AIOps Challenge

基于全方位系统工程的微服务架构系统故障识别与分类

AeroSpaceX -- 陈东辉 谢志鹏 胡彬 张楠 高若姝 李铭 袁俊 张曦

华为技术有限公司

2022 CCF国际AIOps挑战赛决赛暨AIOps研讨会

AIOps Challenge

2022 CCF国际AIOps挑战赛决赛
暨AIOps研讨会

目录
CONTENTS

HUAWEI

第一章节

队伍介绍

队伍介绍

AIOps Challenge

华为2012-庞加莱实验室

庞加莱实验室承载构建国家数字基础设施根技术的使命，致力于将欧拉操作系统打造成极致性能、安全可信的开放操作系统平台。A-Tune团队致力于通过AI技术赋能、改造、颠覆操作系统，提供智能调优、智能运维、智能安全等能力。

华为2012-集成供应链实验室

集成供应链实验室面向泛供应链领域，致力于主导计划、物流、采购、生产等数字化项目的能力建设和业务落地，追踪和应用运筹优化、数据科学、启发式与人工智能等前沿算法和技术，提高供应链智能化水平，实现供应链能力创新。

华为集团IT

华为IT平台服务部-UniAI产品承载华为AI战略，专注实现企业场景AI，深耕销售、服务、供应、制造、财经等20+业务及颗粒化领域900+海量场景，基于“场景、算法、数据、算力”四位一体，建设企业AI解决方案及服务，联接开放生态，践行智能之道。

陈东辉 浙江大学 计算机博士

华为2012-庞加莱实验室

研发工程师

谢志鹏 浙江大学 计算机本科

华为2012-庞加莱实验室

研发工程师

胡彬 德克萨斯大学 计算机硕士

华为2012-庞加莱实验室

研发工程师

张楠 浙江大学 计算机硕士

华为2012-庞加莱实验室

研发工程师

李 铭 南京邮电大学 统计学硕士

华为2012-集成供应链实验室

机器学习运筹优化算法工程师

高若姝 石溪大学 计算机本科

华为2012-庞加莱实验室

研发工程师

袁 俊 湖南大学 计算机硕士

华为集团IT - UniAI

智能决策算法工程师

张 曦 犹他州立大学 统计学博士

华为集团IT - UniAI

智能决策算法科学家

HUAWEI

第二章节

方案介绍

赛题剖析：挑战难度升级--及时性、准确性缺一不可

AIOps Challenge

微服务架构电商系统下的故障识别与分类

挑战

实时故障识别

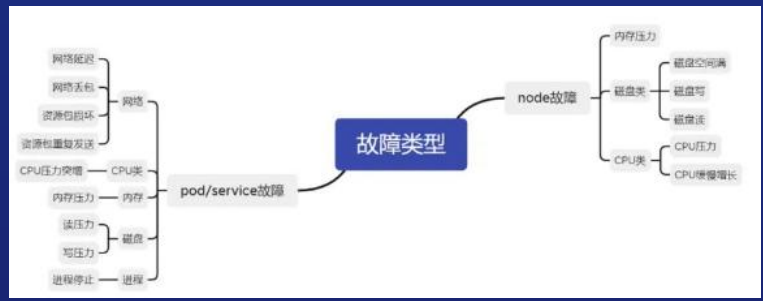
- 如何在海量多模态监控数据：log、trace、metric（400种）、kpi（4种），压缩后~800M/day中快速发现故障？

- 延迟分数$delay_{score}$：
  - $delay_{score}$ 是关于 $(t-T_l)$ 的单调递减函数
  - 约束到 $[\beta,1]$ 之间，β 暂定0.5

$$delay_{score}=\frac{10-(t-T_l)}{10}*0.5+0.5$$

准确故障分类

- 如何在微服务动态部署架构中准确定位三种层级的故障位置，并识别具体的故障类型（1/15)？



有限提交次数

- 如何在每次故障注入后保证提交次数<=k (3->2) 且要考虑其他非故障注入阶段的异常波动造成的误报？

选手提交预算

- 在整个故障检测过程中，选手一共有*k*N*次预算(k暂定为3)
- N为故障的个数，会提前告诉选手
- 如果选手提交答案的次数（合法的submit次数）超过了3*N，则按照时间顺序排序后，超过的部分会被评分系统忽略

应对原则

异常检测算法必须高效

1. 400种Metric指标须进行整合、重构：如识别关键指标、聚合相似指标，保证既高效识别异常，又不丢失关键信息
2. Log预处理、解析、特征提取必须快速
3. Trace有效利用时序和拓扑信息，快速判断问题根节点

故障定位和分类算法必须准确

1. 三类异常检测结果综合判断故障根因节点须结合从离线有标签异常样本中的故障知识
2. 三种层级的故障level定位也离不开从异常样本中学习经验知识
3. 需提取能准确mapping15种故障类型的异常特征

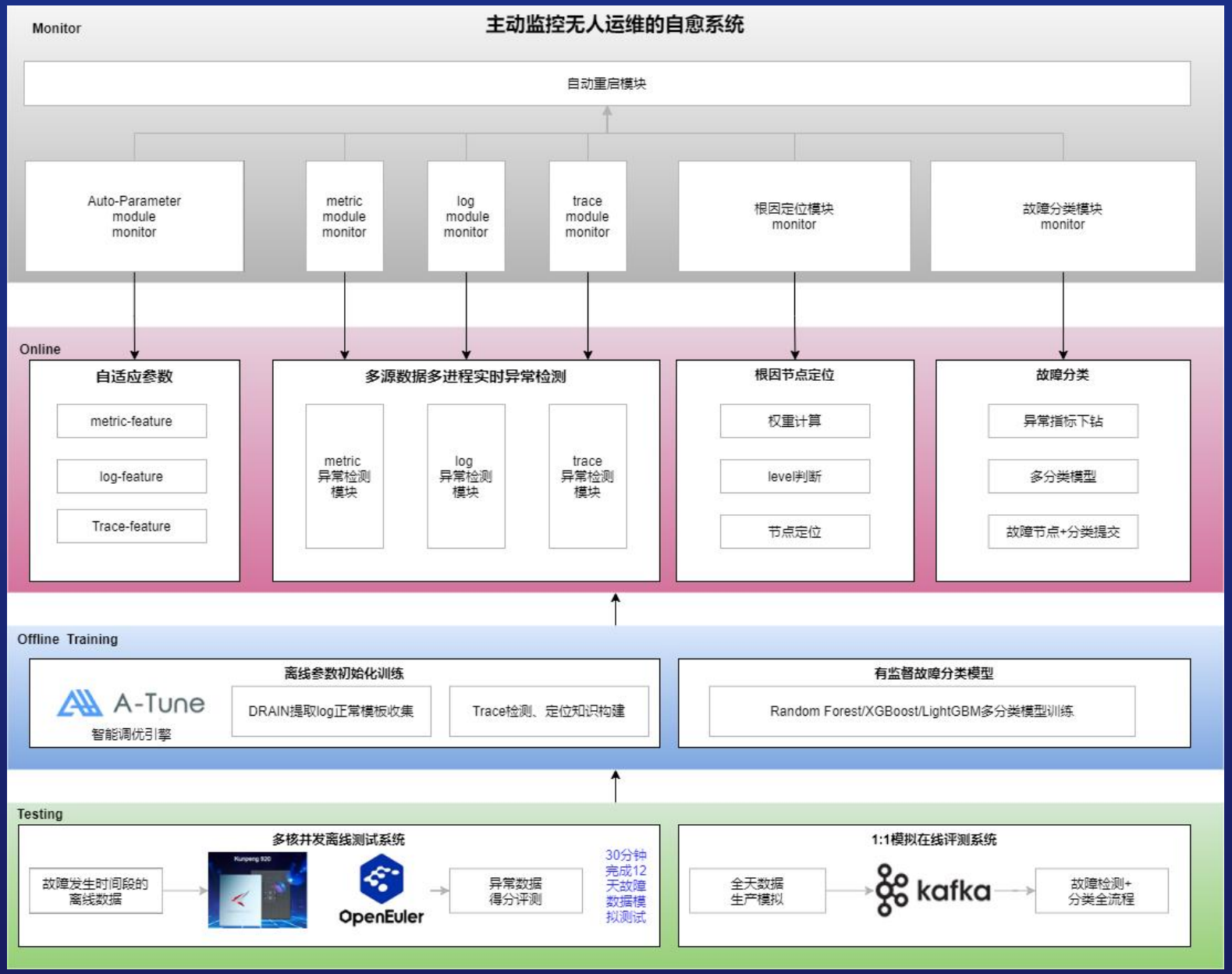
避免产生告警风暴

1. 准确区分系统在正常时段的异常波动和故障注入时段的异常波动
2. 综合考虑故障注入时的各指标的表现、异常程度、异常发生时间，珍惜和把握每一次提交机会

HUAWEI

整体方案：监控系统的监控和自愈

AIOps Challenge



★ 全方位的系统工程

- 整体方案分为监控模块、在线检测模块、离线训练、测试模块
- 监控系统的监控：对运维系统自身进行组件级监控，任何组件崩溃后10s内自动拉起恢复；组件间进程互相解耦
- 离线+在线性能、算法效果测试：基于鲲鹏多核芯片和高性能openEuler系统实现离线并发测试，30分钟完成12天故障时段数据测试；1：1模拟线上测评系统，保障运维系统稳定

★ 基于A-Tune调参的高效异常检测算法

- 充分学习历史故障数据，设置奖惩策略，实现数据类型级、实例级的参数调优，减少误报和漏报
- metric、log、trace独立进程高效全面捕获异常

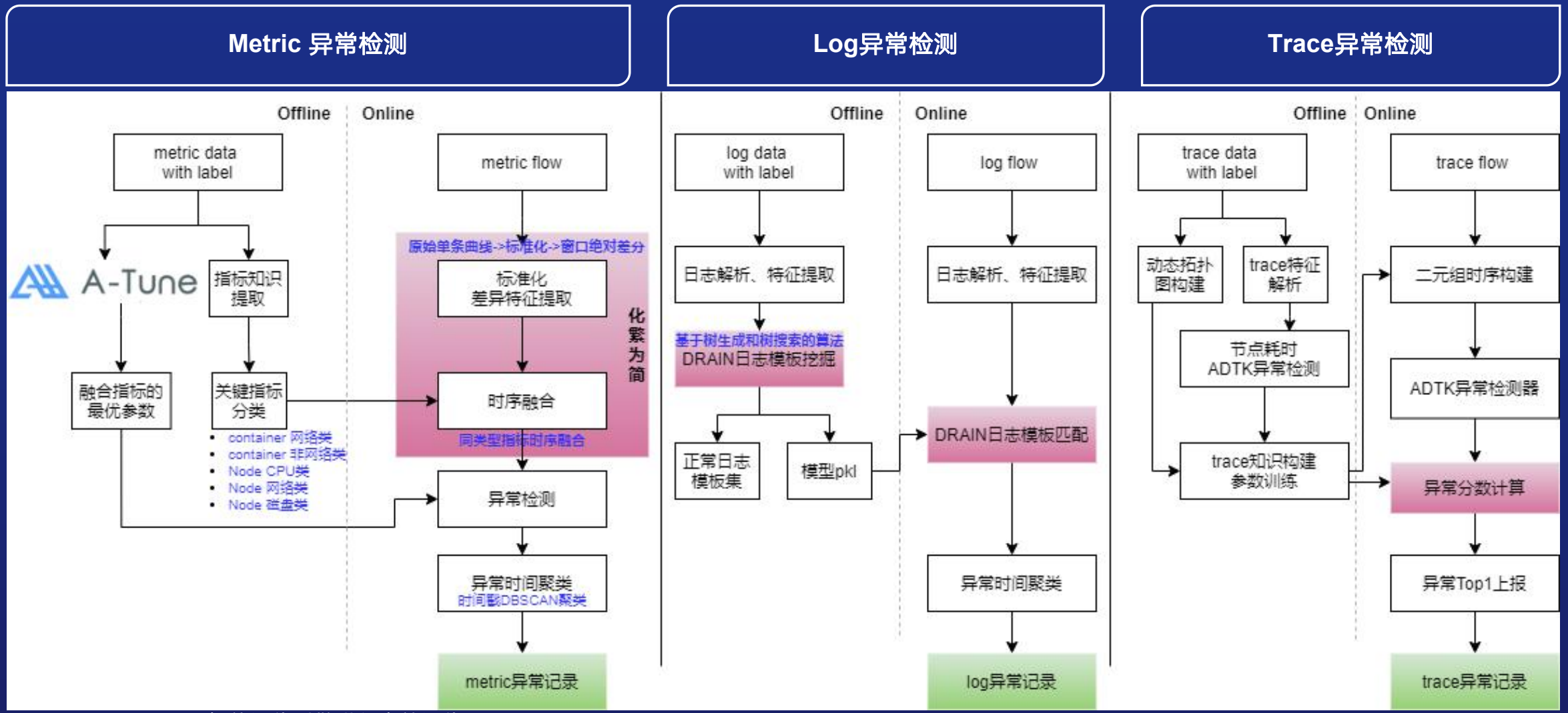
★ 基于时间聚类和有监督模型的故障分类算法

- 利用DBSCAN对多源数据异常检测的结果进行聚类定位根因节点
- 利用历史故障事件信息，提取有效故障模式特征，训练有监督故障训练模型，测试分类f准确率分别为90%+/80%+

HUAWEI

多源、异构数据多进程实时异常检测

AIOps Challenge



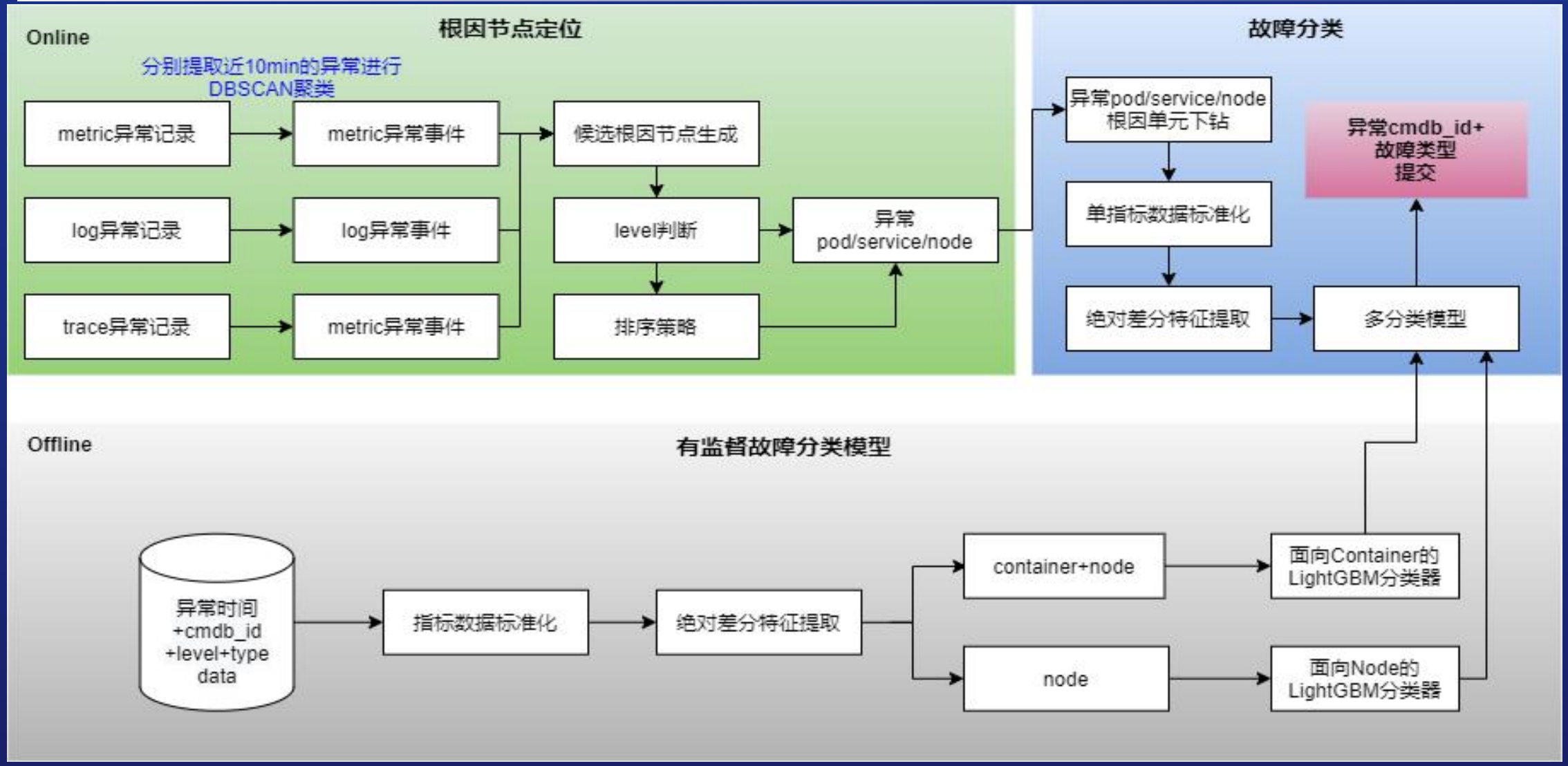
* openEuler A-Tune智能调优引擎进行高效调优 https://gitee.com/openeuler/A-Tune

* 树生成：将日志组存储入相同深度的解析树　树搜索：通过长度、开始token和token相似度进行搜索，如合适则更新解析树；如无匹配则新建日志组。Pinjia He, Jieming Zhu, Zibin Zheng, and Michael R. Lyu. Drain: An Online Log Parsing Approach with Fixed Depth Tree, Proceedings of the 24th International Conference on Web Services (ICWS), 2017.

HUAWEI

根因节点定位及故障分类

AIOps Challenge



- ✓ 将一个时间段的异常记录聚类成异常事件，有效避免故障跨节点传播导致的误报
- ✓ log、trace、metric三种类型异常事件综合决策，有效提升service级别的定位准确率
- ✓ 依据已知异常标签类型，提取历史事件时刻的关键指标、关键特征，面向container和node两种类型分别构建故障多分类模型

HUAWEI

方案创新性和通用性

AIOps Challenge

- **时序融合算法对多维时间序列指标进行重构，实现海量指标序列快速检测，precision：90%+**
  - 对400+metric指标，聚类后，先标准化，再提取绝对差分特征，利用时序特征融合曲线对该类指标进行重新表示
  - 该方法在异常检测的多个项目中得到应用，2021年AIOPS挑战赛初赛第一
  - 该方法还可以应用于资源利用率指标的融合和预测，在华为多个资源预测项目中应用效果良好
- **应用DBSCAN对时间戳进行聚类，有效避免告警风暴**
  - 故障注入发生后的异常发生时间具有集中性特征，利用DBSCAN基于密度聚类的特征，准确识别同次故障
  - 将一个时间段的异常记录聚类成异常事件，有效避免故障跨节点传播导致的误报
  - 该方法已在两次AIOPS挑战赛中应用，效果显著
- **历史故障事件提取时间特征、异常指标特征、异常程度特征等进行有监督故障多分类，测试准确率90%+/80%+**
  - 故障根因具备“二八定律”，多模态运维数据中挖掘历史故障指纹模式用于多分类算法
  - 集成学习融合多种分类算法：Random Forest /XGBoost/LightGBM 针对每种故障类型选择最优分类模型
- **全方位的系统工程：针对封闭系统测评，对监控系统进行监控，并设置自动重启机制，实现无人运维的自愈系统**
  - 封闭测评期间，对组件运行状态进行实时监控，并设定故障后自动重启机制，确保运维系统稳定
  - 1:1模拟在线测评系统，对生产系统进行全方位学习
  - 该方法借鉴了操作系统监控告警机制，具备对任意监控系统适用的通用性

HUAWEI

算法效果评价和应用探讨

AIOps Challenge

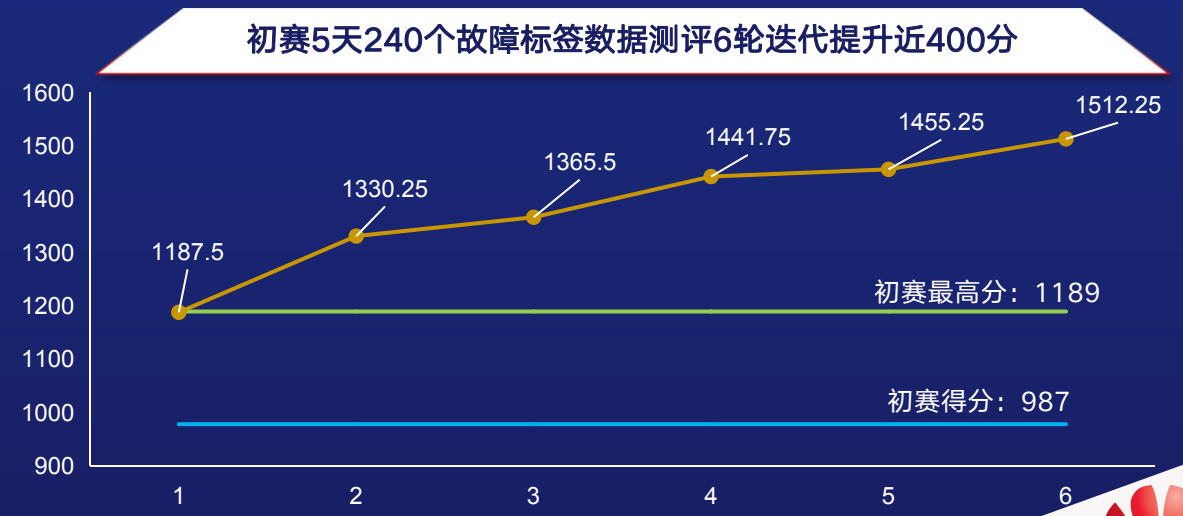
多维度测评系统

算法得分效果(正确、cmdb_id正确、故障类型正确)，提交总次数，提交时间、故障时间、log/metric/trace相应的异常检测结果，供调优使用

```
[root@linux-DGEgvZ metric]# sh check.sh
故障总数:300
全部正确:207
部分正确:38, cmdb_id正确:19,  故障类型正确:19
上报告警:348

未检测到node故障:23
未检测到pod故障:21
未检测到service故障:11
groundtruth-k8s-1-2022-03-20.csv
[score] score cases 29
[{'teamId': 'AeroSpaceX', 'score': 191.5, 'submitNum': 35}]
groundtruth-k8s-2-2022-03-20.csv
[score] score cases 33
[{'teamId': 'AeroSpaceX', 'score': 270.75, 'submitNum': 33}]
groundtruth-k8s-3-2022-03-20.csv
[score] score cases 27
[{'teamId': 'AeroSpaceX', 'score': 197.25, 'submitNum': 31}]
groundtruth-k8s-1-2022-03-21.csv
[score] score cases 35
```

提交时间	分析	故障时间	故障节点	故障类型	提交节点	提交类型	延迟	延迟打折	提交时间戳	cloudbed	故障上报时metric告警信息	故障上报时trace告警信
3/21/2022 5:37	误报	3/21/2022 5:36	shippingservice2-0	k8s容器网络资源包损坏	checkoutservice-1	k8s容器进程中止	60	0.95	1647812223	cloudbed1	cmdb_id alarm_start_time continuous_interval status count has_it_alarmed first_time last_time first_timestamp last_timestamp type 1 checkoutservice-1 "2022-03-21 05:27:00" 2 close 2 no "2022-03-21 05:27:00" "2022-03-21 05:28:00" 1647811620 1647811680 notnet_container	timestamp datestr cmdb_id
3/21/2022 5:40	误报	3/21/2022 5:36	shippingservice2-0	k8s容器网络资源包损坏	checkoutservice2-0	k8s容器网络延迟	240	0.8	1647812403	cloudbed1	cmdb_id alarm_start_time continuous_interval status count has_it_alarmed first_time last_time first_timestamp last_timestamp type 1 checkoutservice-1 "2022-03-21 05:27:00" 8 close 2 no "2022-03-21 05:27:00" "2022-03-21 05:28:00" 1647811620 1647811680 notnet_container 2 checkoutservice2-0 "2022-03-21 05:39:00" 0 open 2 no "2022-03-21 05:39:00" "2022-03-21 05:40:00" 1647812340 1647812400 notnet_container	timestamp datestr cmdb_id



HUAWEI

第三章节

总结展望

加强领域先验知识学习，算法模型自迭代优化

AIOps Challenge

挑战

实时故障识别

- 如何在海量多模态监控数据：log、trace、metric（400种）、kpi（4种），压缩后~800M/day中快速发现故障？

准确故障分类

- 如何在微服务动态部署架构中准确定位三种层级的故障位置，并识别具体的故障类型（1/15)？

有限提交次数

- 如何在每次故障注入后保证提交次数<=k (3->2) 且要考虑其他非故障注入阶段的异常波动造成的误报？

应对方案

一套参数应用于所有cmdb_id，造成node型故障无法识别 ✗

- ✓ 多维时间序列融合算法+N-sigma高效检测算法
- ✓ ADTK针对Trace的时间序列检测算法
- ✓ DRAIN高效识别日志模板

只学习异常指标的指纹模式，丢失正常指标在故障注入时的信息，分类准确率60% ✗

- ✓ **学习历史故障模式，识别对故障分类有意义的关键特征**
- ✓ **学习三种层次的故障特征表现，沉淀领域知识**

基于其他聚类方法如k-means、mean-shift等结果不稳定 ✗

- ✓ 基于故障注入后异常发生的集中表现，利用DBSCAN对时间戳进行聚类
- ✓ 结合领域知识，制定重新提交策略

待改进点

1. 从多模态数据中提取更多专业领域知识用于多维指标分类、日志模板学习、Trace拓扑构建、故障分类识别等
2. 算法模型的自动化调优和自动化增量训练：如日志新模式识别，自动检测出日志新模式，并更新正常模板集，降低误报
3. 故障分类特征提取：加入日志解析特征，利用NLP等相关技术对待标签的训练日志数据进行解析形成特征，构建日志异常检测分类器，和作为故障分类特征输入

HUAWEI

Welcome to join us!

AIOps Challenge

华为2012 庞加莱实验室

庞加莱实验室承载构建国家数字基础设施根技术的使命，致力于将欧拉操作系统打造成极致性能、安全可信的开放操作系统平台。A-Tune团队致力于通过AI技术赋能、改造、颠覆操作系统，提供智能调优、智能运维、智能安全等能力。

华为集团IT UniAI

华为IT平台服务部-UniAI产品承载华为AI战略，专注实现企业场景AI，深耕销售、服务、供应、制造、财经等20+业务及颗粒化领域900+海量场景，基于“场景、算法、数据、算力”四位一体，建设企业AI解决方案及服务，联接开放生态，践行智能之道。

HUAWEI

AIOps Challenge

2022 CCF国际AIOps挑战赛决赛暨AIOps研讨会

THANKS