

故障定位智能体的建设挑战和实践经验

向阳 云杉网络

主办单位：中国计算机学会（CCF）

承办单位：中国计算机学会互联网专委会、中国科学院计算机网络信息中心、中国移动研究院、清华大学

协办单位：华为2012实验室、阿里云、中兴通讯、中国移动九天团队、南开大学、西安电子科技大学、清华大学计算机科学与技术系、神州灵云

目录 CONTENTS

AIOps 故障定位的挑战和困境

解决全栈数据采集和关联的挑战

解决告警收敛和责任关联的挑战

解决持续学习和量身定制的挑战

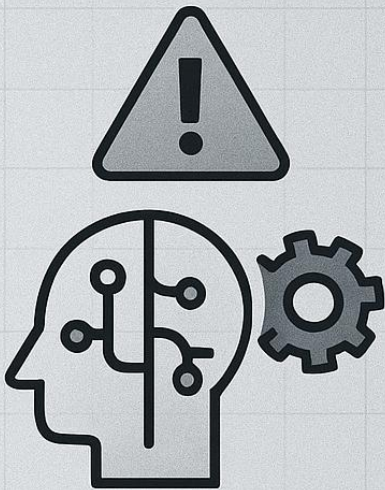
DeepFlow 智能体的探索和经验

传统的 AIops 产品遇到的两大难题



可观测性数据
质量不高

(观测盲点
观测数据孤岛)



问题分析的
智能不足

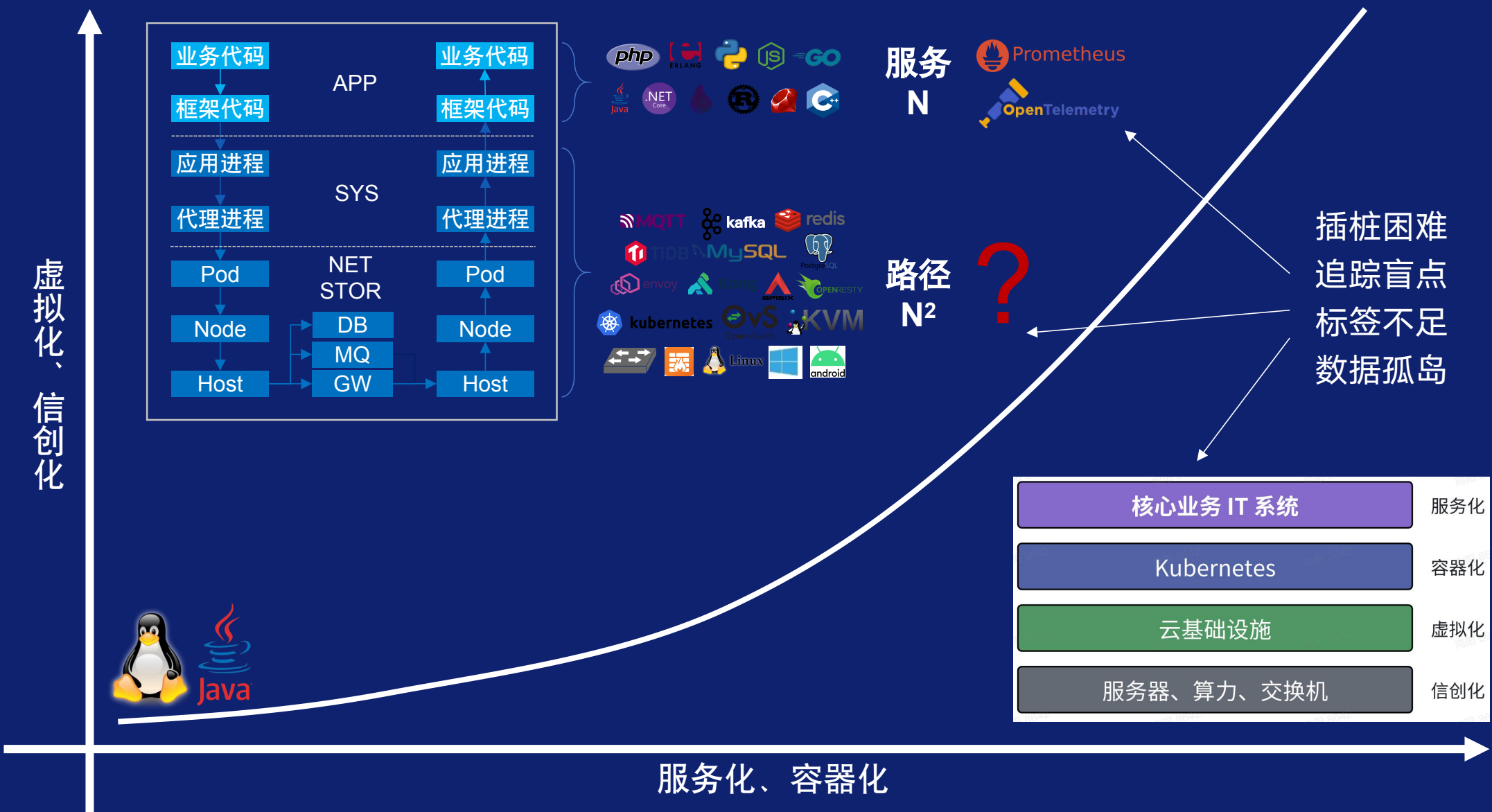
(依靠机器学习、小模型)
和基于规则的专家经验)
缺乏智能

OpenRCA

Leaderboard

Method Name	Model	Org.	Correct ↓	Partial	Date
RCA-Agent	Claude 3.5 Sonnet		11.34%	17.31%	2025/1/23
RCA-Agent	GPT-4o		8.96%	17.91%	2025/1/23
Prompting (Oracle)	Gemini 1.5 Pro		7.16%	23.58%	2025/1/23
Prompting (Balanced)	Gemini 1.5 Pro		6.27%	24.18%	2025/1/23
Prompting (Oracle)	GPT-4o		6.27%	15.82%	2025/1/23

挑战 #1：缺失全栈观测数据



挑战 #2: 对算力的需求

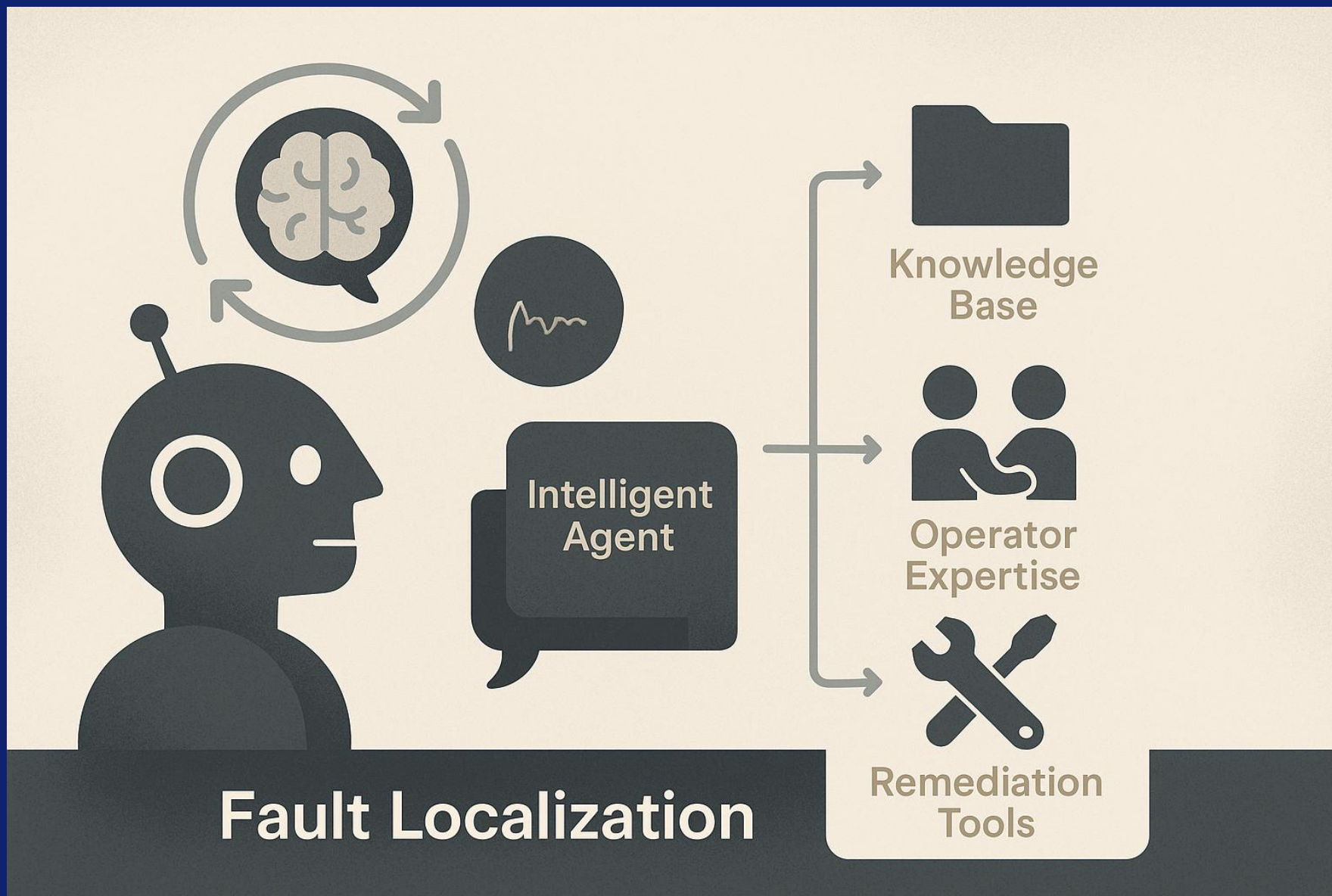
海量观测数据 · 上下文长度



基础设施故障 · 告警风暴

挑战 #3: 为企业提供服务

私有化部署 · 持续学习

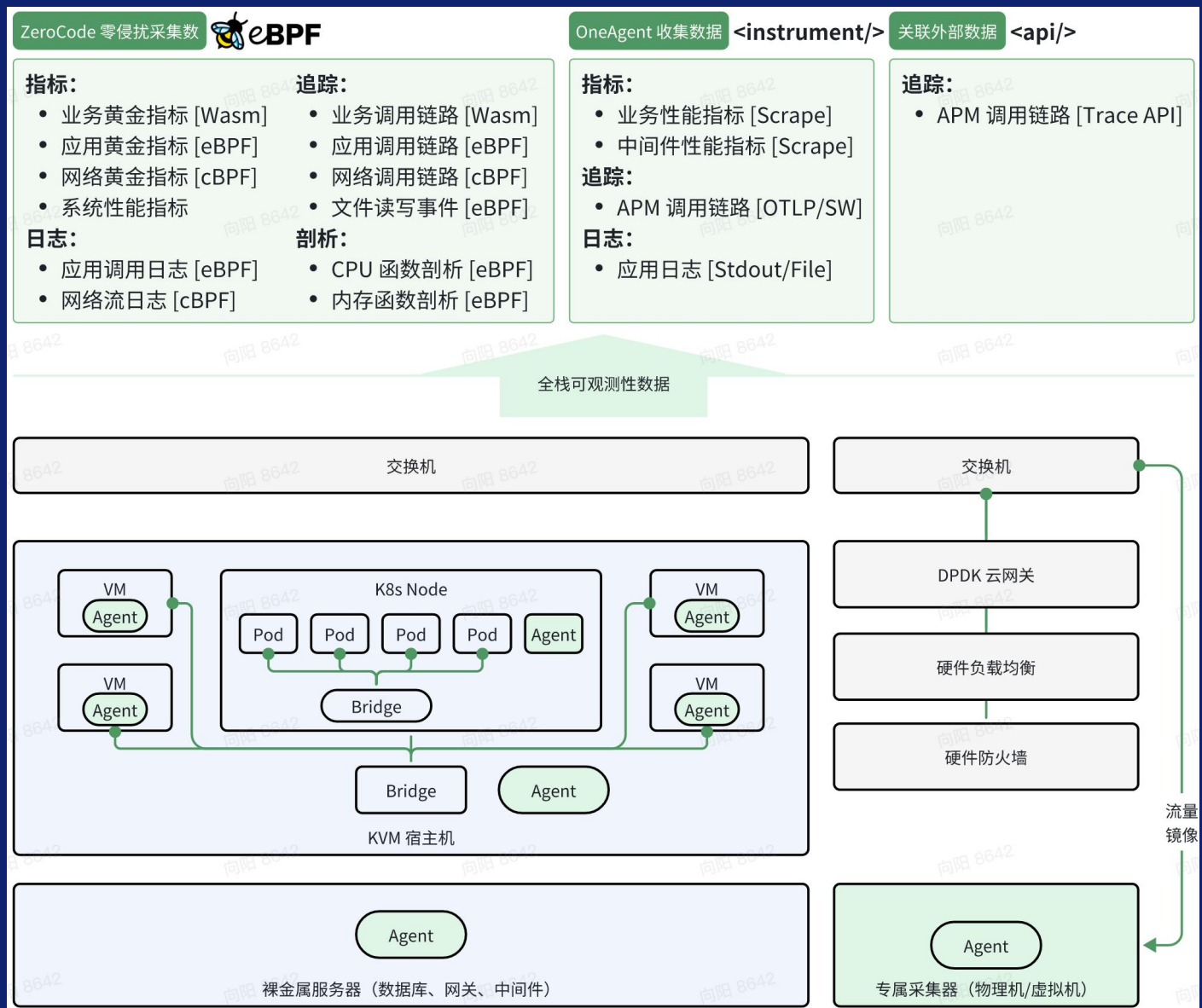


私有数据 · 为企业量身定制

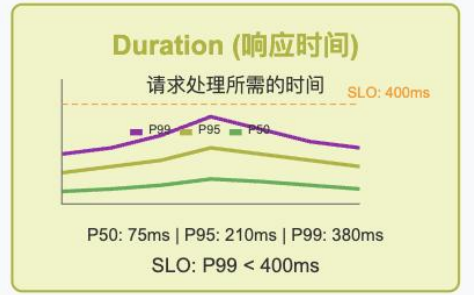
解决数据的挑战：全栈观测数据的采集和统一标注



- 零侵扰：业务无需插码、无需重启
- 全栈追踪：业务、应用、系统、网络
- 智能标注：云资源、容器服务、CMDB



解决算力的挑战：使用多模型的感知、以业务为中心



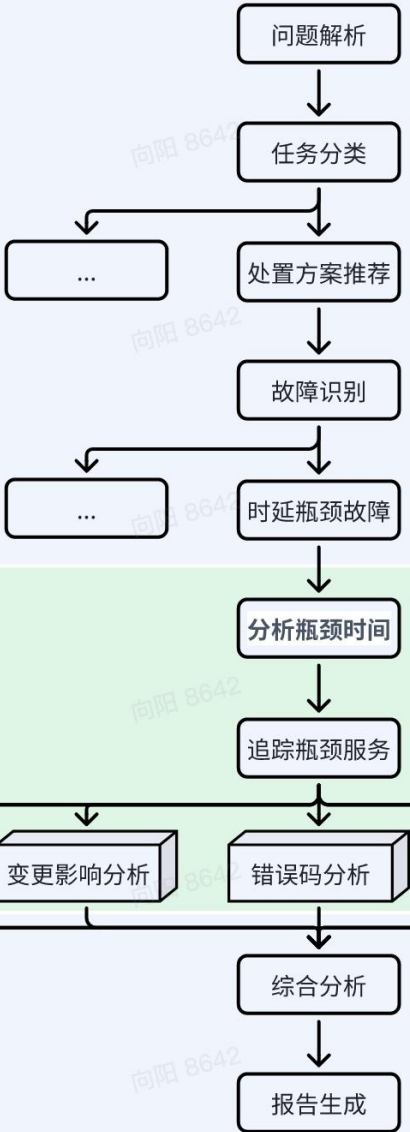
DeepFlow 诊断智能体： workflow 自动规划、智能推理

AIOps

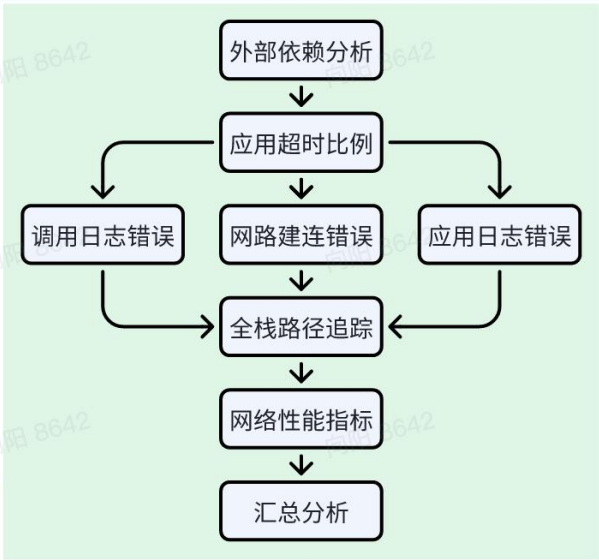
2025 CCF 国际 AIOps 挑战赛
25 CCF International AIOps Challenge



Planning / 规划



Reasoning / 推理



- 基于先验知识自动规划
- 基于实时感知动态推理
- 支持深度定制

DeepFlow 诊断智能体：全栈观测数据的感知、推理

- 感知全栈可观测性数据
- 追踪、指标、剖析、日志、...

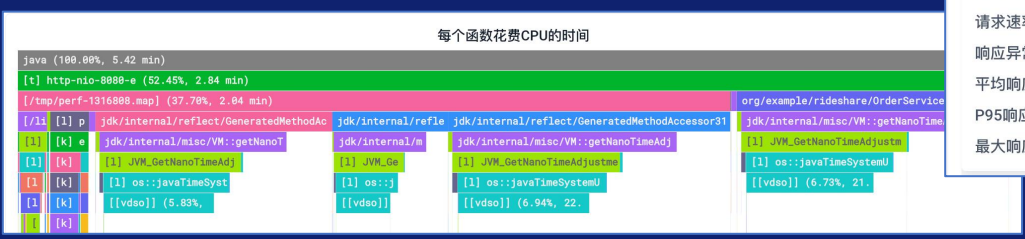


分析全栈路径瓶颈位置 pod_group=svc-order -> pod_service=db-demo · 发现问题

svc-order 访问 db-demo, 全栈路径时延瓶颈在客户端进程到客户端网卡之间 (46.13ms/46.44ms, 99.34%), 绝大部分时延集中在svc-order实例的内核协议栈或IO流程。

客户端 (svc-order) 访问服务端 (db-demo) 的应用请求总时延为 46.44ms, 全栈路径及时延瓶颈信息如下:

- 全栈路径: 客户端进程 -> 客户端网卡 -> 服务端网卡 -> 服务端进程
- 瓶颈路径: 在 客户端进程 到 客户端网卡 之间
- 瓶颈时延: 46.13ms (占总时延的比例为 99.34%)



分析业务拓扑的时延瓶颈 · 发现问题

发现瓶颈服务为pod_service=web-shop, 平均响应时延473.17ms, 自身消耗424.58ms (占比89.73%), 需关注该服务及其上下游依赖。

瓶颈服务: pod_service=web-shop

- 平均响应时延 473.17ms, 服务自身消耗的平均时延 424.58ms (占比 89.73%)
- 需要关注的入向路径: pod_service=ingress-nginx-controller-admission -> pod_service=web-shop;
- 需要关注的出向路径: pod_service=web-shop -> pod_service=svc-user; pod_service=web-shop -> pod_group=svc-item; pod_service=web-shop -> pod_group=svc-order; pod_service=web-shop -> pod_service=kube-dns;



DeepFlow 诊断智能体：处置方案的推理、学习、预测

2025 CCF国际AIOps挑战赛
25 CCF International AIOps Challenge



处置方案推荐

关键信息提取

方案推荐

1.APP-LIMIT-002-启用API限流策略（可能性：80%）

- 匹配理由：电商系统/shop/full-test API在09:00-09:10响应慢，该预案通过限流应对请求量突增，缓解API时延瓶颈。

2.SYS-PARAM-001-调整数据库连接参数（可能性：82%）

- 匹配理由：数据库连接数耗尽导致响应慢，该方案直接提升最大连接数，解决连接耗尽问题。

3.CONT-ROLL-003-调整数据库连接池配置（可能性：81%）

- 匹配理由：/shop/full-test 响应慢由连接池配置与并发负载不匹配引起，优化连接池配置可提升并发承载力。

- 3 分钟从告警->根因->修复
- 从感知到决策全过程可回溯

展开

根因分析报告

电商系统 /shop/full-test 在 02:00-02:10 时段响应慢的根因是 svc-order 服务对下游 db-demo 数据库依赖时，数据库连接数耗尽（MySQL 1040 错误），导致大量请求无法及时获得数据库连接，进而引发整体响应时延显著升高。svc-order 服务本身资源与并发处理能力充足，无资源瓶颈或网络异常。

详情

方案推荐

1.DB-CONN-LIMIT-002-修改数据库最大连接数（可能性：82%）

- 匹配理由：电商系统 /shop/full-test 响应慢因 db-demo 数据库连接数耗尽导致，该方案直接提升最大连接数，解决连接耗尽问题。
- 执行工具：sandbox-mysql-连接数修改
- 工具参数：`cluster:db-demo` `namespace:default` `pod:svc-order-7c6fcd4ddd-twbss` `workload:svc-order`

执行

2.DB-CONN-CONFIG-001-调整数据库连接参数（可能性：82%）

- 匹配理由：db-demo 数据库连接管理不合理引发 /shop/full-test 响应慢，优化连接参数可提升连接效率，缓解连接耗尽。

3.DB-CONN-POOL-004-优化数据库连接池配置（可能性：81%）

- 匹配理由：/shop/full-test 响应慢由连接池配置与并发负载不匹配引起，优化连接池配置可提升并发承载力。

评估 & 反思

Evaluation & Reflexion

展开

精准

消除隐患，优化 MTBF



不间断巡检

实现了"预防为主"的理念

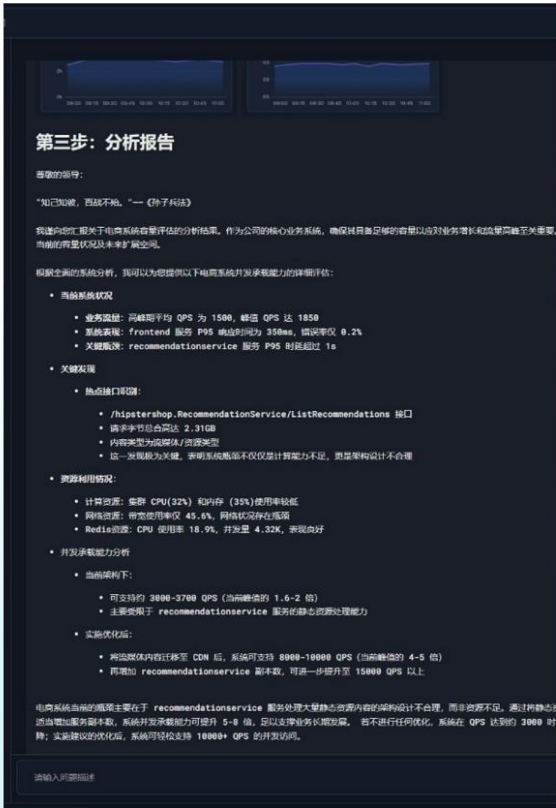
快速处置，降低 MTTR



分钟级诊断

解决了"快速止血"的问题

提升资源和算力利用率



一句话问数

满足了"及时决策"的需求

DeepFlow 智能体 诊断与恢复、巡检与优化



零侵扰采集，数十万部署

OpenAIOps AIOPS | 2025 CCF国际AIOps挑战赛
2025 CCF International AIOps Challenge

THANKS

主办单位：中国计算机学会（CCF）

承办单位：中国计算机学会互联网专委会、中国科学院计算机网络信息中心、中国移动研究院、清华大学

协办单位：华为2012实验室、阿里云、中兴通讯、中国移动九天团队、南开大学、西安电子科技大学、清华大学计算机科学与技术系、神州灵云