

AI 驱动 软件研发 全面进入数字化时代

中国·北京 08.18-19

AI+
software
Development
Digital
summit



阿里云服务器智能异常调度系统构建与实践

朱兆良 阿里云计算有限公司

科技生态圈峰会 + 深度研习 —— 1000+ 技术团队的选择



2023K+
全球软件研发行业创新峰会
上海站

会议时间 | 06.09-10



2023K+
全球软件研发行业创新峰会
北京站

会议时间 | 07.21-22



2024K+
全球软件研发行业创新峰会
深圳站

会议时间 | 05.17-18



K+峰会详情



会议时间 | 08.18-19

AiDD AI+软件研发数字峰会
北京站



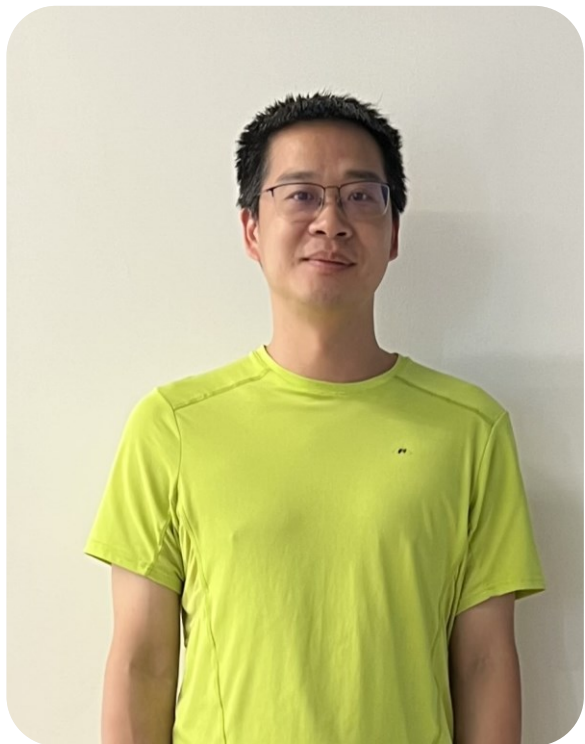
会议时间 | 11.17-18

AiDD AI+软件研发数字峰会
深圳站



AiDD峰会详情

▶ 演讲嘉宾



朱兆良

阿里云计算有限公司 高级技术专家

阿里云弹性计算平台异常调度AIOPS方向负责人。经历淘宝、阿里妈妈、对象存储、弹性计算多个子公司及部门，近8年的运维管控系统建设经验，主导建设存储运维管控系统赤骥，近年来专注弹性计算异常调度平台AIOPS方向，致力于通过AI的方式赋能并提升弹性计算稳定性、运维效率及智能运维能力。

目录

CONTENTS

1. **AIOPS是必然选择**
2. **智能异常调度系统介绍**
3. **AIOPS在异常调度系统的实践**
4. **感悟**

PART 01

AIOPS是云计算发展的必然选择

大规模系统的选择

88
可用区

3200+
网络和CDN节点

29
数据中心区域

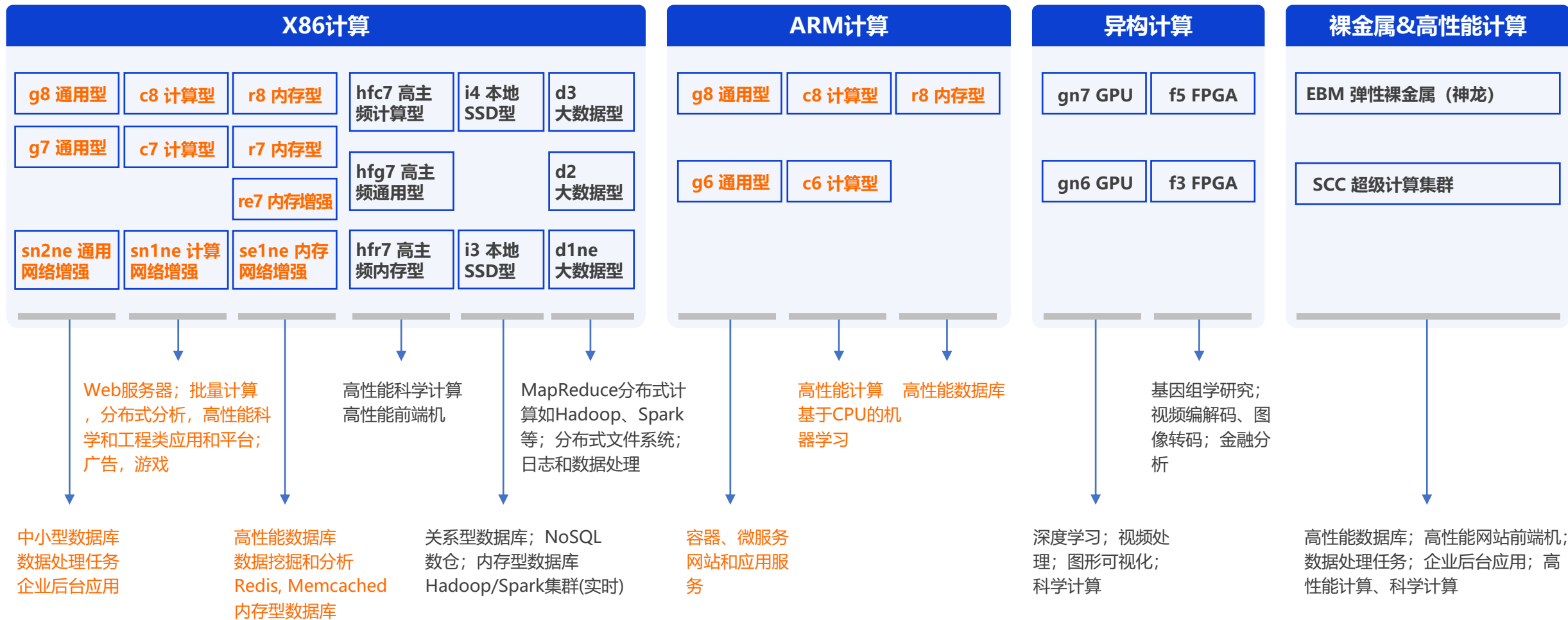


~ 100,000,000+
部件
(CPUs, disks etc.)

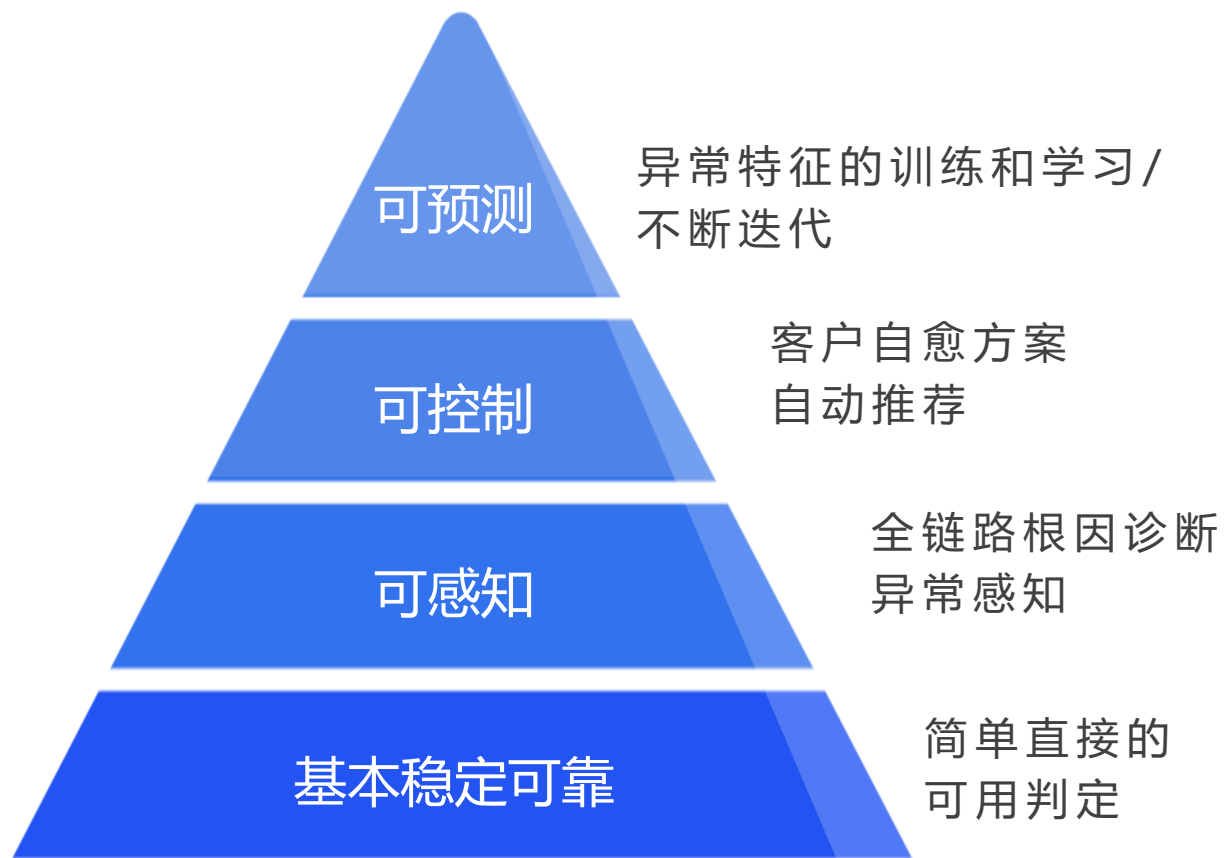
~1,000,000+
设备

~5000+
集群

复杂产品形态的选择



▶ 客户稳定性及运维体验的选择



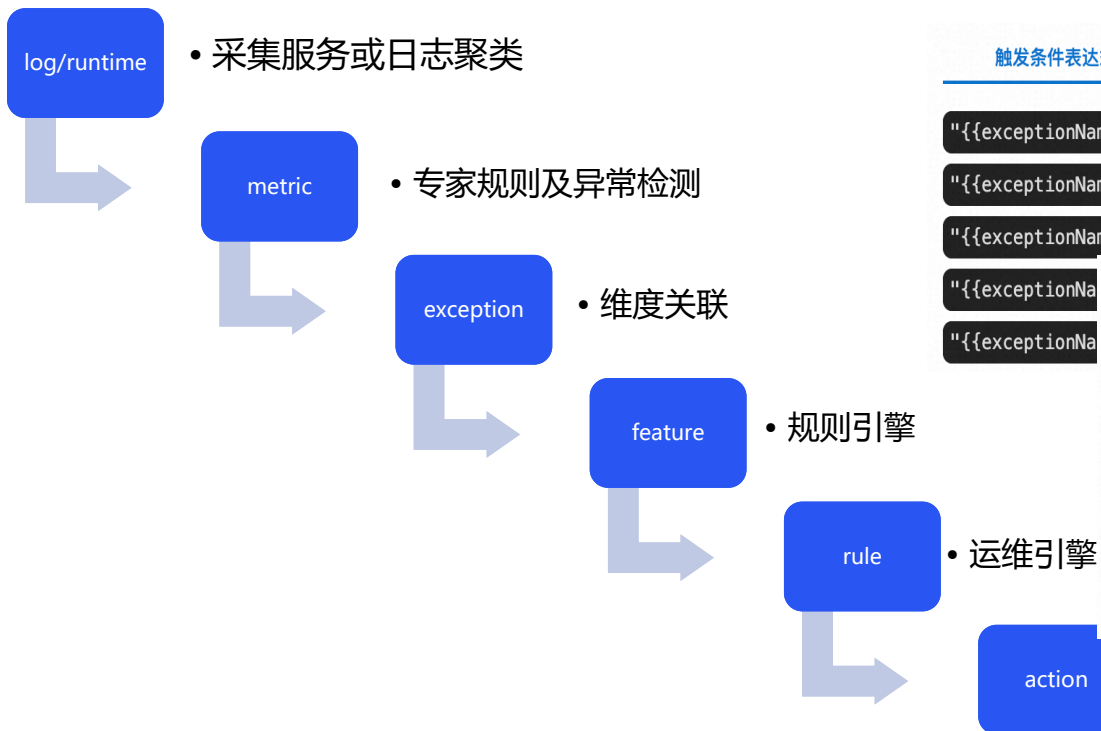
PART 02

智能异常调度系统介绍

系统架构



▶ 关键路径及概念



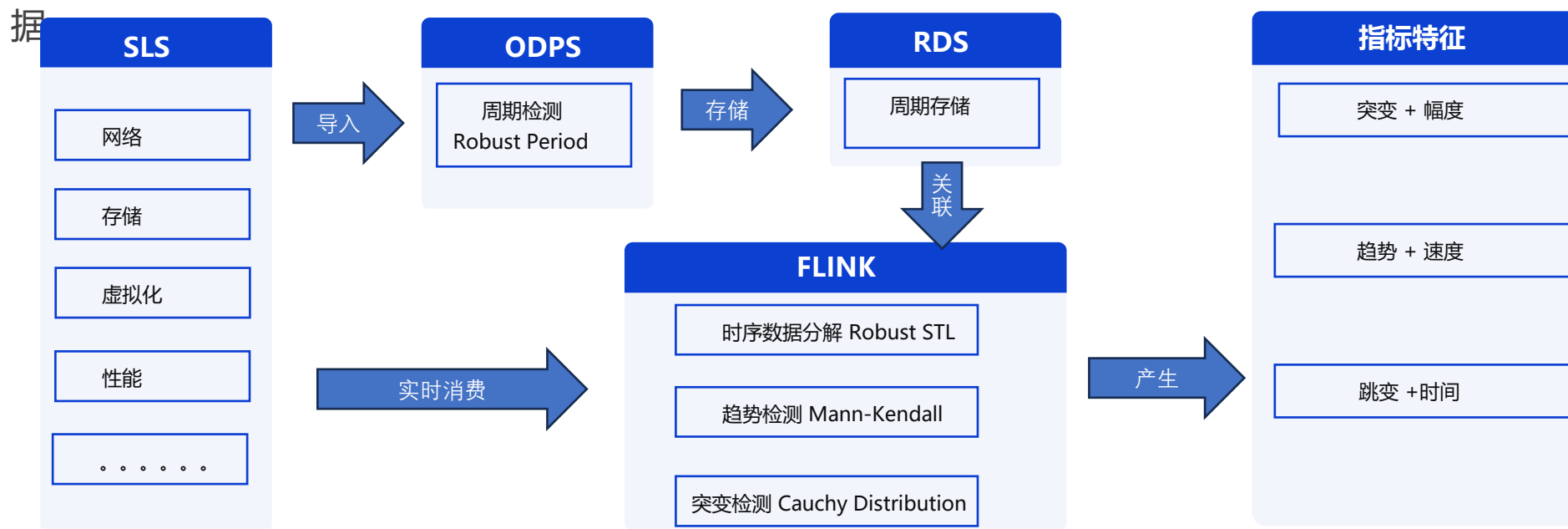
```
触发条件表达式 5 限制条件 0  
"{{exceptionName}}" == "libvir" and "{{warningLevel}}" == "fatal"  
"{{exceptionName}}" == "alarm_agent_ption" and "{{reason}}" == "libvirtd_error"  
"{{exceptionName}}" == "xdragon" and "virsh list cmd is not OK" in "{{additionalInfo}}"  
触发条件表达式 4 排除条件表达式 7  
"{{exceptionName}}"  
"{{exceptionName}}"  
{{cpu}} or {{hardware}} or ( {{motherl}} and not {{mothe}} )  
{{user}} < 2  
{{isLocalDisk}}  
{{hardware}} or {{unfinishe}}_exclude_disk}}
```

PART 03

AIOPS在异常调度系统的实践

智能诊断(指标异常检测)

- 采集服务采集了大量的系统运行时数据(每天800+TB的数据量, 1000+的指标项)
- 专家阈值设定判断异常无法跟随系统变化而变化, 无法适配周期性数据



当前**实时**检测序列超过5000W+条, 预计9月底会达到破亿

智能诊断(指标异常检测)

策略名称: exception_cluster检测策略

指标粒度(s): 1800

过期时间(s): 604800

缺值补值策略: MEDIAN(中值补齐)

平滑窗口: 12

匹配策略: 前缀匹配

异常检测配置

新增

阈值: 1686067200

波动值: 5

+添加详细配置

突增

静默时间(s): 10800

波动率: 2

持续性: 2

波动值: 5

+添加详细配置

+添加异常类型

-删除异常类型

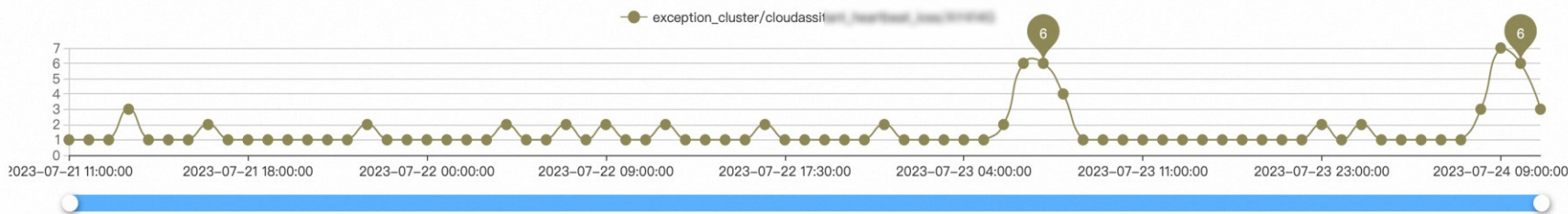
提交审批

关闭

指标名称	周期长度	数据源	异常检测	丢点检测	丢点阈值(s)	丢点率(%)	异常检测策略	指标趋势	操作
exception_cluster/cloudassitant_h...	0	exception_cluster/	已开启	未开启	--			指标趋势	

指标趋势

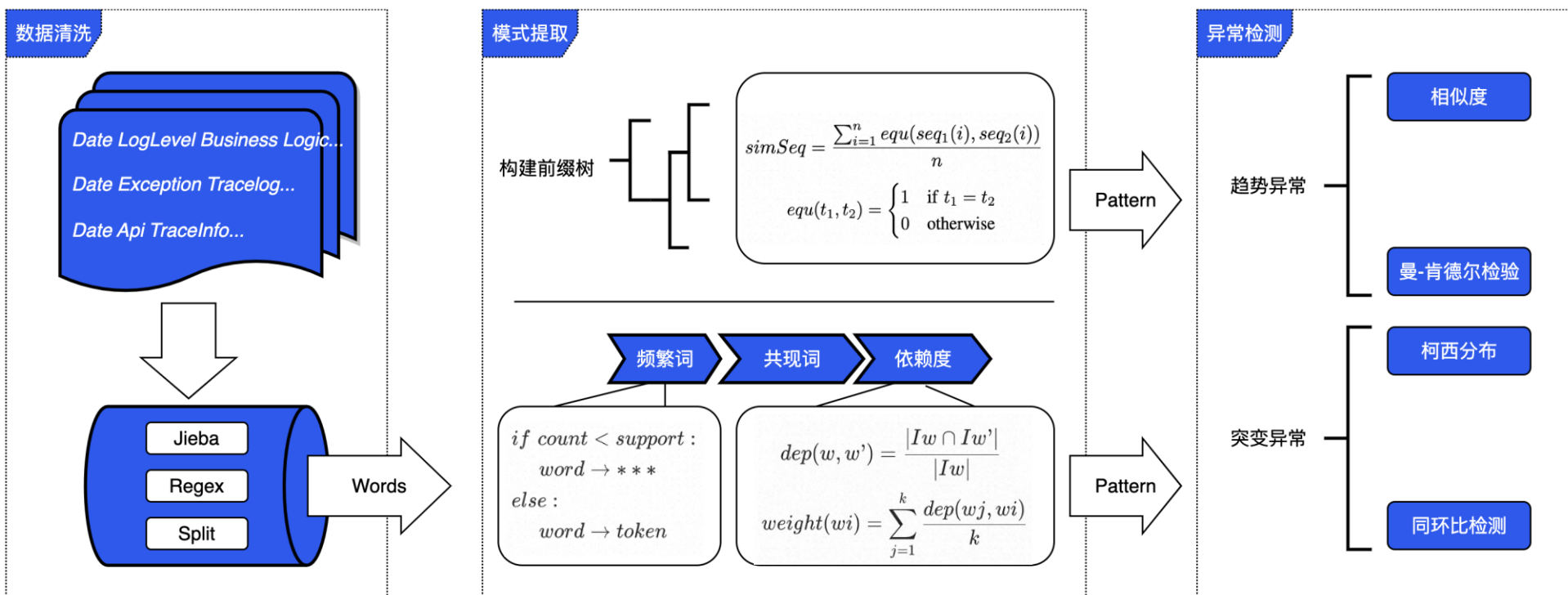
2023-07-21 10:37:03 ~ 2023-07-24 10:37:03



名称	告警时间	趋势开始时间	异常类型	数量	拓展信息	操作
exception_cluster/cloudassitant_h...	2023-07-23 06:30:00	2023-07-23 06:00:00	spike	6		
exception_cluster/cloudassitant_h...	2023-07-24 09:30:00	2023-07-24 09:00:00	spike	6		

智能诊断(日志聚类分析)

- **异常特征检测:** 特征发生异常时 (突变), 底层逻辑变化或真实异常发生变化, 需要人工处理
- **新Pattern:** 当关键日志解析出现新的LogPattern时, 这类问题可能会导致之前的专家规则发生了变化, 需要及时更正
- **新异常特征:** 通过异常模式提取以及模式异常检测, 找到突变异常或趋势异常



智能诊断(日志聚类分析)

基于词频的日志聚类

获取频繁词

统计每个独立单次的词频
如: (w1, 3), (w2, 99)

去除低频词

如: (w2, 99)

相同候选项合并

(interface * down, 1)
(interface * down, 1)
(interface * down, 2)

覆盖候选集合并

(User bob login from 10.1.1.1, 5)
(User * login from 10.1.1.1, 5)
(User * login from 10.1.1.1, 10)

聚类合并

(Interface *down at node router1, 5)
(Interface *down at node router2, 5)
(Interface *down at node *, 10)

实时异常检测新增及突变

```
AisRiskConsumer handlerMessage exception message *** jav  
a.lang.NullPointerException null at com.alibaba.cloudops.servic  
e.risk.RiskMessageServiceImpl.insertRiskMessage RiskMessag  
eServiceImpl.java *** at com.alibaba.cloudops.service.risk.Risk  
MessageServiceImpl$$$FastClassBySpringCGLIB *** invoke gen  
erated
```

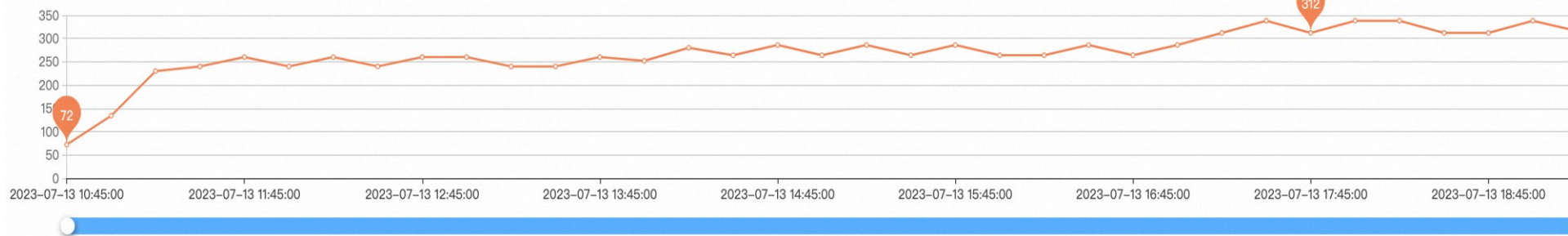
暂无备注标签

TOP 详情

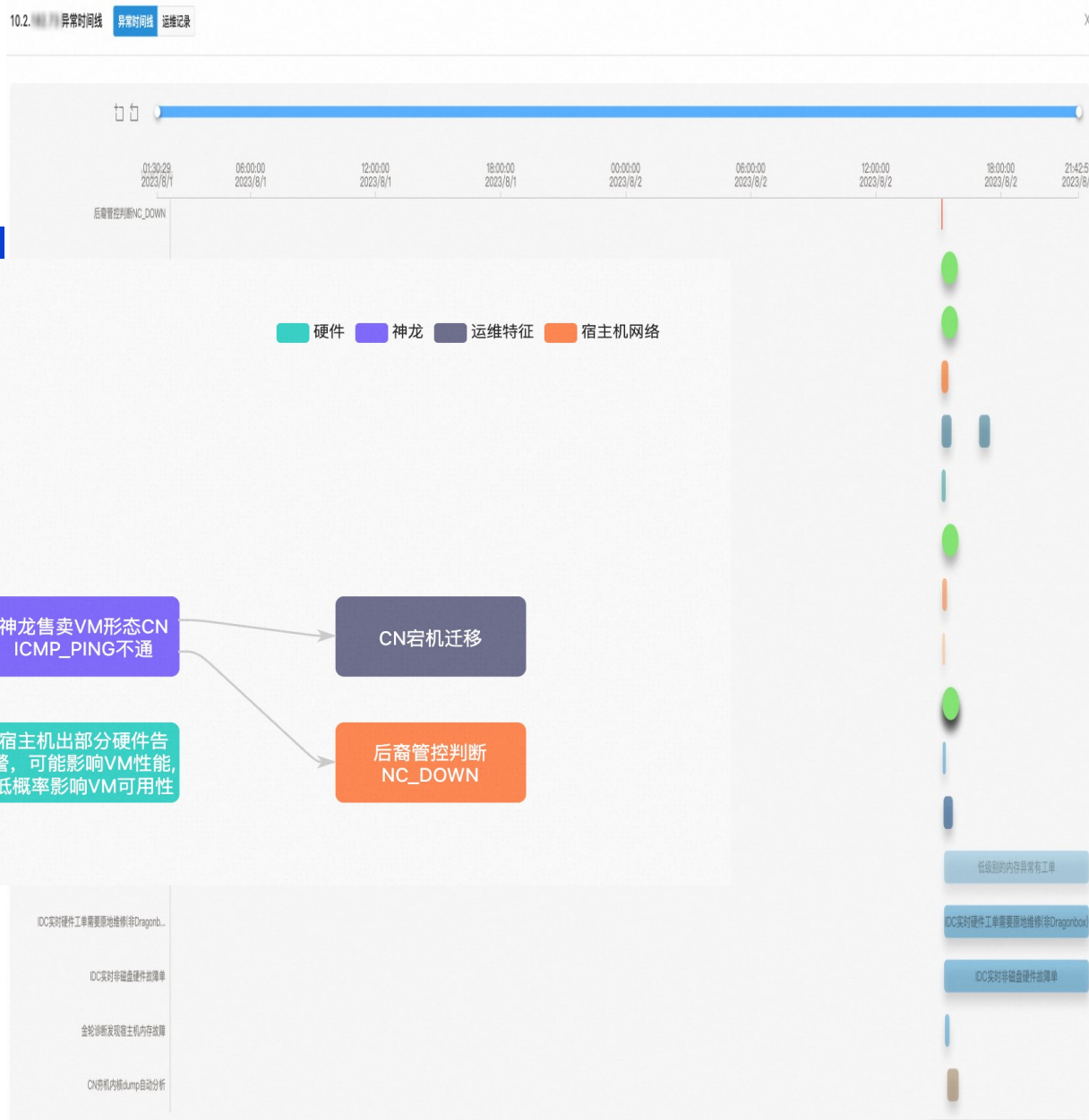
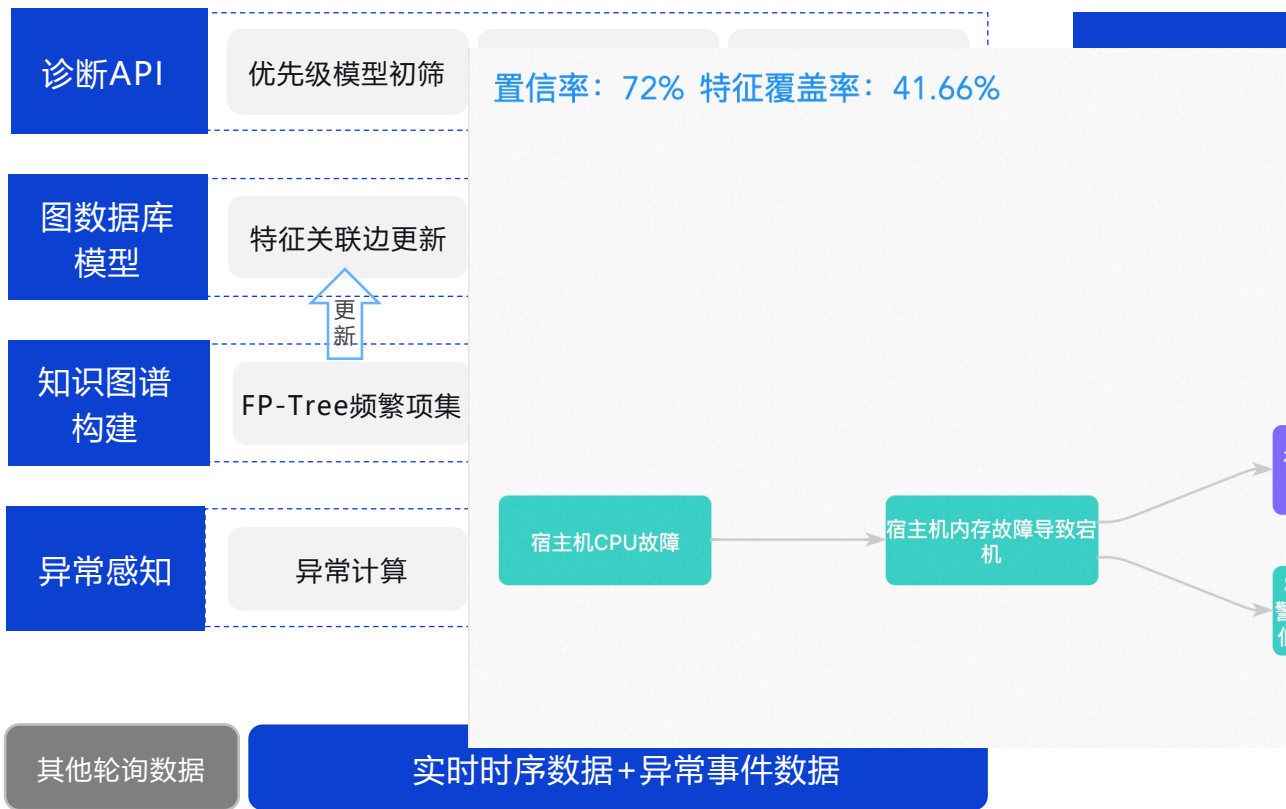
2023-07-13 10:45:00

2023-07-13 19:15:00

9,398



智能诊断(根因诊断)



智能运维决策-概览



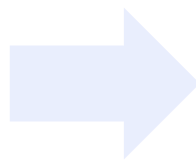
▶ 智能运维决策-运维规则

• 优化现有规则

R: A and B and C \Rightarrow R: (A or A₀) and B and C

- 关联关系挖掘 (FP-Tree) , 发现A与A₀
- A与A₀的关联系数超过阈值, $P(A|A_0) > \theta$
- A₀发生的事件早于A

Sequence_id	Sequence
10	$\langle a(abc)(ac)d(cf) \rangle$
20	$\langle (ad)c(bc)(ae) \rangle$
30	$\langle (ef)(ab)(df)cb \rangle$
40	$\langle eg(af)cbc \rangle$



• 挖掘新规则

针对特定场景X当前存在运维规则

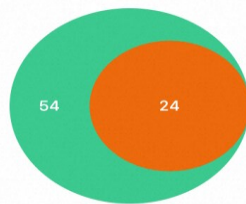
R: A and B and C

- 通过历史上场景X发生时的事件序列集合, 挖掘频繁序列挖掘(PrefixSpan)作为新的规则: $DEF \Rightarrow D \text{ and } E \text{ and } F$
- D/E/F发生的事件早于A/B/C

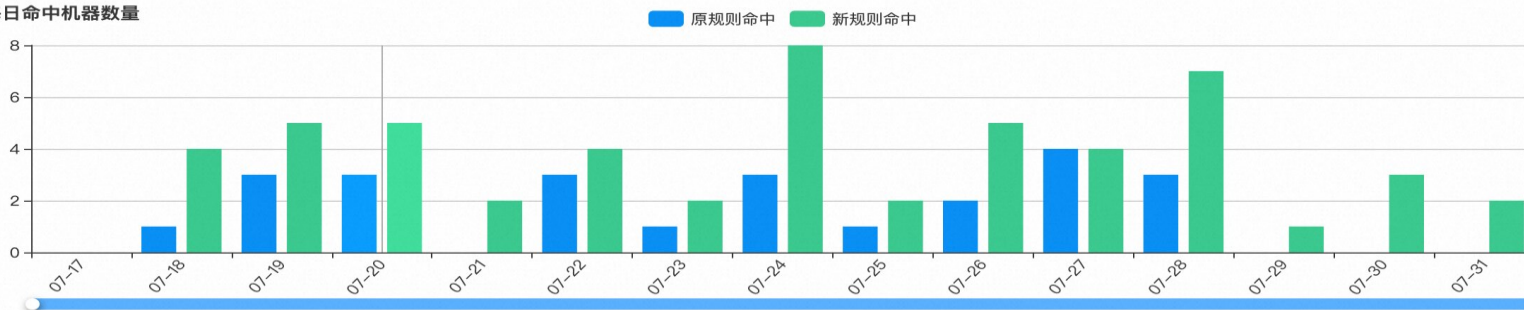
prefix	projected (suffix) database
$\langle a \rangle$	$\langle (abc)(ac)d(cf) \rangle, \langle (d)c(bc)(ae) \rangle, \langle (b)(df)cb \rangle, \langle (f)cbc \rangle$
$\langle b \rangle$	$\langle (c)(ac)d(cf) \rangle, \langle (c)(ae) \rangle, \langle (df)cb \rangle, \langle c \rangle$
$\langle c \rangle$	$\langle (ac)d(cf) \rangle, \langle (bc)(ae) \rangle, \langle b \rangle, \langle bc \rangle$
$\langle d \rangle$	$\langle (cf) \rangle, \langle c(bc)(ae) \rangle, \langle (f)cb \rangle$
$\langle e \rangle$	$\langle (f)(ab)(df)cb \rangle, \langle (af)cbc \rangle$
$\langle f \rangle$	$\langle (ab)(df)cb \rangle, \langle cbc \rangle$

智能运维决策-新老规则回溯

■ 新规则新增数量 ■ 两者都命中数量



每日命中机器数量



新规则漏报 0 新规则新增 30 原规则/新规则 均命中 24 旧规则命中数据 24 新规则命中数据 54

targetId	命中时间	所属集群	productName
10.10.10.17.80	2023-07-18 07:00:43	10.10.10.17	10.10.10.17:2_0
10.10.10.17.215	2023-07-18 14:52:43	10.10.10.17	10.10.10.17:2_0
10.10.10.17.142	2023-07-18 20:37:41	10.10.10.17	10.10.10.17:oc02
10.10.10.17.5	2023-07-19 03:07:46	10.10.10.17	10.10.10.17:2_0
10.10.10.17.168	2023-07-19 11:56:32	10.10.10.17	10.10.10.17:2_0
10.10.10.17.129	2023-07-20 14:16:48	10.10.10.17	10.10.10.17:2_0
10.10.10.17.2	2023-07-20 14:50:01	10.10.10.17	10.10.10.17:2_0
10.10.10.17.208	2023-07-21 01:40:03	10.10.10.17	10.10.10.17:2_5
10.10.10.17.213	2023-07-21 16:41:36	10.10.10.17	10.10.10.17:normal
10.10.10.17.9	2023-07-22 08:53:46	10.10.10.17	10.10.10.17:2_5

新规则漏报 0 新规则新增 30 原规则/新规则 均命中 24 旧规则命中数据 24 新规则命中数据 54

targetId	原命中时间	新命中时间	提前命中时间差(s)
10.10.10.17.30.231	2023-07-18 05:26:36	2023-07-18 05:26:20	15
10.10.10.17.4.240	2023-07-19 00:39:35	2023-07-19 00:39:35	0
10.10.10.17.8	2023-07-19 15:28:53	2023-07-19 15:26:11	161
10.10.10.17.98.77	2023-07-19 19:48:32	2023-07-19 19:44:27	245
10.10.10.17.9.3	2023-07-20 12:06:47	2023-07-20 12:06:47	0
10.10.10.17.9	2023-07-20 16:08:10	2023-07-20 16:06:20	110
10.10.10.17.97.94	2023-07-20 18:38:35	2023-07-20 18:12:13	1581
10.10.10.17.50.47	2023-07-22 04:48:20	2023-07-22 04:48:20	0
10.10.10.17.2.42	2023-07-22 08:11:31	2023-07-22 08:11:21	9
10.10.10.17.7.22	2023-07-22 09:04:03	2023-07-22 09:03:58	4

智能运维决策-运维评价体系

评价诉求

同一问题的不同运维策略哪个好?

实例运维后, 真的更“健康”了吗?

评价度量

不可用

$$KeyMetric_{down} = \frac{\sum_{i=1}^{i=n} F_e * Vm_i Down_T}{\sum_{i=1}^{i=n} Vm_i TotalLifeTime_T}$$

数据面有损

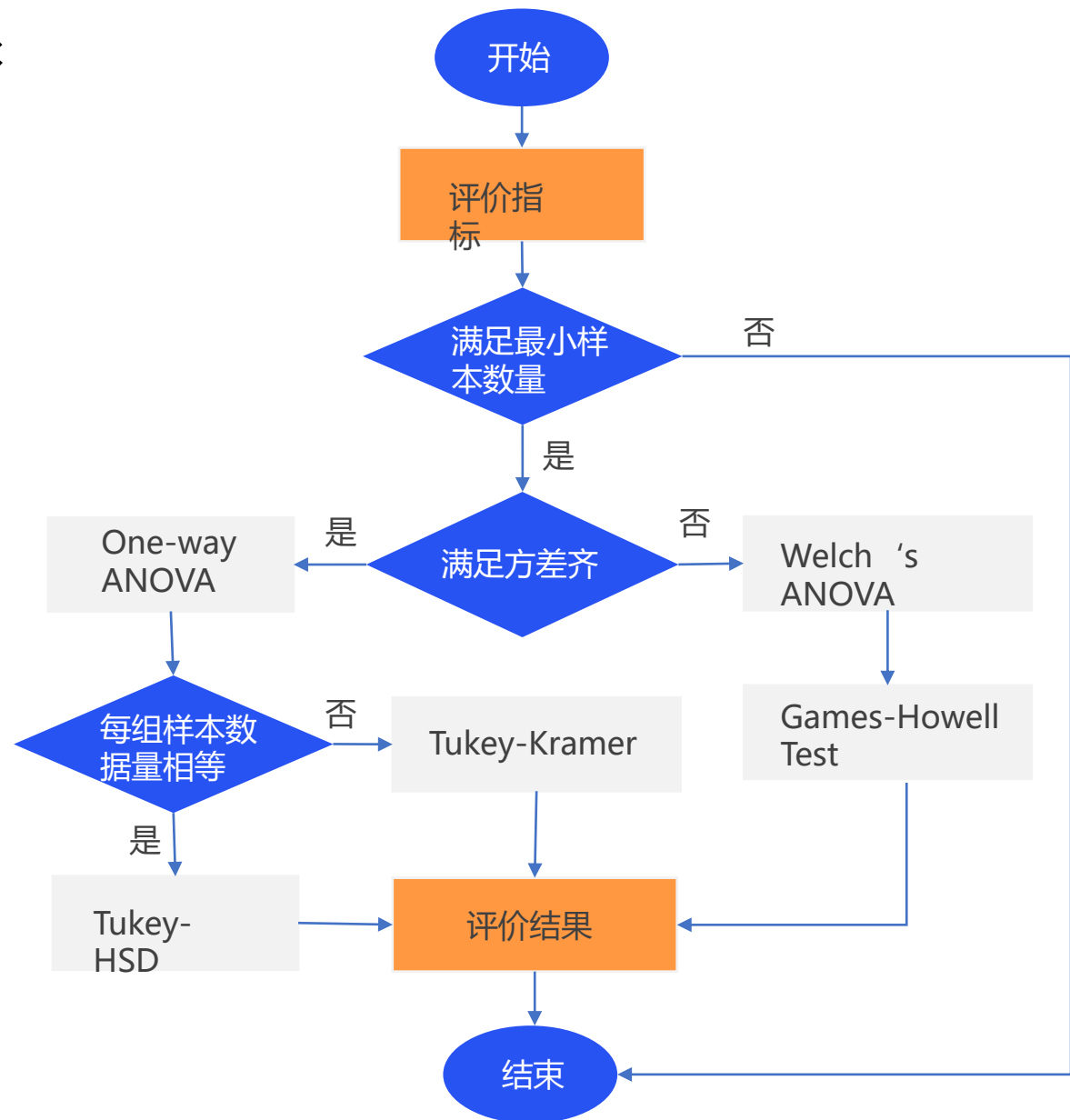
$$KeyMetric_{perf} = \frac{\sum_{l=1}^{l=n} F_l \sum_{i=1}^{i=n} F_e * Vm_i Perf_T}{\sum_{i=1}^{i=n} Vm_i TotalLifeTime_T}$$

控制面有损

$$KeyMetric_{ctrl} = \frac{\sum_{l=1}^{l=n} F_l \sum_{i=1}^{i=n} F_e * Vm_i Ctrl_T}{\sum_{i=1}^{i=n} Vm_i TotalLifeTime_T}$$

差异化分析

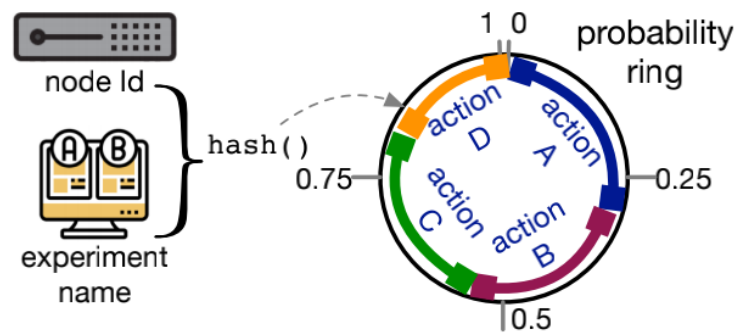
显著性差异检验



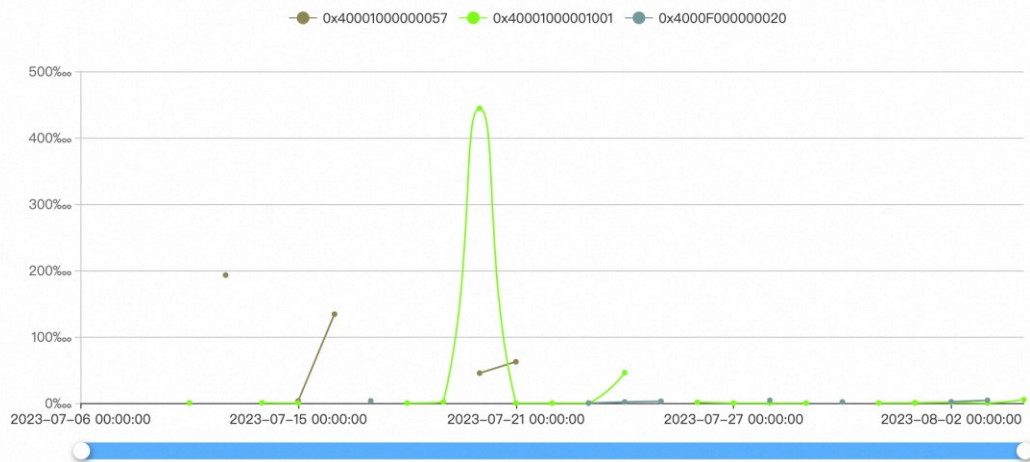
智能运维决策-A/B Test

* 运维规则名称:
 * 运维类型: 下线金刚压测(offlineAndPressTest)
 * 运维对象: NC
 * 是否开启: 是
 * 影响级别: 紧急
 * 运维动作: NC深度下线-金刚压测(0x40001000001001) × 下线观察长周期(0x40001000000057) ×
 运维动作扩展参数: 禁用 启用
 A/B Test: 启用

HashKey: resourceId 无条件的运维动作之间互斥, 只能用同一个HashKey, 请在此选择
 运维动作: NC深度下线-金刚压测(0x40001000001001) 运维比例(%) 65 HashKey: resourceId
 触发条件: 当前运维动作被执行的概率为(满足条件时): 65%
 运维动作: 下线观察长周期(0x40001000000057) 运维比例(%) 35 HashKey: resourceId
 触发条件: 当前运维动作被执行的概率为(满足条件时): 35%



不可用KeyMetric



较优action	比较对象	比较组均值差	比较组差异结果	差异结果
抢占盘迁移后清理	vm(0x4000F000000020)-NC深度下线-金刚压测(0x40001000001001)	-3.319e-3	无差异	
抢占盘迁移后清理	vm(0x4000F000000020)-下线观察长周期(0x40001000000057)	-2.043e-3	无差异	无差异
NC深度下线-金刚压测	(0x40001000001001)-下线观察长周期(0x40001000000057)	1.276e-3	无差异	

▶ 参考文献

- 【1】 Localizing Failure Root Causes in a Microservice through Causality Inference
- 【2】 Anomaly Detection in Time Series: A Comprehensive Evaluation
- 【3】 TcpRT: Instrument and Diagnostic Analysis System for Service Quality of Cloud Databases at Massive Scale in Real-time
- 【4】 RobustPeriod: Robust Time-Frequency Mining for MultiplePeriodicity Detection
- 【5】 RobustSTL: A Robust Seasonal-TrendDecomposition Algorithm for Long Time Series
- 【6】 System Log Parsing: A Survey
- 【7】 LogCluster - A Data Clustering and Pattern Mining Algorithm for Event Logs
- 【8】 SPINE: A Scalable Log Parser with Feedback Guidance
- 【9】 Mining Frequent Patterns without Candidate Generation: A Frequent-Pattern Tree Approach
- 【10】 Mining Sequential Patterns by Pattern-Growth: The PrefixSpan Approach
- 【11】 Predictive and adaptive failure mitigation to avert production cloud VM interruptions
- 【12】 Taking parametric assumptions seriously: Arguments for the use of Welch' s F-test instead of the classical F-test in one-way ANOVA
- 【13】 Relationship between Omnibus and Post-hoc Tests: An Investigation of performance of the F test in ANOVA

PART 04

感悟



▶ 感悟

- 1.AI入门没有想象的那么难
- 2.循序渐进，不好高骛远
- 3.给予结果充分的信任
- 4.做好安全防护

感谢聆听

