睿象云 | DataFun.

可观测系统中的告警管理实践

王金良 北京睿象科技有限公司 技术总监

目录

CONTENT

01 可观测性的整体生态

02 告警在可观测系统中的价值

03 可观测系统中告警管理的核心功能

04 智能告警管理最佳实践

01 可观测性的整体生态

运维可观测性整体生态

近两年，可观测性红遍IT运维领域，火起来的导火索是CNCF(云原生计算基金会)在云原生定义中提到 Observerbility，并声称这是云原生时代的必备能力。加之包括谷歌在内的众多大厂一拥而上，“可观测性”正式出道。谷歌给出可观测性的核心价值很简单：快速排障（troubleshooting）

对于业务系统，尤其是云原生时代的分布式、微服务化、容器化的复杂应用，随着系统业务量日益庞大、内部结构日益复杂、组件间交互日益频繁，传统的监控管理手段已经不足以满足新时代的需求，可观察性就自然而然地被引入IT领域，变成与性能、可用性、可靠性、可扩展性一样的关键维度。

有业界专家一句话总结传统监控与可观测性的区别：“监控告诉我们系统的哪些部分是工作的；可观测性告诉我们那里为什么不工作了。”

运维可观测性整体生态

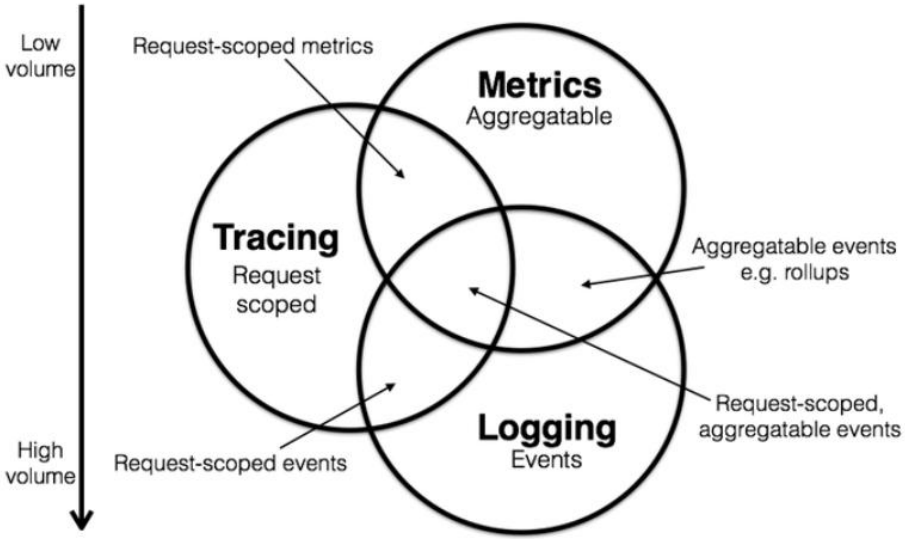
业务可观测性其实就是对一个系统内部状态的测量、观察的能力；在有一些领域也叫可维、可测、可控能力。谈起可观测性的概念，必须要从“三大支柱”这个名词讲起。

2017年，一篇博文总结了可观测性的三大支柱：指标（Metrics）、追踪（Tracing）、日志（Logging），文中将可观测性问题映射到了如何处理指标（metrics）、追踪（tracing）、日志（logging）三类数据上，由此形成了流传很广的业务可观测性三大支柱理论。那么业务可观测性就可以具体化为：如何定义、获取、分析这三个层面的数据。实现对业务系统的运行状态、异常状况、服务质量的可观测、可发现、可管理的能力。三大支柱理论出现后的几年间里，这个观点受到了业内的广泛认可，发展为对可观测性能力的基本要求，并且每一个方面都有了众多成熟的解决方案。例如各类开源工具：

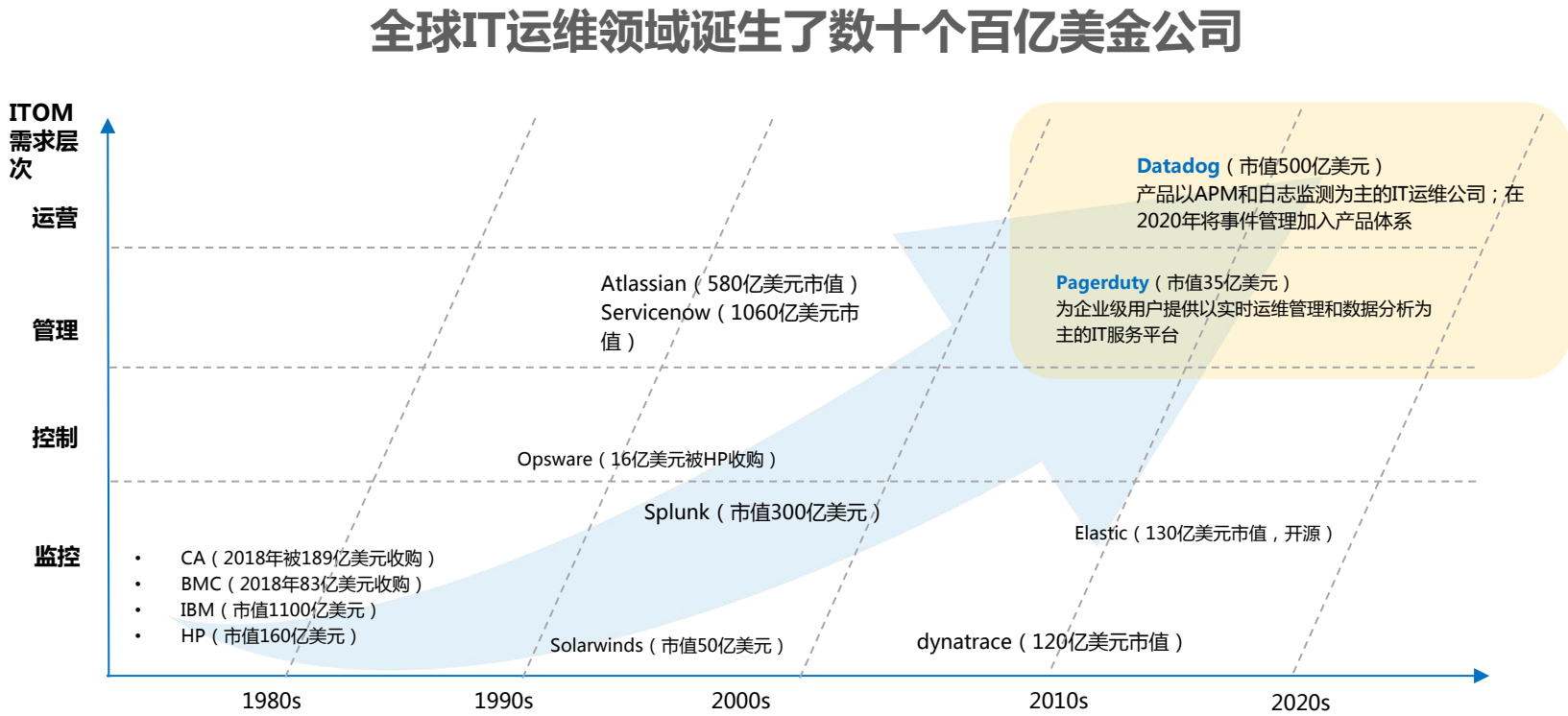
聚焦于Metrics的：Prometheus、zabbix、Grafana等；

聚焦于Tracing的: Skywalking、OpenTracing等；

聚焦于Logging的: ELK、Graylog等。



运维可观测性整体生态



运维可观测性整体生态

国内市场 国内IT 运维的潜在发展空间超过千亿元，其中，金融、制造、能源等领域的大中型客户的年客单价达到几百万元。例如前瞻产业研究院数据显示，2012-2019年，中国IT运维市场规模呈现波动趋势。从增速来看，2014年达到近年来最高增速17.34%，达到了1121.2亿元的市场规模。2019年，中国IT运维市场规模达到2324.3亿元，同比增长15.73%，2020年我国中国IT运维市场规模达到2690亿元左右。

艾瑞咨询数据显示，2021年中国IT服务将突破万亿大关，其中，IT运维市场规模在2021年达到2941.2亿元， 预计2023 年达到3236.4 亿元，2020-2023 年的年复合增长率为 11.7%。

引入业内大佬的一句表述：到目前为止，IT运维赛道已经逐渐的层级化，强者逾强，因为软件领域一旦产品化程度高了，技术壁垒随之建立起来，逐渐的收敛市场。在这样的形势之下，完全初创没有积累的企业再进入，发展难度会越来越大。

02 告警在可观测系统中的价值

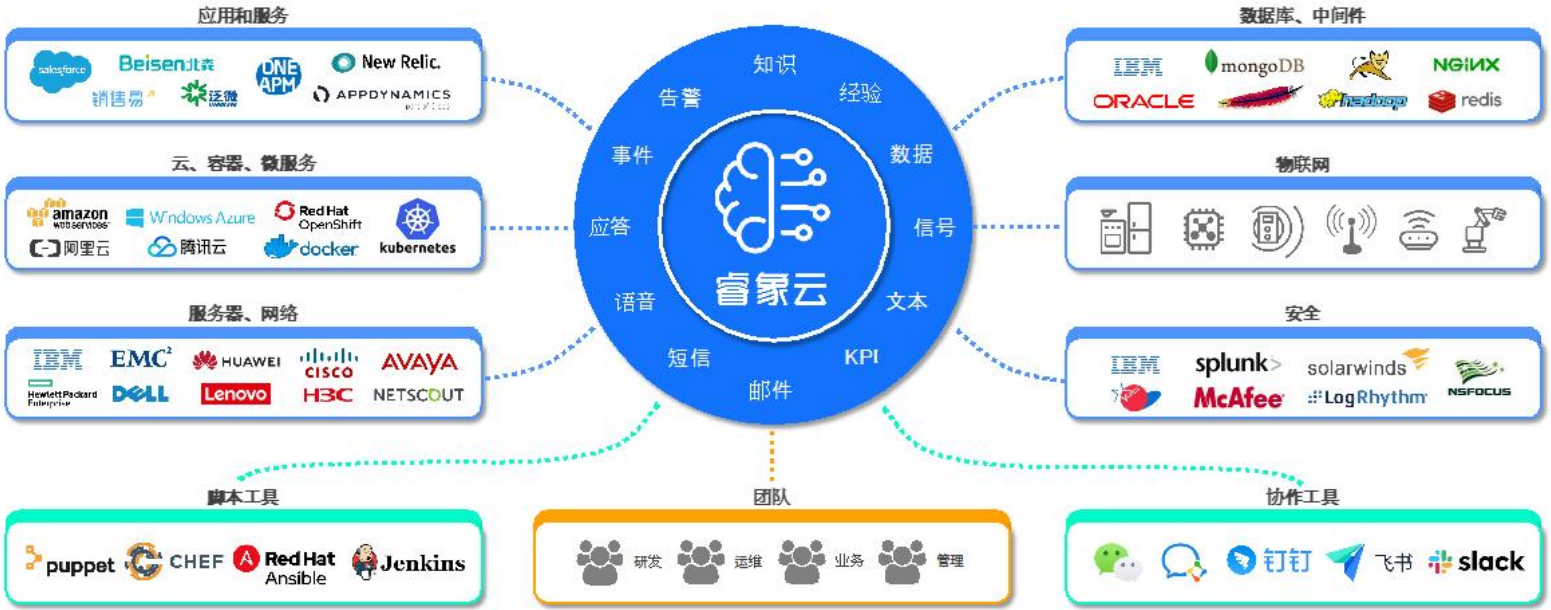
运维可观测性整体生态

可观测性并不是在取代监控，它也不是一种我们通常理解的某一个监控或运维工具的形态。相比较而言，可观测性更像是描述一种属性的范畴，很多时候是种能力的体现形式，越复杂的系统越需要这种属性或能力。可观测性也并非万能的，它可以引导开发人员找到准确的答案，但也只是停留在引导层面，不能不能保证让他们100%找到答案。这个过程当中依旧需要当事人对系统、网络等有着良好的理解甚至直觉，才能让定位问题变得轻松并高效。

告警作为IT运维信号总量的金字塔尖，对于可观测性接下来的流转及判断起着至关重要作用。而统一的、准确的、智能化的告警平台的诞生，为可观测性平台的分析，提供了更为可靠的“现象”，可帮助运维人员接下来分析原因引导方向。

图片引自dockone.io

智能告警管理平台—企业数字化运营的中枢神经



告警管理成熟度模型

告警管理分级	名称	定义	特点
L0	有监控无告警	通过监控工具或日志对IT系统的运行状态进行监控，但未配置告警或缺少告警模块，无法实时获取系统故障信息。	被动感知系统故障，运维效率低下。
L1	告警分散管理	通过不同监控工具的告警模块分别配置告警策略和通知机制，告警管理分散在各个监控工具之中。	任务分派和通知手段单一，管理效率低下。
L2	告警统一管理	将不同监控工具或系统产生的告警接入统一的管理平台，实现告警的统一分派和通知，并能基于规则对告警进行去重和压缩。	丰富分派和通知手段，降低告警处理量。
L3	告警智能管理	通过运用人工智能算法，无需人工参与的情况下，自动识别告警类别和新增类型，对复杂场景下的相似告警进行更高比例的压缩降噪。	极大减少告警处理量，提升告警故障分析效率。
L4	根因告警定位	通过运用知识图谱技术和告警专业领域知识，能够自动推荐各个业务场景下海量告警信息中的根因告警。	自动定位系统故障根因，减低系统运维难度。
L5	告警自愈	针对根因告警，通过结合告警故障知识库和运维自动化工具，对系统故障进行自动恢复。并通过不断地知识沉淀，提升自愈能力。	沉淀运维知识，实现真正的无人值守。

03 可观测系统中告警管理的核心功能

睿象云智能告警平台Cloud Alert简介

睿象云智能告警平台是一套支持 SaaS 和本地化部署的告警管理工具类产品，能够收集企业内部的事件类数据（机器日志、告警等），IT配置信息（业务调用关系、CMDB 等），和知识数据（故障手册、厂家文档、告警处理意见等）等三种 IT 运维数据。通过事件驱动发现异常事件，自动分析事件根因，对未来可能发生的威胁及时预警，并结合解决方案智能推荐形成企业内部智能运维体系闭环。

智能告警平台核心能力：

- 告警整合
- 告警加工
- 告警模式发现
- On-call 管理
- 告警管理
- 告警分析

睿象云 | DataFun.

CA智能告警核心能力：告警整合管理

- 企业数字化运营中枢神经
- 连接 10 大类近 100 种 DevOps 主流工具和平台
- 覆盖开发和运维人员日常工作环境
- 分钟级完成系统和人员的连接

基础设施监控

ZABBIX　Prometheus　Nagios

夜莺　Open-Falcon　elastic

云监控

amazon web services　阿里云　腾讯云

HUAWEI 华为云　观测云 GUANCE.COM　QINGCLOUD

Microsoft Azure　UCLOUD

金山云　百度智能云　京东云

应用和服务监控

ONE APM　dynatrace　DATADOG

New Relic.　监控宝

网络监控

Cacti　solarwinds　Net-SNMP

云监控

vmware　Grafana　Site24x7

协作平台

slack　钉钉　企业微信

飞书　UDESK　简聊

amazon Chime　倍洽　云之家 Cloud-Hub

通知工具

电话　短信　邮件

微信　APP　APP

项目和流程管理

JIRA　WISESIGN 慧铭软件　嘉为蓝鲸

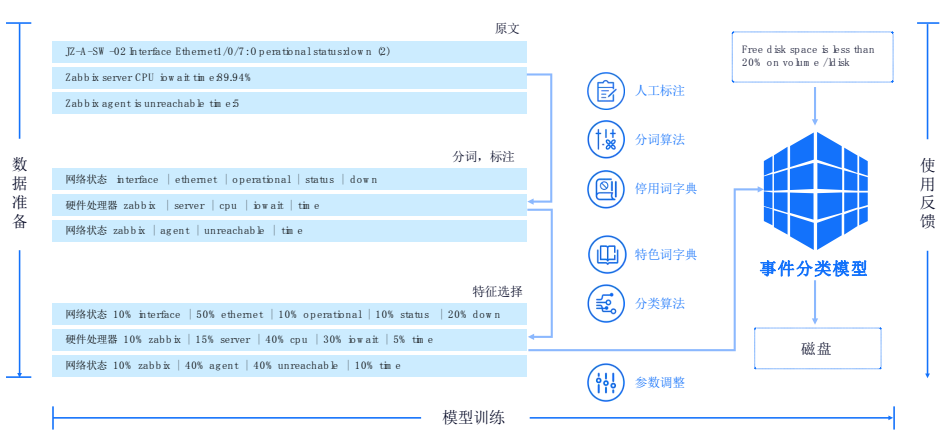
会议协作

腾讯会议　钉钉

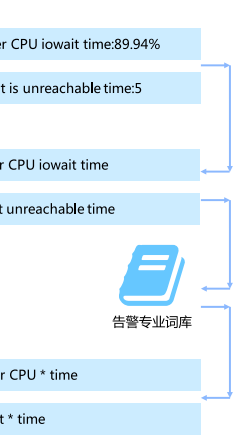
其他

Cloud Alert REST API　Cloud Alert Email　Webhook

CA智能告警核心能力：告警智能化处理



- 1万家 企业、 2亿条 原始数据
- 230万条 人工标注
- 20种 人工智能算法
- 3GB 专业特色词库
- 2年 模型在线训练和迭代
- 95%+ 告警降噪比



- **告警处理** 通过数据格式化，自定义字段提取和内容丰富，为基于算法的模式发现提供数据准备。

  — 告警数据格式化：根据智能事件平台数据规范和事件源的格式对应关系，以键值对的形式分拆原始事件。
  — 自定义数据标签：依据事件特征，赋予事件自定义属性值；或从事件主体中自定义提取字段。
  — 告警数据丰富：通过映射丰富技术在数据映射表中查找关联关系来为原始事件增加新的数据字段和值。

- **模式发现** 基于规则和人工智能算法，对符合特征的告警进行分类、聚合、降噪，自动监测和发现异常情况，降低超过 95% 的 IT 噪音。

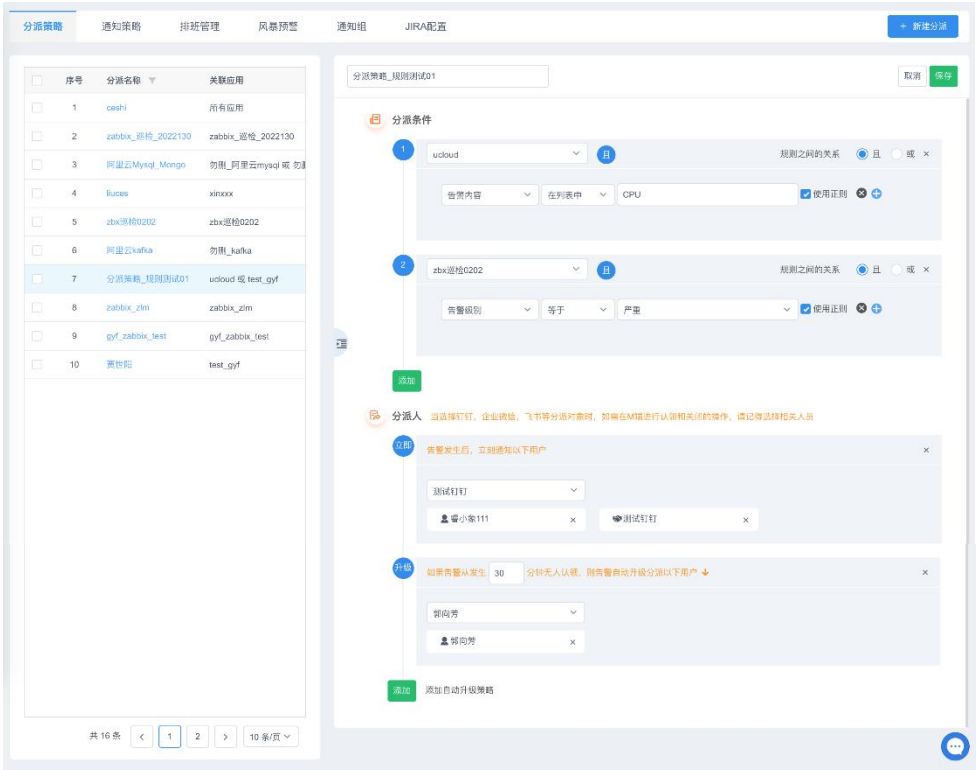
  — 基于规则的模式发现：通过正则表达式，用户自定义告警分类和聚合规则，对特征事件进行归类和压缩。
  — 基于人工智能算法的模式发现：将无监督与有监督算法相结合，自动对告警进行识别分类，并对相似和相关事件进行聚合和压缩。
  — 事件异常检测：通过算法实时检测告警发生频次，基于信息熵，自动发现突发事件和异常事件。

CA智能告警核心能力：分派和协作

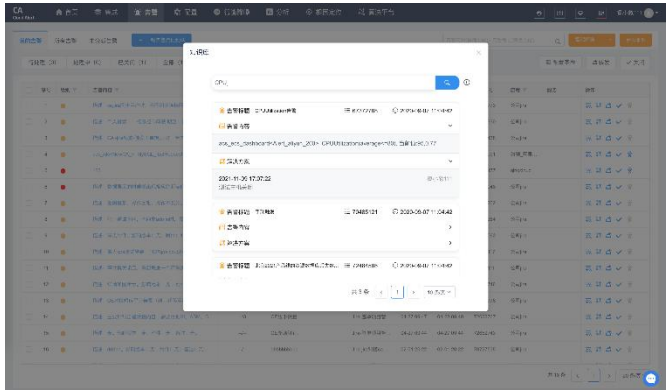
On-Call Management 是智能告警平台的管理控制模块，帮助企业将事件处理流程固化下来，通过分派、升级、转发、协作、排班等操作，确保信息在个人、组、团队间高效的协同。

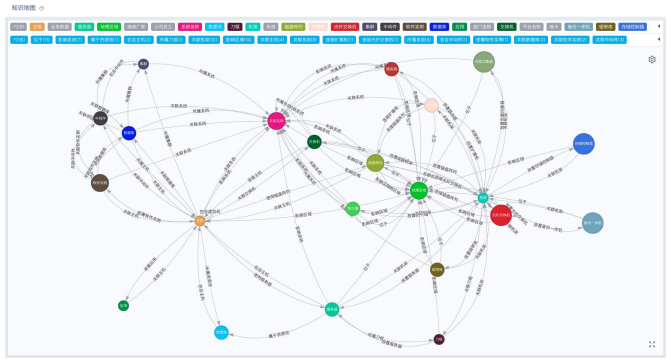
- 分派：以任务为对象，根据一定的限制条件将事件指定给某个具体人员。
- 升级：当前任务在约定时间内未被分派人员认领，该任务将自动升级到上一级事件处理人员。
- 转发：当前任务处理人员无法解决该问题时，可以手动的将该任务转发给其他人员或组。
- 协作：对于需要多人共同处理的任务，可以交由多个个人或组来协作进行处理，也可发起面向第三方协作办公工具的协作。
- 排班：按小时、日、周、或自定义周期制定周期性的多人工作安排计划。

On-Call Management 通过落地企业事件管理最佳实践，将任务分配工作变得更加简单，加强了团队内的责任和质量管理。通过直观、灵活的调度和升级，可以确保重要信息每次都通知到正确的人员。



CA智能告警核心能力：知识库和知识图谱





运维知识库

- 跨用户共用运维行业知识库。
- 单用户私有知识库。
- 超过5万条运维故障分析和处理建议。
- 运维知识自动与告警信息相匹配，主动推送知识给处理人员。

COKG 运维知识图谱模型

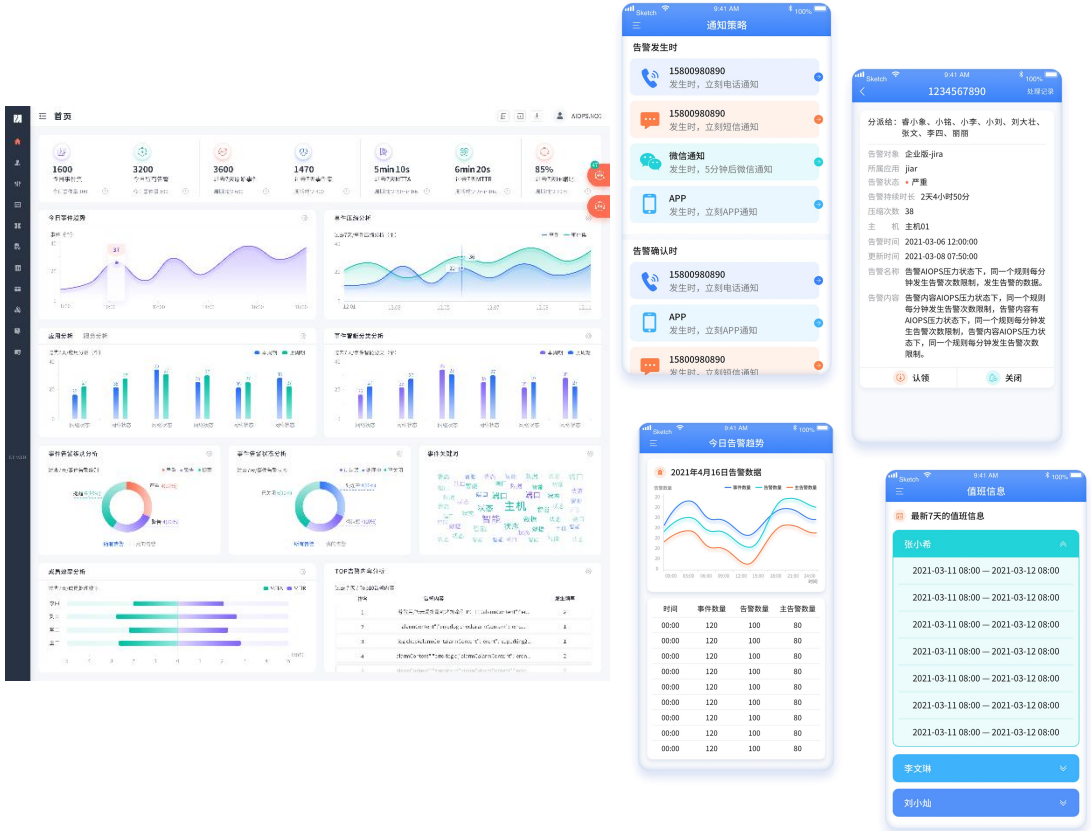
- 2020年9月份联合中山大学计算机学院陈鹏飞教授团队，发布全球首个面向运维领域的中文知识图谱 COKG（Chinese Operation Knowledge Graph）。
- 首批发布图谱包含：30 多万个实体，400 万个实体关系，并成功应用于黑龙江移动、中船重工等用户。
- 知识图谱是人工智能技术的重要组成部分，旨在帮助运维人员描述企业 IT 系统中各种实体或概念及其关系，从而构成一张巨大的语义网络。
- 知识图谱是智能问答、知识推理、决策分析的技术基础。

CA智能告警核心能力：分析及可观测

事件分析 以服务为对象对事件进行多维度实时分析，通过对事件、告警、事件集、服务、团队、相应操作和业务影响的整体视图，达成对告警态势的可观测性，最终实现更智能的实时决策。

- 多维度仪表盘及APP以可视化方式展现运营指标和 KPI。
- 事件智能降噪分析。
- 智能事件异常分析。
- 团队成员绩效分析。
- 事件详情分析。
- 日、周、月报表查询和导出。

事件的归档和分析是形成事件管理闭环的重要组成，帮助团队从海量的信息中实时洞察业务状态，归纳事件特征，沉淀处理经验。



睿象云 | DataFun.

04 智能告警管理最佳实践

助力德电中国打通运维监控最后一公里

客户背景

T-Systems 是全球信息和通信技术 (ICT) 领域的领导者之一，也是德国最大的云服务提供商，业务遍布 20 多个国家。德电（中国）通信技术有限公司是 T-Systems 全资子公司，在北京、上海、香港、深圳和武汉都设有分公司，主要为在华德国企业和国内企业提供传统系统和传统 ICT 服务的安全运营、面向云服务的转型以及新业务模型和未来业务领域的创新项目。

管理挑战

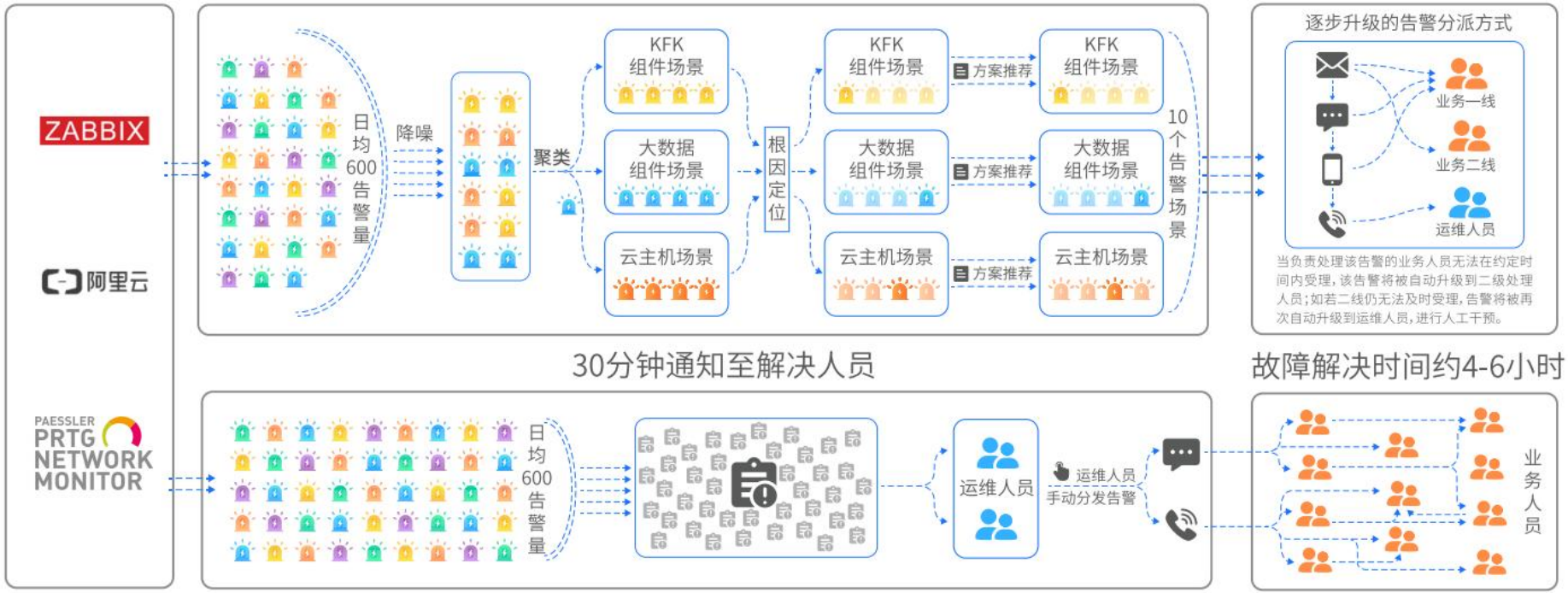
德电中国在各大项目中建立了标准的运维服务管理体系，以及 24 小时的 Service Desk 团队，负责通知工程师项目运行重要事件，但随着项目的不断扩容， 不同监控系统中产生的事件逐渐增多，对告警系统提出了更严格的要求。当前的告警工具和流程无法及时准确的将信息通知到对的人，导致错过了解决问题最佳时机，对项目运维造成了很大困扰。运维部门急需一个统一的告警平台，打通监控最后一公里。

应对方案

- 通过智能事件平台实现对现有监控工具事件信息的接入和存储。
- 通过平台的智能算法对海量告警进行去重降噪归类。
- 通过平台的On-Call机制和通知手段，对告警进行及时分派和推送。
- 推荐相关告警的根因，辅助运维人员快速定位故障。

平台价值	平台上线前	平台上线后
告警汇聚	监控工具多样，告警散落在各个系统，缺少统一管理	一个平台对接处理所有监控工具的告警
告警收敛	告警量大，告警风暴、告警疲劳，易忽略重要事件	去重+降噪+聚类，聚焦重点问题，告警量减少98%
告警分派	告警无差别通知，缺乏自动升级机制和响应机制	每类告警按规则分派到负责人，升级机制保障更及时响应，告警送达率100%
告警通知	通知方式单一，重要告警湮没在海量邮件中	5种通知手段，包括，微信、电话、邮件、短信、APP等
根因分析	人力有限，告警量多且难以定位故障根因，故障恢复慢	帮助定位故障根源，更快解决问题，故障定位时间缩短80%
分析优化	没有数据作为分析回顾，决策靠拍脑门	频发告警统计，告警规则发现，自动推送报告

助力德电中国打通运维监控最后一公里



为某国有综合性投资集团成功搭建IT可观测性平台

客户背景

某国有综合性投资集团(以下简称“该集团”)是一家以金融为主体、涵盖投资与资产经营的国有综合性控股集团，成立于2004年。截止2019年末,该集团合并资产总额达1722.56亿元。随着信息化的不断深入,该集团的T环境日趋复杂，传统的监控手段已经无法满足日常运维的要求。睿象云分两期为该集团成功搭建了IT可观测性平台。

管理挑战

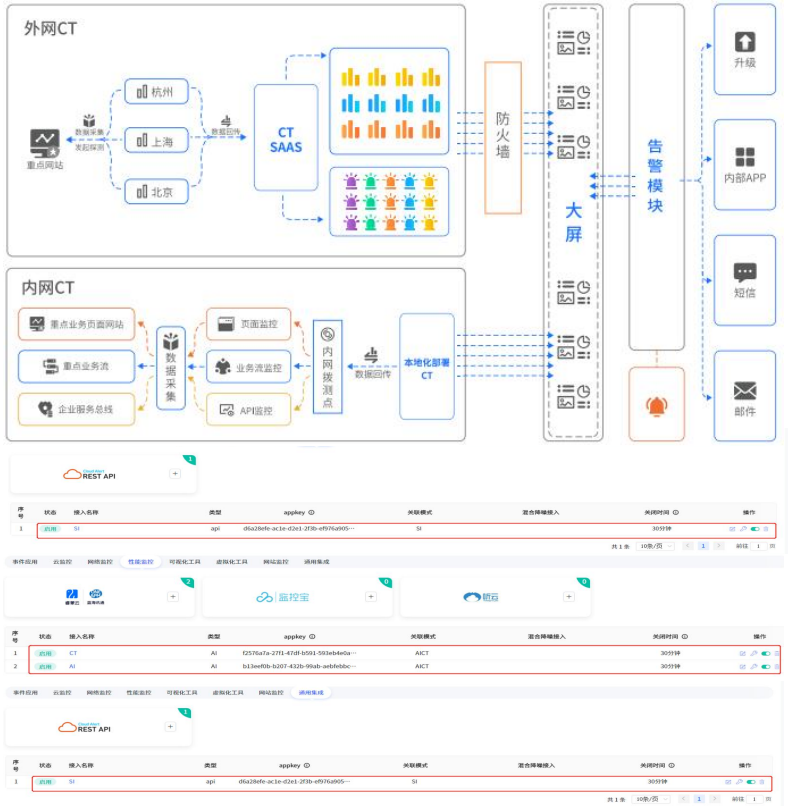
1多种监控工具告警分散：存在多种告警监控工具，包括基础资源监控、应用性能监控工具、模拟拨测系统等，告警策略过于分散，管理较为不便；

2.互联网区的核心业务系统缺乏监控手段：互联网区的核心业务系统缺乏有限监控和观测手段监控实时其相关性能情况；

3.缺乏可视化大屏展示：缺乏可视化大屏视图将一期的IT运行监控系统的数据通过直观的、图形化的方式呈现出来；

应对方案

项目引入集中化告警及可视化运维理念，整合已建成的基础监控、应用监控、模拟拨测、配置管理等多种运维工具所产生的告警信息，使其可集中通过内部APP平台来进行告警统一发送；抽取现有基础监控、应用监控、模拟拨测、配置管理等多种运维工具所产生的数据指标信息，通过监控数据多维度大屏展示等方式，实现运维数据高度可视化，同时针对互联网区的业务系统，通过主动拨测方式实时监控其可用状态。



睿象云 | DataFun.

为某国有综合性投资集团成功搭建IT可观测性平台

可观测性成果

通过可视化视图将信息化建设成果、业务系统组成、系统运行态势以及日常运维工作等信息通过多层次、多维度的展现视图直观、综合呈现出来，其中包含基础网络可视化、服务器、数据库等基础IT软硬件资源监控管理可视化、核心应用可用性呈现、核心应用运行态势感知、告警可视化等内容，通过面向领导、管理者、运维人员构建不同的管理视图，实现信息化环境和工作的可见。

睿象云 | DataFun.

智能告警平台服务的客户

科技互联网

腾讯云
MEGVII 旷视
中寰卫星 ZHONG HUAN
车行易
人人行 rrxw.net
微鲤
团油
NAVINFO 四维图新
雪球

IT 服务

中科软科技 Sinosoft Co.,Ltd
世纪互联 VNET
Analysys 易观智库
欢乐逛
汇通达 HUITONGDA
万翼科技
TEAMSUN 华胜天成
URORA 极光
Neocrm 销售易

游戏传媒

VIVAVIDEO
省钱快报 购物前·来省钱
RayJoy 雷尚
elex
TOPJOY
樊登读书
LRGame 灵刃游戏
IM30
趣头条

保险金融

华夏基金 CHINA ASSET MANAGEMENT
北京中关村银行 ZHONGGUANCUN BANK
哈尔滨银行 HarbinBank
社保卡科技
中证指数 CHINA SECURITIES INDEX
天星数科 小米旗下品牌
TCL 金融 TCL FINANCIAL
SHIE 上海保交所
和融通支付 HRT Payment

制造零售

ERKE
Vanguard 华润万家
BMW
ECOVACS
KYOCERA 京瓷
老百姓大药房 LBX PHARMACY
lota
土巴兔 装修大平台
药明奥测 WuXiDiagnostics

睿象云
DataFun.

非常感谢您的观看

睿象云 | DataFun.