开放·连接·预见

2021 K+

全球软件研发行业创新峰会

可观测平台下的海量告警智能降噪设计实践

主讲人：丁来强

KEYLINK ing

创造关键链接 实现美好未来

——— 专业人力资本服务供应商 ———

20年
成立二十年专注于
人力资本服务

1000+
客户的共同选择

10万
覆盖精英学员

200+
专家智库

关注主办方公众号

目 录

Contents

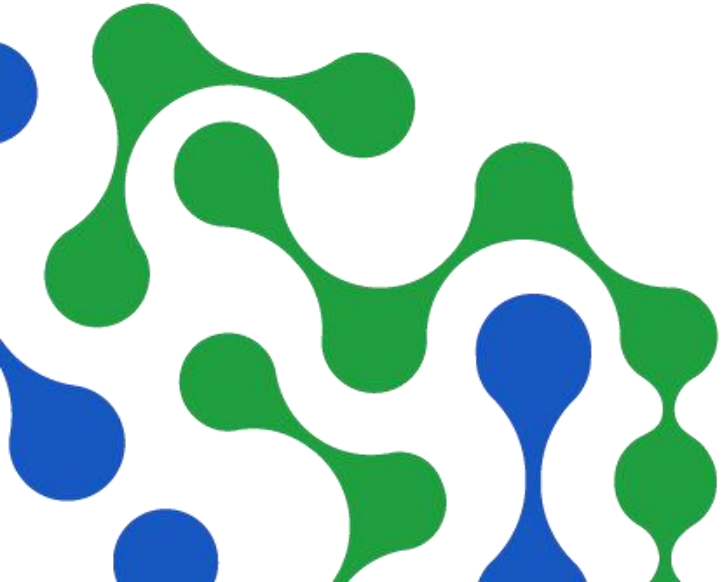
丁来强

阿里云SLS上海研发负责人

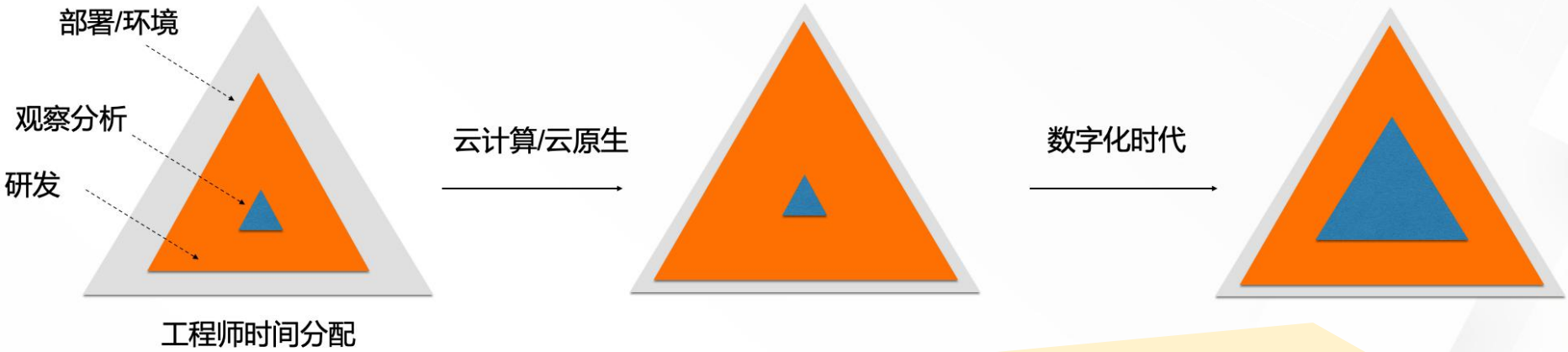
工作10+年，熟悉大数据分析、ITOps、SecOps等领域。2015年起，多次在PyCon、CSDN、TDI、GOPS、云栖等大会或活动上，分享超过30个大数据系统产品技术、架构设计模式等不同议题

01

背景挑战



数字化对工程师的挑战

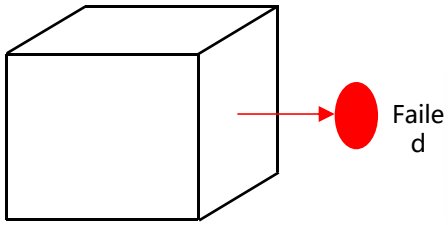


大量实时、碎片化、关键的数据分析工作

研发中	集成中	上线中	上线后	
• 研发 • Debug	• 研发 • 测试 • 联调	• 开发迭代 • 监控 • 运维	• 开发迭代 • 运维 • 运营 • 安全	• Scope变大：服务全球 • 关注面增多：系统，软件，用户... • 多能：承担测试、部分运维与运营工作 • 需要快速响应

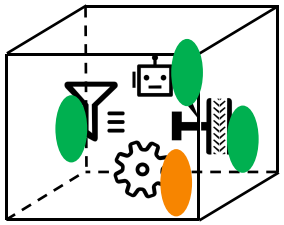
什么是可观测性

海恩法则(Heinrich's Law)指出： 每一起严重事故背后，必然有29次轻微事故和300起未遂先兆以及1000起事故隐患。



监控

- 聚焦在发现
- 确保系统稳定性
- Precision>Recall
- Ops为主



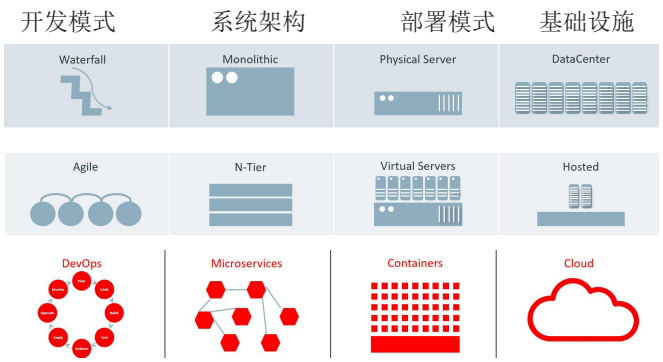
可观测性

- 目标白盒化，多种观测手段
- 确保找到根因，防患于未然
- 注重Recall + Precision
- 贯穿Dev/Tester/Ops等环节

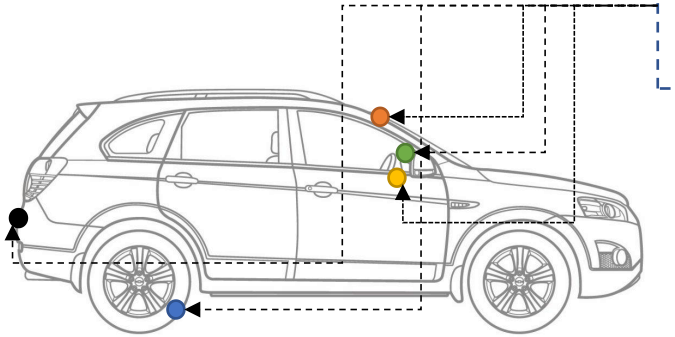
信息论角度

从一大堆低信息量数据中推导Fact过程

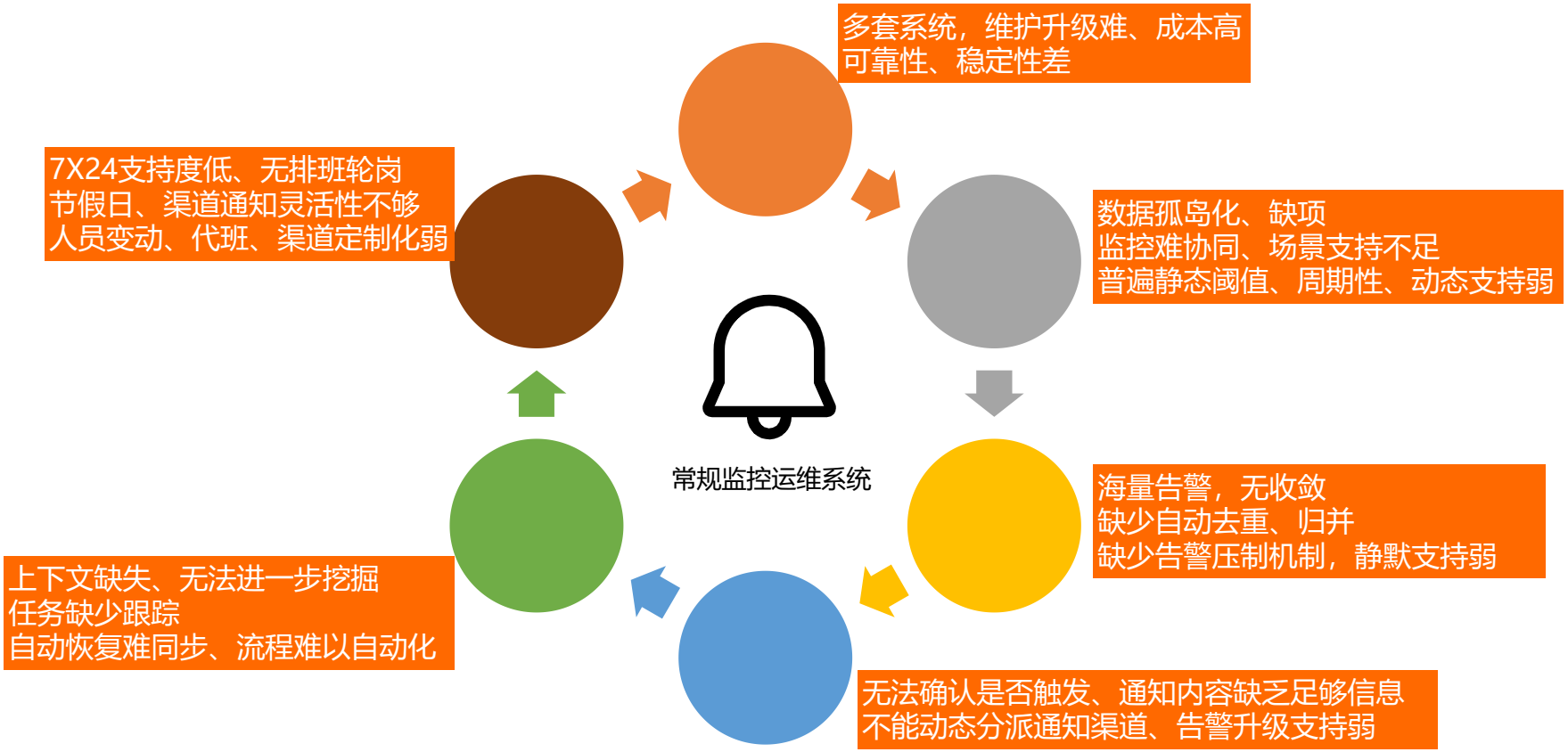
可观测性的挑战


开发模式
系统架构
部署模式
基础设施
Waterfall
Monolithic
Physical Server
DataCenter
Agile
N-Tier
Virtual Servers
Hosted
DevOps
Microservices
Containers
Cloud


influxdb
ZABBIX
Grafana loki
ZIPKIN
splunk>
Nagios
Skywalking
fluentd
sumologic
loggly
OPENTSDB
telegraf
dynatrace
New Relic

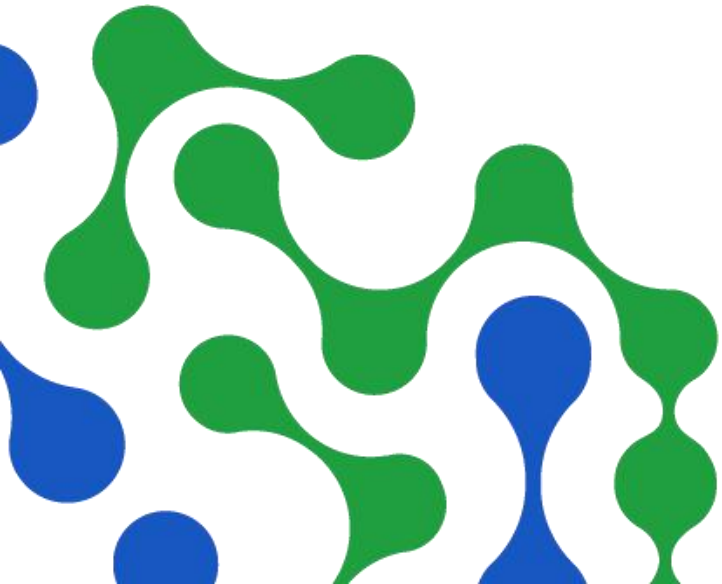


传统告警运维系统的痛点



02

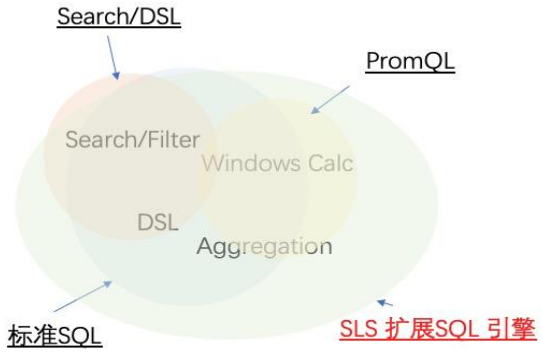
方案概述



典型可观测平台解决方案

数据统一接入与管理

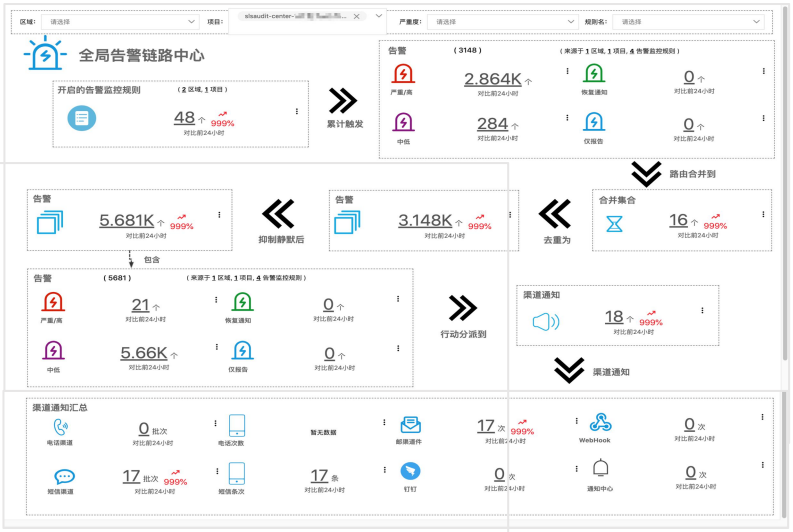

查询分析
日志
可视化
指标
其它场景
Trace
监控告警
AI


Search/DSL
PromQL
Search/Filter
Windows Calc
DSL
Aggregation
标准SQL
SLS 扩展SQL 引擎

典型可观测平台的智能告警运维系统

全局态势大盘

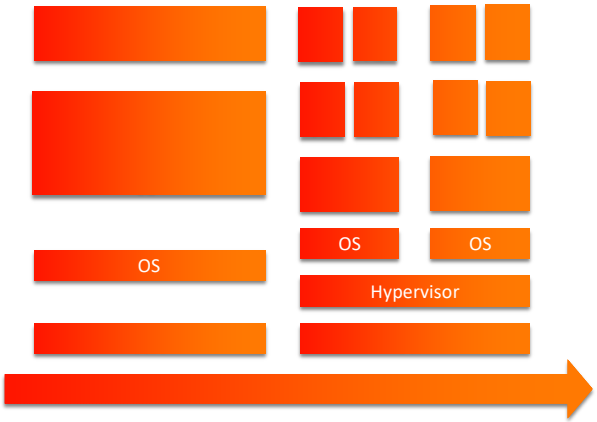

全局告警监控规则中心
开启的告警监控规则
299
999%
对比前24小时
状态分布
出错 134
触发 38
未触发 126
未知 1
告警监控规则最新评估状态
各合并集合中告警的最新状态（前1000条）

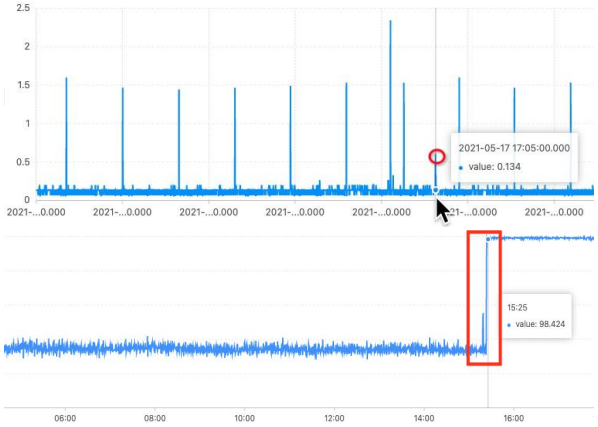

全局告警链路中心
开启的告警监控规则
48
999%
累计触发
告警
2.864K
对比前24小时
284
对比前24小时
恢复通知
0
仅报告
路由合并到
合并集合
16
999%
5.681K
抑制静默后
3.148K
去重为
包含
21
5.66K
行动分派到
渠道通知
18
渠道通知汇总
电话渠道
0 批次
短信渠道
17 批次
电话次数
短信条次
暂无数据
邮箱邮件
17 次
钉钉
WebHook
通知中心

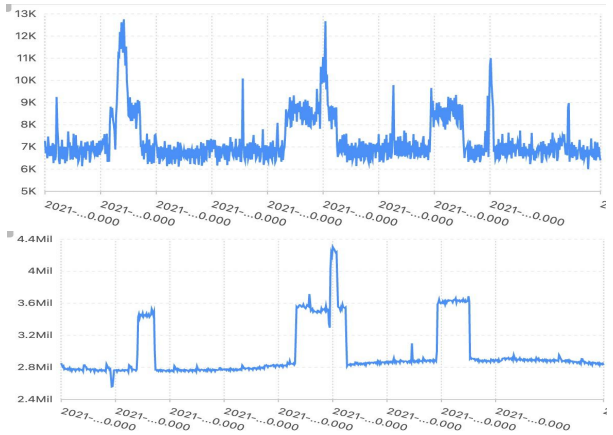
03

智能监控

传统监控的困难和挑战







监控对象爆炸式增长

服务部署从以“主机”为中心向“容器化”方向转化，容器本身轻量化和短生命周期等特点，导致监控对象和指标急剧增加。

监控规则无法自适应

基于人工定义的规则的监控会有漏报、误报、规则阈值无法自适应等常见问题。

监控规则泛化能力弱

不同的业务、甚至同一业务的不同版本，指标的规律性、阈值都有可能是不同的。

智能监控1：智能巡检

智能前置

从告警触发后的智能管理，到告警触发前的智能监控，提升智能能力，挖掘数据潜在价值

优势一

监控自适应

基于历史数据自动学习，可进行动态阈值判断，告警更精准；实时检测，异常发现更及时

优势二

动态反馈

根据用户反馈，基于人的经验进一步修正和完善模型，减少误报

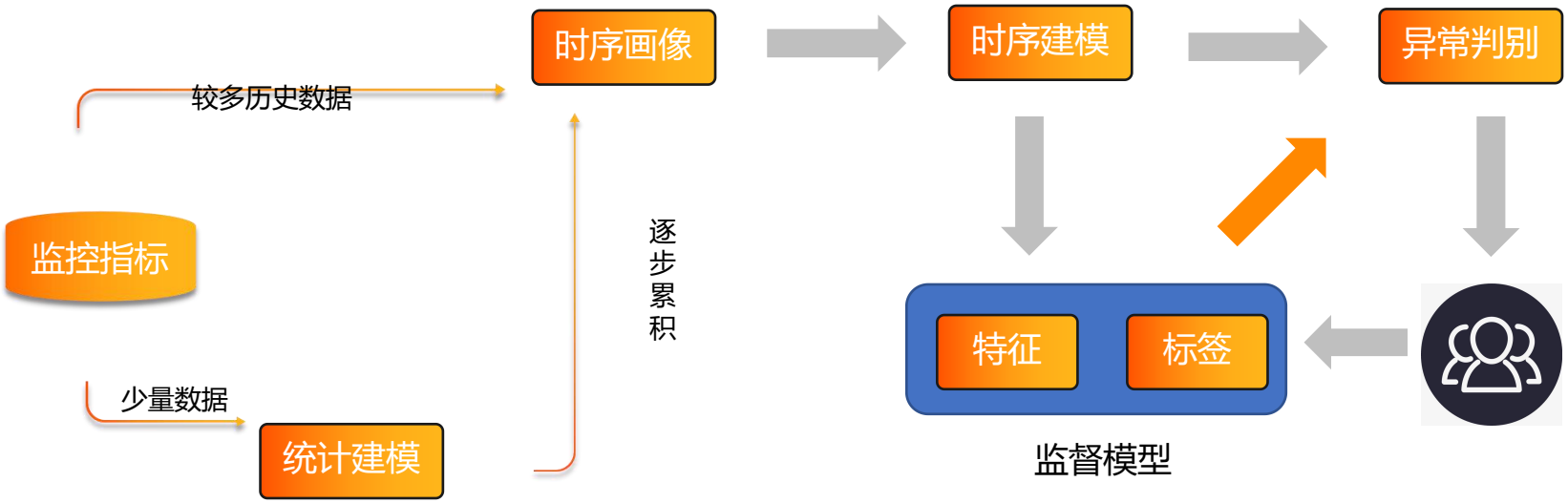
优势三

智能巡检：场景与案例

- 场景：某些指标没有固定阈值，需要根据变化趋势判断数据异常
- 案例：某个服务新版本线上灰度过程中，由于新版本bug导致网络流量异常抖动

智能巡检：实现思路

算法原理：采用无监督学习算法，自动识别实体的数据特征，根据数据特征选择不同的算法组合，针对数据流实时建模，完成异常检测任务。并根据用户的打标信息，训练监督模型，实现算法的不断优化，提高准确率。



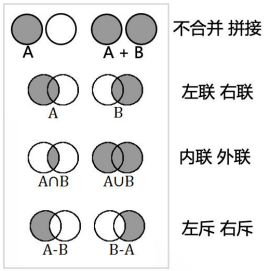
智能巡检：算法说明

	流式图算法	流式分解算法
适用场景	适用于一般性时间序列的异常检测场景，包括： • 机器级别的监控指标的异常巡检，例如CPU占用率、内存利用率、硬盘读写速率等。 • 业务指标的异常巡检，例如QPS、流量、成功率、延时等。 • 黄金指标的异常巡检。	适用于对具有周期性的数据序列进行巡检，且要求数据的周期性较为明显。例如游戏的访问量、客户的订单量。
相关论文	Time-Series Event Prediction with Evolutionary State Graph	RobustSTL: A Robust Seasonal-Trend Decomposition Algorithm for Long Time Series

智能监控2：基于增强规则引擎监控

多源协同

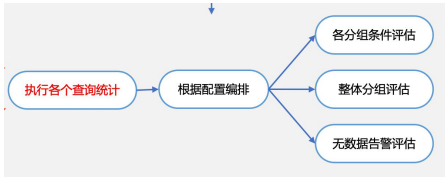
直接使用SQL、同比环比、多数据源全局协同关联



充分发挥SQL能力

数据感知

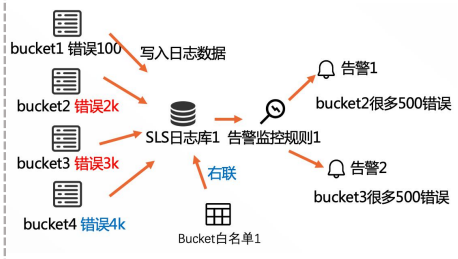
关键字、无数据、数据恢复场景、内置模板库等



经典模式

简单灵活

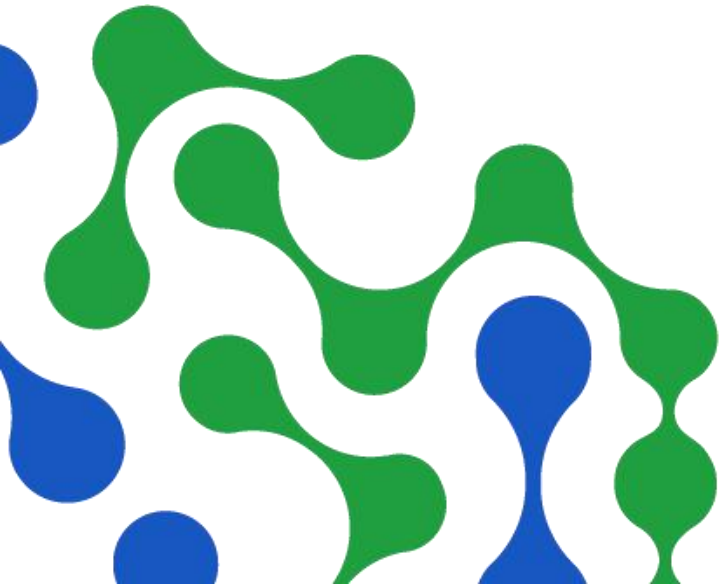
多目标监控、动态条件判断、黑白名单等



经典策略

04

智能管理



告警智能管理

智能降噪

告警自动去重，相关告警合并发送，抑制和静默机制。

案例：代码扫描&依赖检查，同一个Commit的多个问题合并发送通知。

不啰嗦，减少打扰

动态分派

基于告警属性、工作时间等信息，使用不同的通知渠道通知到不同的人。

案例：不同模块的CICD或线上告警，通知给相应的负责人。

把问题分派给合适的人

值班管理

基于日历的值班管理，法定节假日和国际化时区自动感知。

案例：团队成员轮流负责CICD流水线或线上问题。

三个和尚没水喝

告警智能降噪

自动去重

根据告警指纹与自动去重

案例：某主机每一分钟触发CPU使用率过高告警，1小时触发60次；通知设置为30分钟重复，则一共发送两次通知

智能合并

相关告警进行合并，一并发送通知

案例：根据告警所在集群进行合并；假如某集群短时间内产生了10个告警，则只会发送一条通知，包含这10个事件

关联抑制

告警之间的相互影响

案例：某一k8s集群发生OOM严重告警，可以暂时忽略同一集群的低级别告警

灵活静默

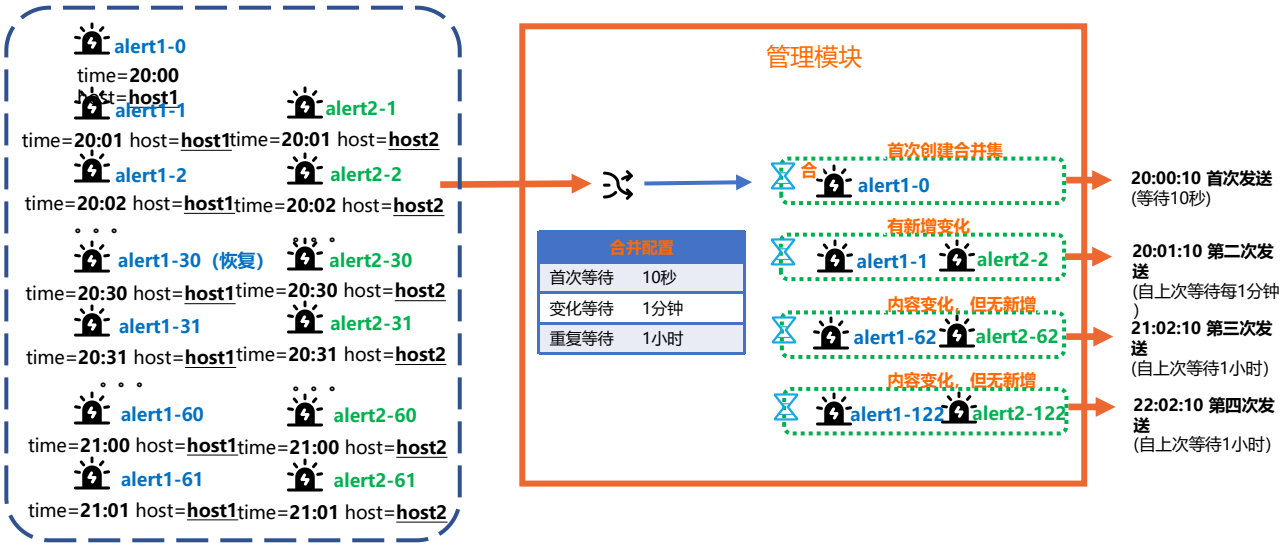
特定告警无需通知

案例：测试集群在凌晨有计划内变更，期间服务会有短暂不可用，触发预期内告警，该告警可以忽略

智能降噪：智能合并

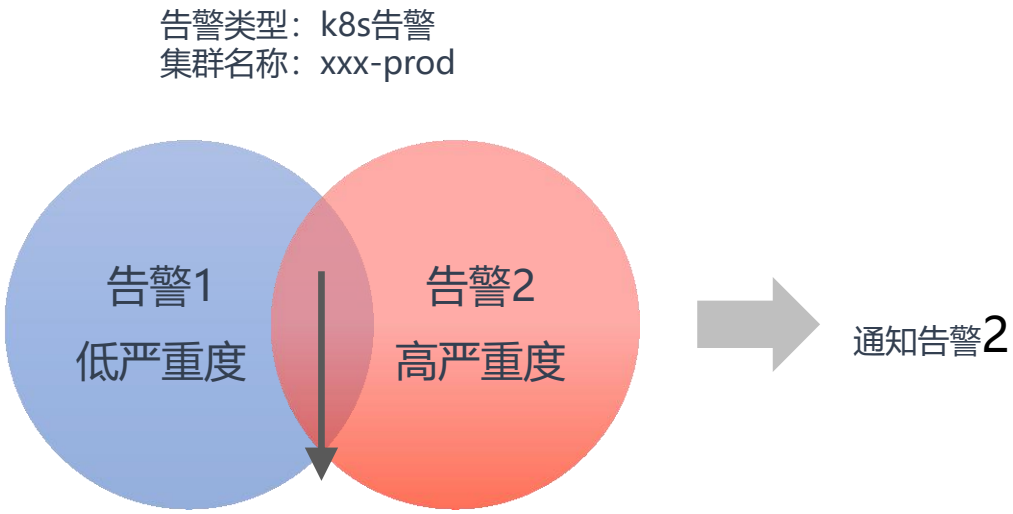
- 根据告警消息属性路由到bucket中，智能去重（基于文本精准或相似算法）并合并通知，避免告警风暴的困扰。在合并的告警首次出现以及变化时发送，在不变时延迟发送。

案例：某服务的2个主机从20:00和20:01分别每分钟不停触发高CPU告警，收到2小时125条告警，经过降噪仅发送4次。



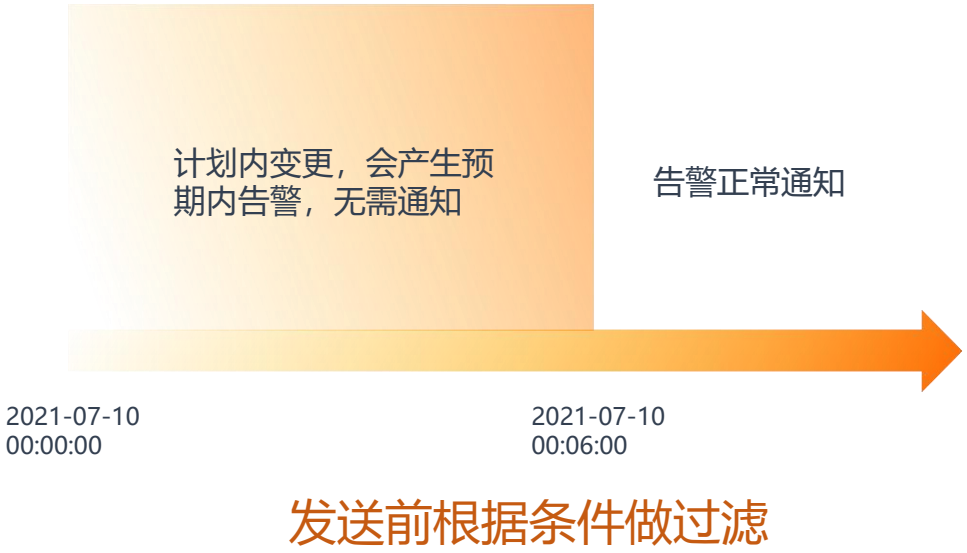
智能降噪：关联抑制

- 故障之间往往是相互影响和依赖的，例如：
- 案例1：主机宕机，造成运行在该主机上的服务访问异常，同一主机上的其余告警可以忽略
- 案例2：某一k8s集群发生OOM严重告警时，可以暂时忽略同一集群的低级别告警



智能降噪：灵活静默

- 符合条件的告警无需通知，例如：
- 案例1: 预期内的告警：测试集群在凌晨有计划内变更，期间服务会有短暂不可用，触发告警
- 案例2: 所有没有按照规范设置 owner 的告警



动态分派

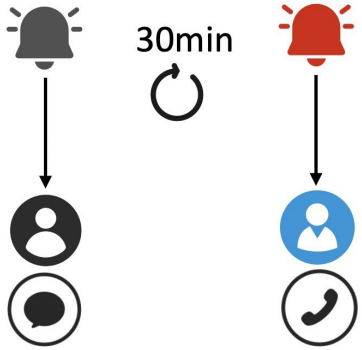
多渠道

支持多种通知渠道及扩展

动态通知

根据告警属性动态分派通知

通知升级



长时间未解决告警通知升级

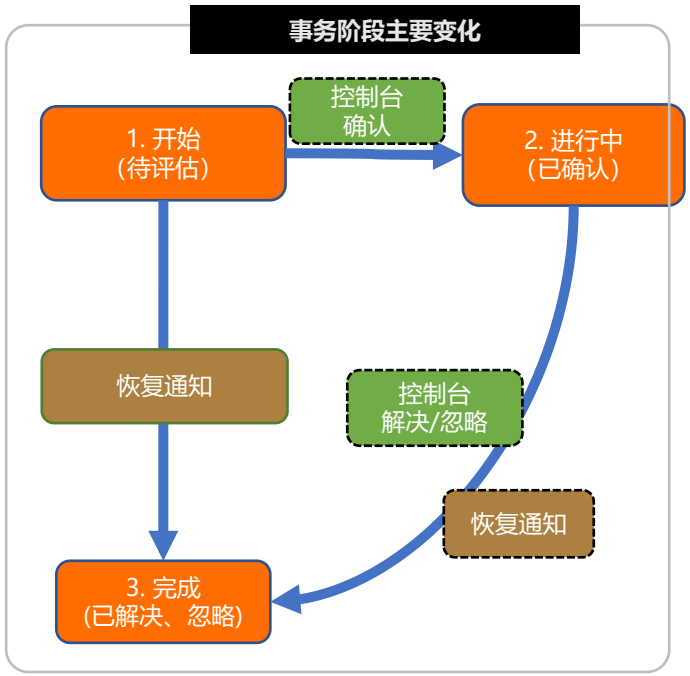
通知升级：场景与机制

通过多套告警通知策略，在事务一直处于特定未完成状态时触发通知升级

案例1：告警常规通过钉钉、邮件方式发送，15分钟内一直未确认，升级为电话和语音。

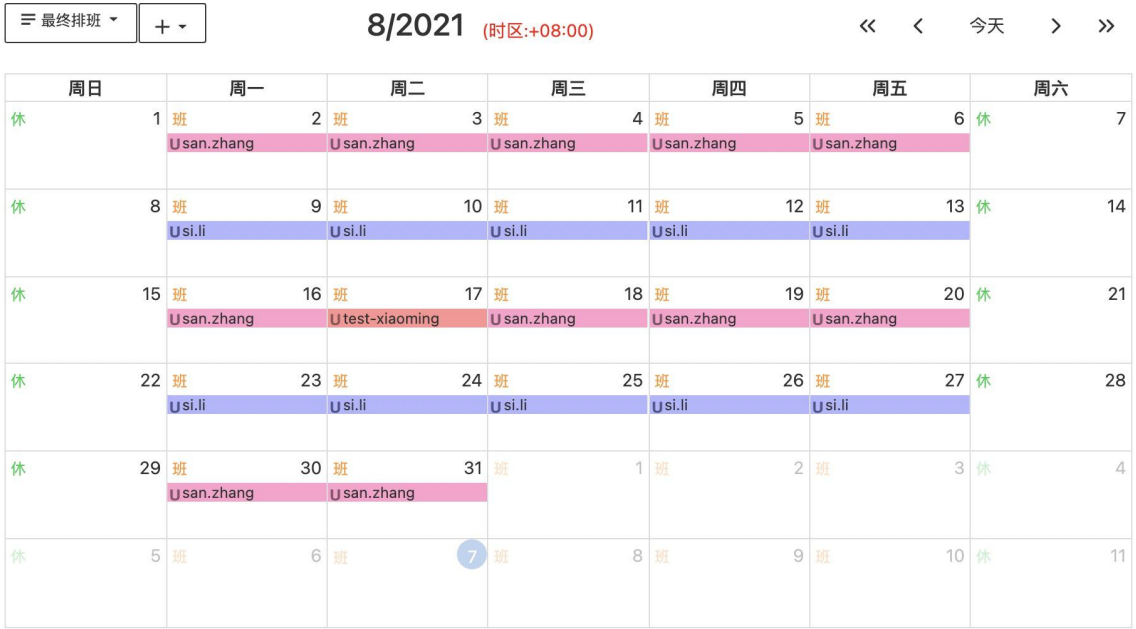
案例2：运维组收到的告警在4小时内一直未解决，则会通知组长。

告警消息进行多状态管理，通过状态机实现告警消息处理工作流



值班和代班

- 案例：2021年8月由张三和李四值班（每班一周，仅工作日值班），首个工作日交班；8月17日张三请假，由小明代值班



THANKS

2021 K+

全球软件研发行业创新峰会