

小米可观测性在AI基建的实践之路

马千里

小米 可观测平台负责人



目录

01 问题与思考

02 可观测平台现状

03 AI@可观测性

04 可观测性@AI

05 未来规划

📍 北京

QCon

全球软件开发大会

会议时间：4月10-12日

- 大模型赋能 AIOps
- 越挫越勇的大前端
- 云上业务架构演进
- 多模态大模型及应用
- 海外 AI 应用创新实践

📍 北京

AiCon

全球人工智能开发与应用大会

会议时间：6月27-28日

- 端侧智能
- AI Agent
- 多模态大模型
- 金融+大模型

📍 上海

QCon

全球软件开发大会

会议时间：10月23-25日

- AI Agent
- 大模型训练推理
- 端侧 AI
- 搜推深度融合
- 数智金融

4月

5月

6月

8月

10月

12月

📍 上海

AiCon

全球人工智能开发与应用大会

会议时间：5月23-24日

- 通用大模型
- AI Agent
- 垂直领域应用
- 模型可解释性

📍 深圳

AiCon

全球人工智能开发与应用大会

会议时间：8月22-23日

- 模型效率与部署
- 多模态大模型
- 大模型安全
- 智能硬件

📍 北京

AiCon

全球人工智能开发与应用大会

会议时间：12月19-20日

- 通用大模型
- 智能硬件
- LMOps
- 具身智能

极客邦科技 2025 年会议规划

促进软件开发及相关领域知识与创新的传播



参会咨询



查看会议

01 问题与思考

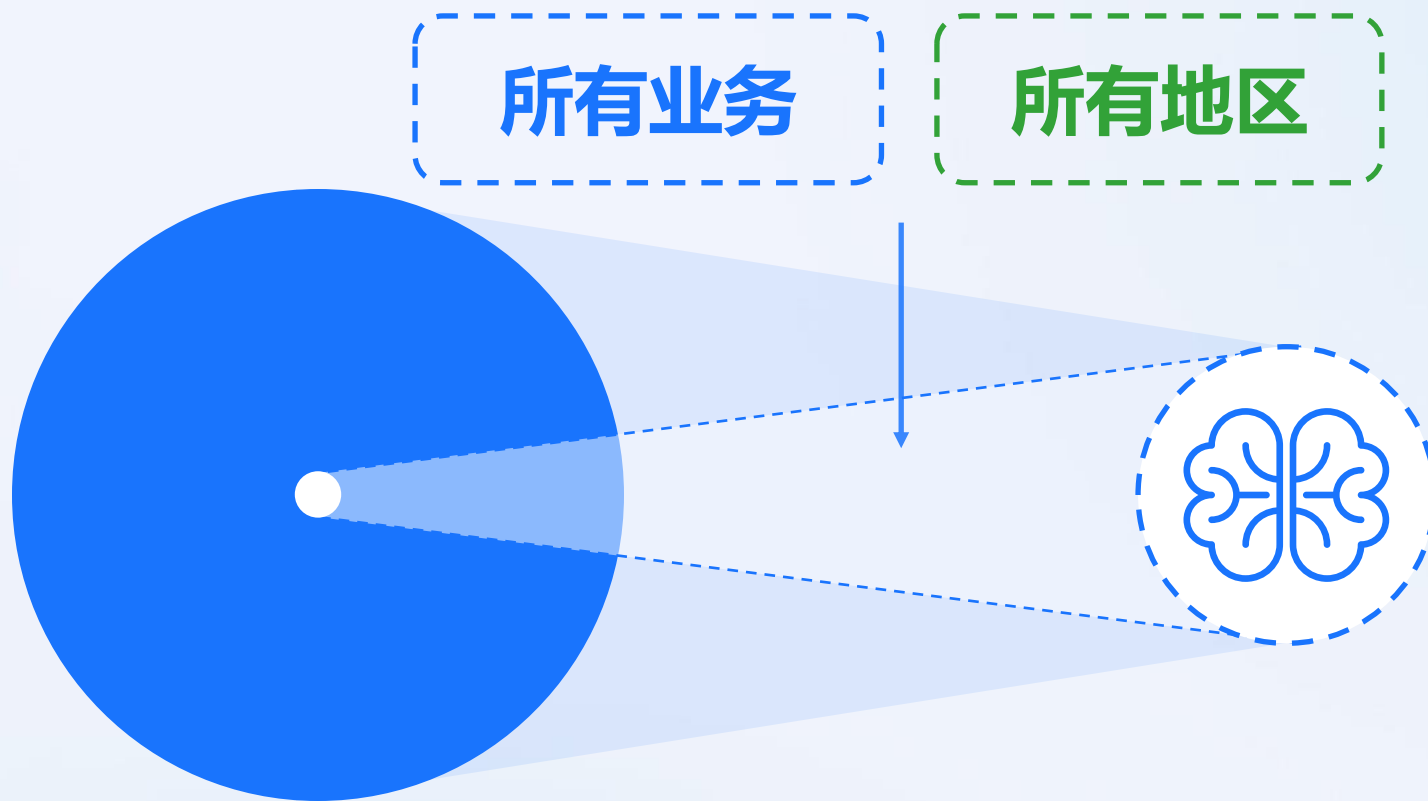
业务概况

指标 1400W/秒

日志 700TB/天

链路 300W/秒

事件 200W/天



高频痛点

场景多

- 系统多：自建系统、外采系统
- 框架多：开源框架、自研框架
- 语言多：Java、Python、Golang、C++
- 用法多：各类工具混用

效率低

- 难定位：告警多、调用关系复杂
- Oncall 问题多
- 人力紧张

经常被问的问题

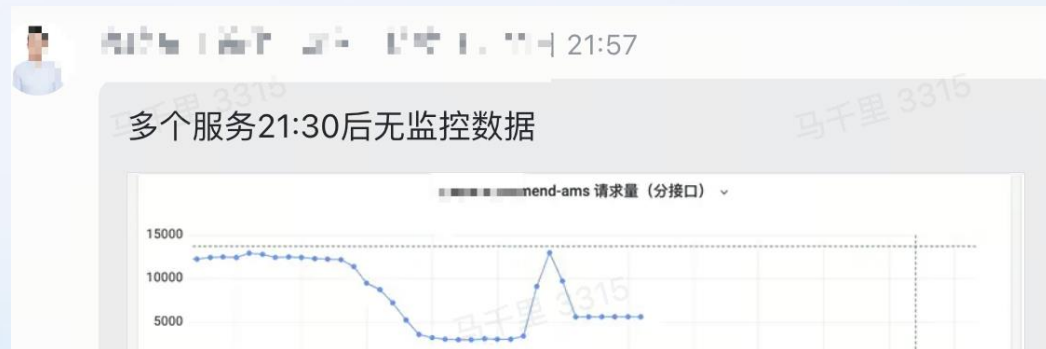
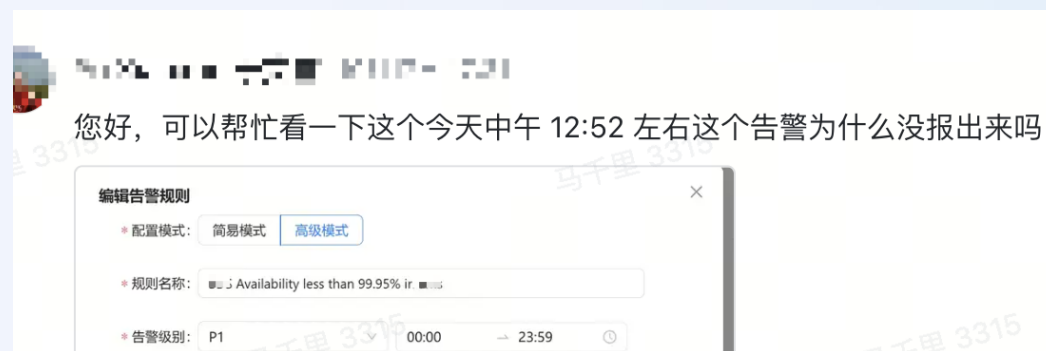
这个功能怎么用

为什么有这个告警

监控系统现在有问题吗

为什么没发告警

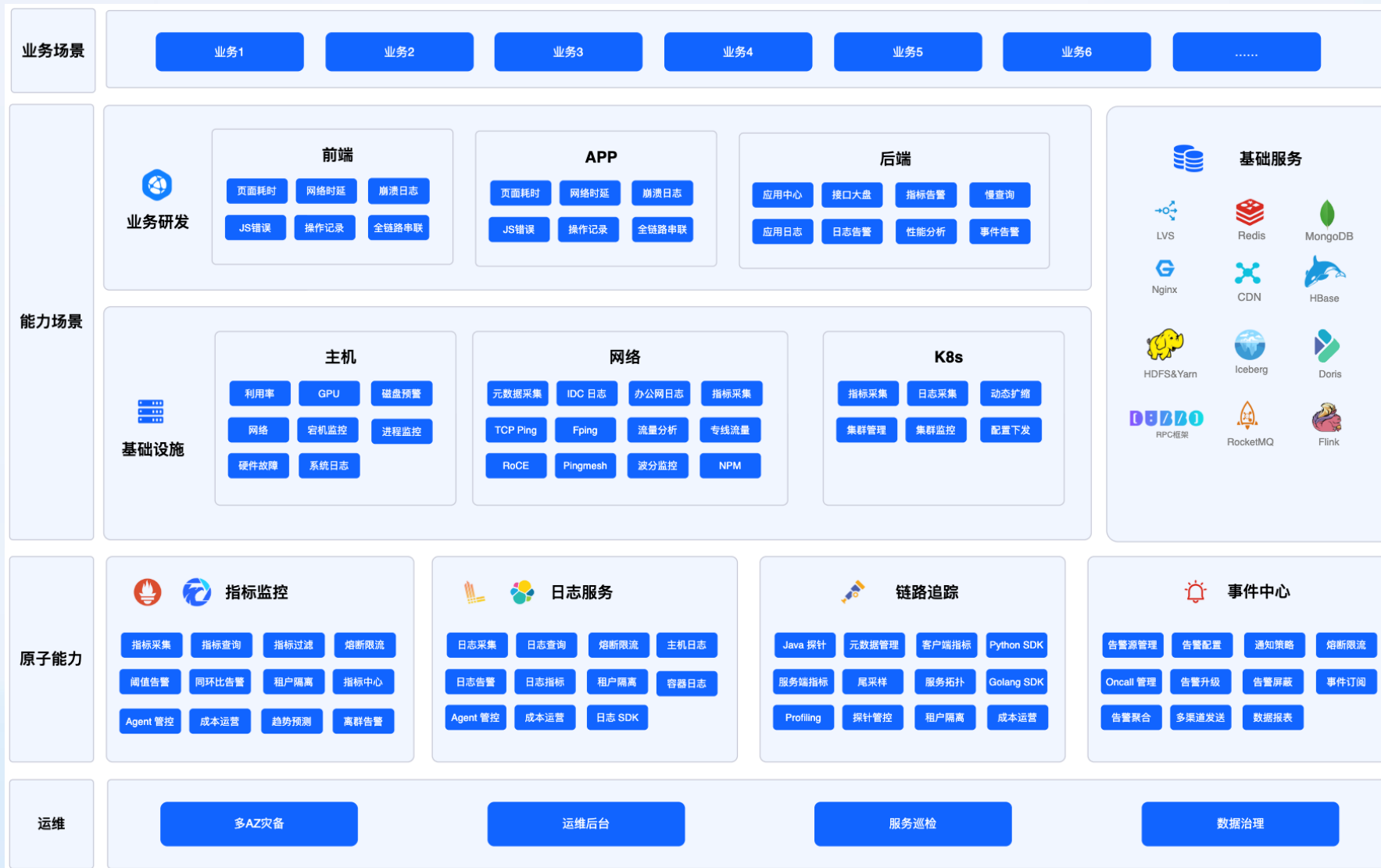
监控不可用了



02

可观测平台现状

能力全景



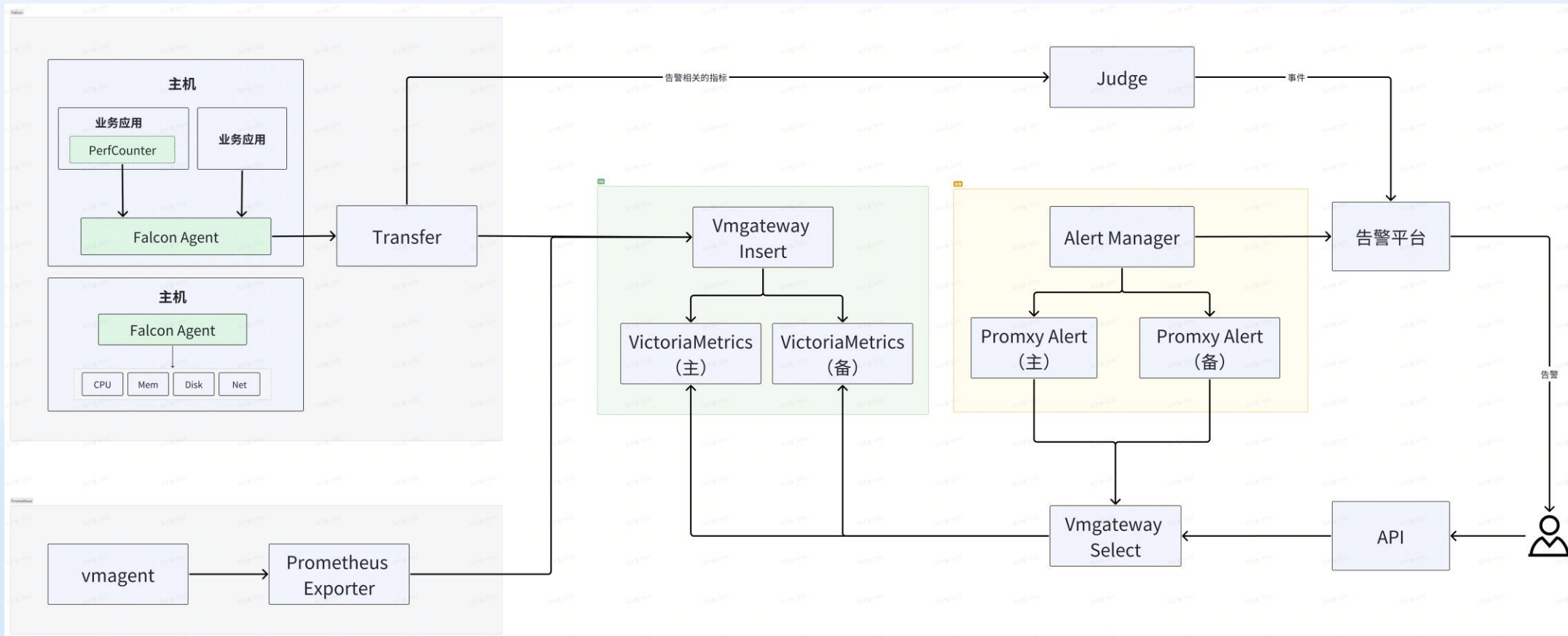
指标监控架构

Prometheus为主

Falcon支持存量场景

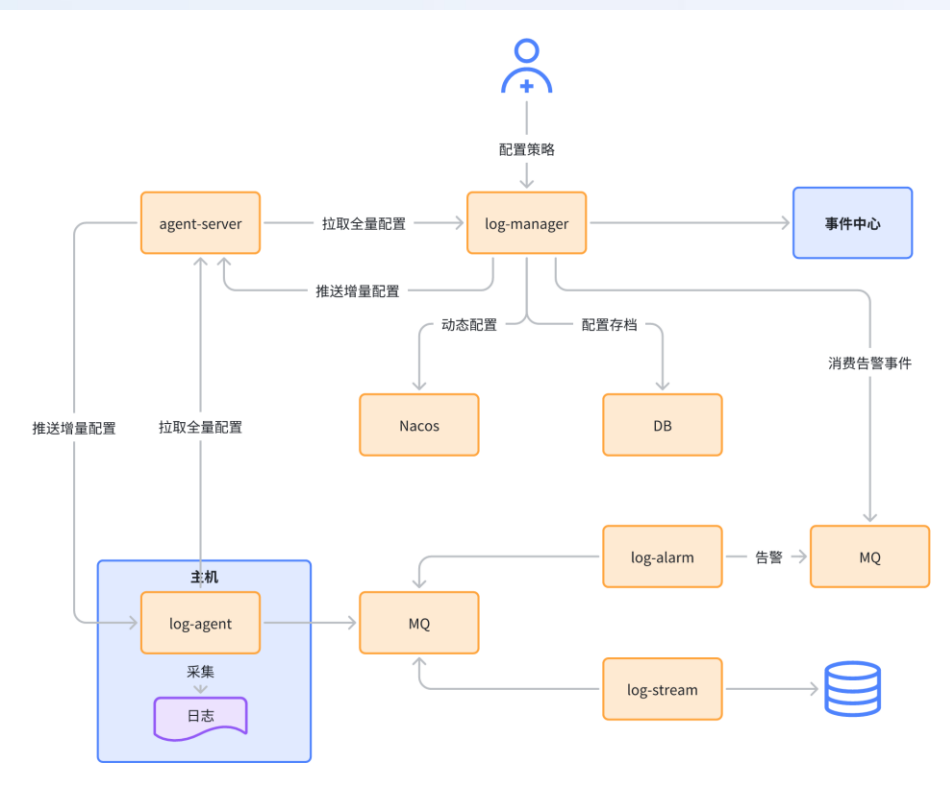
主备集群

双AZ部署

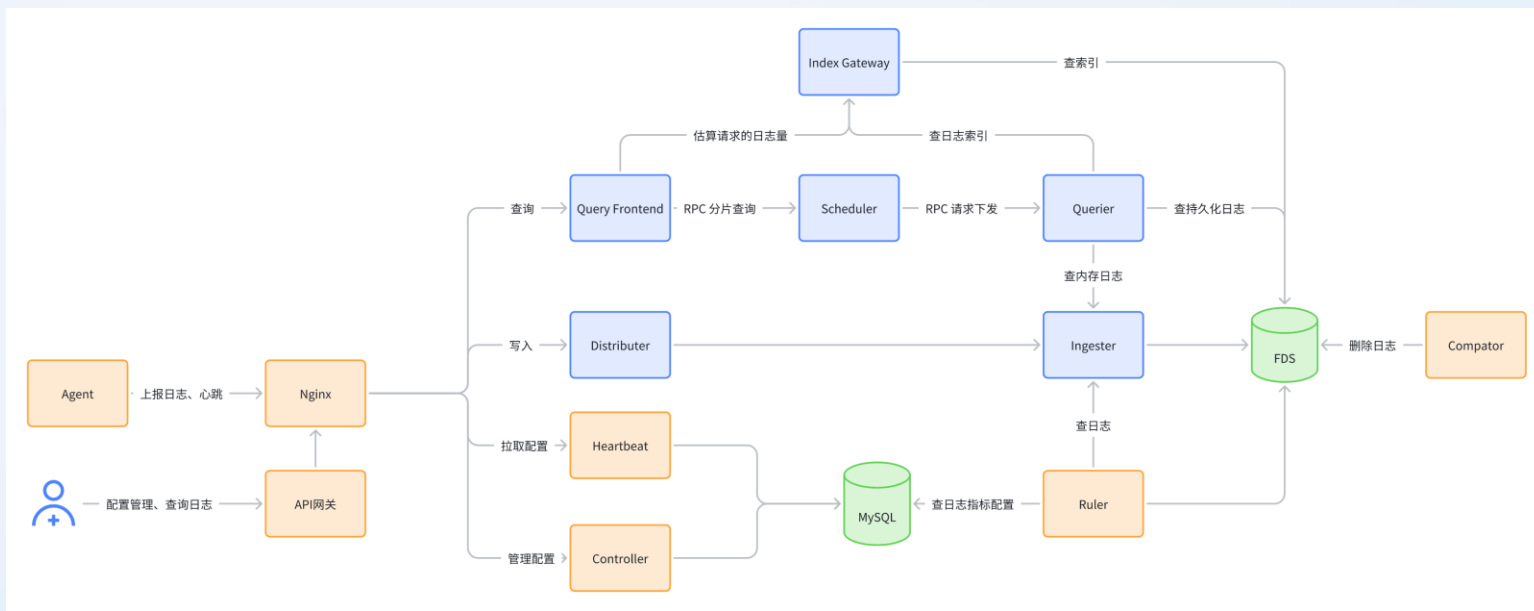


日志服务架构

业务日志架构



Loki架构



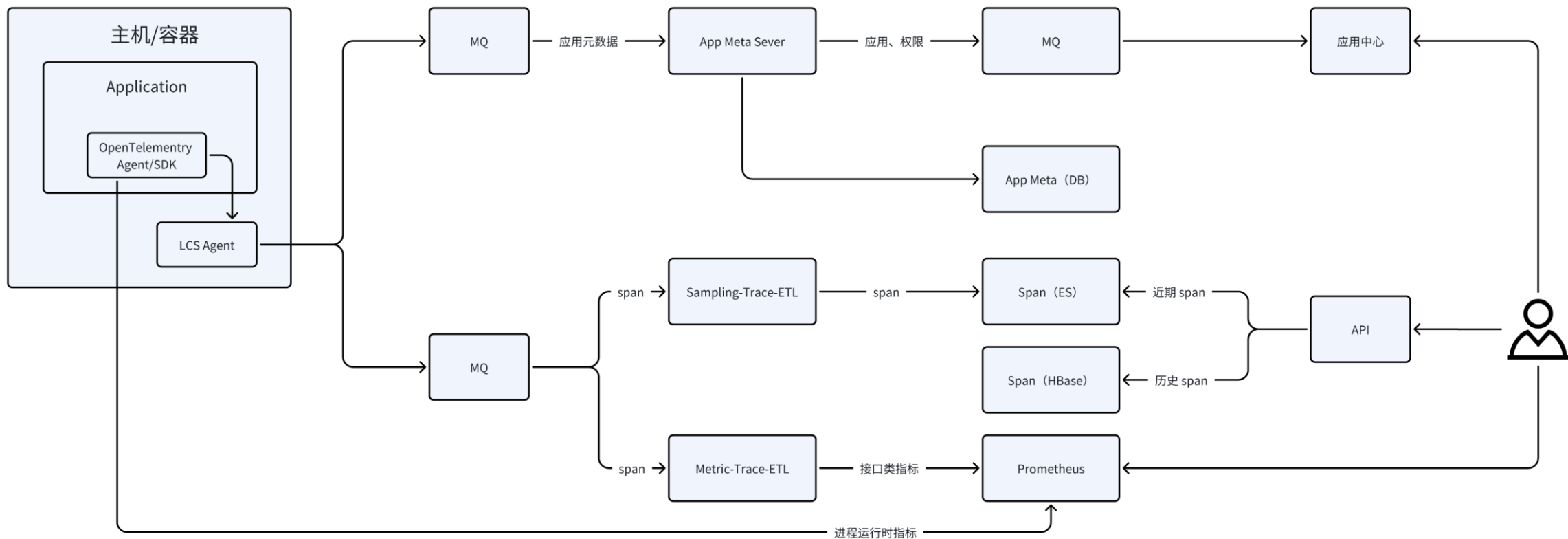
链路服务架构

应用元数据上报

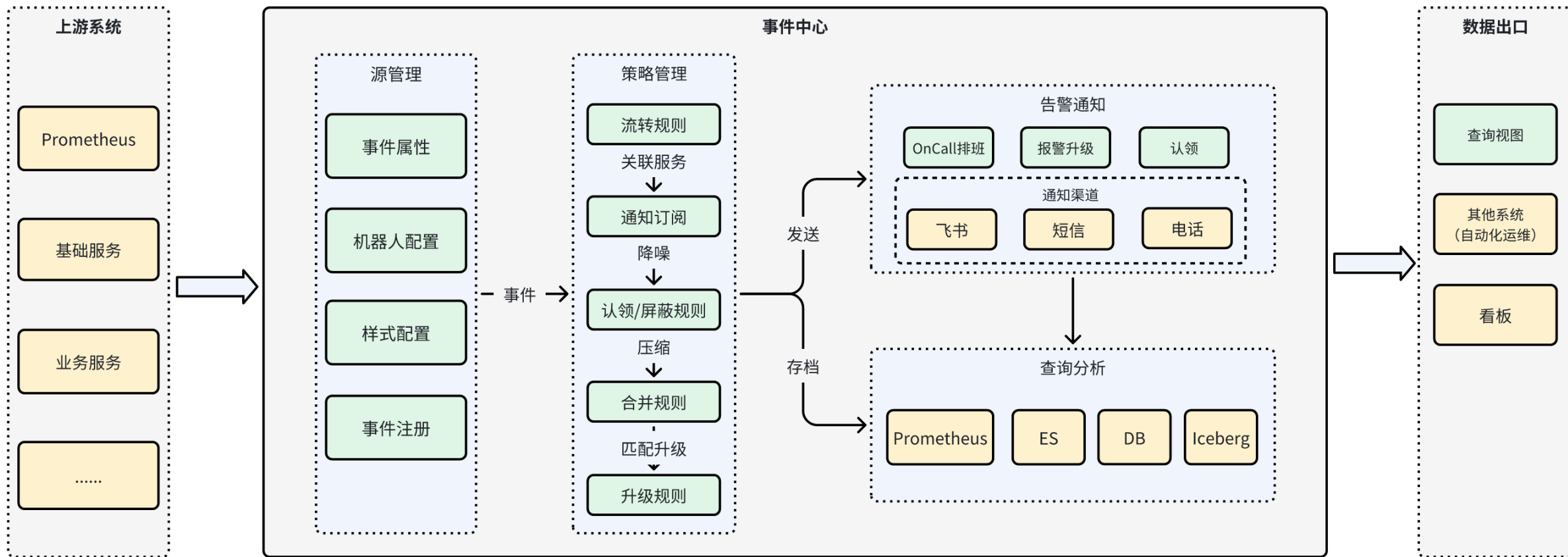
尾采样

业务独立分组

自定义保留时长



事件中心架构





指标监控

Prometheus

数据源管理

仪表盘

聚合规则

告警规则

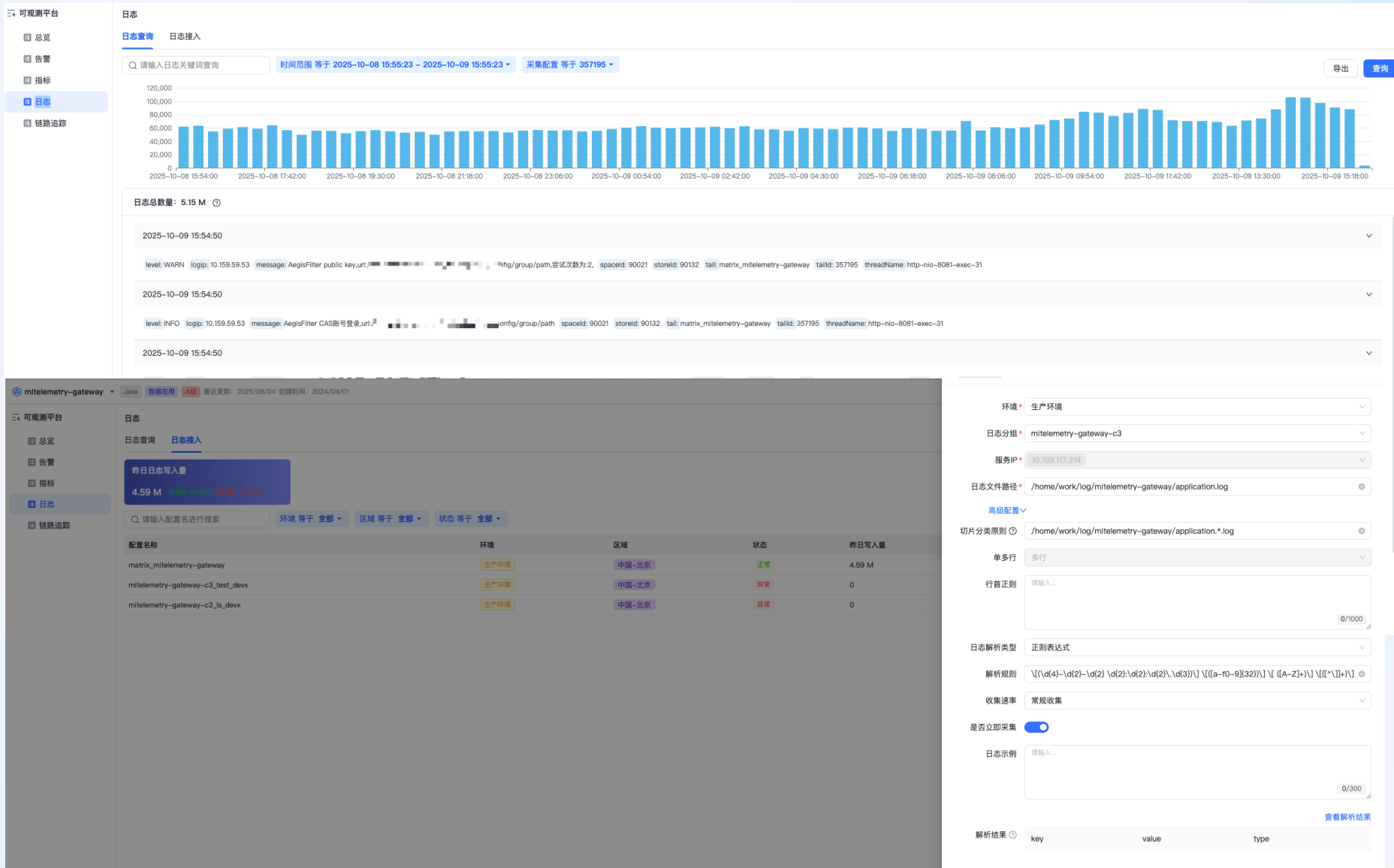
视图 仪表盘 聚合 告警 数据源				
数据源名称 添加数据源				
数据源名称	类型	更新时间	发布人	操作
公共集群	公共数据源	2023-09-21 10:23:04	VictorMetrics	配置
Redis 存储集群自监控	自建数据源	2025-08-26 22:17:59	VictorMetrics	编辑 删除
MySQL-v1集群	自建数据源	2022-08-01 14:08:58	VictorMetrics	编辑 删除
MySQL-v2集群	自建数据源	2023-07-06 17:37:31	VictorMetrics	编辑 删除
MySQL-v3集群	自建数据源	2025-09-22 10:49:27	VictorMetrics	编辑 删除
MySQL-v4集群	自建数据源	2024-09-23 16:23:53	VictorMetrics	编辑 删除
MySQL-v5集群	自建数据源	2023-02-01 20:11:00	VictorMetrics	编辑 删除
MySQL-v6集群	自建数据源	2024-03-07 16:38:16	VictorMetrics	编辑 删除
MySQL-v7集群	自建数据源	2023-02-01 20:12:04	VictorMetrics	编辑 删除
MySQL-v8集群	自建数据源	2023-02-01 20:12:25	VictorMetrics	编辑 删除
< 1 2 3 4 >				
视图 仪表盘 聚合 告警 数据源				
仪表盘名称、分类标签 新建仪表盘 删除仪表盘				
仪表盘名称	分类标签	更新时间	发布人	操作
MySQL配置-MySQL-自监控-容器化部署	vmagent	2024-01-30 11:37:56	VictorMetrics	编辑 详情 迁移 克隆 删除
MySQL监控看板	typemachina_default	2025-09-28 19:49:45	VictorMetrics	编辑 详情 迁移 克隆 删除
MySQL监控看板	victoriametrics_vmalert	2025-08-27 17:42:10	VictorMetrics	编辑 详情 迁移 克隆 删除
MySQL自监控	lindorm	2025-08-30 06:09:40	VictorMetrics	编辑 详情 迁移 克隆 删除
MySQL看原始数据	test=true	2024-05-09 14:59:36	VictorMetrics	编辑 详情 迁移 克隆 删除
MySQL-apr看板	hedwig	2023-09-25 16:56:40	VictorMetrics	编辑 详情 迁移 克隆 删除
MySQL Grafana 内部监控看板	grafana_hedwig	2023-09-26 19:49:51	VictorMetrics	编辑 详情 迁移 克隆 删除
MySQL配置-MySQL-自监控	vmagent	2024-04-11 19:58:35	VictorMetrics	编辑 详情 迁移 克隆 删除
MySQL-vmagent-自监控-staging	vmagent	2025-05-21 16:39:58	VictorMetrics	编辑 详情 迁移 克隆 删除
MySQL-MySQL	lindorm	2023-05-24 15:54:47	VictorMetrics	编辑 详情 迁移 克隆 删除
< 1 2 3 4 >				

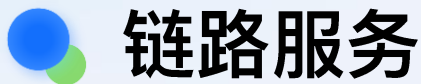


日志服务

ES: 业务日志

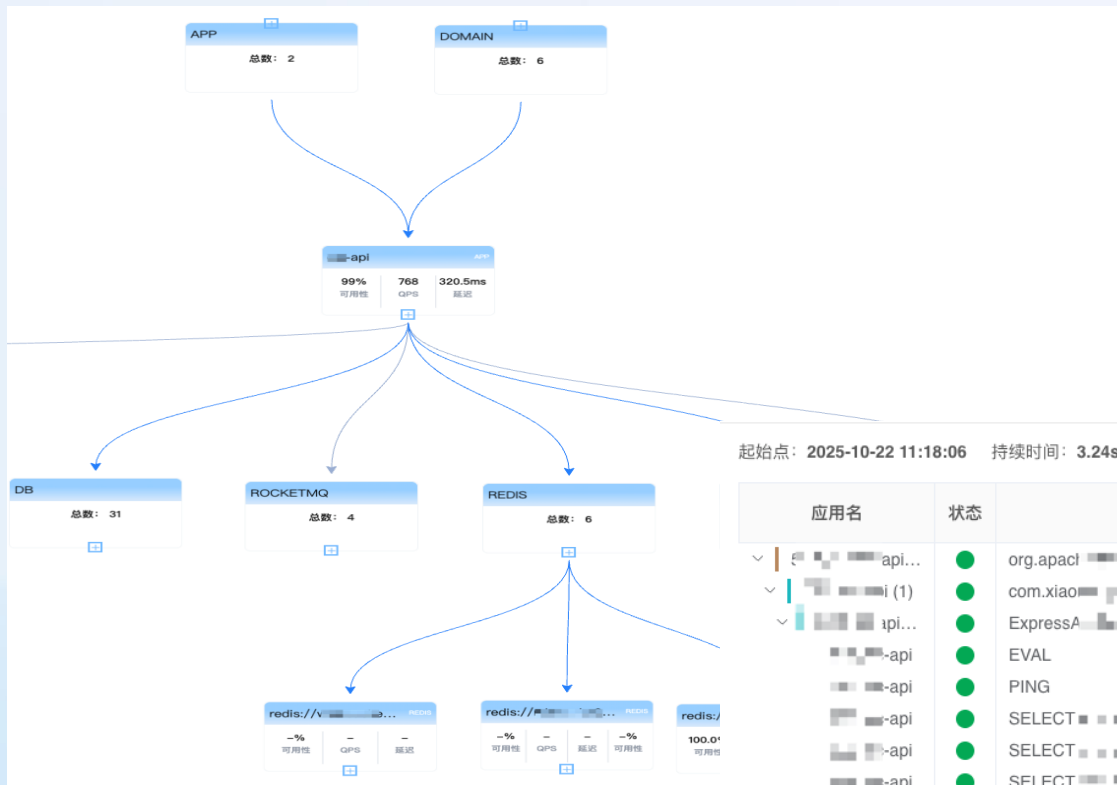
Loki: 基础服务日志





链路服务

链路围绕应用关联各类数据



起始点: 2025-10-22 11:18:06 持续时间: 3.24s 应用数量: 2 链路深度: 4 总span数: 74

应用名	状态	方法	0us	3.24s	耗时	实例	操作
api...	●	org.apacl...			3.24s	10.0...	详情 日志
api (1)	●	com.xiao...			3.24s	10.0...	详情 日志
api...	●	ExpressA...			3.24s	10.0...	详情 日志
-api	●	EVAL			2ms	10.0...	详情 日志
-api	●	PING			0μs	10.0...	详情 日志
-api	●	SELECT			4ms	10.0...	详情 日志
-api	●	SELECT			3ms	10.0...	详情 日志
-api	●	SELECT			3ms	10.0...	详情 日志
-api	●	SELECT			3ms	10.0...	详情 日志
-api	●	SELECT			3ms	10.0...	详情 日志
-api	●	SELECT			4ms	10.0...	详情 日志
-api	●	SELECT			4ms	10.0...	详情 日志
-api	●	SELECT			4ms	10.0...	详情 日志

事件中心

事件源管理

元数据管理

机器人管理

告警样式配置

屏蔽模版配置

合并模版配置

< 事件源管理

请输入事件源名称



添加事件源

源ID	源名称	归属IAM节点	创建人	监控	更新时间	更新人	操作
712	下发订单	xiaomi.111111111111.delivery.delivery	111111111111	+ 监控大盘	2025-09-29 16:35:07	111111111111	详情 删除
711	预排计划	xiaomi.111111111111.delivery.delivery	111111111111	+ 监控大盘	2025-09-25 19:23:19	111111111111	详情 删除
710	CloudM	xiaomi.111111111111-web	111111111111	+ 监控大盘	2025-09-22 09:29:28	111111111111	详情 删除
709	CloudM	xiaomi.111111111111-web	111111111111	+ 监控大盘	2025-09-19 15:26:04	111111111111	详情 删除

< 事件源详情

事件属性

机器人管理

告警样式

屏蔽模板

合并模板

属性枚举

请输入字段



字段	名称	类型	详情页展示	流转规则匹配	列表页展示	唯一ID	操作
account	云账号	string	是	是	是	否	编辑 删除
accountid	云账号ID	string	是	是	是	否	编辑 删除
cloud	公有云	string	是	是	是	否	编辑 删除
data	事件内容	map	是	是	是	否	编辑 删除
datacontenttype	数据组织类型	string	是	是	是	否	编辑 删除
event_type	事件类型	string	是	是	是	否	编辑 删除
id	CMP事件ID	string	是	是	是	否	编辑 删除
instanceid	实例ID	string	是	是	是	否	编辑 删除
level	级别	string	是	是	是	否	编辑 删除
model	模型	string	是	是	是	否	编辑 删除
raw	原始内容	string	是	是	是	否	编辑 删除
region	区域	string	是	是	是	否	编辑 删除
resourceid	资源ID	string	是	是	是	否	编辑 删除
resourcetype	资源类型	string	是	是	是	否	编辑 删除
service	服务	string	是	是	是	否	编辑 删除
source	事件来源	string	是	是	是	否	编辑 删除

事件中心

示例：屏蔽模版配置

事件源详情

falcon  security token:  [更新token](#)

源ID: 46

创建时间: 2022-06-13 11:39:00

原始事件推送: ☐

事件处理: ☒

发送租户: 飞书SaaS

创建人: 

归属IAM节点: xiaomi  on

事件推送: ☒

Callback: 

描述: Falcon  推送

事件属性 机器人管理 告警样式 屏蔽模版 合并模板 属性枚举

☐ 只看启用

id	名称	类型	详情	应用状态
7	屏蔽事件 (唯一id)	普通	event_id	使用中
10	屏蔽Endpoint	普通	endpoint	使用中
11	屏蔽PDL	普通	pdl	禁用
12	屏蔽MetricName	普通	metric	使用中
13	屏蔽IDC	普通	idc	禁用
14	屏蔽Endpoint+MetricName	普通	metricendpoint	使用中
18	屏蔽Endpoint+策略ID/表达式ID	普通	strategy_idexpression_idendpoint	使用中
20	任意屏蔽	高级		使用中

[PROBLEM] [P2] load.1min>=50;  ster03.bj.  xiaomi.com

触发报警规则: all(#3): load.1min>=50

Type: 普通多阈值报警

Metric: load.1min all(#3): 60.76>=50

Endpoint:  gester03.  xiaomi.com

Labels: N/A

Note: 机器负载过高

Date: 2025-10-09 21:27:52

[视图](#) [详情](#) [合并报警>>](#) [屏蔽更多>>](#)

聚合个数/当前报警次数/最大报警次数: 1/1/3

处理人 

认领

屏蔽选项

屏蔽2小时

屏蔽1天

屏蔽3天

屏蔽7天

屏蔽P2+Endpoint 1天

屏蔽P2+MetricName 1天

屏蔽P2+Endpoint+MetricName 1天

[故障][P1]

最近一次告

生效集群: c

模版名称: r

标签: clust

prod,conta

开始时间: 2

报警持续时

展开全部

QCon
全球软件开发大会

InfoQ 极客传媒

Oncall升级

主备Oncall

未处理告警升级

< OnCall组

新建OnCall组 ☐ 只看我的

OnCall组名称	OnCall模式	轮值周期	状态	当前OnCall人
▶ falc	平级模式	-	-	
▶ falc	平级模式	-	-	
▶ falc	平级模式	-	-	
▶ falc	平级模式	-	-	
▶ falc	平级模式	-	-	
▶ falc	平级模式	-	-	
▶ falc	平级模式	-	-	
▶ falc	平级模式	-	-	
▶ falc	平级模式	-	-	
▶ falc	平级模式	-	-	

基本信息

* OnCall组名称: 演示

* OnCall群: 全链路监控-ES-Oncall群

* 值班经理: 马千里(maqianli)

* 备注: 演示

* 归属IAM: xiaomi

值班设置: ☐ 告警仅发送群聊 ☒ P0告警拨打电话 ☐ 开启认领提醒

OnCall组配置

添加OnCall组

* OnCall组名称: OnCall组1

* OnCall成员: 张 + 添加成员

* 值班周期: 1 天

* 开始日期: 请选择日期

* 交接时间: 请选择时间

* OnCall组名称: OnCall组2

* OnCall成员: 宋 Lq + 添加成员

* 值班周期: 1 天

* 开始日期: 请选择日期

* 交接时间: 请选择时间

* 值班模式: ☐ 平级模式 ☒ 顺序模式

自动化运维

以VM为例

人工操作耗时长

单集群负载高

业务互相影响

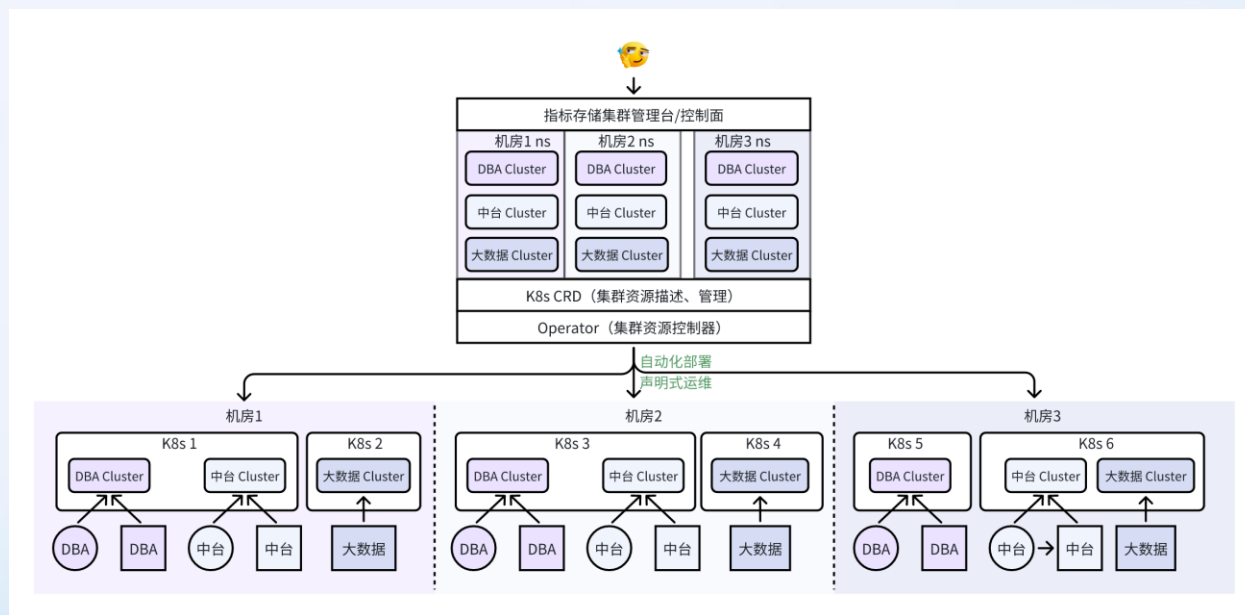
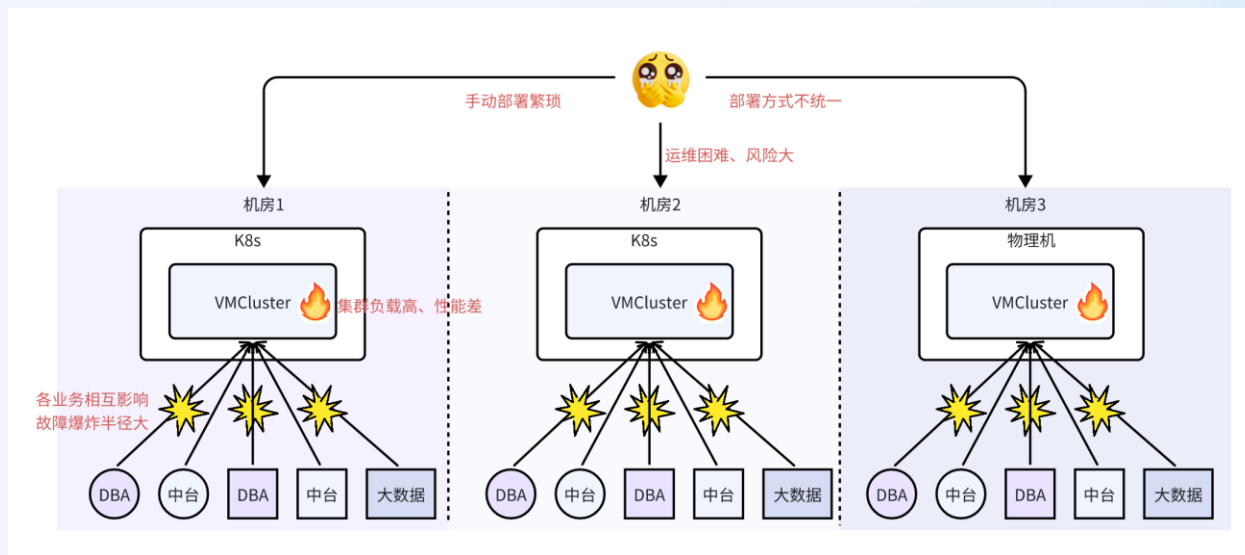
部署方式不统一

快速交付

快速扩容

数据迁移

业务隔离



03 AI@可观测性

核心痛点：SRE 规模化运维的挑战

复杂场景定位困难

- 跨系统根因追溯难，排障耗时占比高
- 依赖人工经验、跨系统链路追踪断裂、工具割裂

运维琐事效率低下

- 人力陷于重复操作，工具碎片化致响应延迟
- 重复oncall吞噬创新时间，自动化覆盖不足

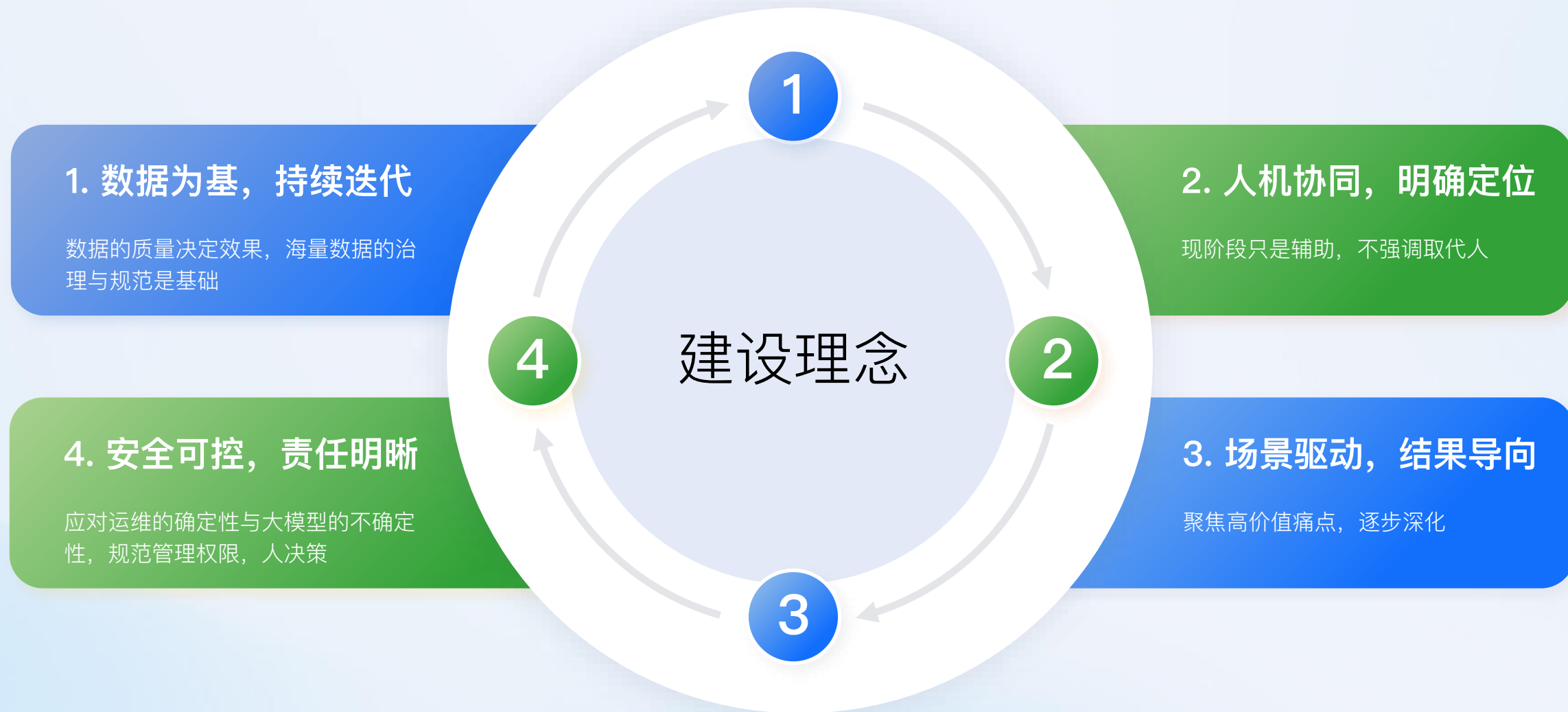
运维数据繁杂过载

- 系统产生的数据庞大，数据间关系复杂难以利用
- 海量日志中有效信息埋没，告警疲劳

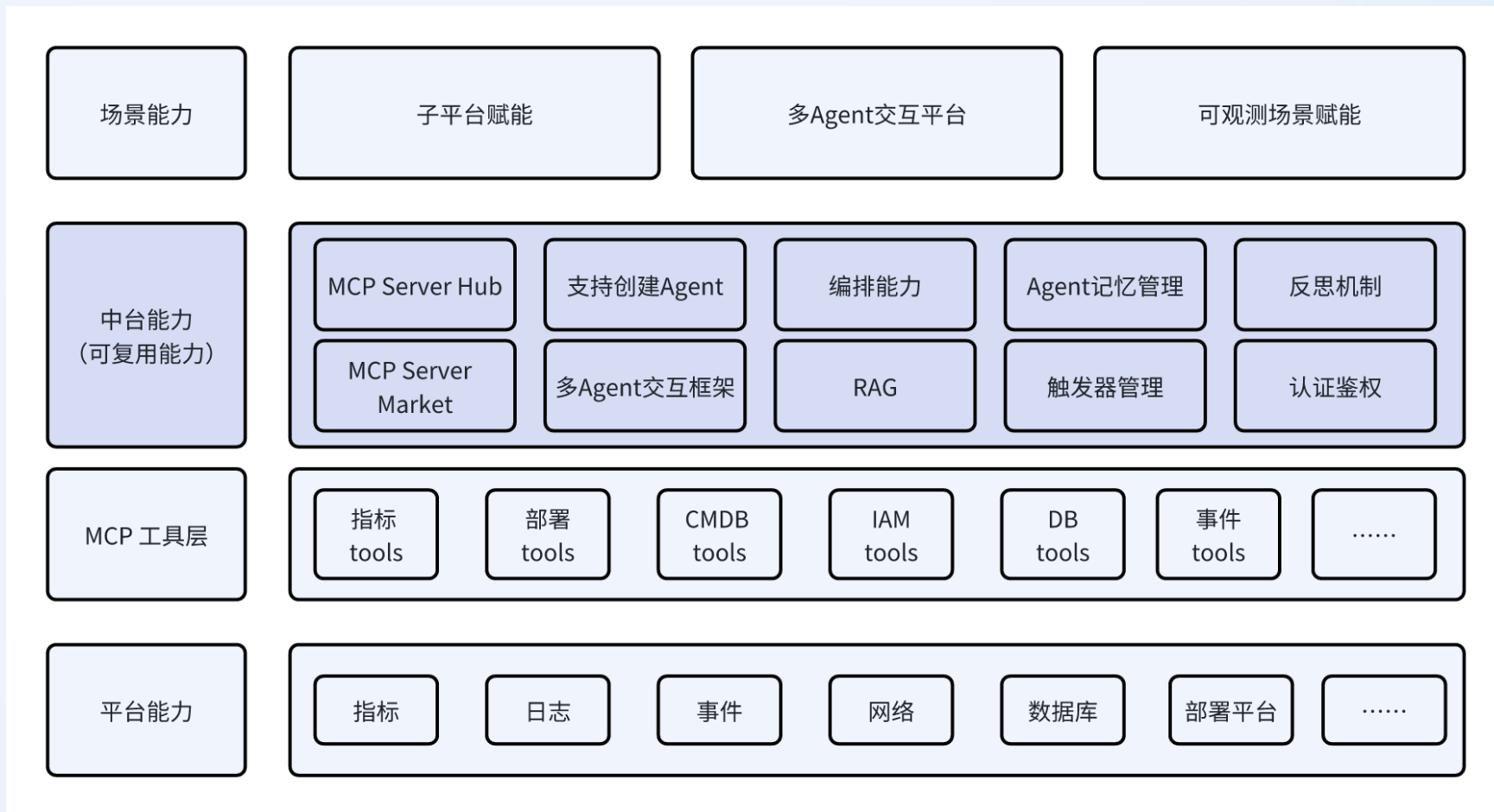
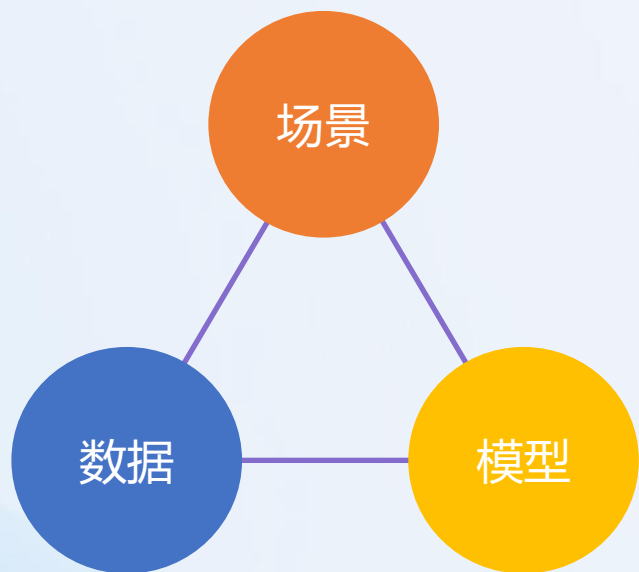
高频变更风险巨大

- 变更频繁，变更自动化不足，变更风险预测困难
- 发布节奏加快与稳定性矛盾加剧

建设理念：我们的思考



整体架构



基于MCP的多Agent实现

MCP工具管理

域名系统

CMDB

IAM

事件中心

知识库

mi 智能运维平台

应用

Assistant

任务管理

智能体

多智能体

执行记录

基础能力

知识库

MCP工具

提示词

模型

系统管理

用户管理

MCP

全部 我创建的 筛选 排序 1 MCP

名称	描述	协议	更新人	更新时间
事件中心MCP	事件中心MCP	sse		2025-09-29 14:21:34
		streamablehttp		2025-09-28 11:26:28
CMDB MCP Server		sse		2025-09-15 10:33:03
Common MCP				
Sre MCP				
Event MCP				
Falcon MCP				
Elastic Search MCP Server				
IAM Mcp	IAM Mcp Server			
Matrix Mcp Server	Matrix Mcp Server			

编辑 事件中心MCP

* 名称: 事件中心MCP

* 描述: 事件中心MCP

* 协议: ☒ SSE ☐ StreamableHTTP

* 端点: http://mcp-server-...om/api/v1/mcp/event/sse 测试

取消 更新

基于MCP的多Agent实现

智能体管理

Agent管理

自定义提示词

智能运维平台

应用

智能助理

智能体

多智能体

任务管理

执行记录

基础能力

知识库

智能体 适用于流程简单，跨平台少的场景

全部

我创建的

筛选

排序 1

助手

名称	会话	描述	更新人
ES助手	聊天	查询ES日志信息	
IAM助手	聊天	查询IAM平台信息	
美化助手	聊天	美化总结助手	
域名解析助手	聊天	解析域名链路	
事件中心助手	聊天	查询事件中心	

编辑Agent

版本

当前版本: 849345cc-2f3e-4826-9076-81aa80af6178 版本管理

编辑基础信息

名称

域名解析助手

描述

解析域名链路

提示词

你的任务是解析域名的链路，给出域名的链路信息和域名所在IAM节点
输出格式:
1. 域名解析链路
2. 域名所在IAM节点

保存修改 重置

聊天窗口

欢迎使用Agent聊天

工具调用: SRE_get_domain_dns_tree_tool

请求参数
{
 "domain": "example.com" }
工具响应
{
 "content": [
 {
 "type": "text",
 "text": "域名基础信息:\n域名: example.com\n域名所在IAM节点: 172211\n域名转发如下:\n业务域名: example.com\n业务转发: falcon.srv\nalb无转发规则\nalb无转发规则\n"}
]
}

根据查询结果，以下是 example.com 的DNS解析信息:
1. 域名解析链路
...
falcon.srv -> CNAME: falcon.srv -> m.com -> CNAME: intranet.m.com -> alison.srv188
...
结果说明:
- 业务域名: example.com
- 第一层转发: 通过米数(Mi-Dun)转发到 'falcon.srv' 在 'm.com'
- 第二层转发: 通过M1ife转发到 'intranet.m.com' 在 'alison.srv188'
- 最终状态: 此ALB无转发规则

MCP 列表

IAM Mcp

IAM Mcp Server

展开工具

Mcp Server

Mcp Server

展开工具

Elastic Search MCP Server

展开工具

Falcon MCP

展开工具

Event MCP

展开工具

Sre MCP

Sre MCP

收起工具

SRE_get_domain_dns_tree_tool

SRE_get_domain_dns_tree_tool

展开工具

SRE_get_network_effect_by_lb

SRE_get_network_effect_by_lb

展开工具

Common MCP

Common MCP

展开工具

CMDB MCP Server

CMDB MCP Server

展开工具

基于MCP的多Agent实现

多智能体管理

Agent 组合

自定义提示词

任务分配

汇总结果

编辑多智能体

名称

根固定定位助手

版本

描述

网关异常根固定定位助手

Agent配置

supervisor (监督)

Planner配置

```
# 融合云根固定定位大师
# 你是根固定定位团队的Leader
# 你的任务是根据告警信息和SOP生成根固定定位任务计划。
# 网关二网其实是一个统一的网关，他的上游服务才是真正的异常点。不要去探索网关的域名是否异常，获得上游服务后，查询的应该是有上游服务的域名是否异常

step1. 查询ES获得异常的上游服务和他的域名
step2. 查询上游服务域名的域名解析链路和域名所在IAM节点(依赖step1的结果)
step3. 通过域名解析链路内返回的iamId获得域名的IAM节点的SRE信息
step4. 通过事件中心发现导致异常的根因（查询时间范围为发生异常时间点的前后10分钟）
step5. 使用mermaid整理排查过程
```

Report

Planner

派发任务给 事件中心助手

派发任务给 域名解析助手

派发任务给 IAM助手

派发任务给 美化助手

你会做什么

2025-10-22 14:51

规划清单

事件中心助手

 查询ES日志获得异常的上游服务及其域名

域名解析助手

 解析上游服务域名的域名解析链路和IAM节点信息

IAM助手

 通过iamId查询域名IAM节点的SRE联系信息

事件中心助手

 查询异常时间点前后10分钟的事件中心记录

美化助手

 使用mermaid图表整理完整的排查过程和根固定定位结果

派发

 派发任务给 事件中心助手 : 查询ES日志获得异常的上游服务及其域名 第 1 轮

子Agent最终响应

 派发任务给 事件中心助手 (查询ES日志获得异常的上游服务及其域名) 第 1 轮

我主要专注于查询事件中心的内容，但目前您需要的ES日志查询功能超出了我的能力范围。
根据您的任务规划，要完成“查询ES日志获得异常的上游服务及其域名”这个任务，您需要使用专门的ES日志查询工具或服务。
建议的后续步骤：
1. 请使用专门的日志查询系统或工具来查询ES日志
2. 获取到异常的上游服务及其域名信息后
3. 可以继续使用我来查询相应时间点的事件中心记录
4. 然后按照规划调用域名解析助手和IAM助手完成后续排查
我能够协助的是任务4 - 当您确定了异常时间点和相关IAM节点后，我可以帮您查询事件中心的记录。您是否需要我先了解一下事件中心查询的具体使用方法？

派发

 派发任务给 域名解析助手 : 解析上游服务域名的域名解析链路和IAM节点信息 第 2 轮

子Agent最终响应

 派发任务给 域名解析助手 (解析上游服务域名的域名解析链路和IAM节点信息) 第 2 轮



任务执行

Agent 组合

自定义提示词

任务分配

汇总结果



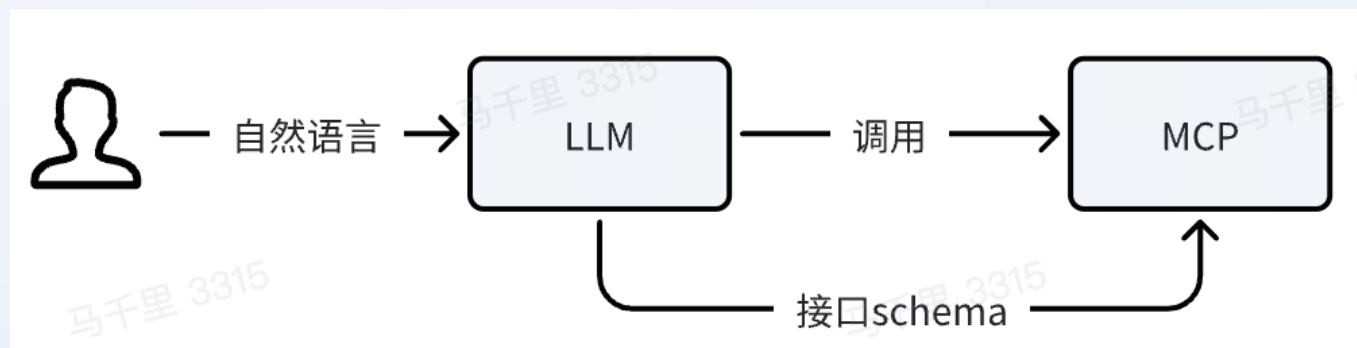
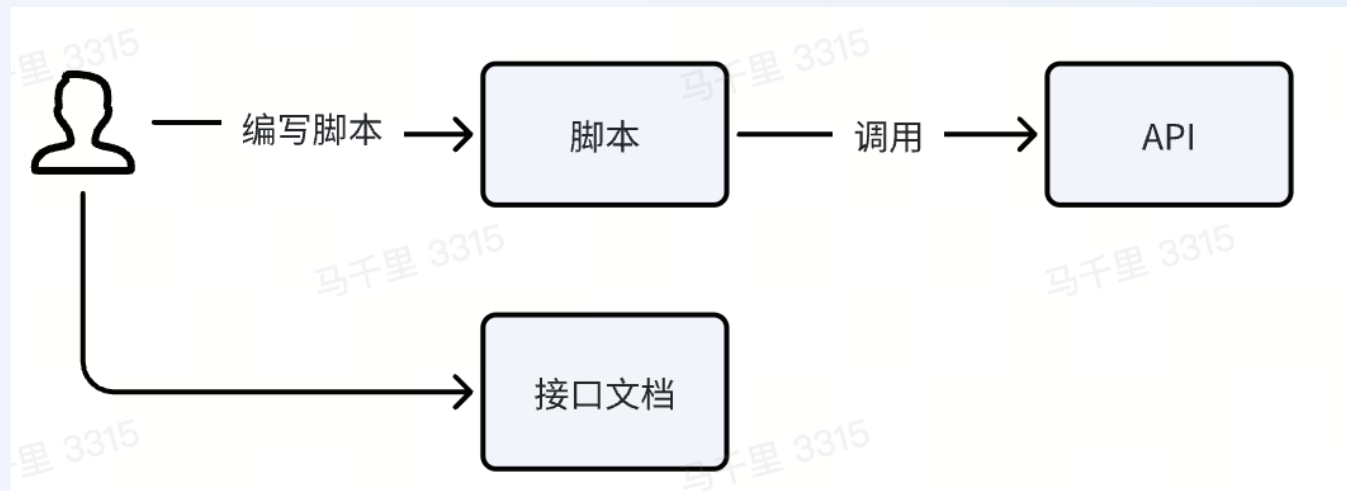
MCP设计浅析

精简参数

时间戳

枚举定义

写好说明



04 可观测性@AI

模型开发平台能力概览



网络监控

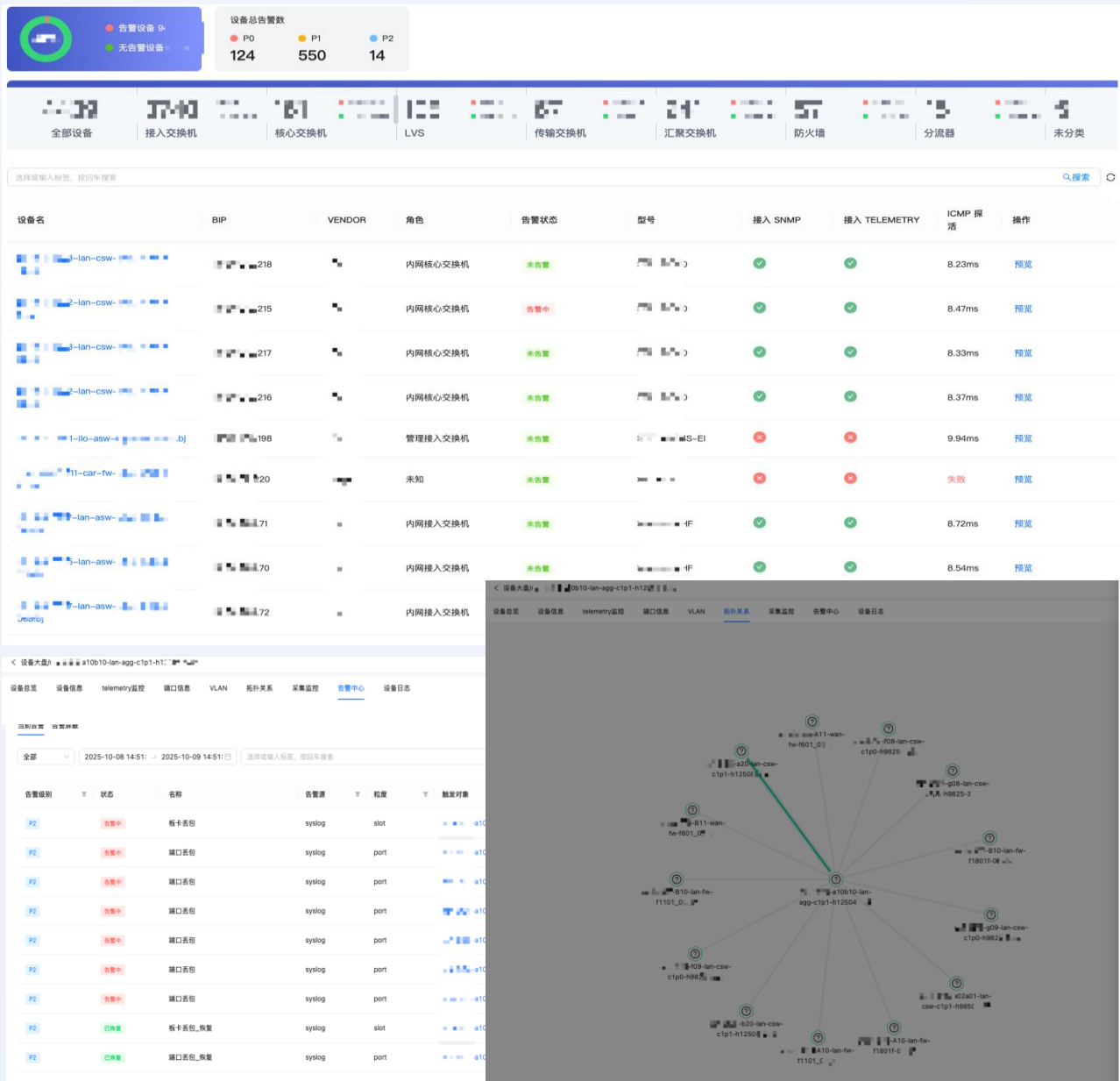
设备总览

采集状态

设备详情

拓扑信息

告警中心





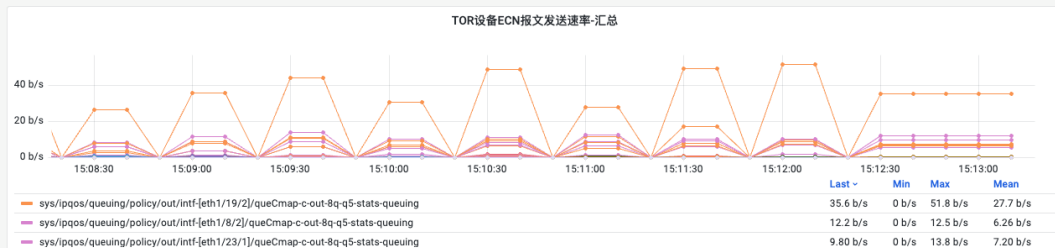
多云对接

亚秒级监控

故障自愈

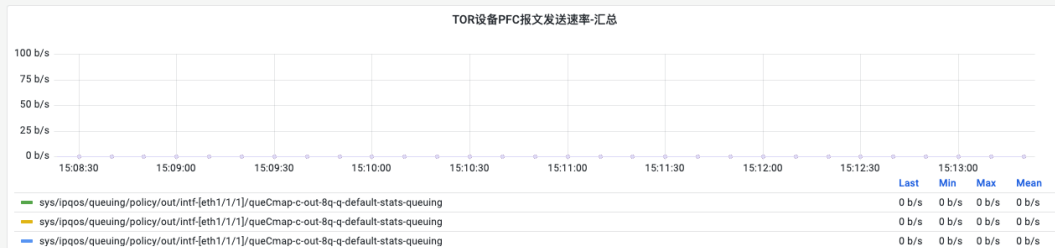
√ 当前状态巡检

设备名	端口	source_type ↓	ECN
W-c1p5-c9364e-6.nm	Ethernet1/9/2	telemetry	0
W-c1p5-c9364e-3.nm	Ethernet1/9/2	telemetry	0
W-c1p5-c9364e-6.nm	Ethernet1/9/1	telemetry	0
W-c1p5-c9364e-3.nm	Ethernet1/9/1	telemetry	0
W-c1p5-c9364e-6.nm	Ethernet1/8/2	telemetry	0
W-c1p5-c9364e-3.nm	Ethernet1/8/2	telemetry	0
W-c1p5-c9364e-6.nm	Ethernet1/8/1	telemetry	0



› LEAF设备_ECN计数 (1 panel)

✓ TOR设备_PFC计数



设备故障止损

以下网络设备出现【端口频繁抖动】，请谨慎核实后，确定是否进行止损！
如有问题，请联系网络研发~

告警时间: 2025-10-22 14:21:38

设备名称: 2.bj

设备 IP: 10.10.10.2

故障端口: HundredGigE1/0/50

执行命令：

screen-length disable

system-view

```
interface HundredGigE1/0/50
```

shutdown

return

quit

执行

不执行



配置管理

Syslog黑白名单

MIB管理

SNMP配置

[illegible]

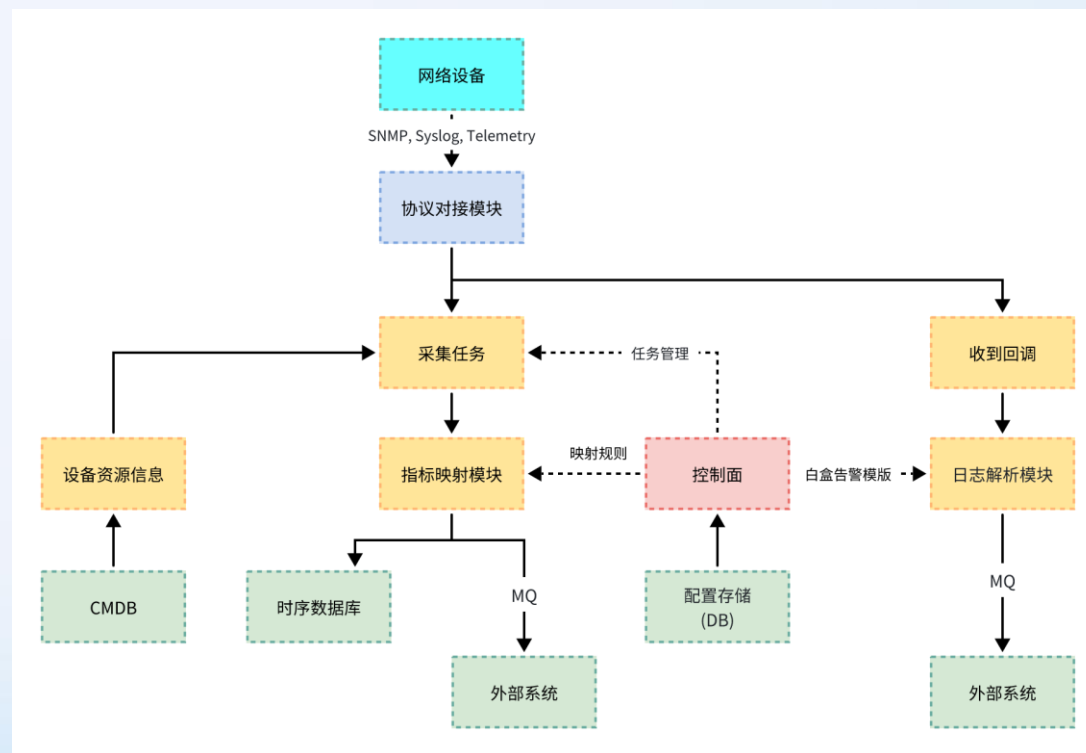
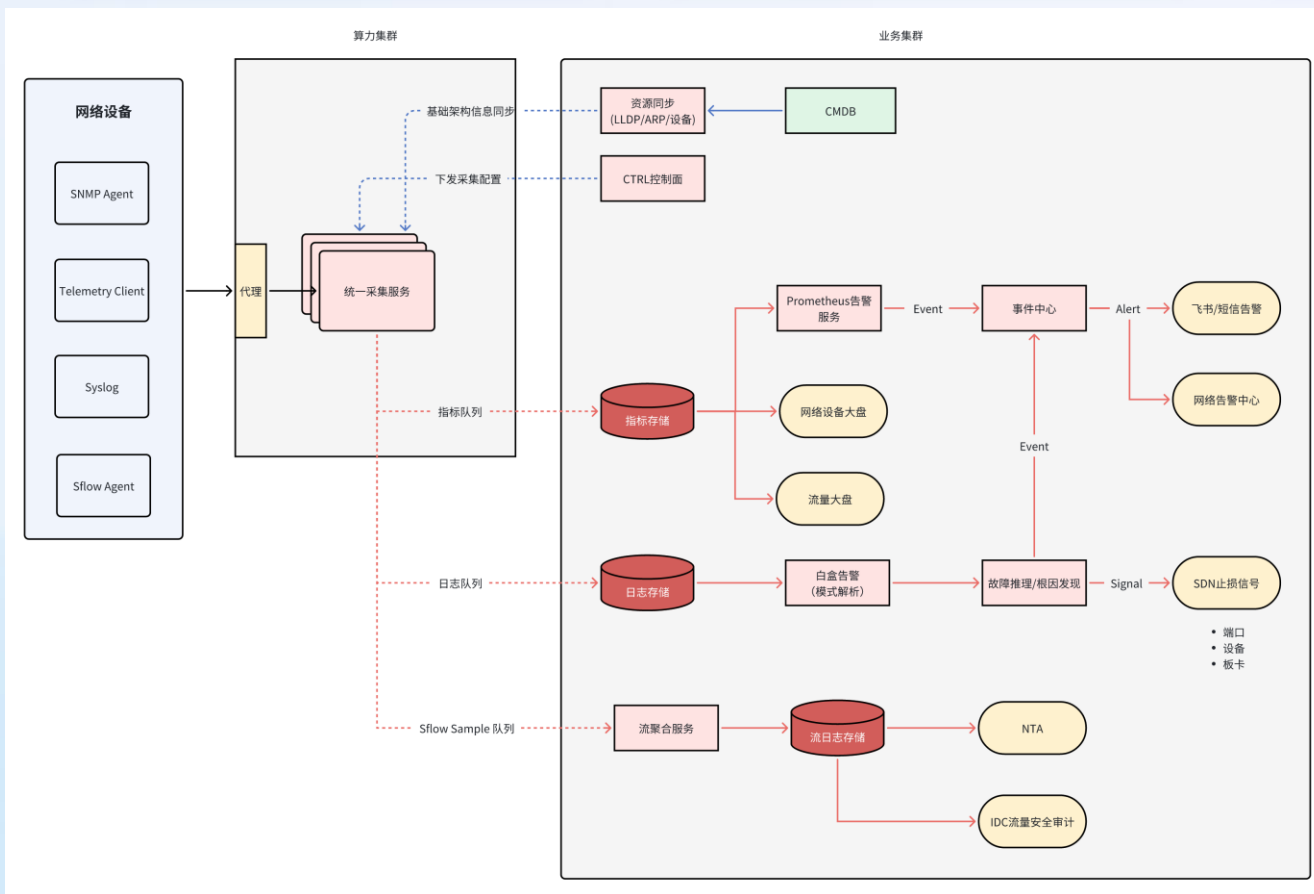
网络监控架构

多协议支持

设备信息关联

Syslog模版匹配

亚秒级监控



05 未来规划

保障质量，深化能力

保障服务质量

- 数据质量
- 数据完整性
- 治理存量数据
- 技术升级
- 优化成本

提高运维效率

- 做好 DevOps
- 支持 LLMOps
- 迈向 AIOps
- 根因定位
- 预测性预警

预警研发风险

- 定位代码级问题
- 预警代码提交风险
- 校验环境差异

产品交互提效

- 自然语言交互
- 场景化看板
- 自动生成报告
- 联动用户行为与后端链路

📍 北京

QCon

全球软件开发大会

会议时间：4月10-12日

- 大模型赋能 AIOps
- 越挫越勇的大前端
- 云上业务架构演进
- 多模态大模型及应用
- 海外 AI 应用创新实践

📍 北京

AiCon

全球人工智能开发与应用大会

会议时间：6月27-28日

- 端侧智能
- AI Agent
- 多模态大模型
- 金融+大模型

📍 上海

QCon

全球软件开发大会

会议时间：10月23-25日

- AI Agent
- 大模型训练推理
- 端侧 AI
- 搜推深度融合
- 数智金融

4月

5月

6月

8月

10月

12月

📍 上海

AiCon

全球人工智能开发与应用大会

会议时间：5月23-24日

- 通用大模型
- AI Agent
- 垂直领域应用
- 模型可解释性

📍 深圳

AiCon

全球人工智能开发与应用大会

会议时间：8月22-23日

- 模型效率与部署
- 多模态大模型
- 大模型安全
- 智能硬件

📍 北京

AiCon

全球人工智能开发与应用大会

会议时间：12月19-20日

- 通用大模型
- 智能硬件
- LMOps
- 具身智能

极客邦科技 2025 年会议规划

促进软件开发及相关领域知识与创新的传播



参会咨询



查看会议

THANKS

大模型正在重新定义软件

Large Language Model Is Redefining The Software

