

AIOps智能化数据体系的构建和应用

孟静磊 孙长辉 / 字节跳动可观测基础设施团队

精彩继续！ 更多一线大厂前沿技术案例

📍 北京站

AiCon

全球人工智能与机器学习技术大会

时间：2021年11月5-6日

地点：北京·国际会议中心

扫码查看大会
详情>>



📍 深圳站

ArchSummit

全球架构师峰会

时间：2021年11月12-13日

地点：深圳·大中华喜来登酒店

扫码查看大会
详情>>



📍 深圳站

GMTTC

全球大前端技术大会

时间：2021年11月19-20

地点：深圳·大中华喜来登酒店

扫码查看大会
详情>>



团队介绍-字节跳动可观测基础设施团队

- 隶属于字节跳动--基础架构部
- 覆盖字节跳动后端可观测性数据源（指标，调用链路，日志，事件）
- 提供监控、链路分析、日志检索、报警、排障归因等能力
- 提供一站式的后端可观测性平台

可观测性平台

可观测性 是基于“**事先未定义**”的模式探索
来帮助团队有效“**调试**”其系统

希望被回答的问题

- 服务部署是否合理，是否有跨机房流量？
- 服务调用链深度是多少，有没有优化空间？
- 针对海量服务黄金指标，报警太多怎么办？
- 服务出错了，是资源负载？下游问题？发布变更？

目录

- 智能化数据体系构建的重要性和挑战
- 智能化场景（成本运营、智能报警&归因修复）
- 智能化数据体系构建实践
- 智能化数据体系未来展望

智能运维实施关键技术

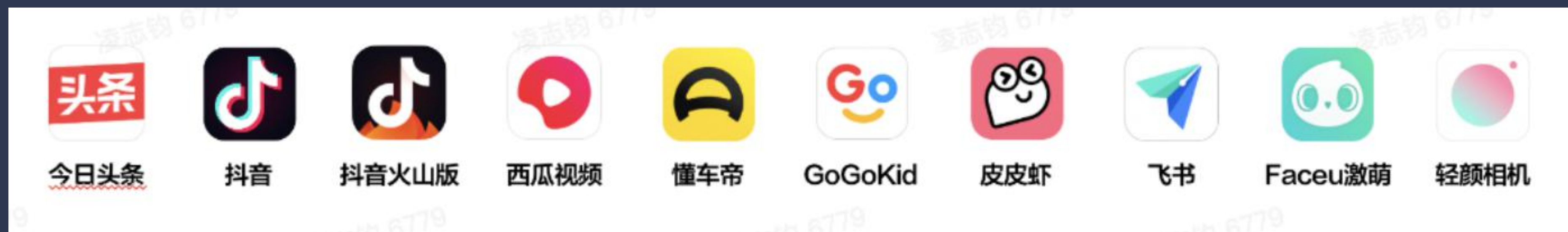
- 数据是基础和关键

- ✓ 接入
- ✓ 计算
- ✓ 存储
- ✓ 管理和治理
- ✓ 分析
- ✓ 建模



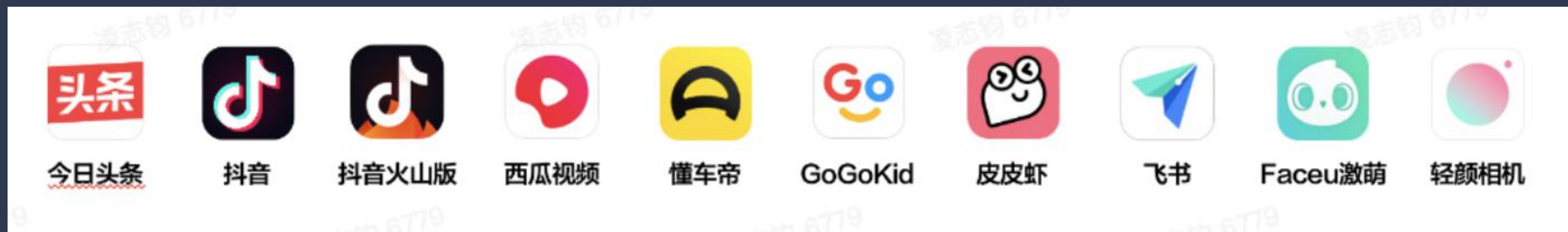
数据化智能体系 – 业务挑战 (1)

- 服务集群: 200K+
- 全网装机量: 几十万
- AB实验: 日均1K
- 服务发布: 日均15K
- 在线微服务装机量: 1/3总量
- mesh流量调度: 日均3K
- 配置变更: 日均35K
- 离线大数据装机量: 1/3总量
- 引流操作: 日均4K
- 总微服务实例数: 10M+



数据化智能体系 – 业务挑战 (2)

- | | | | |
|---------------|-------|-----------------|-----|
| • 后端主流语言: | 4种 | • RPC框架: | 10+ |
| • 用户中心QPS: | 20M+ | • 抖音feed调用服务: | 500 |
| • 抖音feed链路深度: | >50 | • 用户中心容器数量: | 20K |
| • 用户中心P99延迟: | 10 ms | • 抖音feed调用服务接口: | 700 |



数据化智能体系 — 数据挑战

- 业务多样，数据来源多
- 架构庞大
- 技术演进频繁

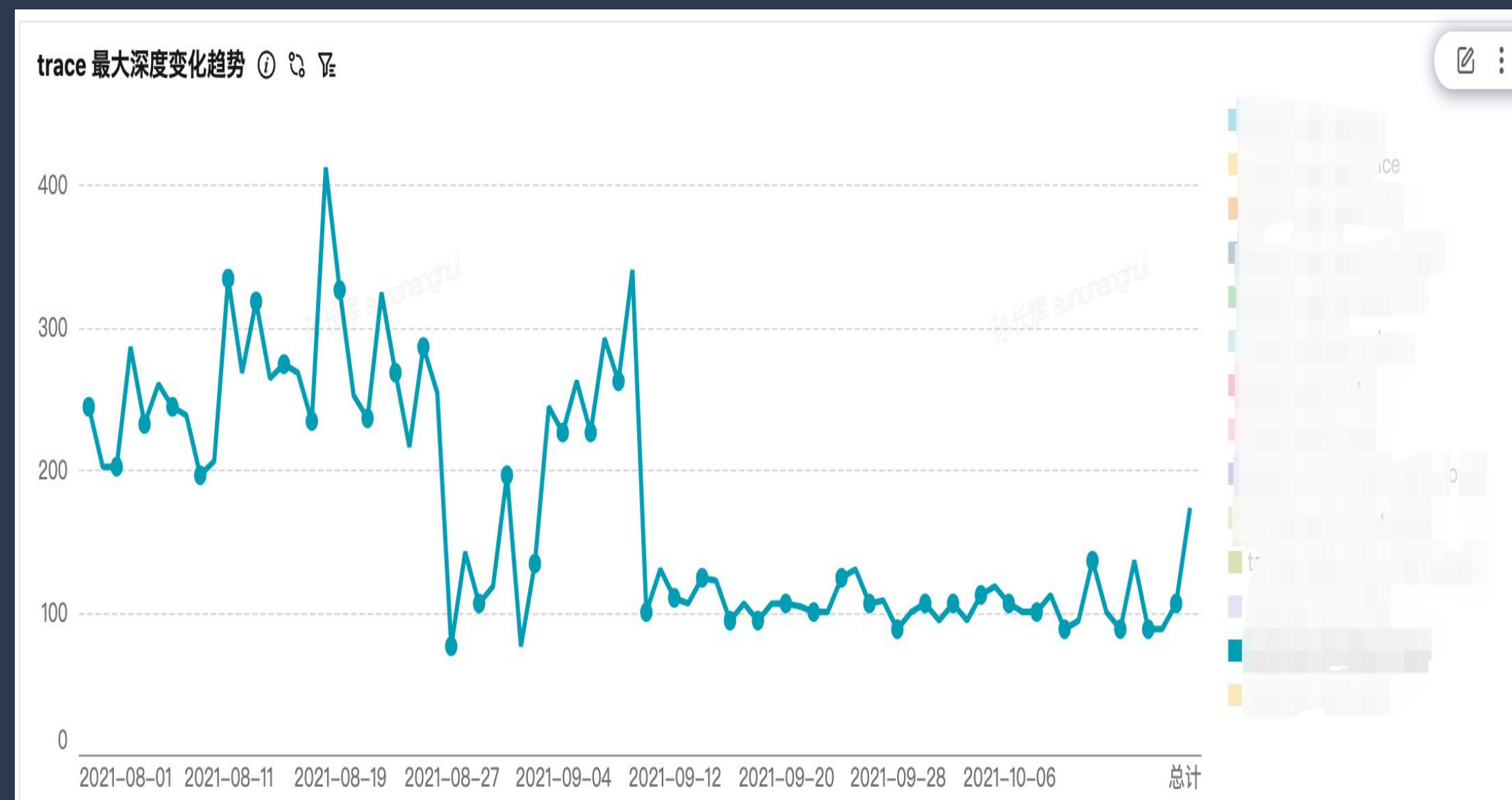


目录

- 智能化数据体系构建的重要性和挑战
- 智能化场景（成本运营、智能报警&归因修复）
- 智能化数据体系构建实践
- 智能化数据体系 未来展望

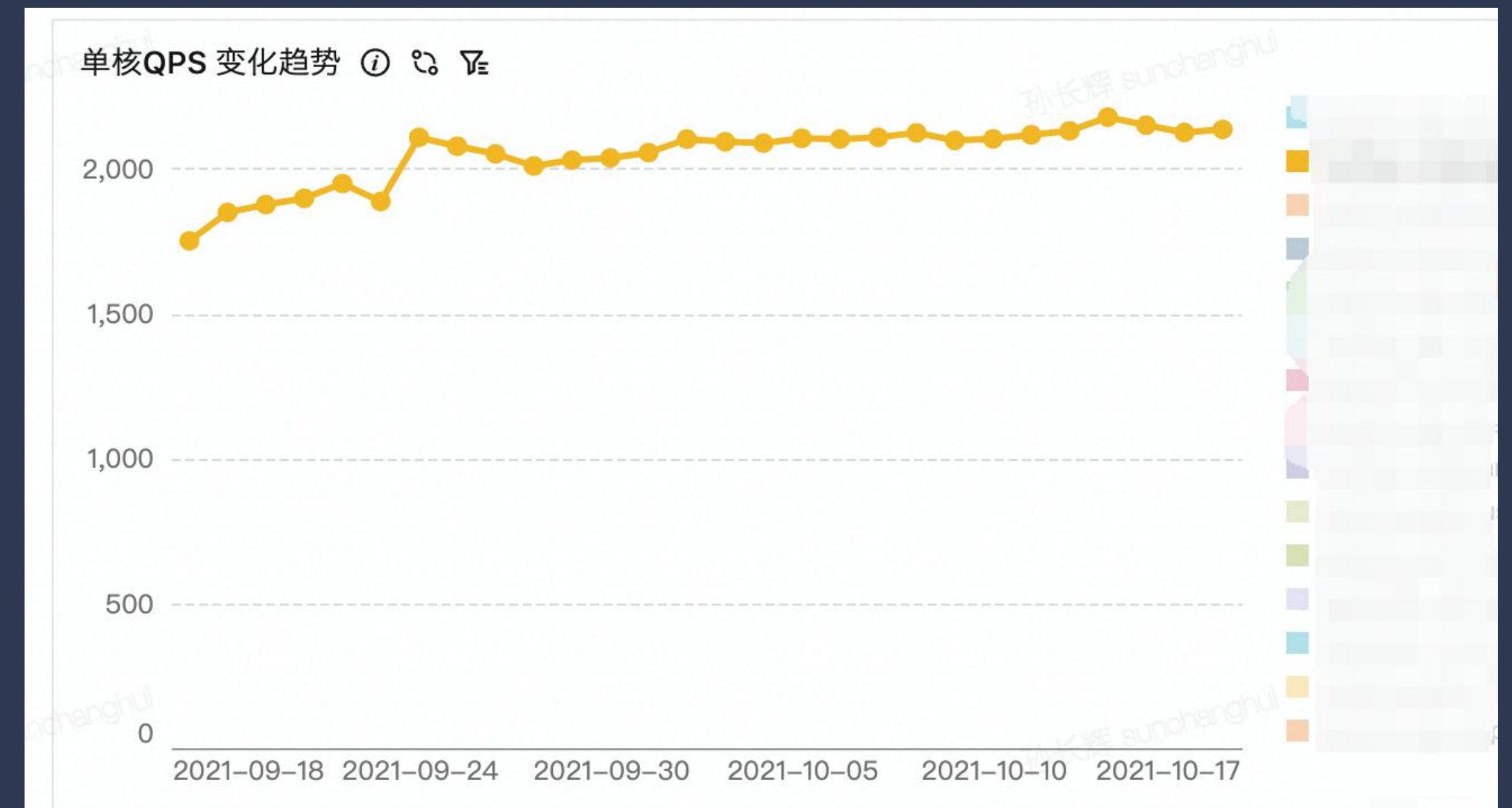
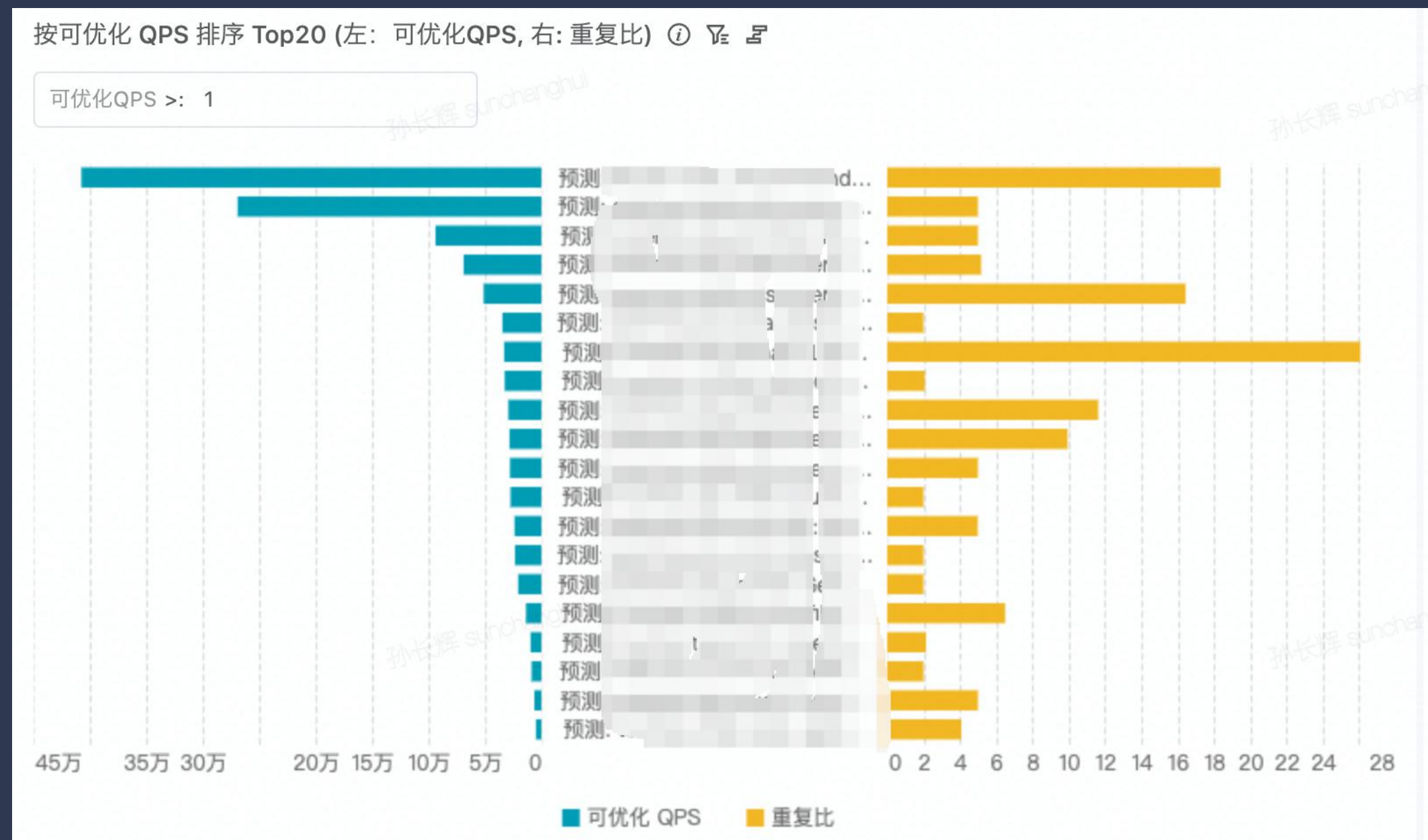
场景：智能运营

- 带宽：跨机房流量过大&服务链路调用深度过深



场景：智能运营

- 性能：服务单核QPS过低



场景：智能报警

- 海量业务黄金指标告警，指标类型多样
- 告警规则多，维护固定阈值成本高
- 各个业务线SLO要求不一
- 对于周期性指标，早晚高峰阈值不固定



场景：智能报警

- 自适应阈值，不用手动配置阈值
- 灵活调整敏感度
- 过滤抖动
- 使用智能引擎之后，原始报警抑制比例 50% ~ 80%
- 报警通知抑制比例最高为70%



场景：归因修复

- 问题平均修复时间MTTR高
 - ✓ 单服务容器运行时问题
 - ✓ 服务黄金指标异常问题
 - ✓ 服务上下游关联问题

场景：归因修复

- 针对服务单实例的诊断每天10W次左右
- 迁移成功次数在1K次左右

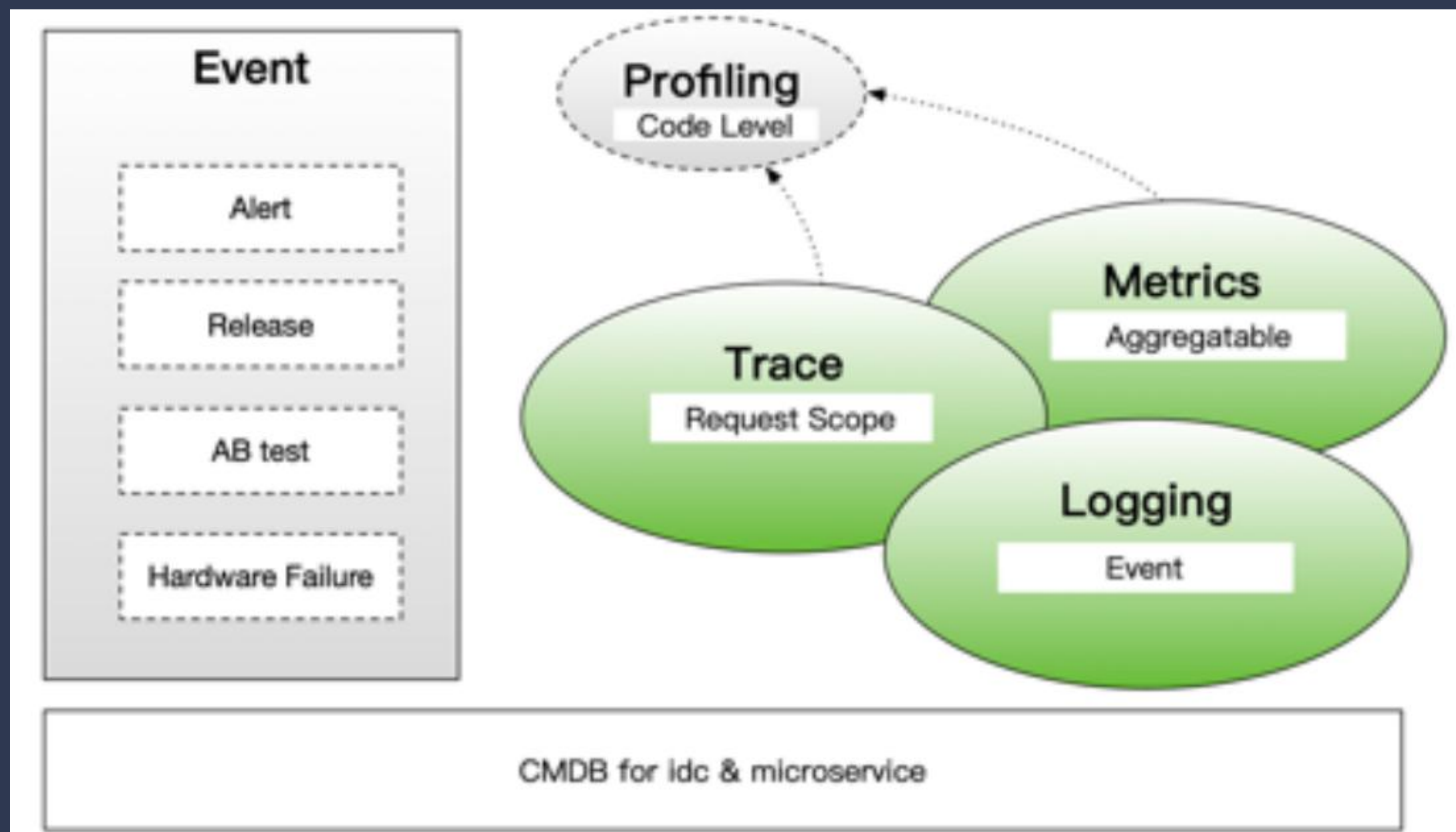


目录

- 智能化数据体系构建的重要性和挑战
- 智能化场景（成本运营、智能报警&归因修复）
- 智能化数据体系构建实践
- 智能化数据体系未来展望

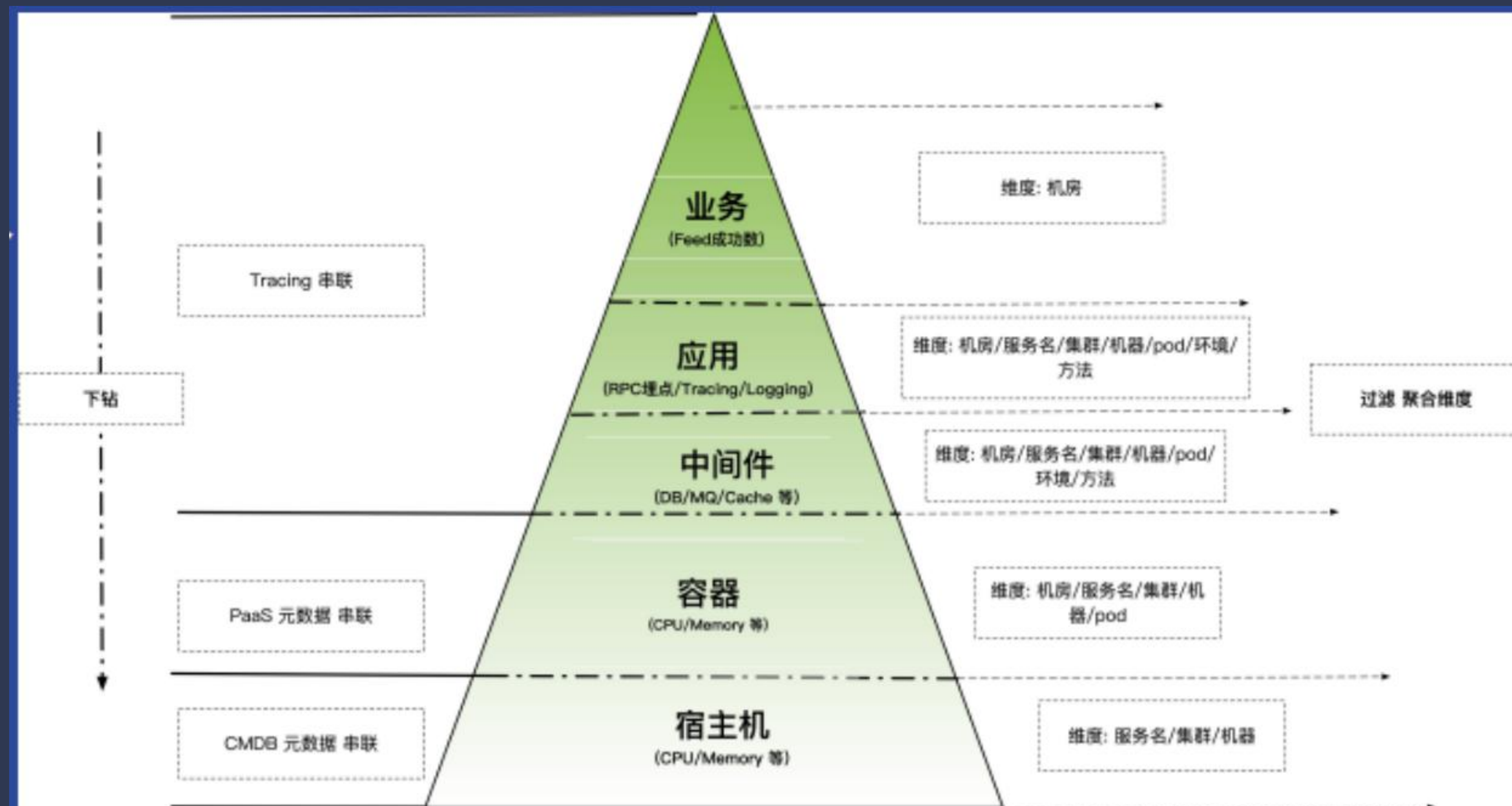
构建：数据来源

- 时序数据
- 日志数据
- 调用链数据
- 事件&CMDB



构建：数据分层&串联

- 业务数据
- 应用数据
- 中间件数据
- 容器数据
- 宿主机数据



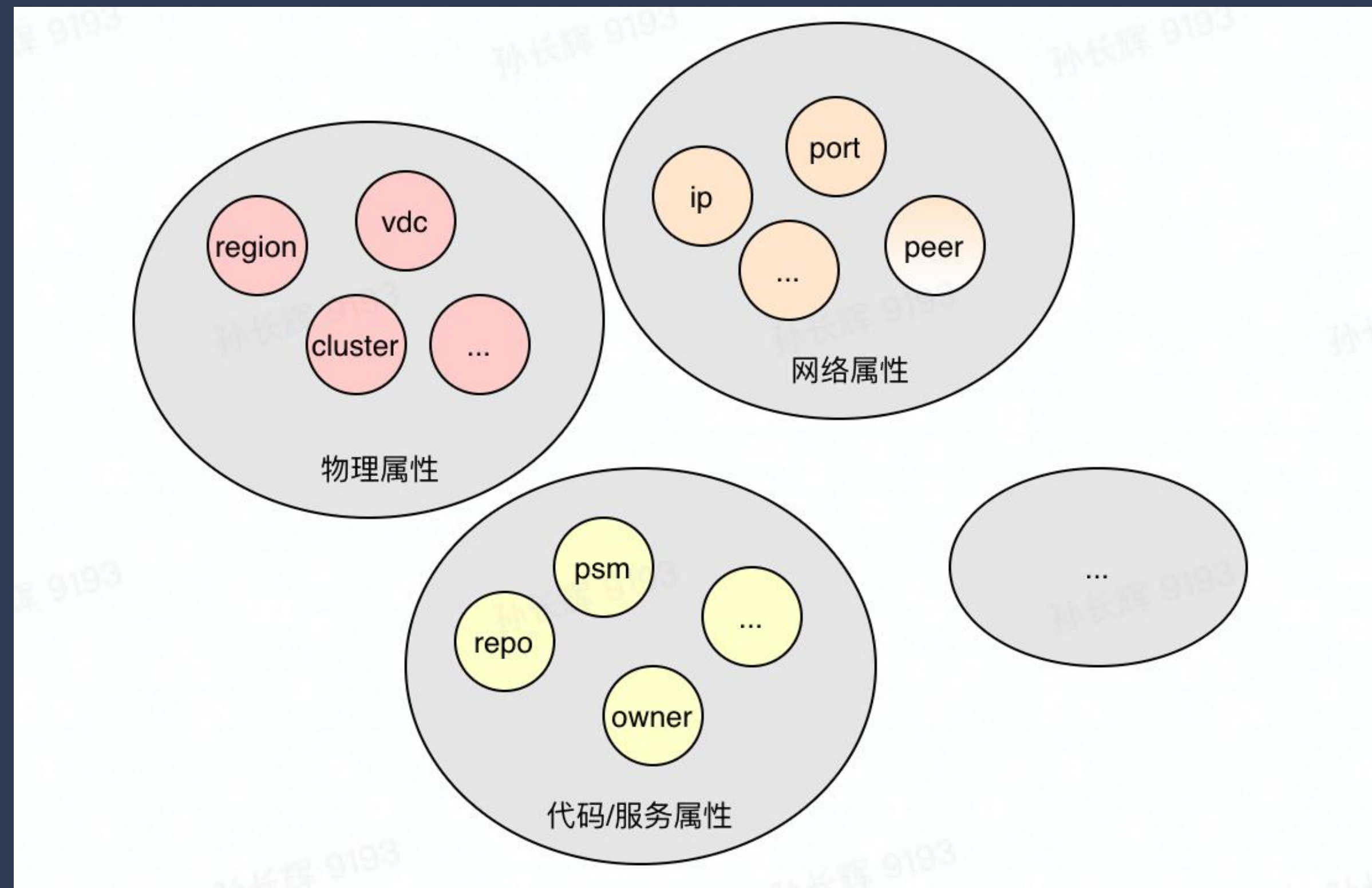
构建： 数据分类

- 性能数据： 延迟，流量
- 质量数据： SLO，稳定性指标，错误率，跨机房流量，调用链深度
- 运行时数据：运行时指标（GC，OOM，panic），错误日志
- 事件： 报警，发布，AB测试，硬件故障

构建：数据建模 — 规范

- 基础术语统一

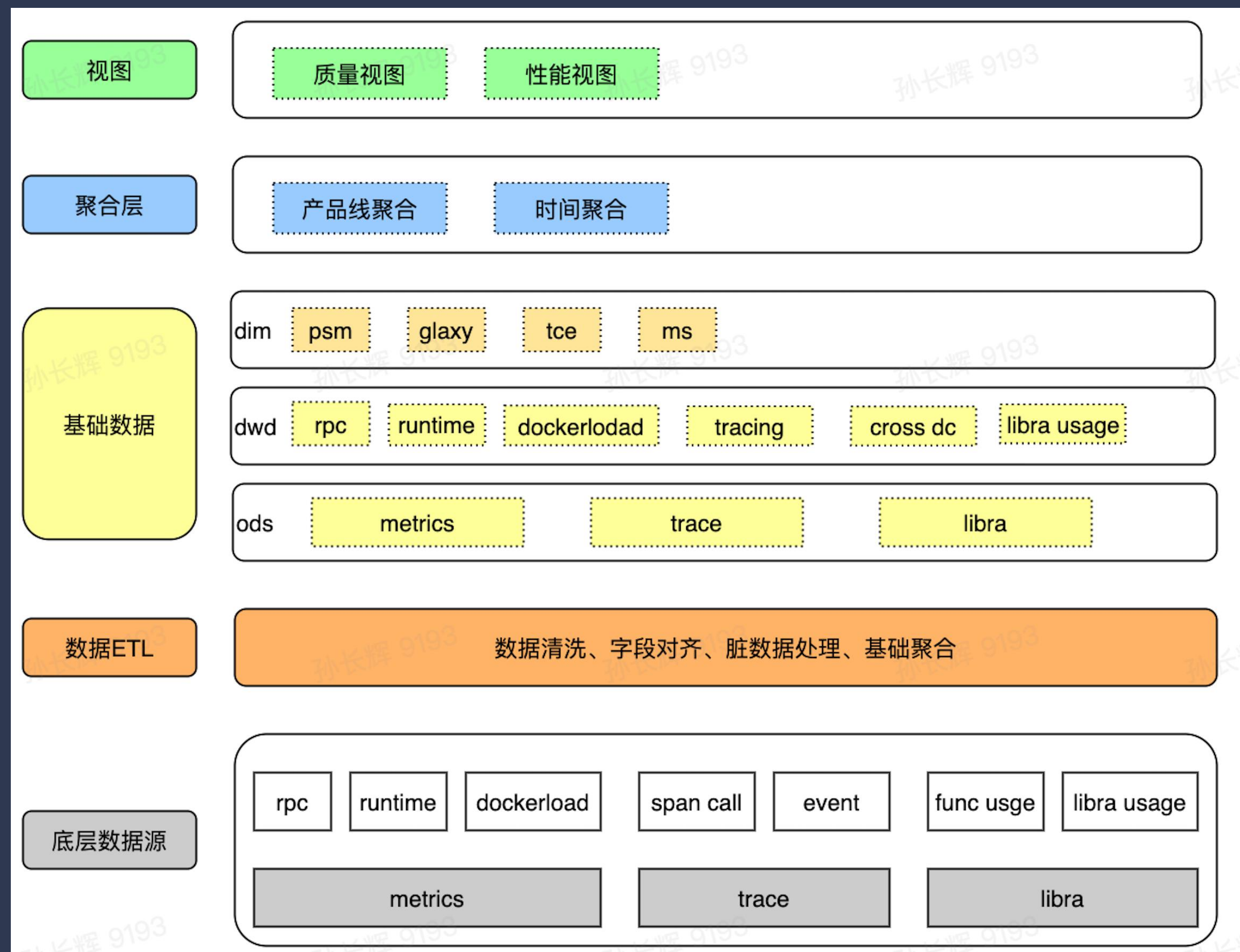
- ✓ 物理属性
- ✓ 网络属性
- ✓ 代码/服务属性
- ✓ 程序运行时属性
- ✓ 统计属性
- ✓ 时间属性



构建：数据建模 - 主题划分

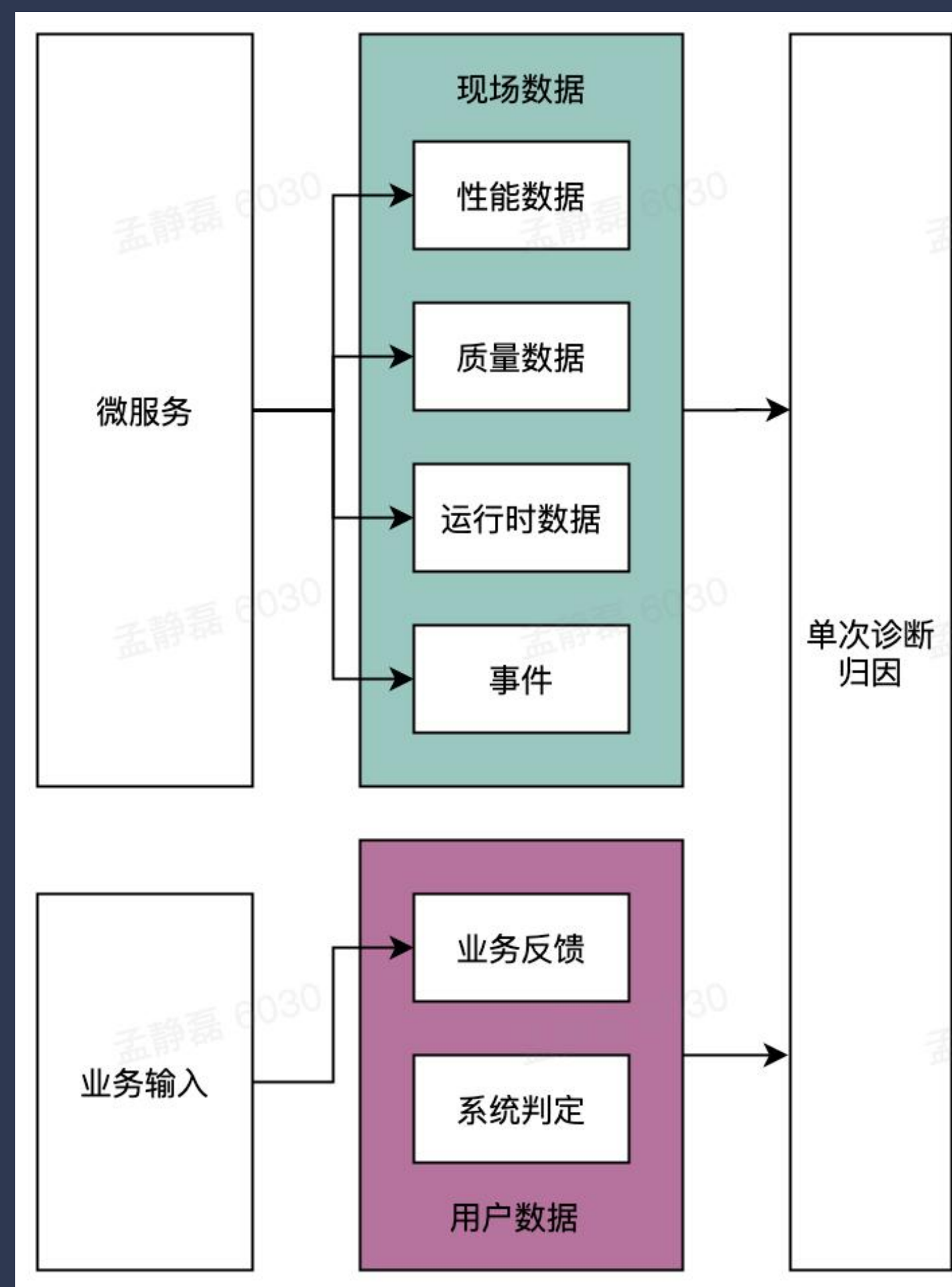
微服务黄金指标 RED & 运行时指标	microservice	微服务的延迟、错误率、流量以及 GC & GoRoutine 指标
微服务容器负载	containerLoad	微服务的容器负载 (CPU usage, Memory usage, Network, DiskIO)
链路分析	tracing	链路的深度 & 最大贡献率的节点
服务 代码热点	hotspot	服务在性能快照下的代码热点

构建：数据建模 - 分层建模



构建：数据建模 - 样本库数据

- 以单次诊断ID串联
- 服务链路上的多个服务共同构成现成数据



构建：策略算法 - 异常检测

- 离线训练（增量训练）

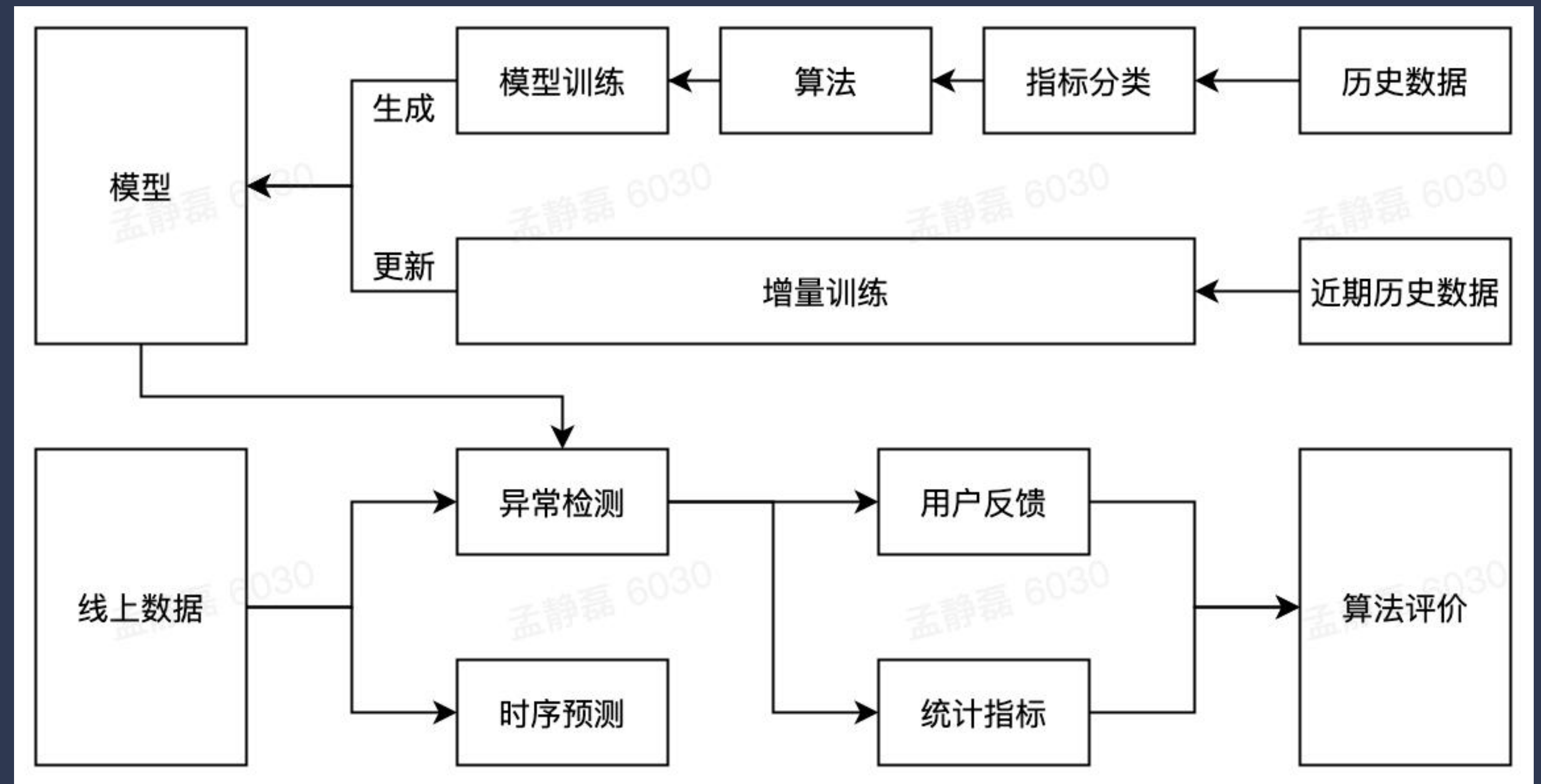
- ✓ 历史数据生成模型
- ✓ 近期数据更新模型

- 在线检测

- ✓ 异常检测
- ✓ 时序预测

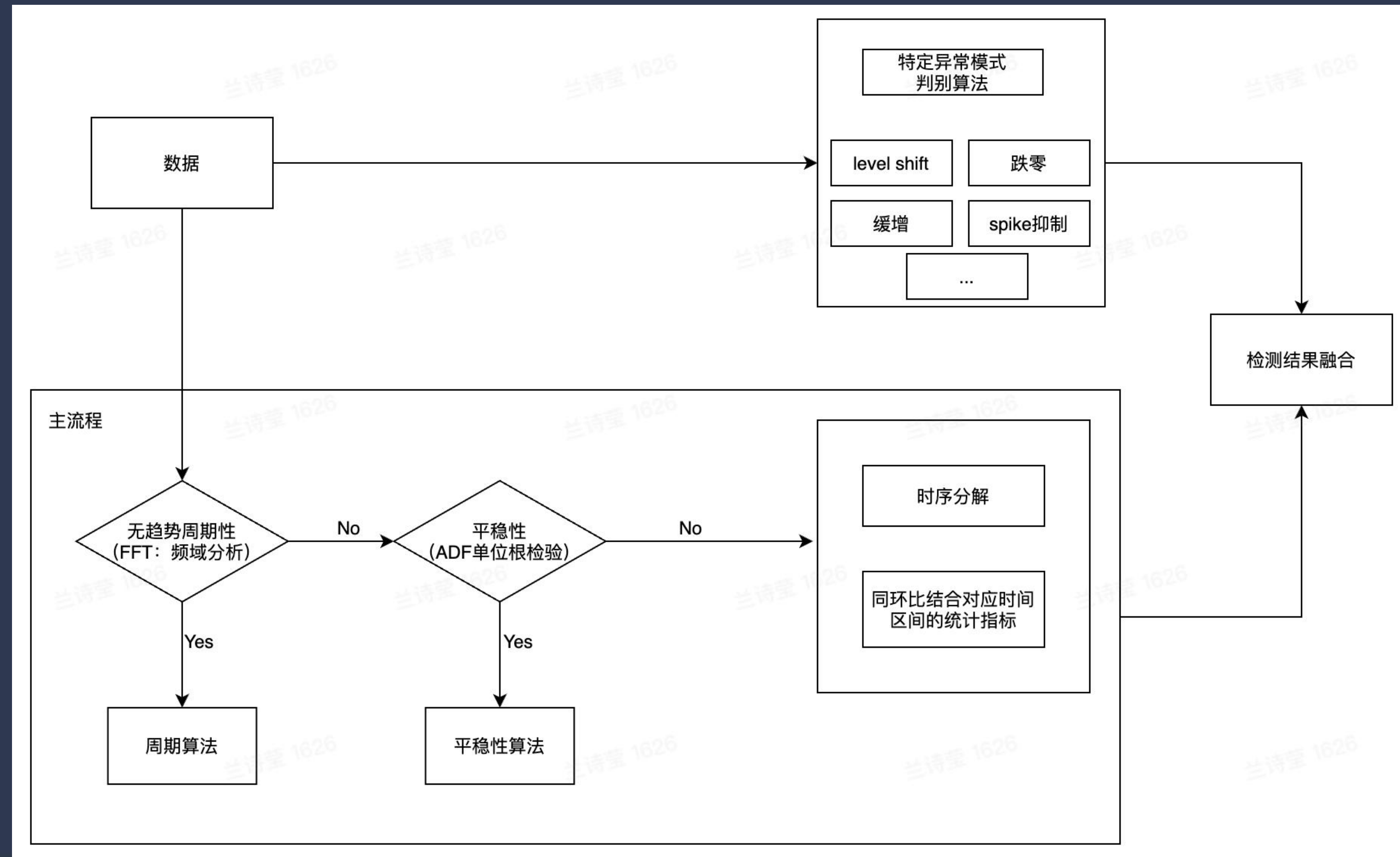
- 算法评估

- ✓ 用户反馈
- ✓ 指标统计



构建：策略算法-异常检测

- 从技术角度：主流程
- 从业务角度：特定异常



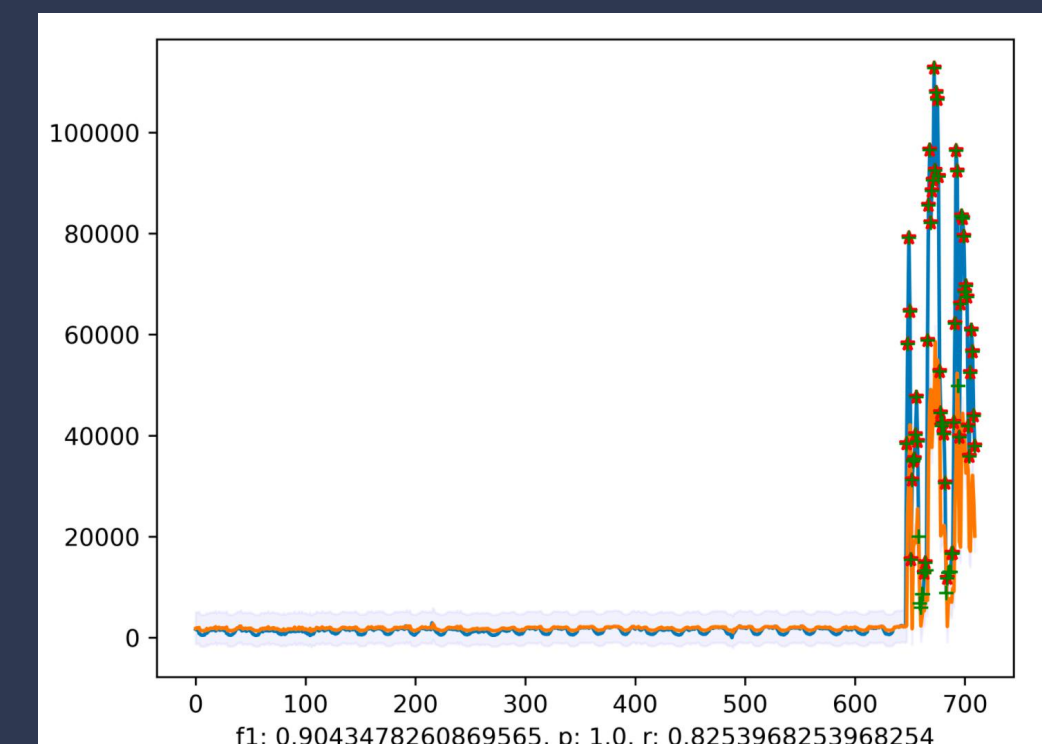
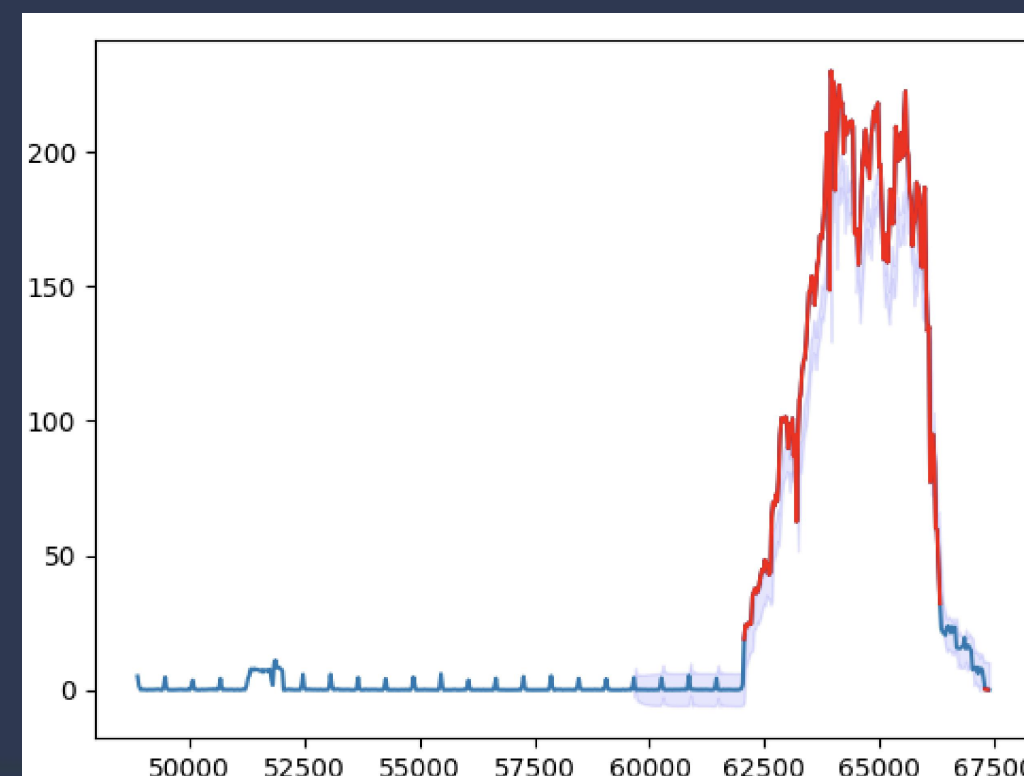
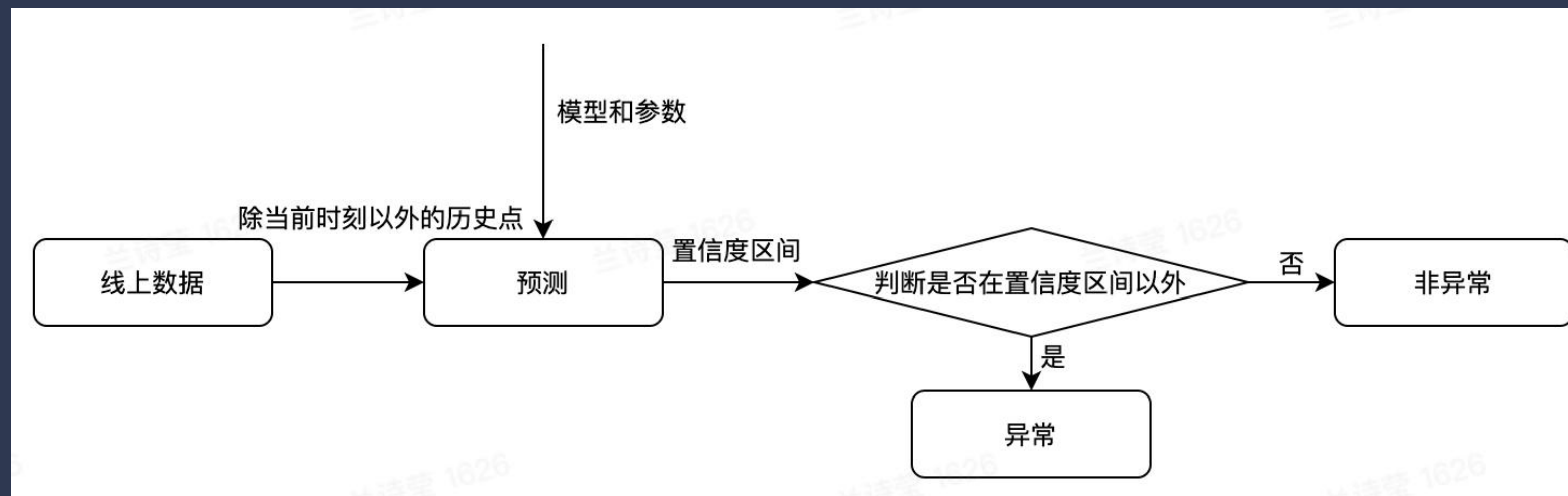
算法：平稳时序异常检测-ARIMA

优点

- 模型简单
- 只需要内生变量，无需手动调参

缺点

- 要求时序是平稳的
- 本质上只能捕捉线性关系



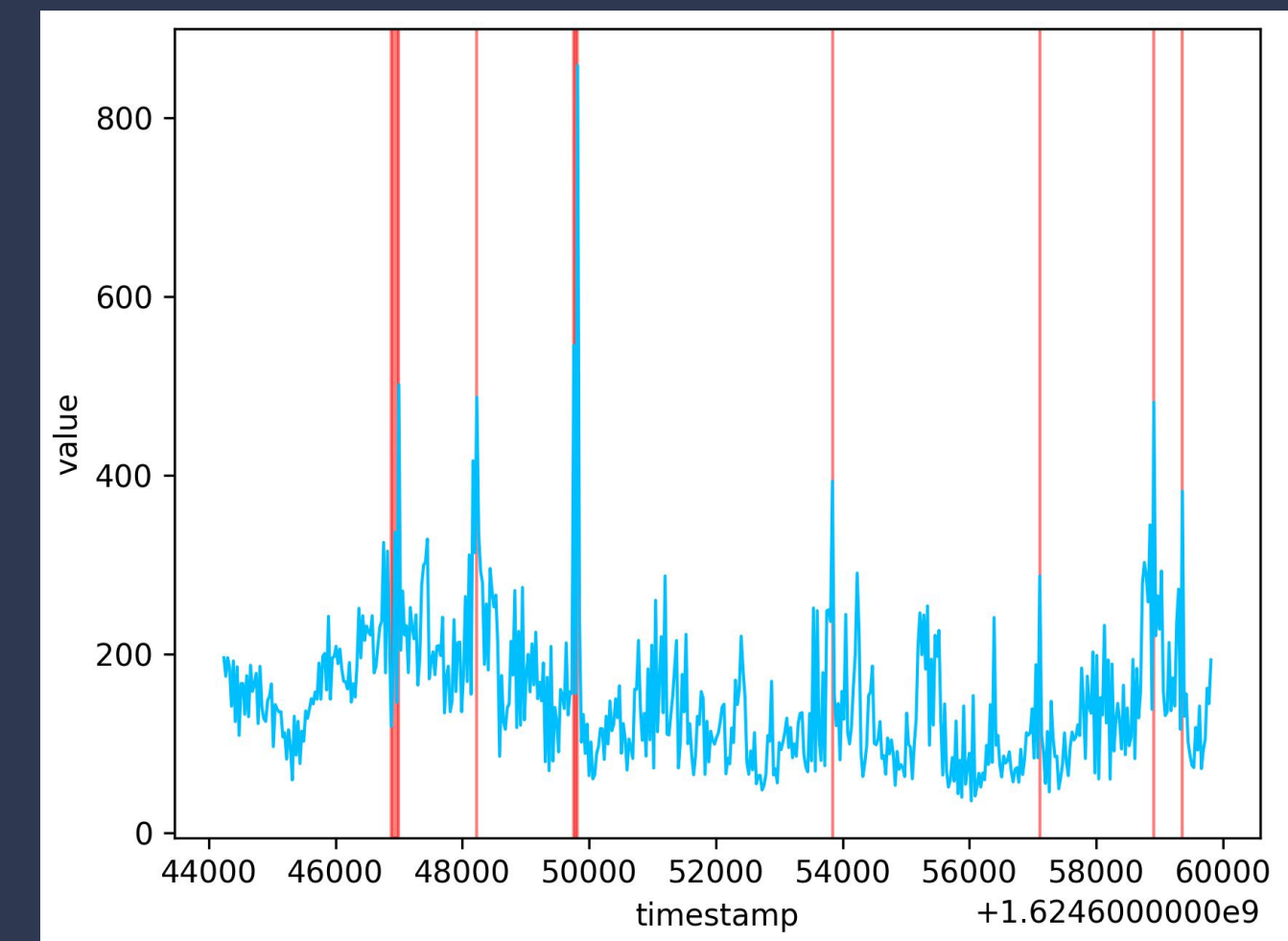
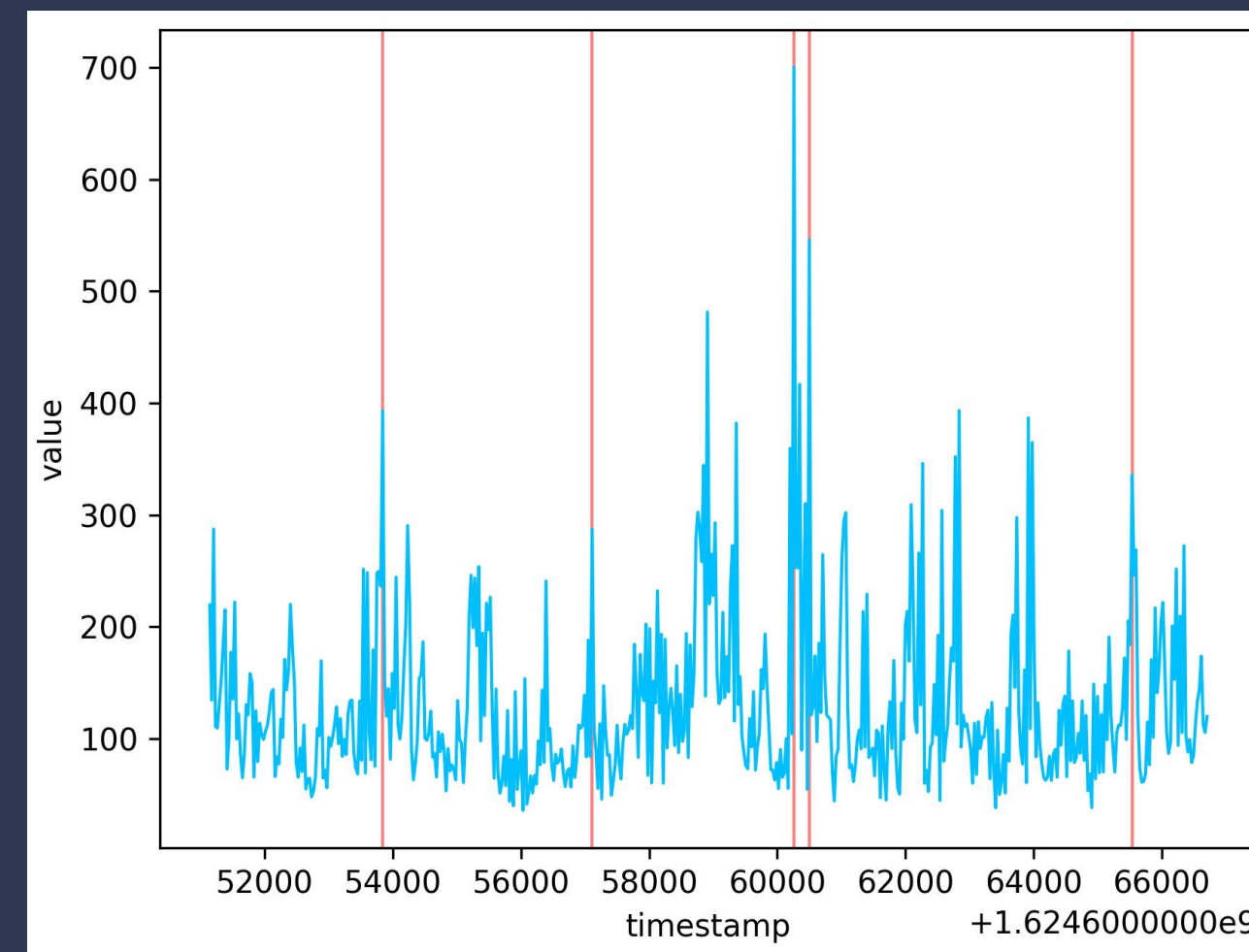
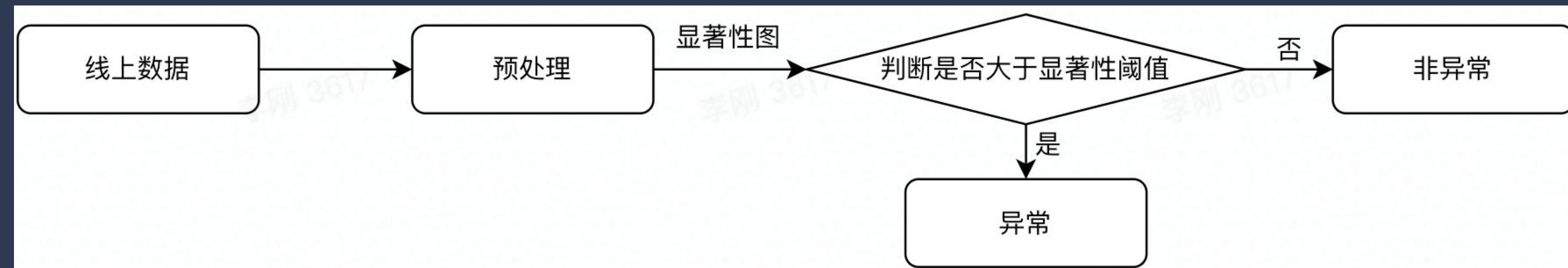
算法：无趋势周期检测算法SR

优点：

- 模型简单，适合于在线检测，对于无趋势周期性数据检测效果较好

缺点：

- 对算法的参数设置要求高



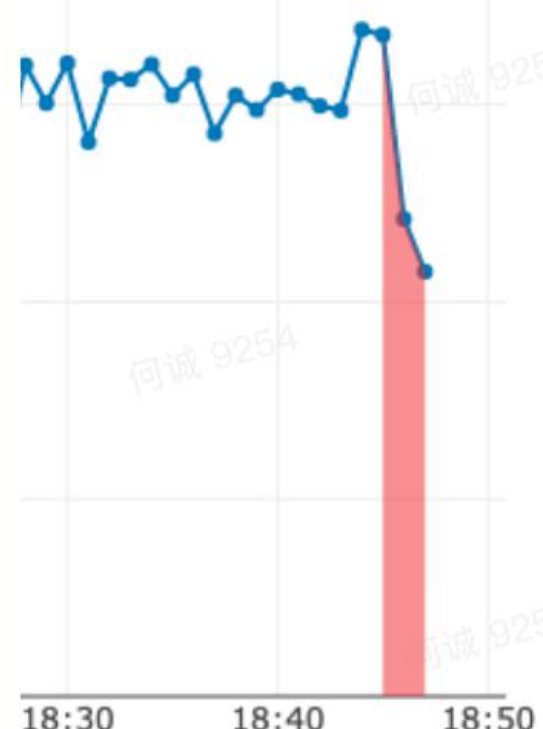
算法：异常模式识别

突然上升 (Rise)



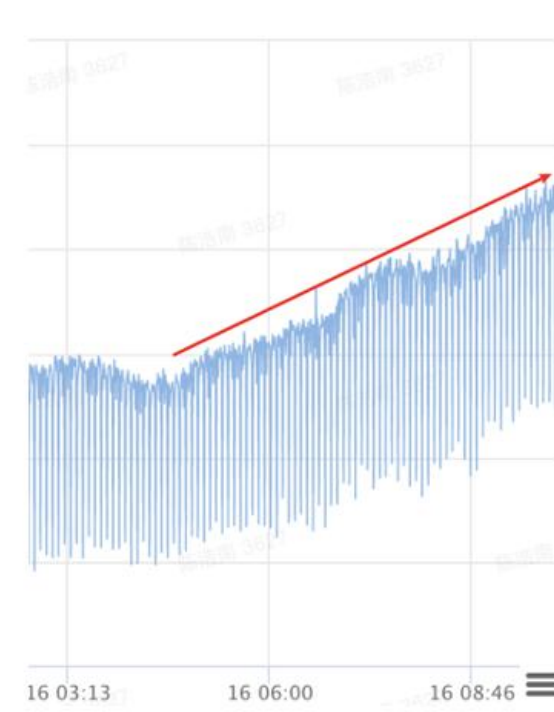
- 相对于正常值突然升高

陡降 (Drop)



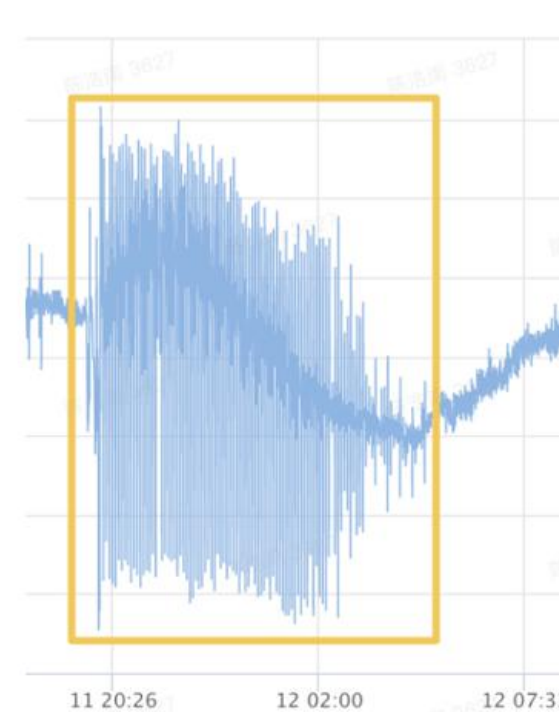
- 相对于正常值突然下降

缓增 (Increase)



- 时间窗口内持续上升

波动 (Fluctuation)



- 时间窗口内方差剧烈变化

突变恢复 (Spike)



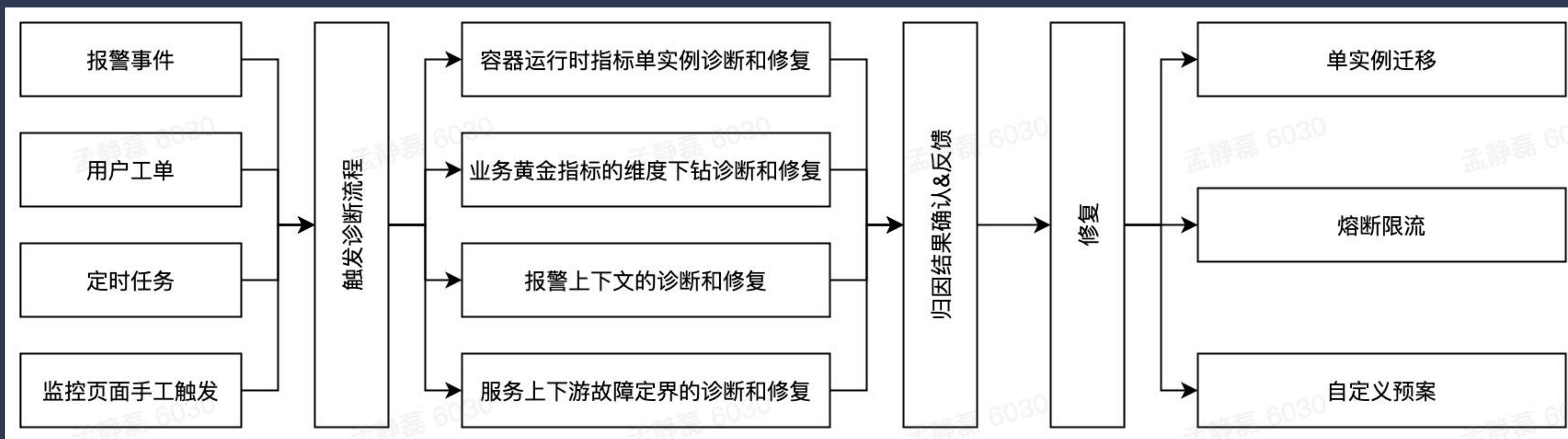
- 时间窗口内突变并且恢复 (包括向上/向下突变后快速恢复)

跌零 (Drop-to-0)



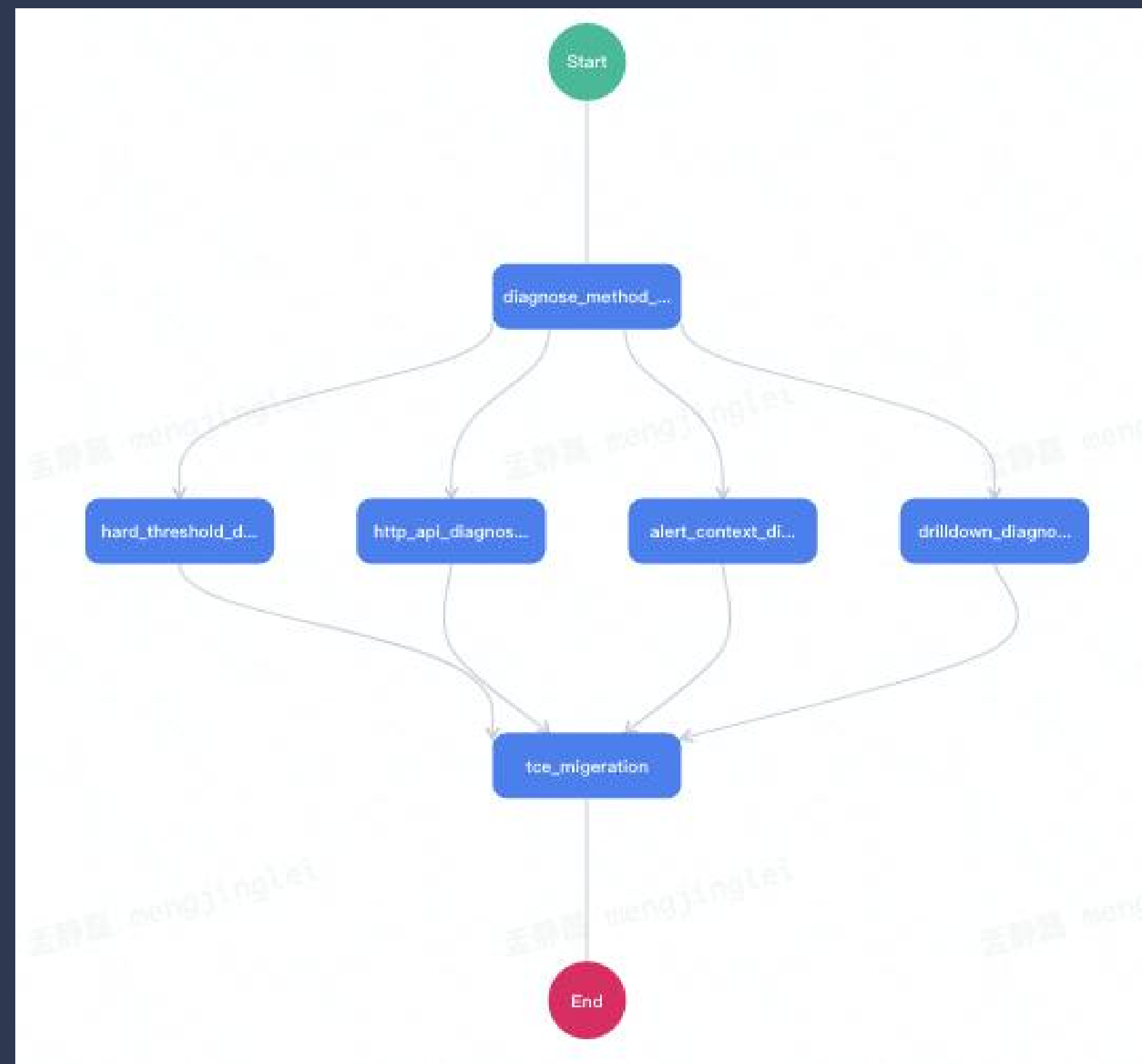
- 监控数值快速跌0

构建：策略算法—归因修复



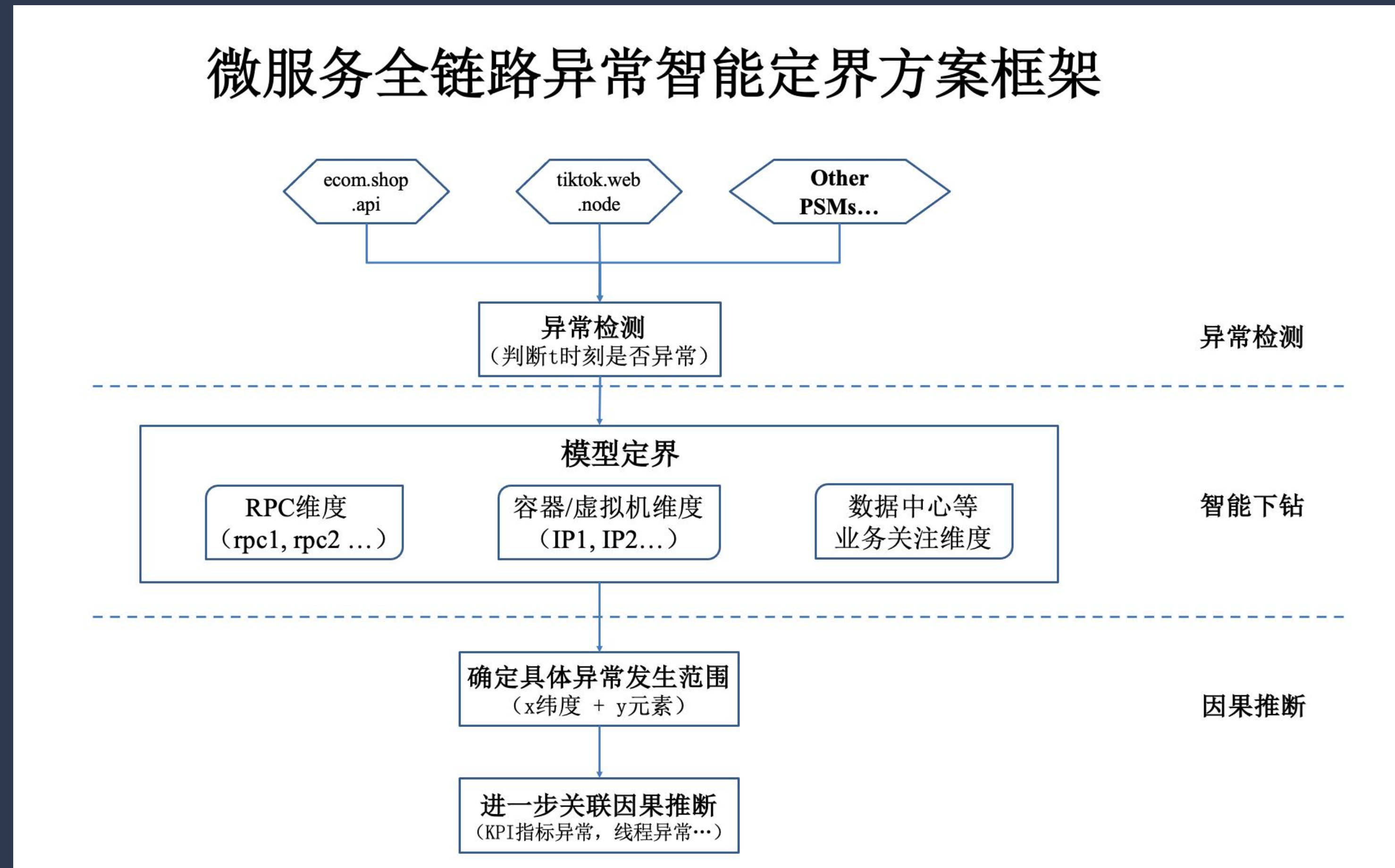
构建：策略算法—可编辑的工作流

- 可编辑工作流
- 静态拓扑沉淀专家知识
- 原子能力节点化（异常检测，关联分析，维度下钻）
- 输入输出标准化
- 支持原子工作流（工作流原子化）



算法：微服务异常智能定界方案

- 异常检测
- 维度下钻
- 因果推断



目录

- 智能化数据体系构建的重要性和挑战
- 智能化场景（成本运营、智能报警&归因修复）
- 智能化数据体系构建实践
- 智能化数据体系总结和未来展望

总结

- 多层次、类别数据的串联是基础
- 按照业务场景对数据规范化建模是关键
- 成本运营
 - ✓ 解决跨机房带宽问题
 - ✓ 解决服务单核QPS过低问题
- 智能运维
 - ✓ 智能报警解决海量服务、指标的报警风暴问题
 - ✓ 归因分析解决服务单实例故障&上下游归因定界问题

未来展望（1）数据流

- 构建实时（近实时）的数据流
- 预聚合前置到存储，加速计算
- 吞吐量大 VS. 查询效率高：Lambda架构统一存储

未来展望（2）数据平台化（开放赋能）

- Schema和权限管控
- 按需构建Data Cube
- 自定义归因诊断修复流程

未来展望（3）算法离线仿真环境构建

- 环境隔离
- 数据回放
- 模型评估和反馈

Q&A

有任何问题也可以
个人微信离线交流



为一线互联网公司核心技术 人员提供优质内容

☑ TGO专访

☑ 技术干货

☑ 每周精要

☑ 行业趋势



关注 InfoQ 公众号

THANKS

软件正在改变世界

SOFTWARE IS CHANGING THE WORLD

QCon