



GOPS 2020  
Shanghai

# GOPS

# 全球运维大会

2020  
-AIOps 风向标



指导单位：



主办单位：



大会时间：2020年11月27日-28日

大会地点：上海中庚聚龙酒店



# 智能化、可视化技术助力 中国移动一级IT云成熟度提升

郭志斌 中国移动信息技术中心



# 郭志斌

信息技术中心  
AIOPS研发项目经理

IT领域专家，AIOPS研发项目负责人

# CONTENTS 目录

- ① 中国移动一级IT云概况
- ② 一级IT云智能化、可视化能力建设思路
- ③ 机房可视化、精益大屏成果分享
- ④ AIOPS能力建设成果分享



# 中国移动一级IT云概况

## IT云

选址布局

8大区：黑龙江 内蒙古 河北 河南  
湖南 浙江 江苏 广东

## 网络云

选址布局

8大区：黑龙江 河北 河南 江苏 陕西  
四川 浙江 广东

## 移动云

选址布局

6大区：北京 河南 湖南 江苏 四川  
广东

信息技术中心统筹投资建设，  
负责运营，负责运维的组织  
和管理。

定义了统一的资源目录，提供全网78种IaaS 服务，其中：云主机服务19种，裸金属服务器10种，镜像服务36种，存储服务5种，网络服务8种。

提供自服务门户和运营，运维模块，提供对全网范围的资源服务进行申请审批、调度管理、分配回收、监控通报、运营分析等功能。



规模

运营

运维

平台

截止到目前，服务器规模超过10万台，虚拟机超过10万台，应用超过6000个。

2019年日常作业工单（包括网络策略、资源管理、账号管理、问题处理等）接近1万张，变更、割接工单超过1600张。

2019年依托移动研究院专家团队，开展了一级IT云成熟度摸底。

云规划

云资源  
提供

云服务

云安全

承担“**云战略**”的功能，负责对云服务的战略规划、技术规划和服务能力改进管理

对各类云资源使用情况信息的管理，衡量资源池能够提供的**资源服务能力**

云服务管理包含面向用户的服务交付相关的**运营工作**和相关的**运维工作**，体现私有云的整体服务水平

私有云的云安全是保障服务可靠运行的重点，包含安全管理及制度，**云环境下需增强的虚拟化安全、网络安全、数据安全及安全服务等**

高优先级：

- 建议增加业务规模，开展合理容量规划，提升资源利用率

中优先级：

- 建议提供资源池内VM粒度的手动资源调度的能力
- 建议增强云配置管理，提供IT可视化
- 建议故障管理增加故障处理初步自动化能力

低优先级：

- 建议运营交付流程增加虚拟计费管理能力



随着一级IT云的规模的迅速扩大，在日常的建设、运营、维护过程中，一些**问题**也随之出现。

### 资源管理精细化程度需提升

- 大量资产定位依赖人工，资源可视化程度有待提升，需要更加直观，能够管理到U位的技术手段
- 温度突变、风扇故障等局部故障很难发现
- 监控、巡检靠人，频率低，很难及时发现问题
- 机房环境巡检、资产核查均靠人工，巡检周期长，自动化水平低
- 外来人员的安全管理难度大，工作负担重，效果有待提高



### 运维工作量巨大

- 设备数据巨大，监控项阈值设置依赖人工，工作量大，工作重复，固定阈值难以准确体现业务特征
- 告警风暴溯源困难，导致处理时间长，告警响应速度需进一步提升
- 大量设备日志的价值没有得到充分挖掘
- 亟需通过科学合理的容量预测工具，引导租户提升资源利用率



# 一级IT云智能化、可视化能力建设思路

# 提升AIOPS成熟度：

不断迭代**升级运维工具**，**丰富智能运维场景体系**，**提升运维效率**

从一级IT云运营、维护实际出发，引入大数据、人工智能等关键技术，强化**顶层设计**，加强数据治理，提升各类软件管控水平，**夯实AIOPS基础**，从保障质量、降低成本、提升效率、实现资源可视化等多个场景入手，构建智慧云运维体系。

- 多维度展现IT资源建设量、分配量、库存量
- **机房可视化**：3D展现机房、机架、设备情况
- **精益大屏**：通过智慧大屏，动态展现资源利用、告警监控、指标分析和工单管理等数据，实现精细化运营

- 通过**智能低效资产评估**，精准淘汰低效、老旧资产，智能预测维保费用变化趋势
- 通过**智能容量预测**，为租户提供精准的资源使用量规划建议

## 加强数据治理能力建设

- 推动主机/网络/存储等基础设备的监控指标标准化
- 提升设备监控的覆盖率
- 实现各类监控数据、日志数据、告警数据、工单数据的集中存储



- 提升**自动化能力**，建立脚本库，实现批量修改密码、批量软件安装、自动化系统巡检等功能
- 通过**日志异常检测**场景，智能化模式提取和日志聚类，方便用户基于少量的模板来检测异常

- 通过**智能异常检测**和**告警溯源**提升告警质量
- 通过**告警自愈**和**智能网络资源优化**等提升服务质量

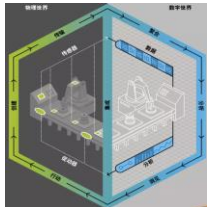
## 提升IT云软件管控水平

- 逐步收敛操作系统类型，降低维护复杂度，减少维护成本
- 从产品能力、服务能力、使用体验、系统架构、安全防控等多个方面提升云管平台建设水平



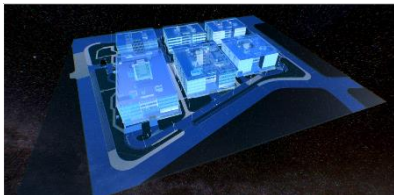
# 机房可视化、精益大屏成果分享

# 数字孪生3D机房



将数字孪生技术应用于数据中心管理，在数字世界中构建物理对象3D模型，实现对数据中心园区、楼宇、机房、机柜、设备、端口和配线的多层级管控，总体呈现一级IT云整体运营情况，打造集中化、精细化和可视化的管理能力。

环境可视



配线可管



容量可控



## 多数据中心运维难点

一级IT云数据中心分布广泛（8区域中心），设备规模大（10万量级），运维及管理人员主要集中在北京、深圳、广州三地，需要一种技术手段打破时间和空间的壁垒，实现对全网资源的“一站式”可视、可管和可控。



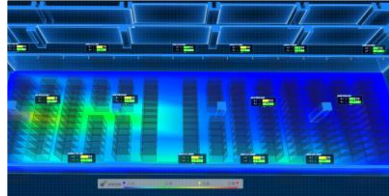
## 数据中心数字孪生应用价值

- 提升数据中心机房管理效率，填补机房管理空白。
- 对标数字-物理世界，提升资产数据准确性。
- 可视化容量统计信息，提高资源容量查询速度。
- 拉齐安防、动环、性能、告警等数据，提升数据标准化和规范化水平。
- 提高日常巡检工作效率，动静结合提升汇报效果和真实感。

资产可视



温度可管



安防可控



# IaaS层端到端监控



## 面临挑战

随着IT技术发展、DevOps理念与容器等技术的推广应用，云运维面临着管理规模大、架构变化快和跨层数据复杂等问题。日常运维中如何快速定位故障成为了难点。



## 解决之道

在云运维工作中精准把握快速变化的架构是关键，以端到端监控为出发点的IT云架构管理是业务变更、故障定位的关键，可视化、自动化和智能化的架构管理有效提升了云运维效率。



逻辑



物理

**智能故障关联**，创新地实现物理-逻辑架构的数字孪生，打通跨资源池、跨网络的不同层面数据壁垒，实现跨层故障定位，有效缩短MTTR。

**自动架构管理**，通过构建弹性资源模型，实现IT架构的自动化管理，数据驱动的架构图更新，做到业务变化随时掌握，有效提升架构管理效率。





# 精益管理大屏



## 一级IT云全景视图

建立“全网-资源池-机房”的多维度、可下钻管理能力，改变了传统大屏交互性差的问题，进一步提升了大屏的实用性和扩展性。在全量资源数据的基础上，将告警、性能、动环、监控等实时数据与分析结果进行集成，实现对跨地域、多数据中心实时动态的“一站式”掌控，打造了一级IT云全景视图。



## 语音控制

系统集成语音识别模块，支持通过PAD发出语音指令进行主题切换和资源定位，有效简化了操作的步骤，提升大屏交互性。



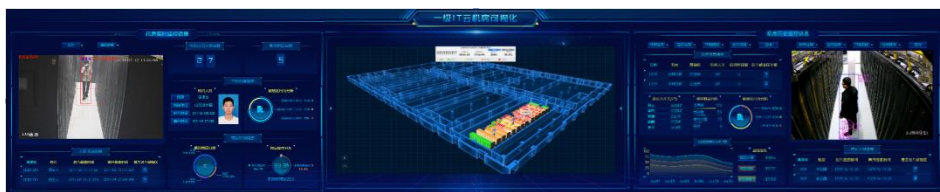
全网视图



资源池视图



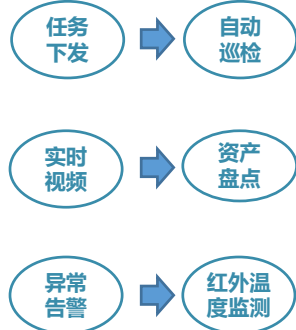
自动巡视图



智能安防视图

# 机房自动巡检

将智能机器人应用于数据中心运维，结合5G通信，实现自动巡检、资产盘点及红外热成像温度监测等功能，同时与精益管理大屏结合形成一体化呈现与交互，提升运维效率，降低人工成本。



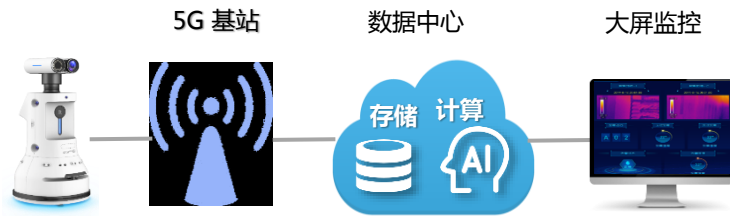
1080P可见光摄像头



640×512热成像



助力 "5G + AI"





# AR智能运维

运维工程师佩戴AR眼镜，选择登录账户，指定巡检员，查询巡检任务，完成巡检任务

## 分权登录



1. 运维人员第一次使用或更换使用者后，需使用账号密码登录进行身份验证
2. 登录不同权限的账号，查看不同级别的信息

## 现场巡检



1. 眼镜端虚拟图形化方式展示巡检任务，辅助现场人员完成巡检计划
2. 机房环境及设备的巡检

## 巡检服务



1. 巡检执行任务时支持拍照及视频录入功能
2. 巡检执行任务时支持图片、PDF及视频资料查看

# 可视+智能融合应用场景

## 人脸特征识别

采用深度学习的人脸特征分析算法提取出人脸的性别与年龄信息，同时提取出人脸的128维特征向量，进行人脸识别。



## 目标检测

采用端到端的目标检测算法YOLO-v3结合机房实际场景数据进行优化训练，实现对机房人员携带物品进行检测。



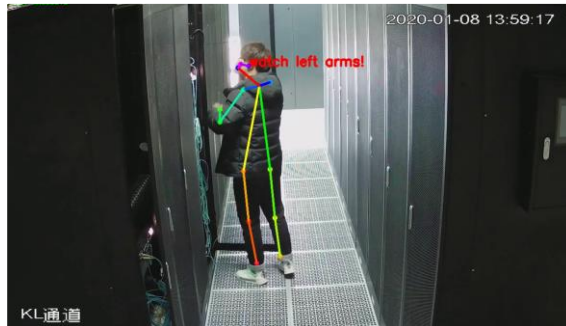
## 轨迹跟踪

使用Yolo算法检测人员位置，并使用DeepSort算法勾画人员运动轨迹，从而判断人员是否进入敏感区域。



## 动作识别

使用openpose肢体识别库对人体18个特征点进行提取，将提取出的特征点送入ST-GCN算法中进行动作分类，识别是否是危险动作。





# AIOPS能力建设成果分享

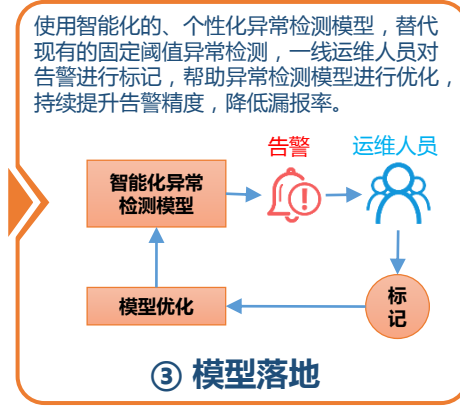
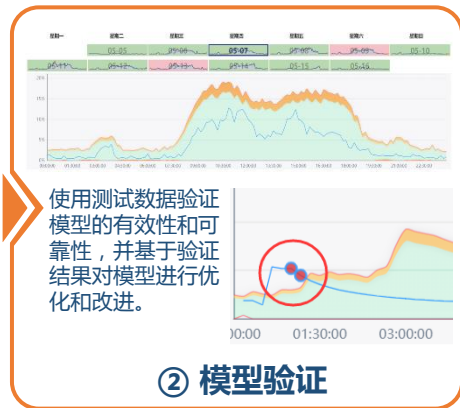
# 智能化指标异常检测

## 大型资源池监控带来的挑战

- 监控指标数量巨大，阈值梳理非常困难
- 基于固定阈值的告警精度低
- 传统阈值无法适应指标的动态性



## 智能化方案



# 智能化告警压缩与溯源

## 告警管理中故障定位和溯源的挑战

- 故障发生时，往往触发大量告警形成风暴
- 维护人员受干扰，难以识别导致故障的根因
- 因此故障定位和故障排除周期都比较长

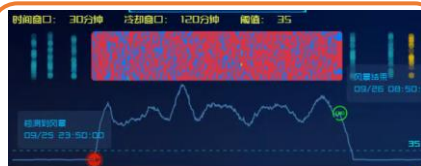


## 智能化方案



基于历史告警，一方面，通过频繁项挖掘算法，寻找告警之间的相关性模式，生成关联规则；另一方面，学习告警的密度、分布模式等特征，建立模型用于识别突发的告警风暴。

### ① 告警特征分析



接入实时告警，基于训练生成的规则和告警特征模型，检测告警风暴和相关性告警。

### ② 风暴识别与告警压制

通过复现告警风暴的时序，结合告警关联规则，为用户推荐根源告警，帮助运维人员理解并定位故障。



### ③ 辅助告警定位与溯源



# 智能化日志异常检测

## 运维中日志管理的痛点

- 各类设备、应用都会持续产生大量的日志
- 日志内容繁杂，缺乏标准，依靠人工解读信息十分困难
- 在查找故障时，难以跨设备复原相关的日志上下文



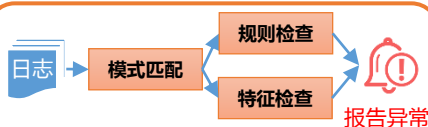
## 智能化方案

```
TIME LEVEL User = ${WORD} FROM ${IP} ACTION
TIME LEVEL User = admin FROM 10.10.10.1 ACTION
TIME LEVEL User = test FROM 10.10.10.1 ACTION
TIME LEVEL User = admin FROM 10.10.10.1 Login
TIME LEVEL User = test FROM 10.10.10.2 Login
TIME LEVEL User = test FROM 10.10.10.2 Logout
TIME LEVEL User = admin FROM 10.10.10.1 Logout
```

多层聚合  
逐级范化

利用历史日志训练模型，自动抽取日志中的文本模式，并通过层次聚类建立多粒度模板。基于所生成的日志模板，再进一步学习日志中模板的出现频度、密度、日志相关性等信息，形成日志异常检测模型。

### ① 日志模式与特征提取



接入实时日志后，日志异常检测模块跟踪每一个日志模式的发生频度和特征，再基于维护人员定义的规则，联合报告异常。



### ② 异常检测

在故障发生时，通过故障发生时点，或者基于维护人员排查需求所人工指定的时点，再配合设备拓扑关系，自动复现指定时点前后的日志现场。

确定故障时点

对相关日志特征进行可视化

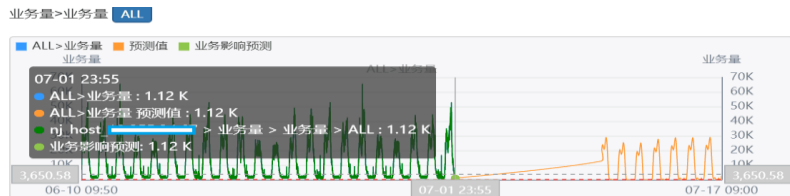
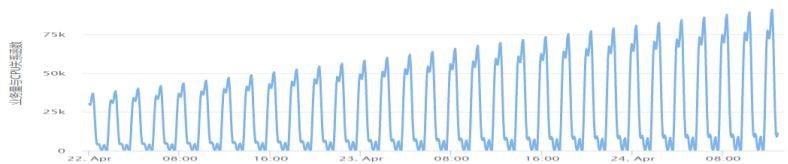
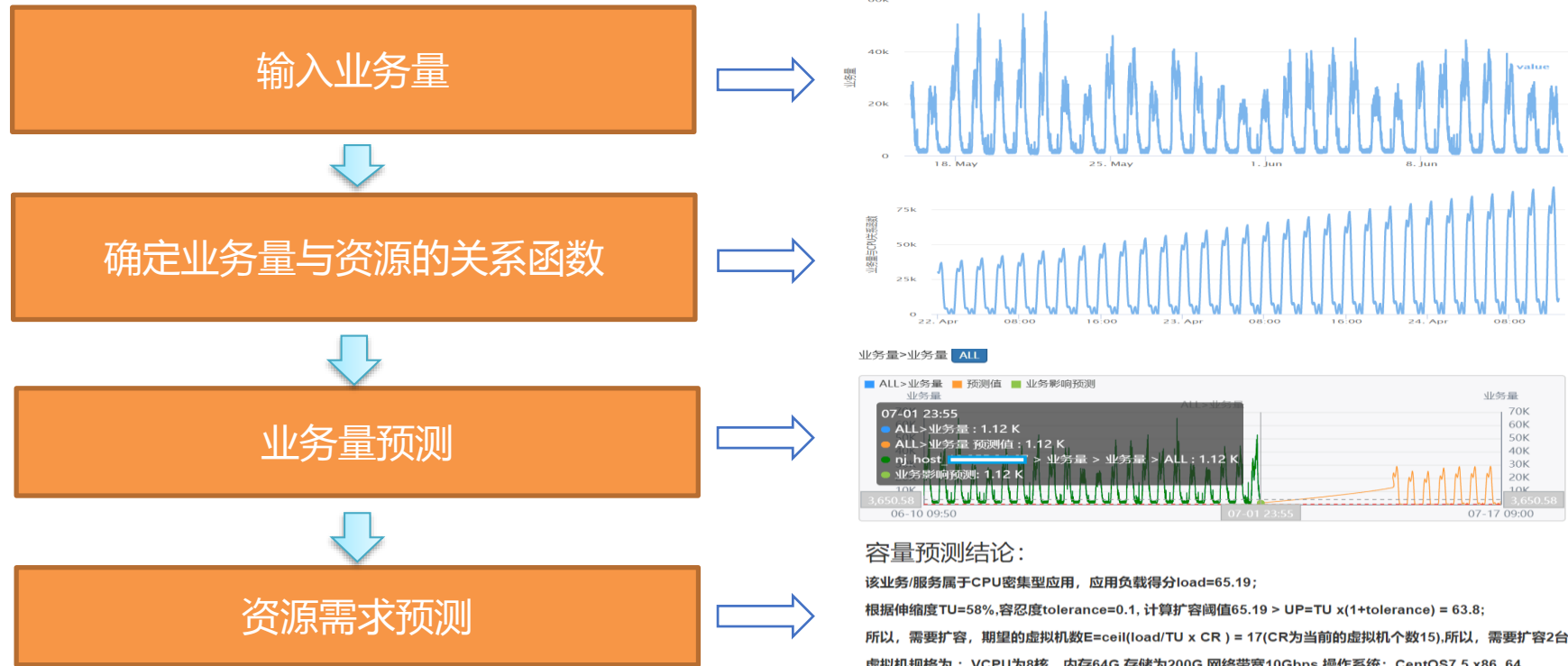


网络、防火墙、存储、主机等相关日志上下文

### ③ 日志上下文复原

# 容量预测：基于业务历史数据预测未来的资源需求

对业务的资源使用情况进行学习，分析业务量、时间对资源用量的影响情况，从而对系统未来的资源容量需求进行预测，给出容量规划建议，提高资源利用率。



## 容量预测结论：

该业务/服务属于CPU密集型应用，应用负载得分load=65.19；

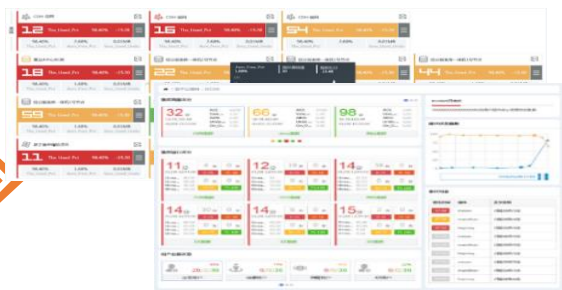
根据伸缩度TU=58%，容忍度tolerance=0.1，计算扩容阈值 $65.19 > UP = TU \times (1 + tolerance) = 63.8$ ；

所以，需要扩容，期望的虚拟机数 $E = \text{ceil}(load / TU \times CR) = 17$  (CR为当前的虚拟机个数15)，所以，需要扩容2台虚拟机。

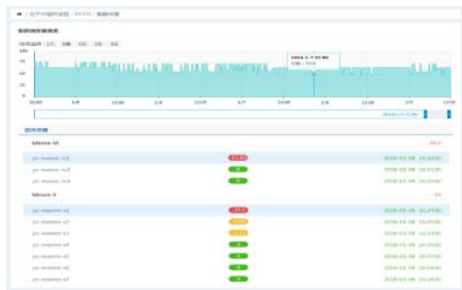
虚拟机规格为：VCPU为8核，内存64G，存储为200G，网络带宽10Gbps，操作系统：CentOS7.5 x86\_64

- 监控指标数量多
- 监控分析专业度高
- 缺少全局监控视角

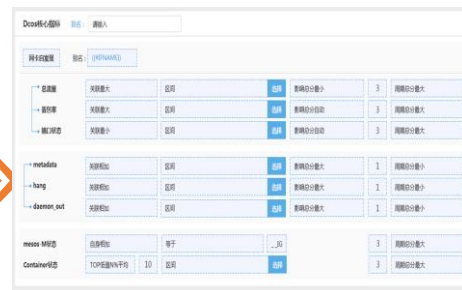
建立**多维健康度打分模型**，实现网元级运行状态的抽象化展示。让一线人员通过全局健康度分数了解设备的整体运行情况，快速定位故障，实现专家经验平民化



## 设备级&服务级监控，简单直观展示系统健康程度



## 历史趋势随时回顾 指标详情一览无余

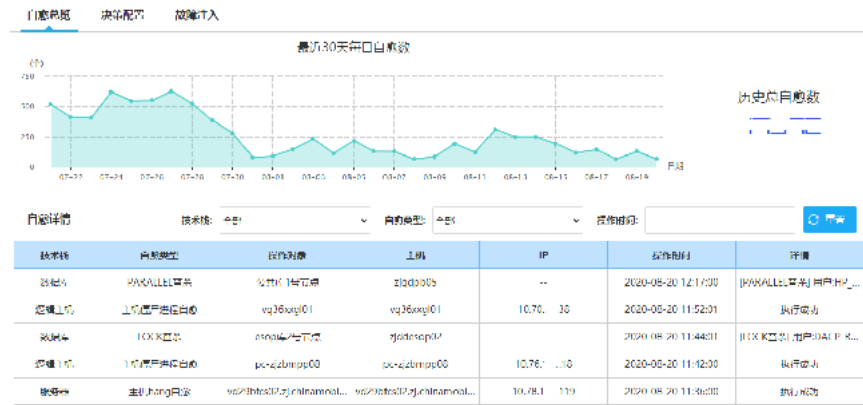
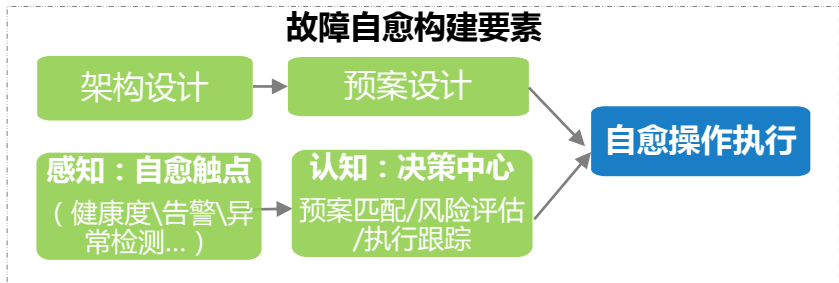


## 可视化模型编排 内置时序预测的动态基线



# 故障自愈：实现感知、认知、操作的联动

建设了故障四级自愈体系，通过感知体系构建自愈触点，通过决策中心实现预案匹配、执行环境风险评估和执行后跟踪，通过自动化操作平台实现自愈操作执行。实现感知、认知和操作的联动，大幅降低故障平均恢复时间。



## 01.业务级（半自动、辅助决策）

- 手段：**业务级切换
- 场景：**遇到数据中心级故障，原生产已无法恢复，故切换到备份系统，实现核心业务连续性
- 原理：**基于业务双平面实现

## 02.服务级（应急、半自动、辅助决策）

- 手段：**容灾切换
- 场景：**服务级故障切换
- 原理：**基于服务级灾备体系实现

## 03.网元级（常规、自动决策）

- 手段：**重启
- 场景：**实例级故障隔离
- 原理：**基于实例所属服务自身或应用多写等高可用能力实现

## 04.进程级（常规、自动决策）

- 手段：**查杀
- 场景：**OS级别异常进程查杀，快速消灭异常
- 原理：**基于应用进程自动重连、自动启动手段实现

# AIOPS能力建设所带来的成效



实施规模

5000+ 台设备接入

## 指标异常检测设备覆盖

目前从呼和浩特和信息港资源池各选择了一个pod的作为接入试点，共覆盖5000+台设备，其中包括网络设备和裸金属服务器。



告警精度

30% 漏报率降低

## 指标异常检测相比固定阈值

在生产环境下实测，智能化指标异常检测模型，可有效检出低负载、低流量场景下的指标异常波动，相比固定阈值，漏报率降低约30%。



告警压缩

23% 告警压缩比

## 告警压缩和溯源

系统上线后至2020年10月，共帮助运维人员发现告警风暴5次，结合关联规则对告警进行压缩和溯源，可实现约23%的告警压缩比。



故障日志排查

29% 速度提升

## 日志异常检测

通过模式提取算法，可将百万行的日志提取为数百条的日志模板，大大降低配置工作量，日志上下文可有效提升基于日志的故障排查速度。



# Thanks

高效运维社区  
开放运维联盟

荣誉出品

想第一时间看到高效运维社区  
的新动态吗？

