



2021 国际AIops挑战赛决赛
暨AIops创新高峰论坛

云环境下商业银行应用系统的故障实时 检测与根因定位

联想研究院企业服务云计算研究室智能运维团队 LR-AIops

报告人：高绍阳 2021-5-13



第一届国际互联网产业科技创新大会暨互联网创新产品展览会
The First International Internet Industry Science And Technology Innovation Conference & Internet Innovation Product Exhibition

目 录

01

团队介绍

02

挑战和应对方案

03

总体框架

04

算法效果



联想研究院 智能运维团队

队伍名称：LR-AIops

队员：高绍阳 张闯

复赛排名：1

联想研究院官网

<http://research.lenovo.com/webapp/view/index.html>

目 录

01

团队介绍

02

挑战和应对方案

03

总体框架

04

算法效果

挑战1 要求快速的发现故障

难点： 1. 实时的流式数据； 2. 数据密集

措施1： 选择计算量少的算法

例如， z-score， 核密度估计， 卡方统计等

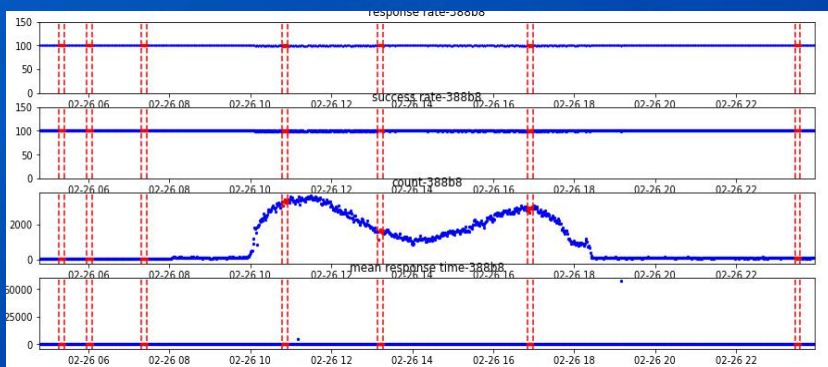
措施2： 多进程并行计算方法

每个进程负责接受特定的数据， 并做及时的处理

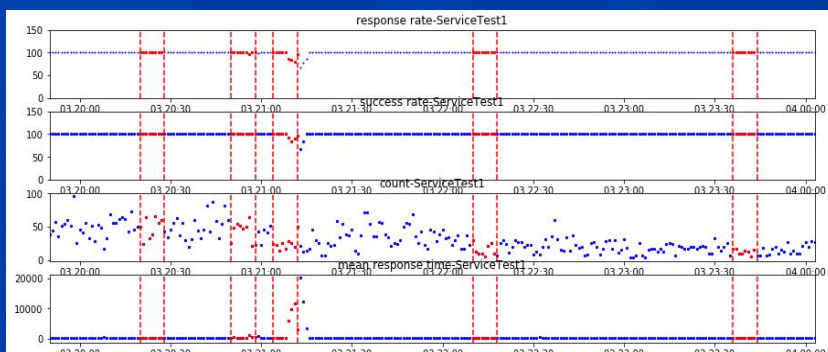
挑战2 要求准确的定位根因

数据特点：1. 故障不会刻画到所有的数据中，即有些故障与黄金指标和调用链数据不具有强相关性；2. 根因故障不是独立出现的，还伴随这其它非根因节点和指标的异常。

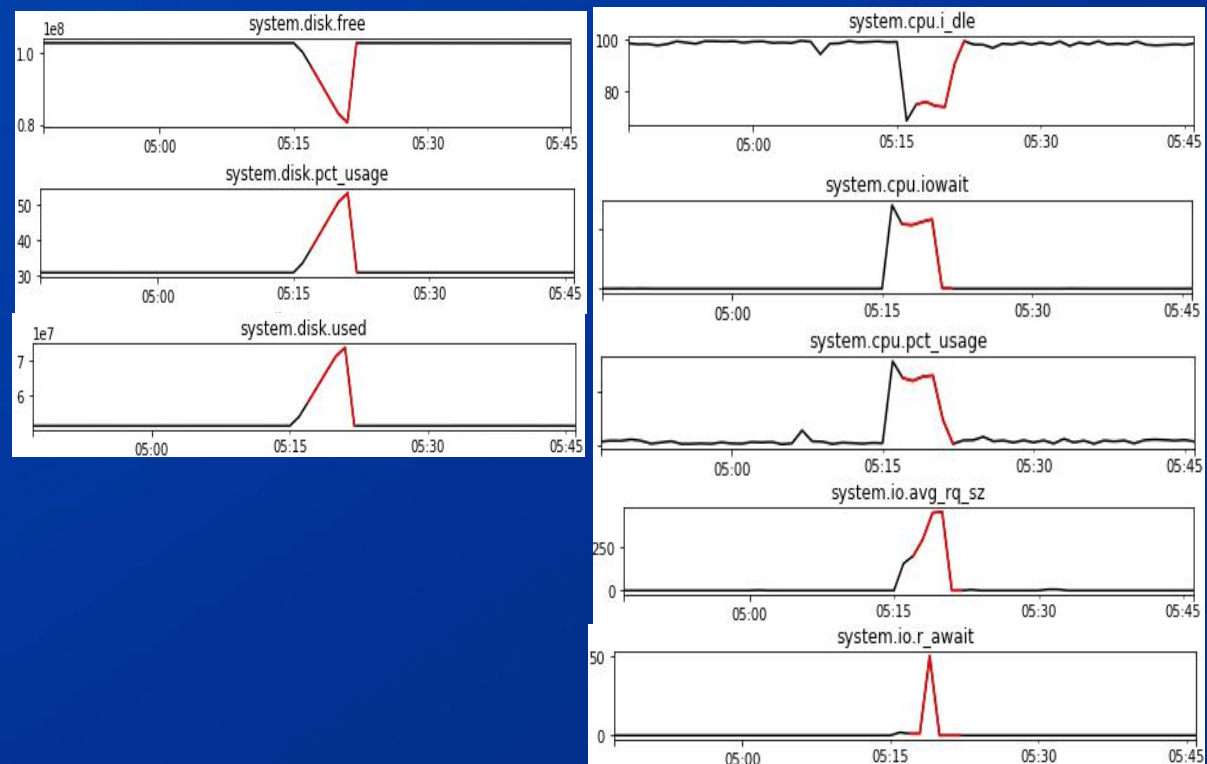
系统A部分标签故障



系统B部分标签故障



系统A， gjjcore8磁盘空间使用率过高， 还伴随着cpu指标异常， io指标异常



措施1 提取已知故障特征，建立领域知识库，根据知识库判断根因

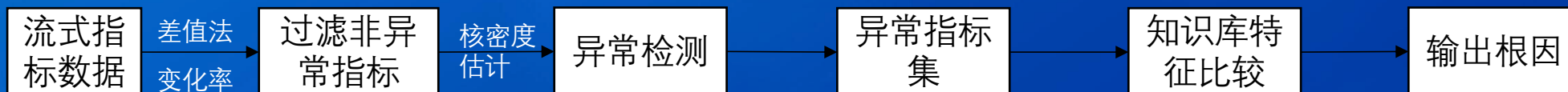
已知故障特征提取：

1. 故障节点异常指标的类型。类型可以根据指标名称判断，例如，disk，cpu，io等
2. 每个类型异常指标的数量
3. 异常指标是单点突变还是连续突变

例如：已知的某一次资源故障，磁盘空间使用率过高，特征内容如下：

1. Disk，cpu，load等指标类型异常
2. 指标与disk、cpu、load等有关的指标数量
3. 除了某个指标是单点异常，其它的都是连续异常

已知故障根因定位流程：



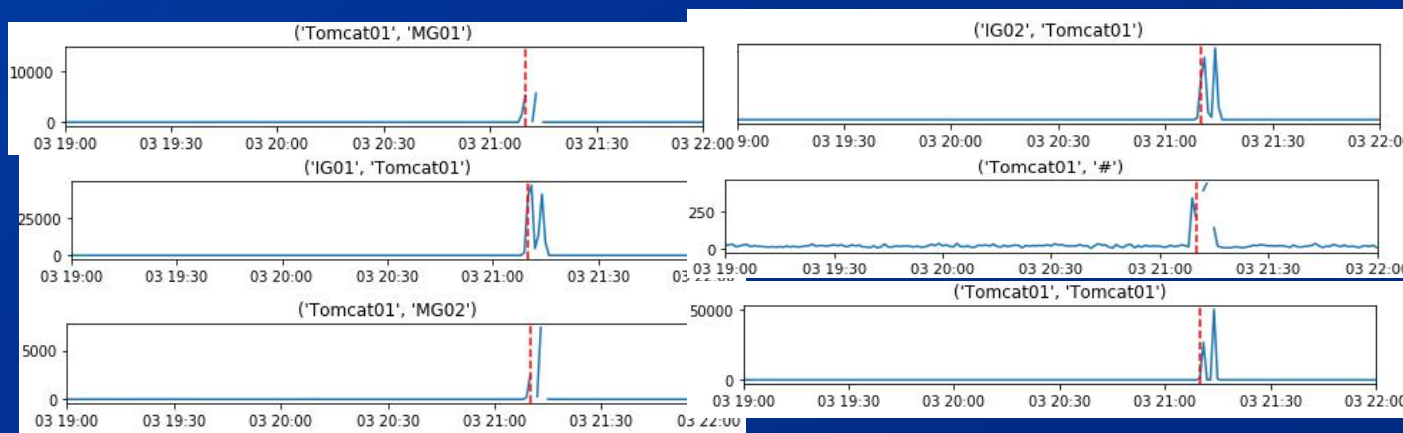
措施1优点	措施1不足
对于已知故障，检测的准确率特别高	不能检测未知故障

措施2 应用异常分数集成的方法定位根因

主要思想：通过集成trace异常检测得分，log异常检测得分，和metric异常检测得分，根据节点得分排名，定位根因

Trace检测：

1. 把调用链的时延转化成为时序数据。即以父子节点对作为时序指标，父节点减去子节点的延时做为时序值
2. 根据时序数据定位故障节点



Log检测:

1. 检测模板是否符合模板序列
2. 提取日志中的一些参数信息, 转化为时序数据, 做一个时序异常检测

Metric指标检测:

卡方统计量异常打分。对于时序的数据 $\{x_k\}_{k=m-p+1}^m$, 我们计算 x_m 的异常得分为

$$soe = \frac{(y_m - \bar{y})^2}{S \cdot \chi_{0.005}^2}$$

其中 $y_m = x_m - x_{m-1}$, \bar{y} 和 S 是 $\{y_k\}$ 的样本均值和方差。

未知根因定位:

Trace异常检测
异常节点得分等于各个异常指标的均值

$W_{t\&e}$

Log异常检测
异常节点得分等于各个异常日志得分之和

W_g

Metric异常检测
异常节点得分等于得分最高的异常指标组

$W_{m\&e\&c}$

$$soe = W_{t\&e} \cdot S_{t\&e} + W_g \cdot S_{lg} + W_{m\&e\&c} \cdot S_{m\&e\&c}$$

在某一个异常时间段，得分最高的节点，对应的日志名和指标名就是根因。如果节点在trace数据或log数据或metric数据没有检测到异常，则相应的得分为0

解决方案效果

效果	原因
适应两个系统	异常分数集成的方法判断根因，我们不需要考虑系统的类型
根因定位准确率高	1. 对于已知故障，采用领域知识库定位，具有极高的置信度；2. 对于未知故障，根据异常得分最高定位根因。 总的来说，推断出的根因准确率很高
故障发现速度快	我们采用多进程的方法分析数据，提高了推理的速度。并且，采用的算法计算量少

目 录

01

团队介绍

02

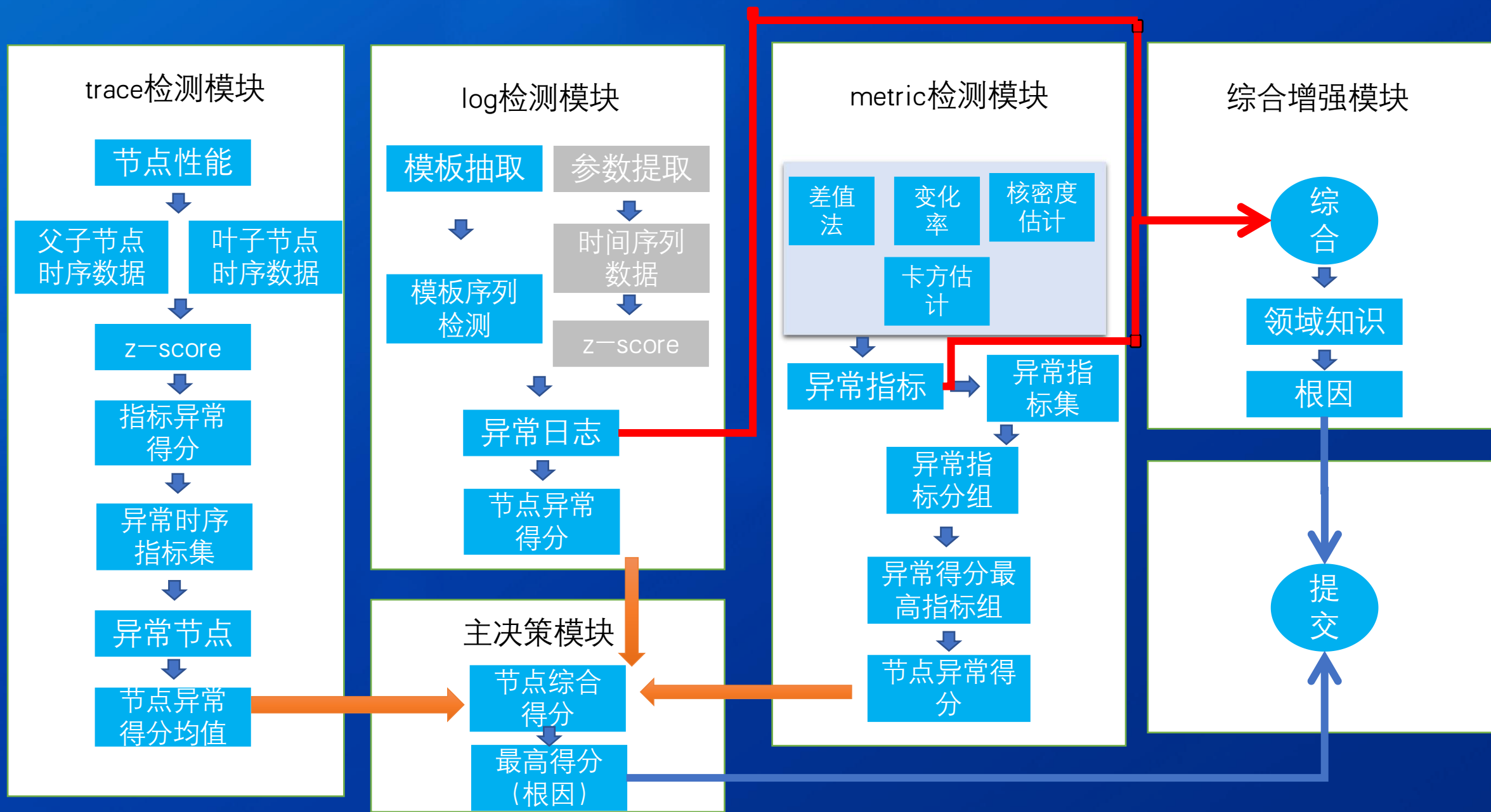
挑战和应对方案

03

总体框架

04

算法效果





01

团队介绍

02

挑战和应对方案

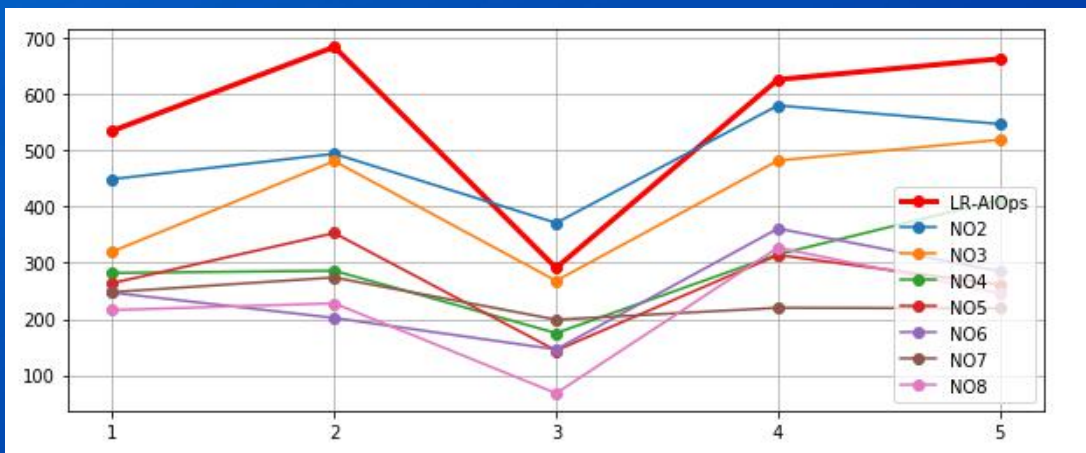
03

总体框架

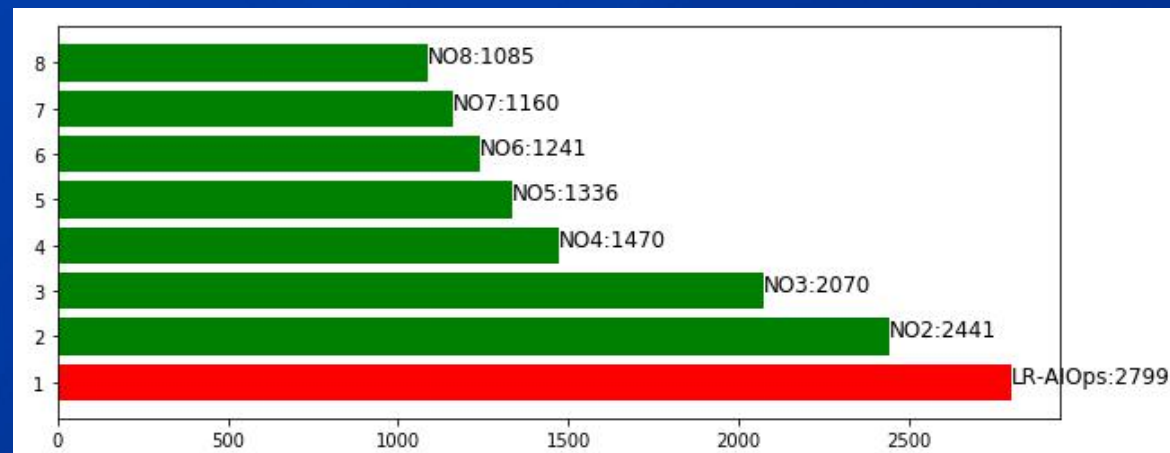
04

算法效果

复赛每天的评测成绩



总成绩



改进地方	方法
指标检测	计算各指标分布，根据分布选择算法
日志检测	1. 完善日志时序检测方法；2. 通过对日志语意的理解，对日志内容做上下文的向量转换，通过NLP算法做故障的预测
根因定位	通过因果算法，判断异常指标或节点的因果关系，确定最终根因



2021 国际AIOps挑战赛决赛暨AIOps创新高峰论坛

THANKS

谢谢观看



有问题可私下讨论



第一届国际互联网产业科技创新大会暨互联网创新产品展览会
The First International Internet Industry Science And Technology Innovation Conference & Internet Innovation Product Exhibition