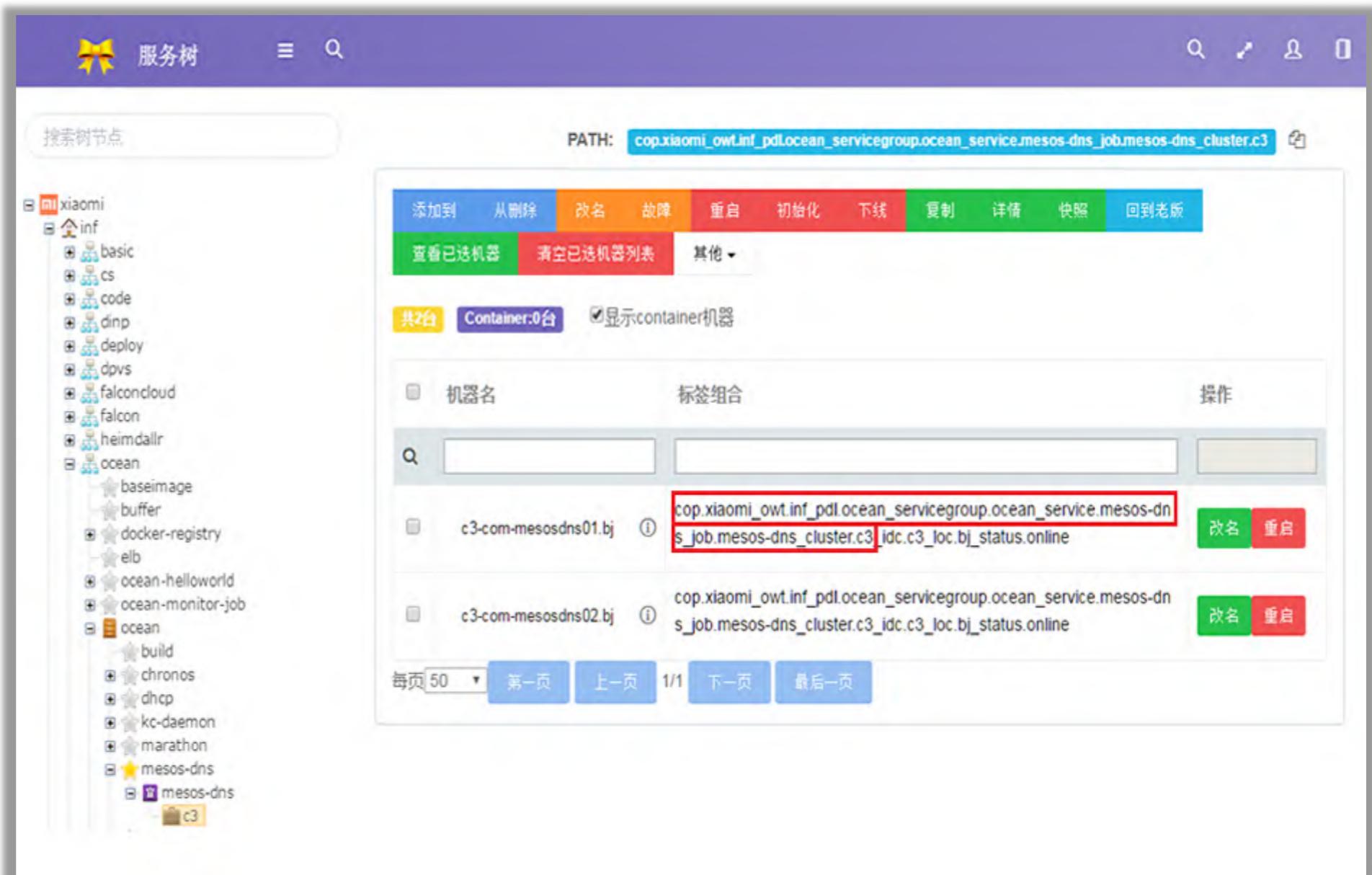
id	策略基本信息(metric/tags [n ote])	触发条件	最大报警次 数	报警级 别	生效时间 段
1724 0	me1 / a=b [无]	all(#3) >= 0	3	3	00:00 - 23: 59
1724 1	me2 / c=d [无]	all(#3) >= 0	3	3	00:00 - 23: 59

At the bottom of the modal, there are buttons for '取消' (Cancel) and '确定' (Confirm). To the right of the modal, a list of operations for the 'tname' template is shown, each with a 'test' creator and options to 'Clone', 'View Binding', 'Modify', and 'Delete'.

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



- 为主机打tag
- 8 tag定义唯一的服务

全周期自动化 - 主机交付

交付即监控

XMAN预算系统

第一次使用，请务必认真阅读使用说明

待审批 新增服务器 服务器借用

待审批

采购中

已上架

采购服务器

申请单名称: Ocean扩容 产品线: cop.xiaomi_owt.inf_pdl.ocean

ID	类型	机型	数量	机房	操作
1	世纪互联物理机	i1	1	C3	修改 删除

添加 取消 完成

XBOX 监控 高级

产品名: falcon

最近的模板 最近的机器 应用商店 金属商店

模板管理 AlarmDash Dashboard 模板库

当前配置的tag: cop.xiaomi_owt.inf_pdl.ocean

inf.common (create on: Nanigans) [编辑]

Falcon

Metrics Tags Expressions Objects Alarms History AlarmDash Dashboard 帐户管理

welcome, sunyn

模板基本信息

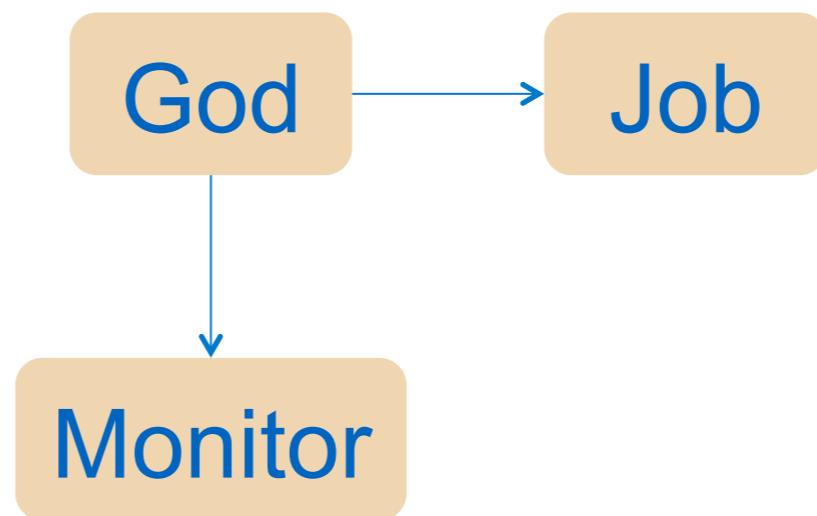
模板名称: inf.common
模板版本: v1.0 (从DCC模块自动创建)

模板中的策略列表

ID	策略基本信息 (metric/tags note)	规则条件	最大报警次数	报警级别	生效时间
5274	cpu.idle	all(%1)<5	3	1	
5275	df.bytes.free.percent/mount=/	all(%1)<10	3	2	
5276	df.bytes.free.percent/mount=/home	all(%1)<5	3	1	
5277	df.bytes.free.percent/mount=/home	all(%1)<10	3	2	
5278	df.inodes.free.percent/mount=/	all(%1)<10	3	2	
5279	df.inodes.free.percent/mount=/home	all(%1)<5	3	1	
5280	disk.io.await	all(%1)>2000	3	2	
5281	disk.io.rvwm	all(%1)>50	3	2	
5282	mem.memfree.percent	all(%1)<10	3	2	
5283	mem.memfree.percent	all(%1)<5	3	1	
5284	net.if.out.bytes	all(%1)>100000000	3	1	
5285	net.if.total.errors	all(%1)>20	3	2	
5286	pmem	all(%1)=0	3	2	
5287	sys.diskrun/mount=/home	all(%1)=0	3	1	
5288	sys.diskrun/mount=/	all(%1)=0	3	1	

全周期自动化 - 产品发布

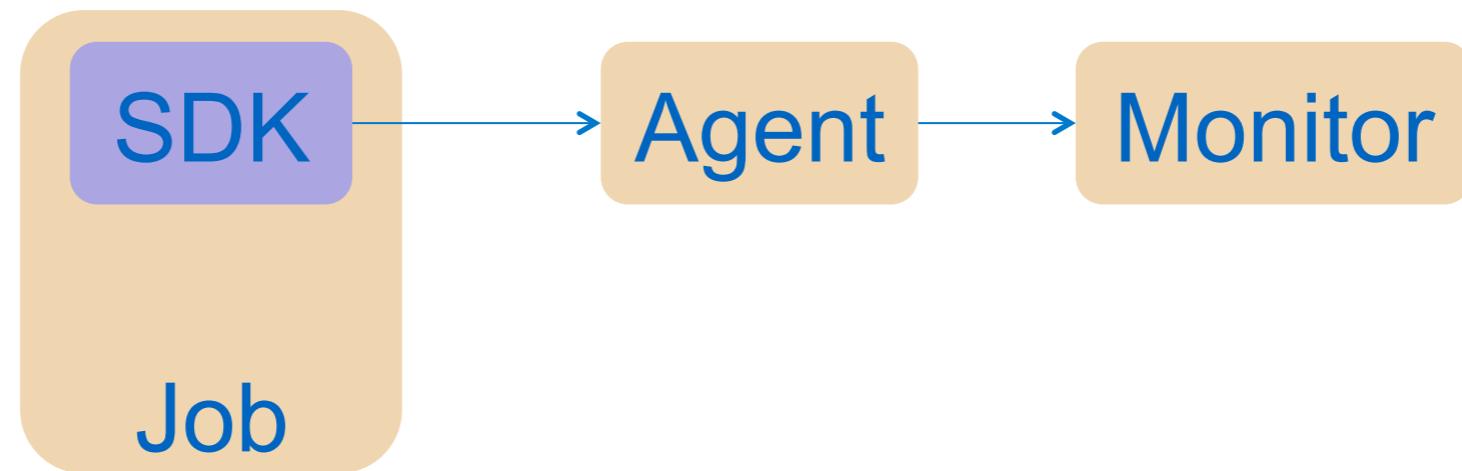
发布即监控



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

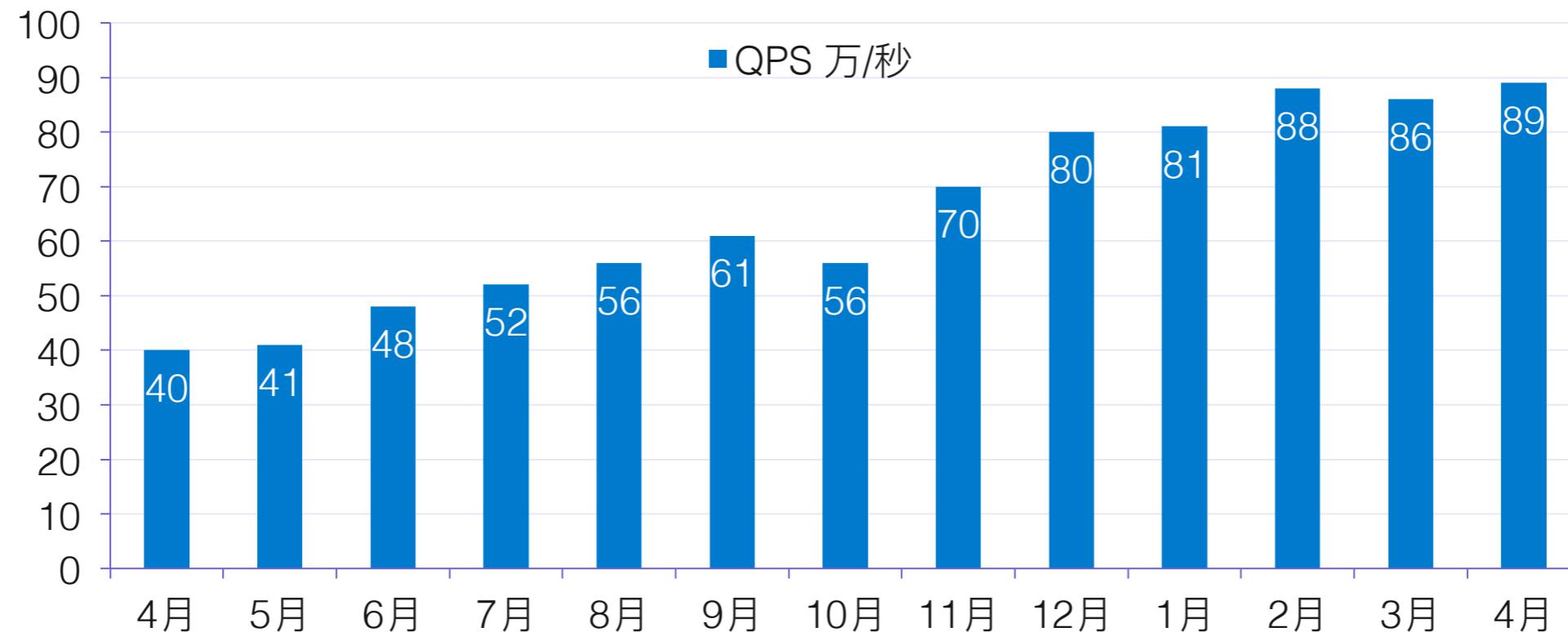


- 集群QPS总和
- 集群平均值
- 集群百分比

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: nodata.test1

endpoint选择: tag
cop.xiaomi_owt.inf_pdl.falcon

metric: nodata.test

tags:

type: GAUGE

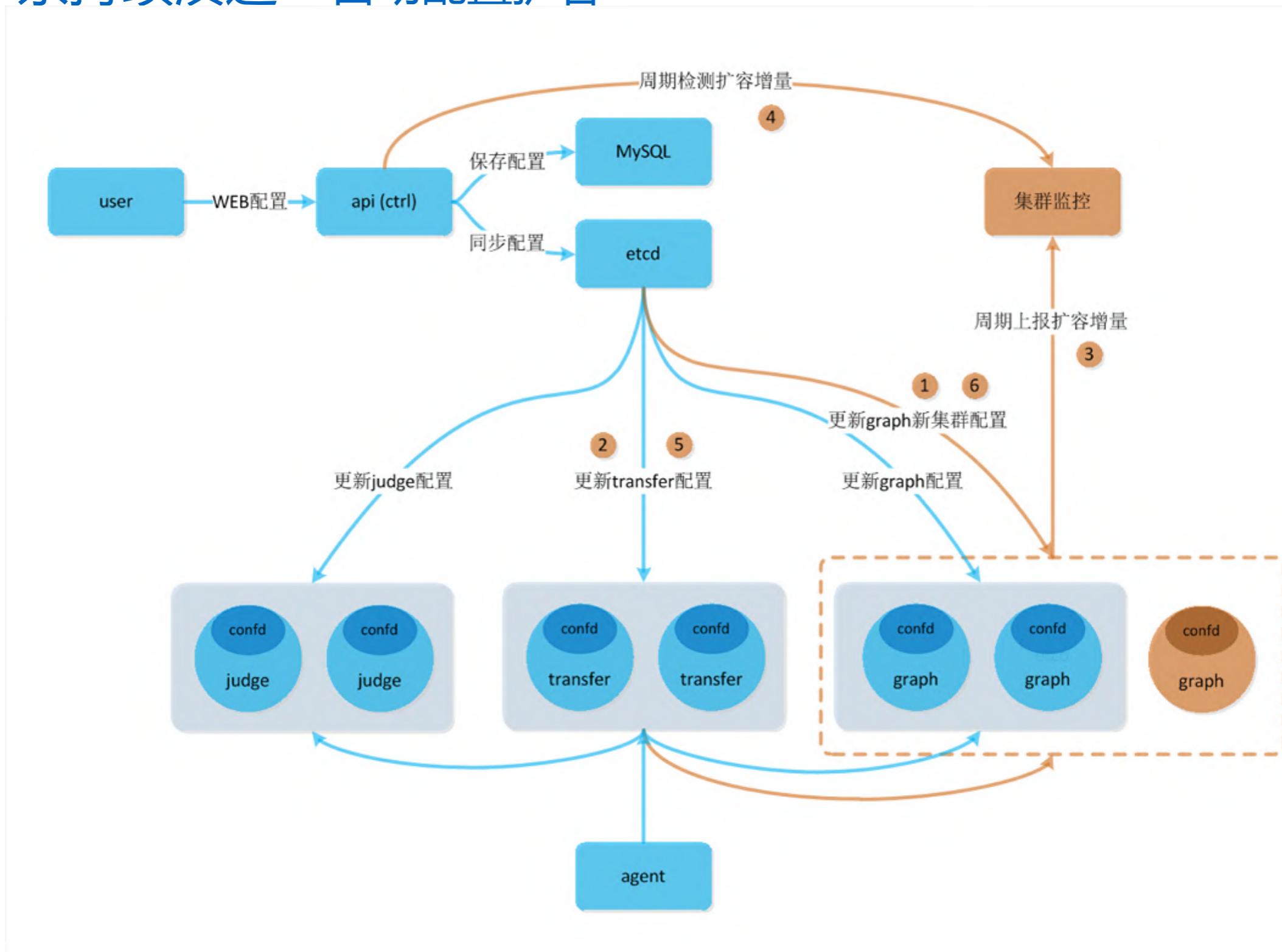
周期(秒): 30

■ 数据上报中断时, 补发如下值:
-1.0

Submit **Back**

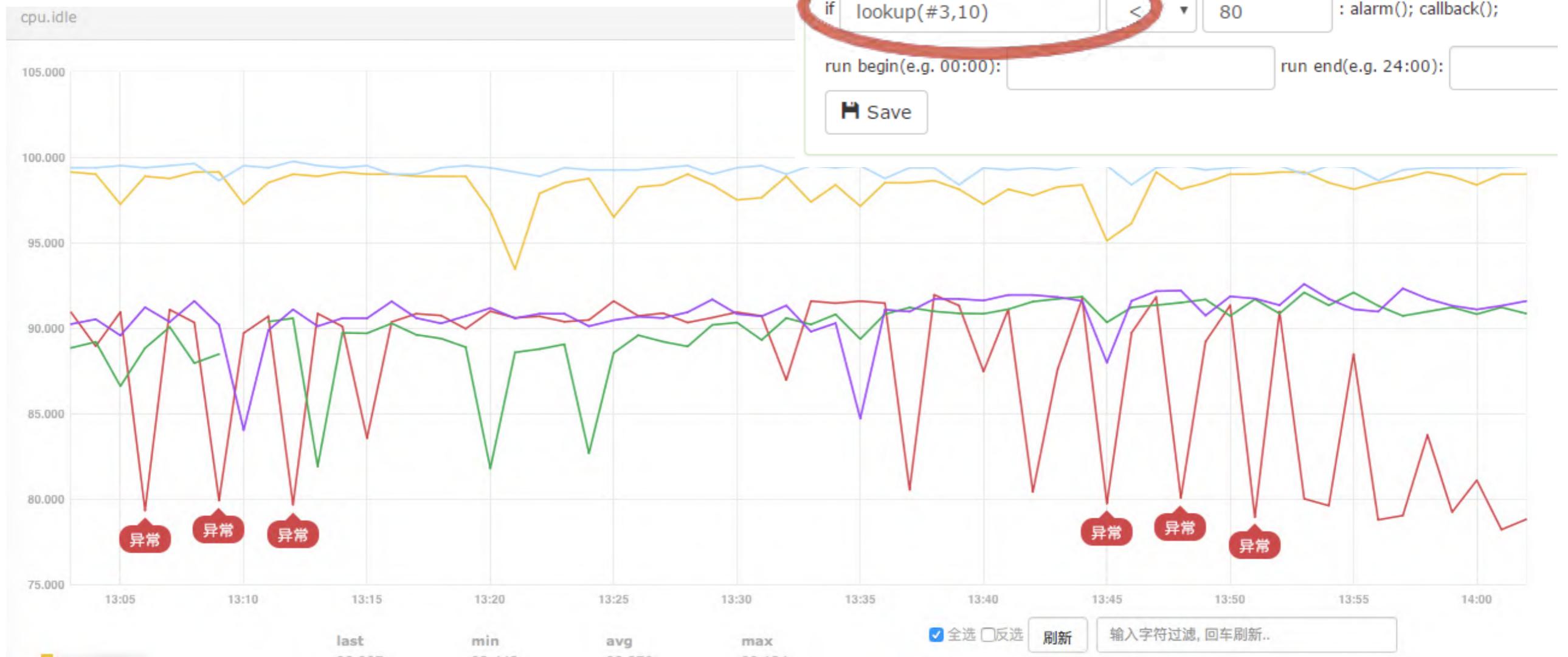
[Nodata监控用户手册](#)

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

The screenshot shows the 'Dynamic Scheduling Deployment System' interface. The main title is 'job.ocean-monitor-job_service.ocean-monitor-job_cluster.production-lg_pdl.ocean_owt.inf_cop.xiaomi'. The navigation bar includes search, refresh, and other system icons. The left sidebar has a 'Main Navigation' section with links: 仪表盘 (Dashboard), 宿主机管理 (Host Management), 产品线仪表盘 (Product Line Dashboard), 任务管理 (Task Management), 数据库管理 (Database Management), XRedis, 模板管理 (Template Management), and ELB.

The central content area is titled 'Automatic Scaling' (自动伸缩). It contains three main sections:

- Scaling Indicator (伸缩指标):** Set to 'CPU Average Usage' (CPU平均使用).
- Scaling Configuration (伸缩配置):** Includes:
 - Instance Upper Limit (实例数上限): 5
 - Instance Lower Limit (实例数下限): 1
 - Step Increase/Decrease (实例增减步长): 1
 - Notify (是否通知): Checked (checked)
 - Short Message Notification Group (UIC) (短信通知组(UIC)): sre_ocean
- Trigger Conditions (触发条件):** Includes:
 - Continuous Occurrences (连续次数):
 - Increment Instances (加实例): 3
 - Decrement Instances (减实例): 3
 - Comparison Operators (> and <):
 - Threshold Values (阈值): 80 and 60 (单位: 百分比)

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

网络管理

设备详情

名称	c1-06-a19-wan-asw-h5500-01.bj
IP地址	10.150.12.20
节点	cop.xiaomi_owt.sa_pdl.net
机房	c1
房间	06
机柜	a19
级别	asw
Zone	wan
设备类型	h5500
系统版本	Software Version 5.20 Release 5206
活跃路由条目	4

CPU 使用情况

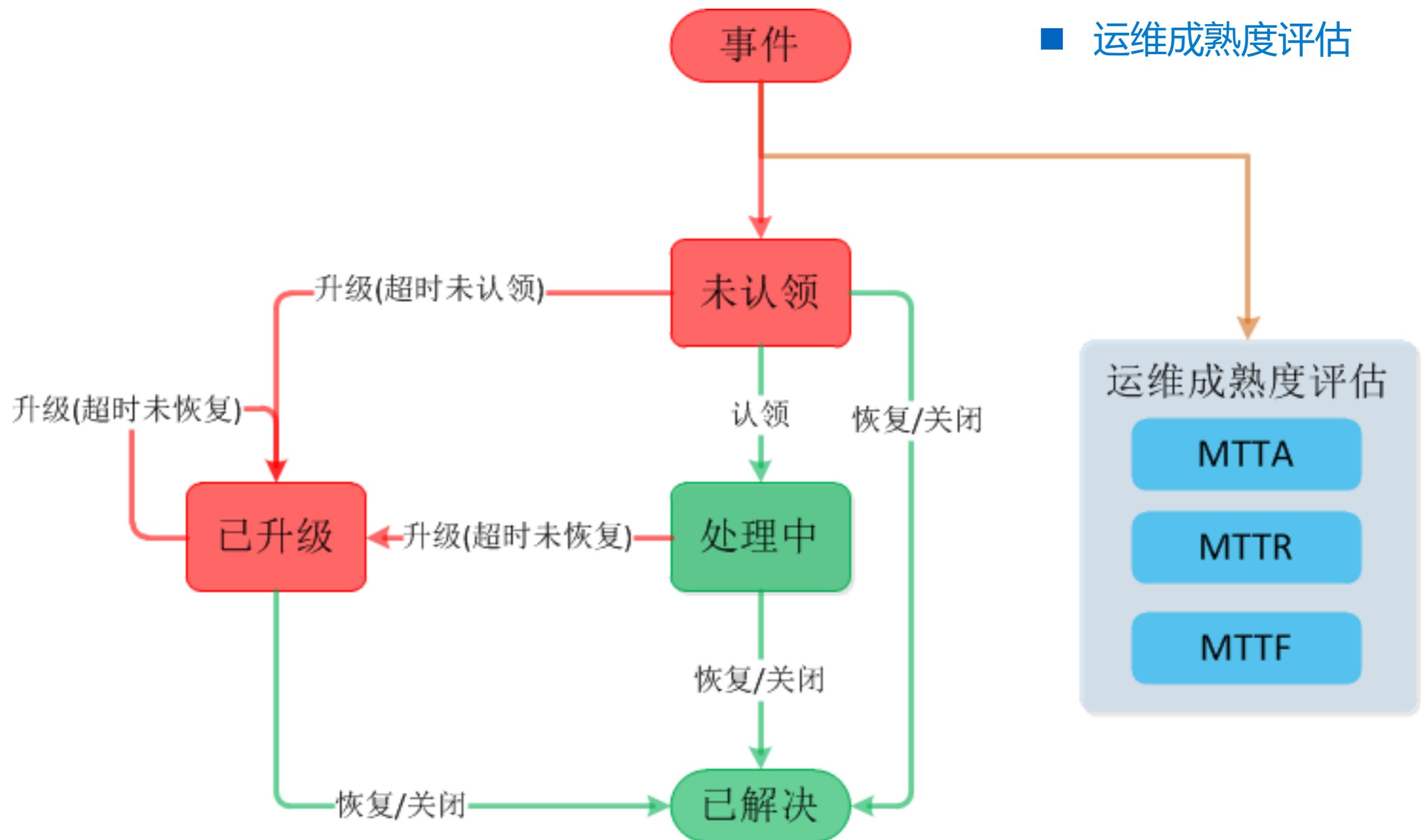


端口(c1-06-a19-wan-asw-h5500-01.bj)

GigabitEthernet1/0模块



体系持续演进 - 告警管理



体系持续演进 - 告警管理

Xbox >> 监控

搜索树节点

2017-06-03 ~ 2017-06-09

3.8min MTTA

9.3min MTTR

116个 总告警个数

按产品线分析

次数(次)

2017.06.3 2017.06.4 2017.06.5 2017.06.6 2017.06.7 2017.06.8 2017.06.9

告警次数

Highcharts.com

告警接收组ta 22270439 sgp2-op-mon-tran01.alisgp agent.alive all(#8) 2 PROBLEM 2017-05-26 02:24:00 +0800 CST

告警接收组fa

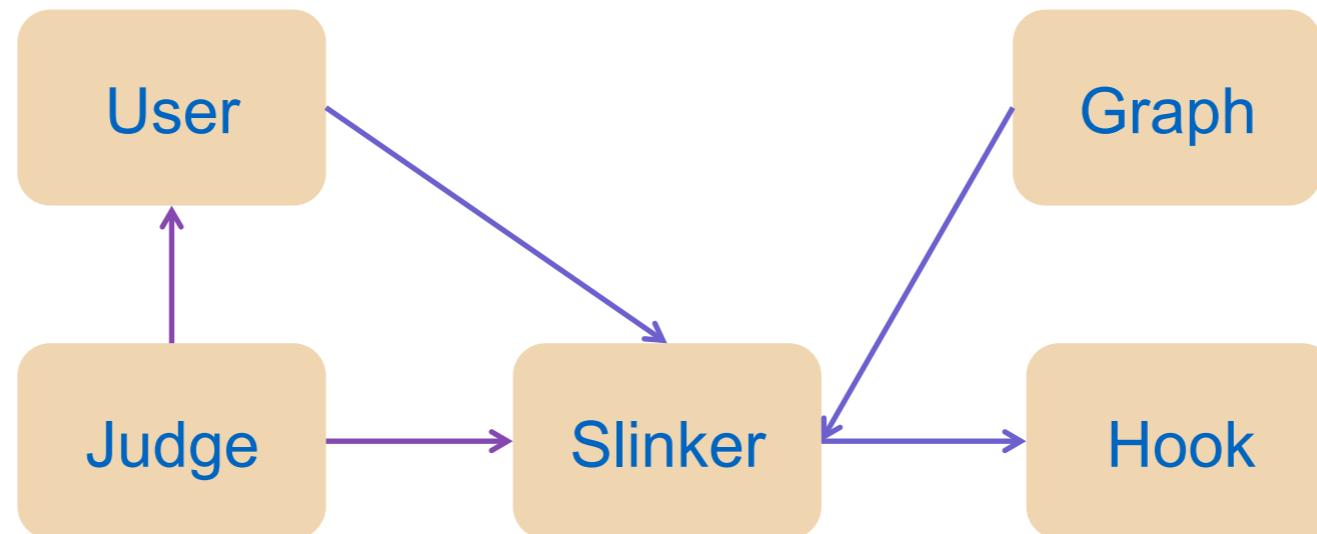
告警接收组fa

< 1 >

跳至 1 页

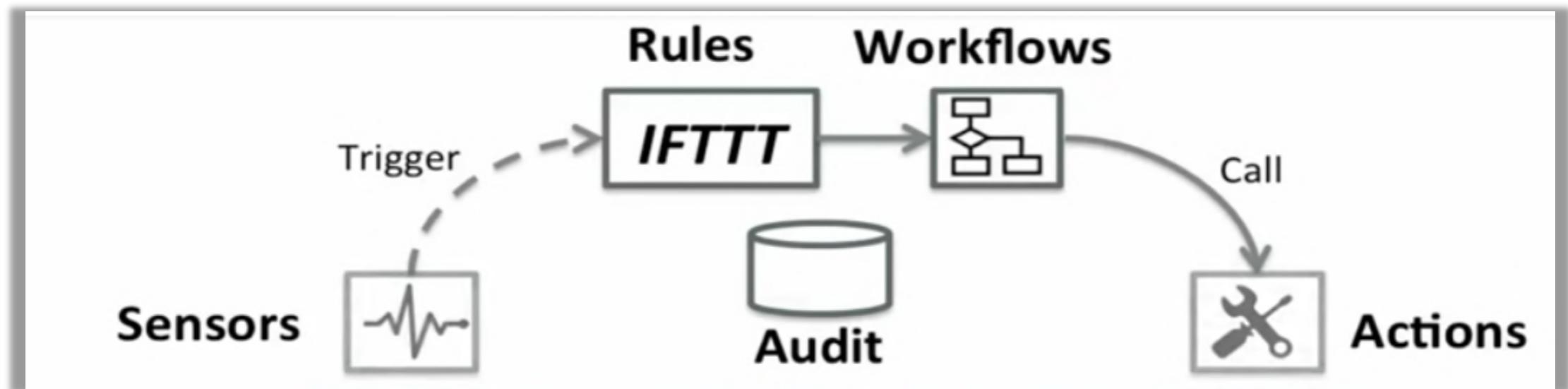
体系持续演进 - 智能监控

- 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



专注中高端技术人员的技术媒体



高端技术人员学习型社交平台



实践驱动的 IT 教育平台

