

流量管控在阿里的 最佳实践

笛 墨（花名）

❖ 限流降级

❖ 异常机器流量调度

❖ 总结

流量洪峰：

- 零点峰值交易创建49.1万笔/秒
- 21秒，成交额突破10亿
- 2分05秒，成交额超100亿

流量洪峰未拦截，可
能导致系统崩溃

TMAIL 天猫

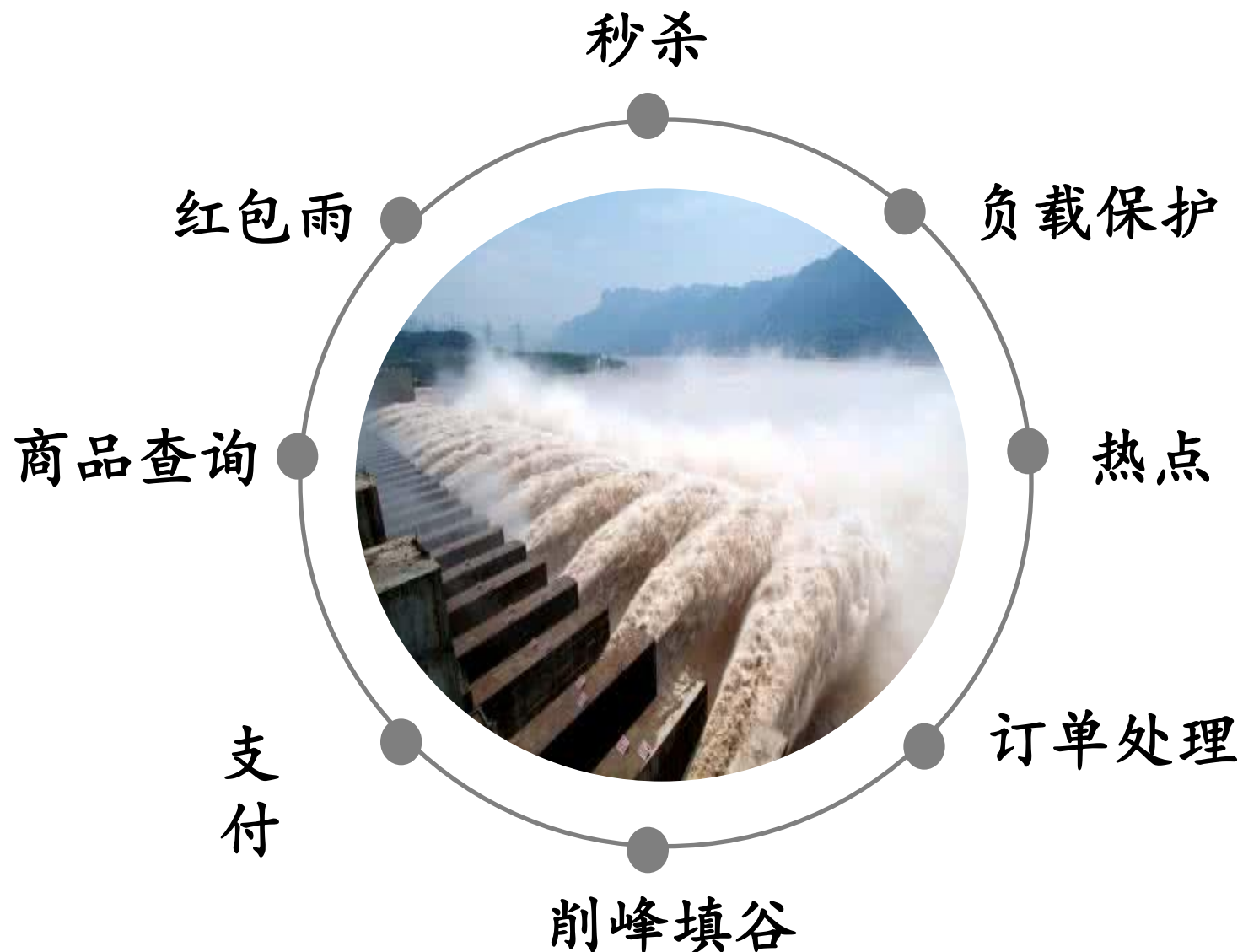
搜索 天猫 商品/品牌/店铺



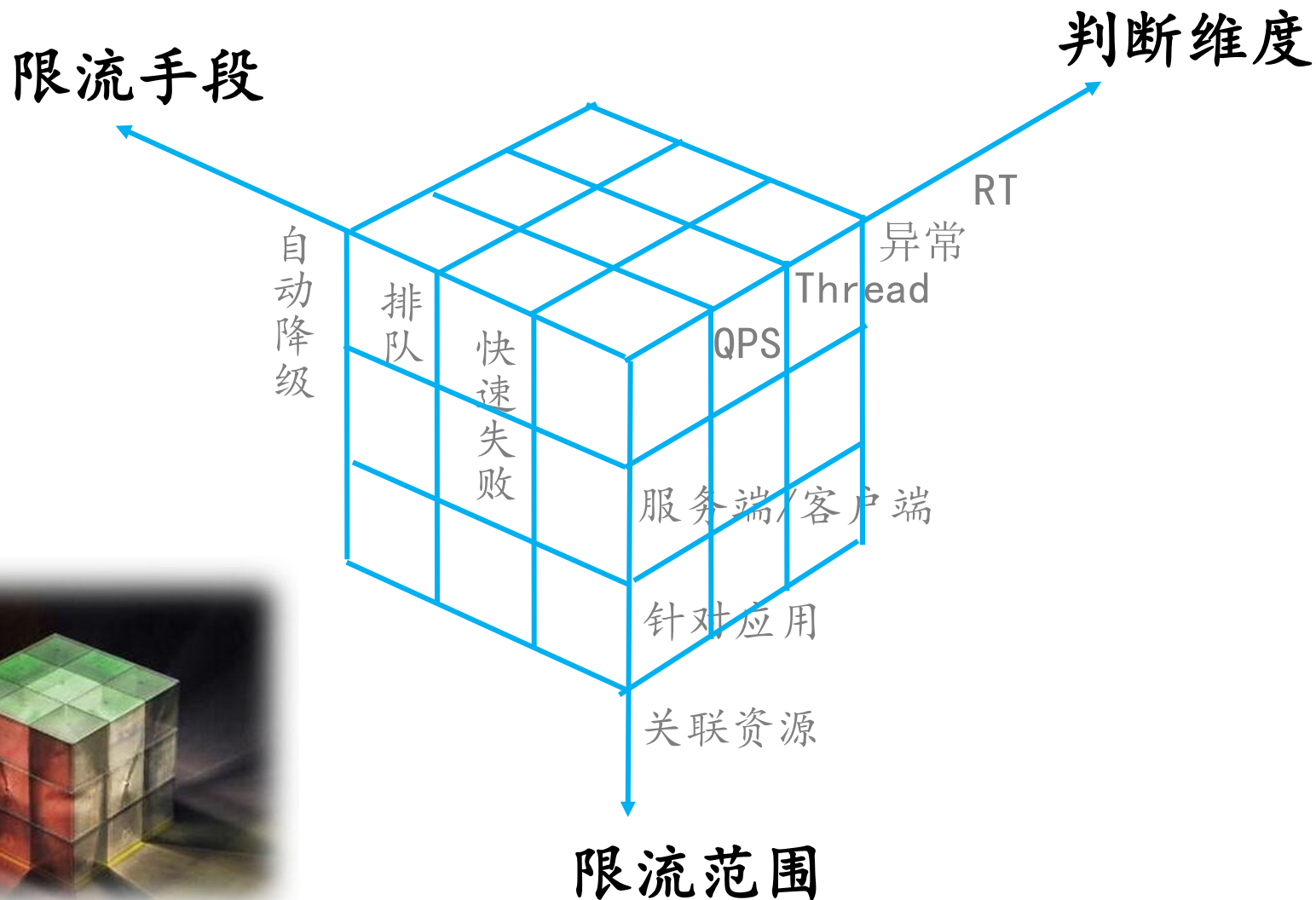
一瞬间，疯起人涌，交通拥堵..
客官别急，马上就好。
我们就算忙到挂水也要让你进去。
2 秒后重试...

购物小贴士

限流场景



限流体系

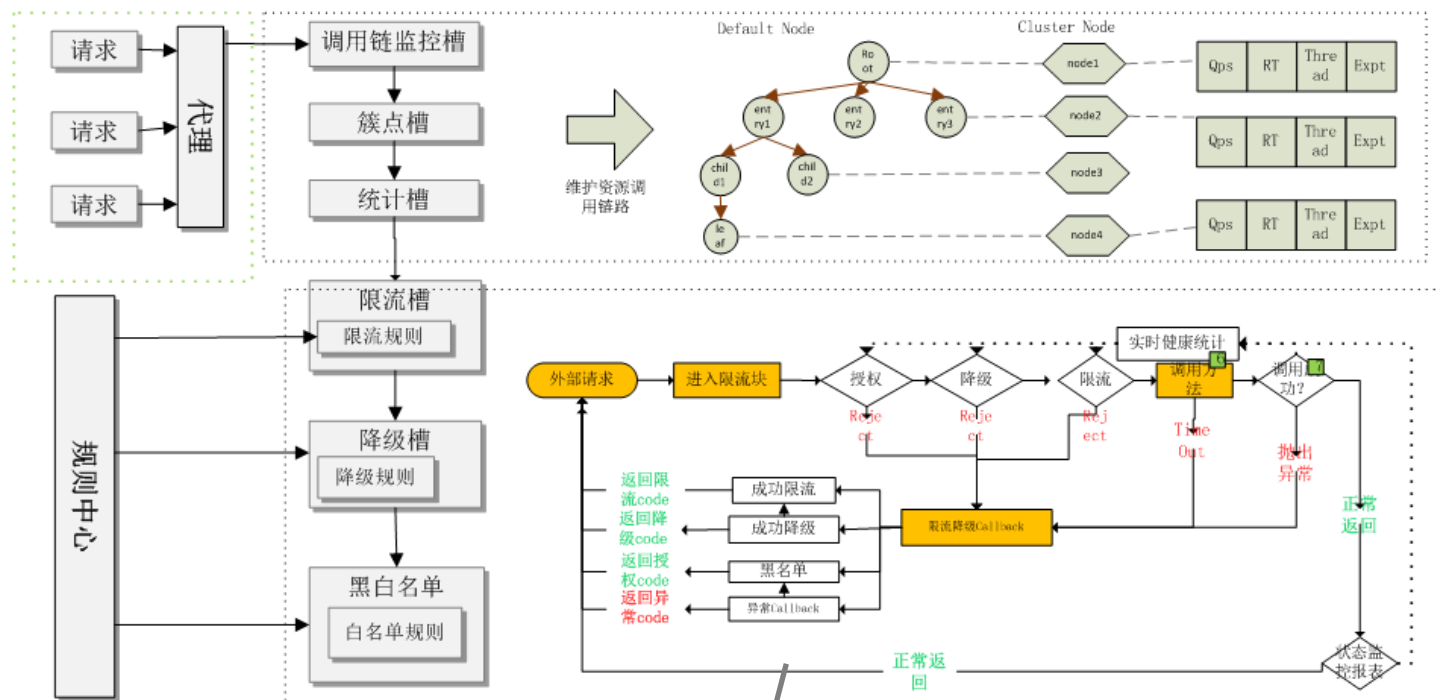


限流体系-整体架构

1. 自动埋点

2. 调用关系

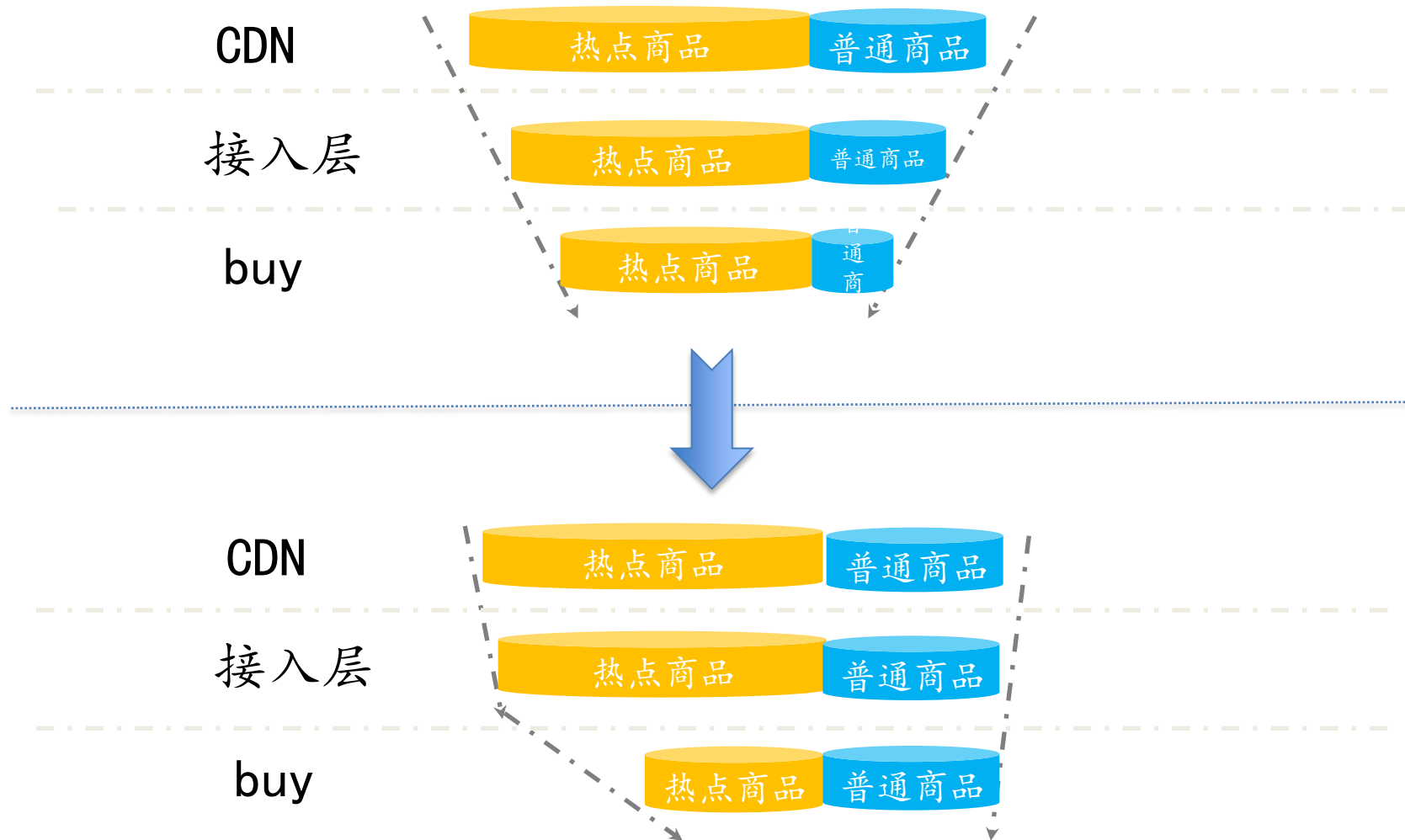
3. 运行状态



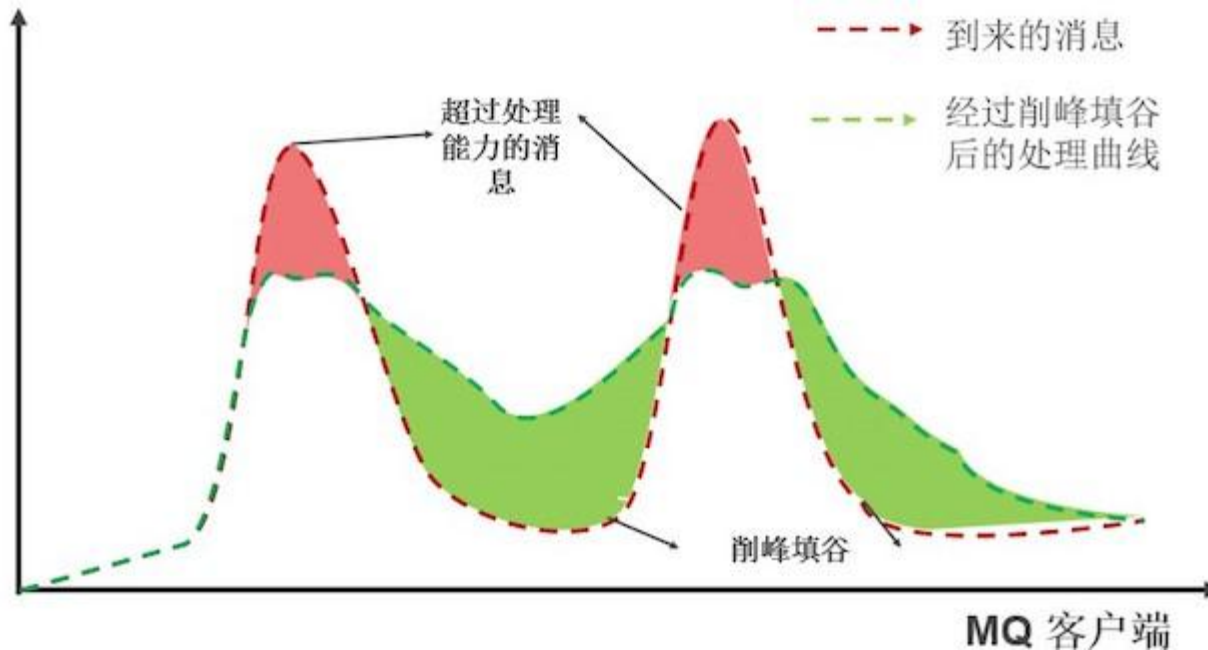
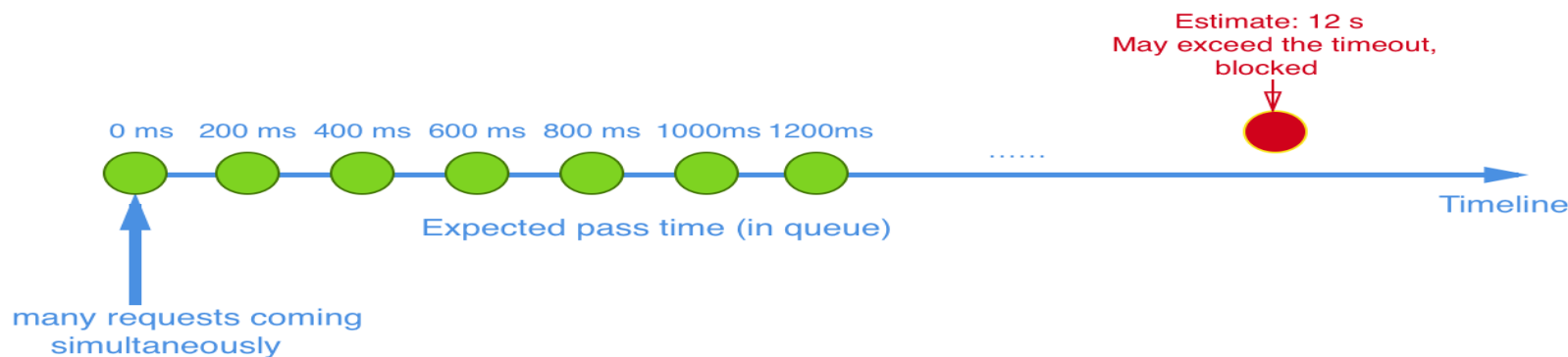
4. 实时生效

5. 采取限流策略

场景 - 热点保护



场景 - 削峰填谷



<https://github.com/alibaba/Sentinel/wiki/Sentinel-%E4%B8%BA-RocketMQ-%E4%BF%9D%E9%A9%BE%E6%8A%A4%E8%88%AA>

在无法预估系统状态的情况下，如何进行系统保护？

1

流量未到预期，系统已经抗不住

2

流量只是相对均匀

3

根据当前系统状态，快速恢复

4

只能保护http

限流降级 - 开源



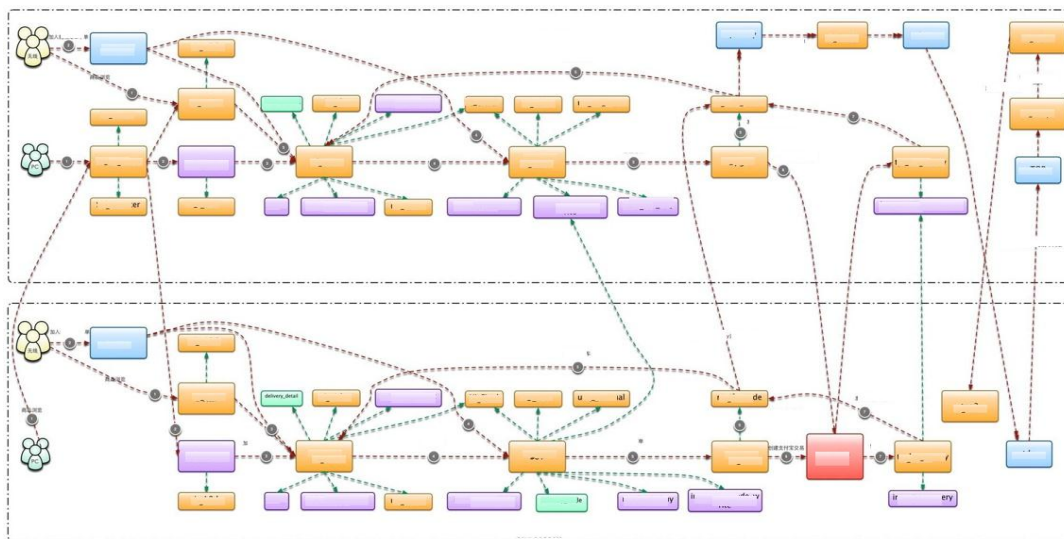
<https://github.com/alibaba/Sentinel>

整体可用性、局部可用性

限流、降级保障
系统整体可用性，
但这就够了吗？

单点、局部问题
影响：

- 调用链路的局部问题会被放大到整个链路
- 单点局部问题会被放大成面



异常机器流量调度 - 局部异常

单点故障:

- 超卖带来的资争抢问题
- 启动时部分机器load飙高
- 网卡问题导致重传高
- 受宿主机影响load飙高
- JVM full gc
- 业务代码bug, CPU打满



分布式环境中，软件、硬件、网络等因素导致机器的实时服务能力有差异

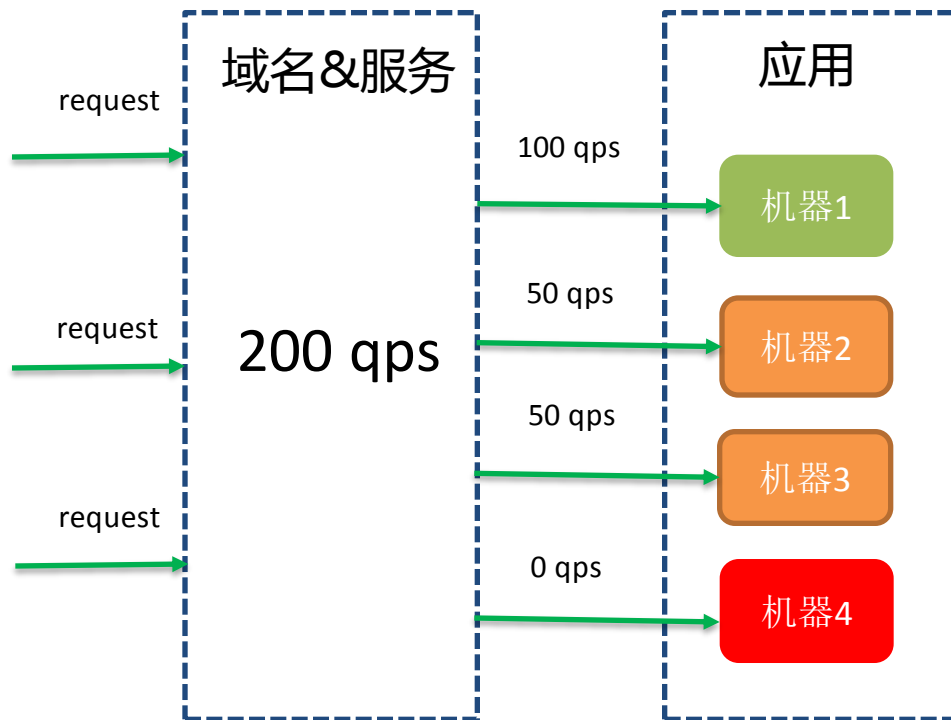


局部异常 - 解法

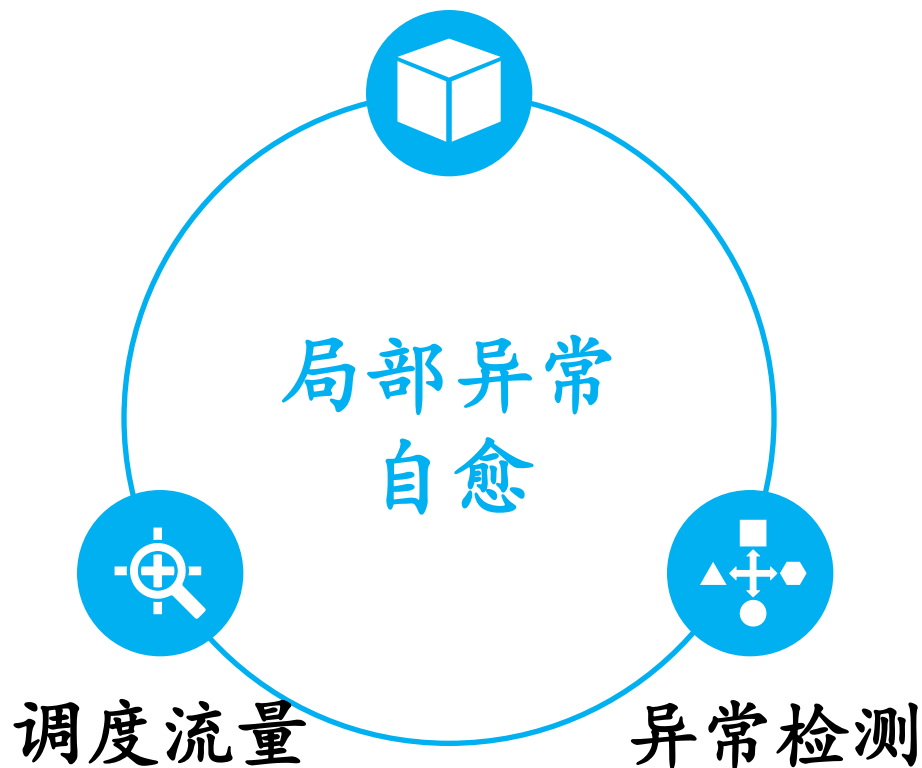
被动解法:

- 配置单机监控
- kill机器
- 重启机器
- 单机系统保护(限流)

主动解法:按机器实时
服务能力调度集群内部
流量



集群状态监控



集群状态监控 - 挑战

指标多、数据量大

系统状态、调用明细
指标要细化到单机粒度
聚合维度多

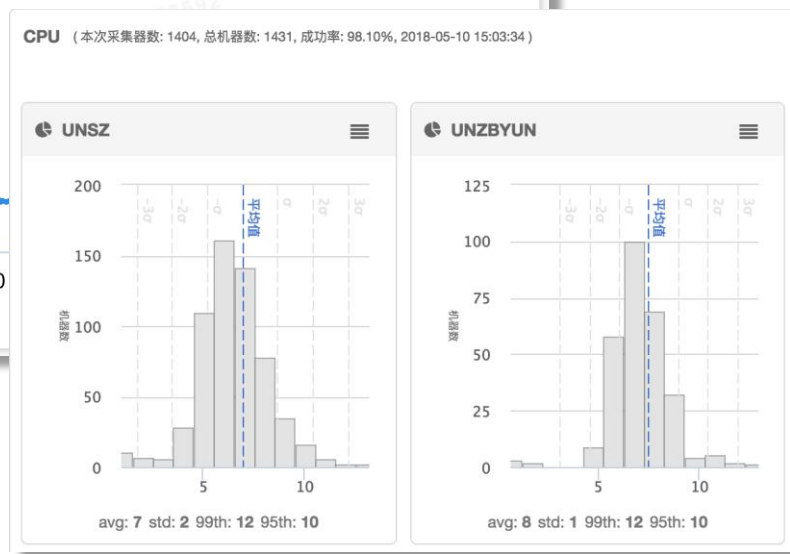


秒级数据、延时低

秒级数据
采集延时要求低

成本

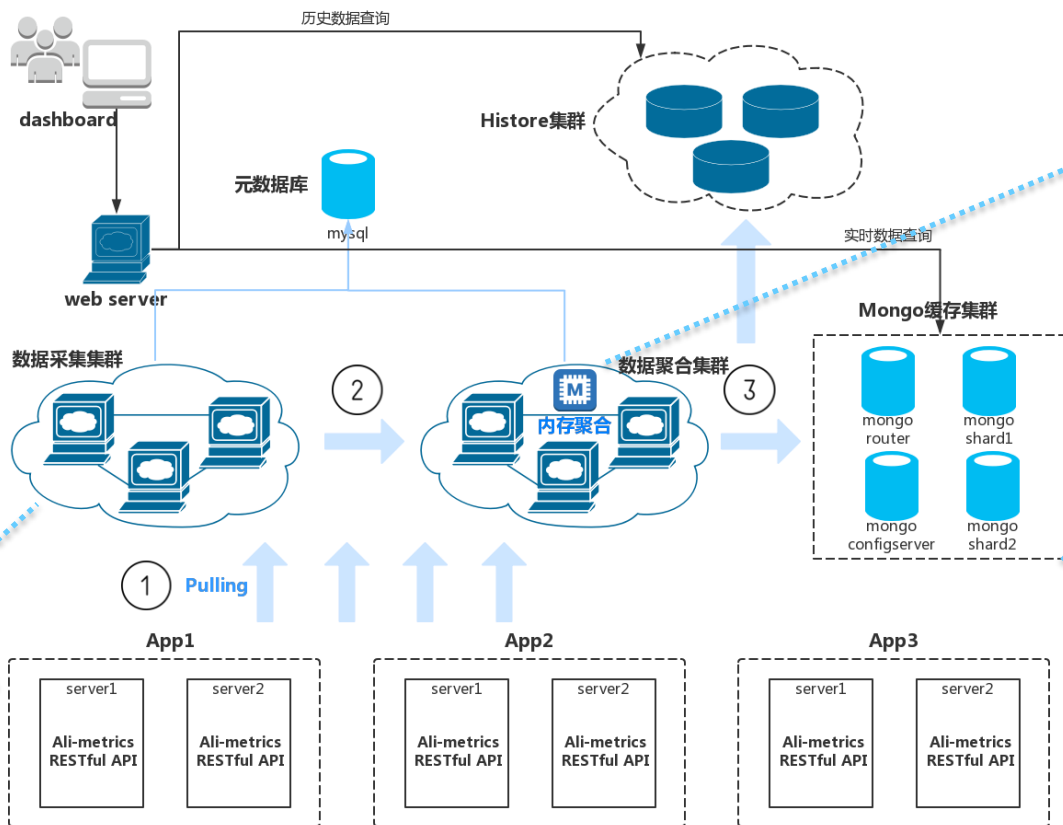
机器成本



集群状态监控 - 系统架构

客户端计算

HTTP 拉模式



内存数据聚合

数据分层存储

集群状态监控 - 数据拉取

自适应任务切分

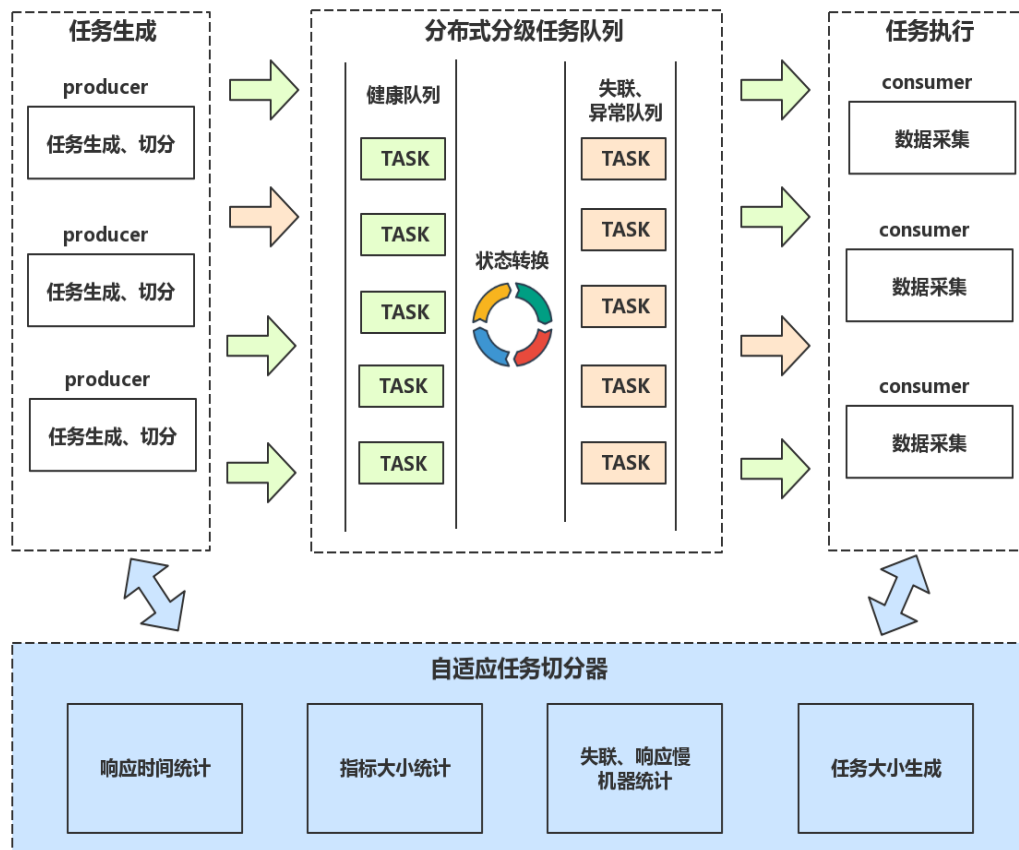
- 根据采集速度、数据大小，自动调整任务大小

分级任务队列

- 任务分级，提升采集速度

无单点、水平扩展

- 集群能力可水平扩展



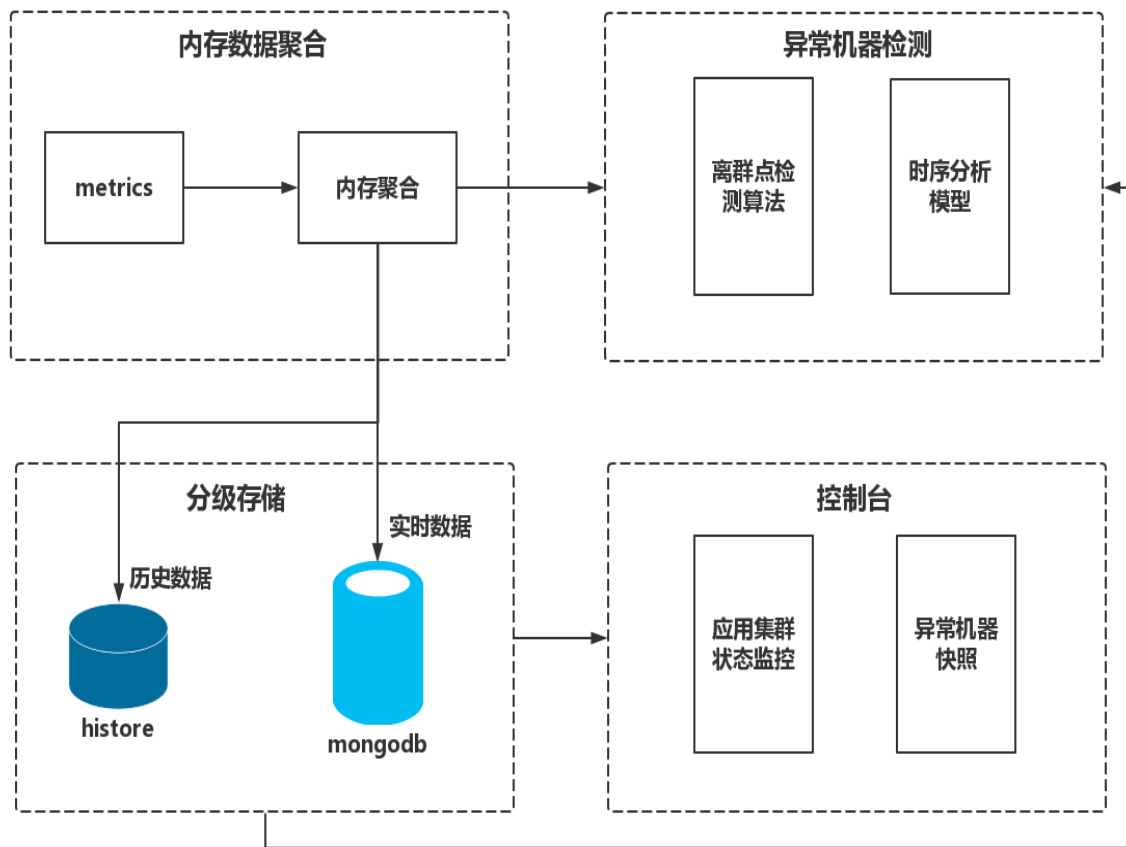
集群状态监控 - 数据处理

内存聚合

- ❑ 拉模式能清楚的知道聚合时间点
- ❑ 内存聚合速度快
- ❑ 内存聚合后的数据第一时间做异常检测

分层存储

- ❑ mongo 存储全量实时数据, 查询速度快、维度多、容灾
- ❑ 历史数据全量写 histore, 支持历史查询、离线分析



无法明确定义异常

不同应用、不同指标、不同时刻、不同压力状态下，机器是否异常无明确定义

异常表现无法穷举

磁盘故障、网络抖动、JVM FGC、资源争抢，不同异常原因导致机器异常时的表现无法穷举

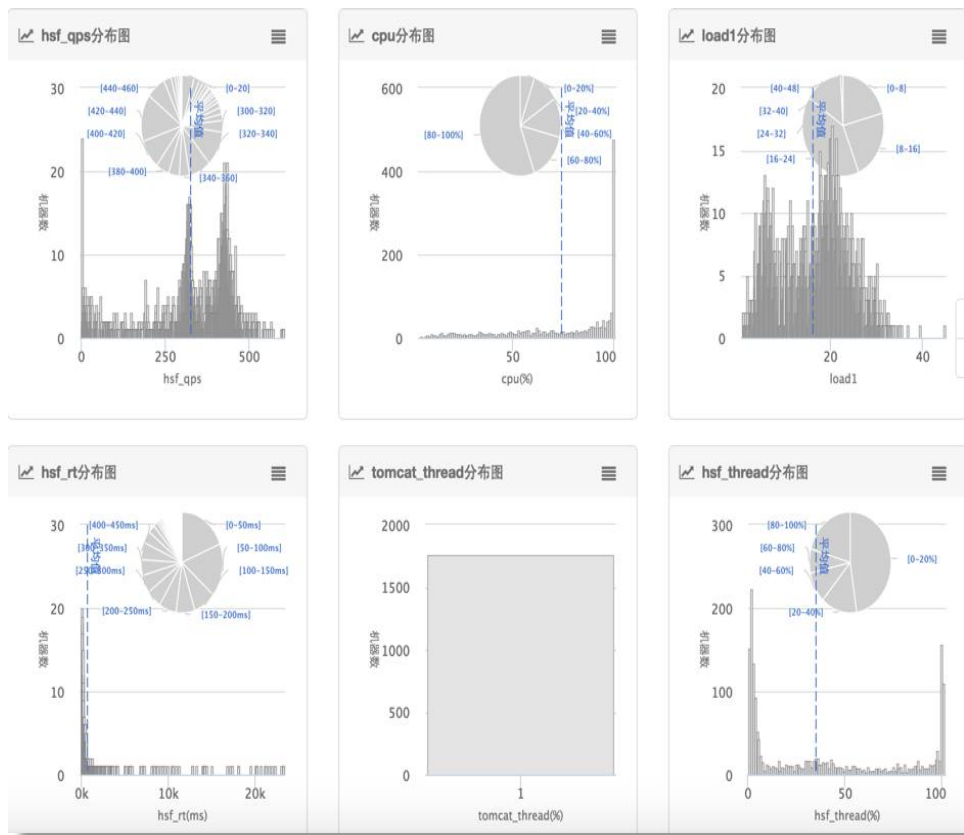
异常检测 - 方案1: 固定阈值

优点:

- 逻辑简单
- 输出结果确定

缺点:

- 每个应用不同, 每个指标要设置不同值, 指标/应用笛卡尔集
- 集群压力高的时候, 容易导致雪崩



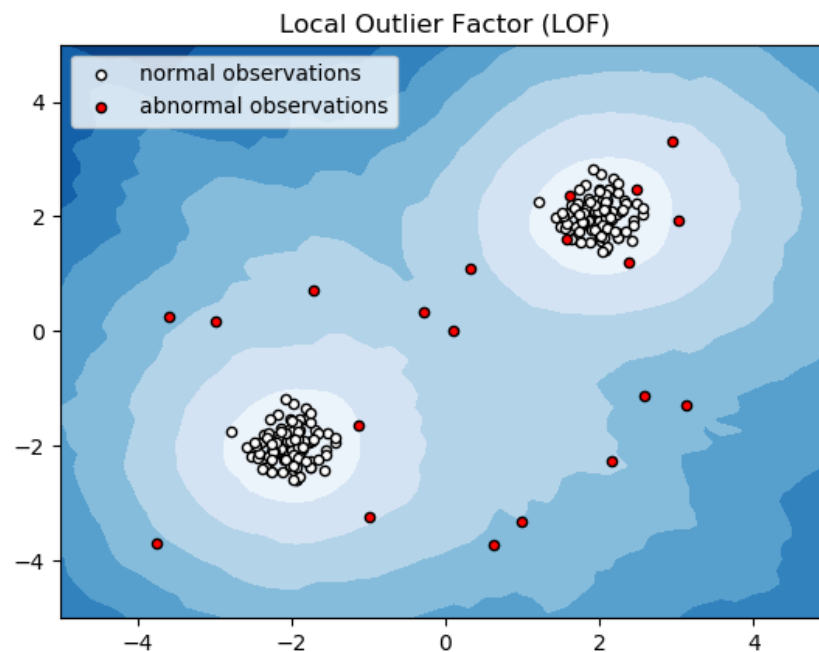
异常检测 - 方案2: 离群点检测

优点:

- 由集群整体状态决定哪些机器是异常机器

缺点:

- 输出结果具有不确定性



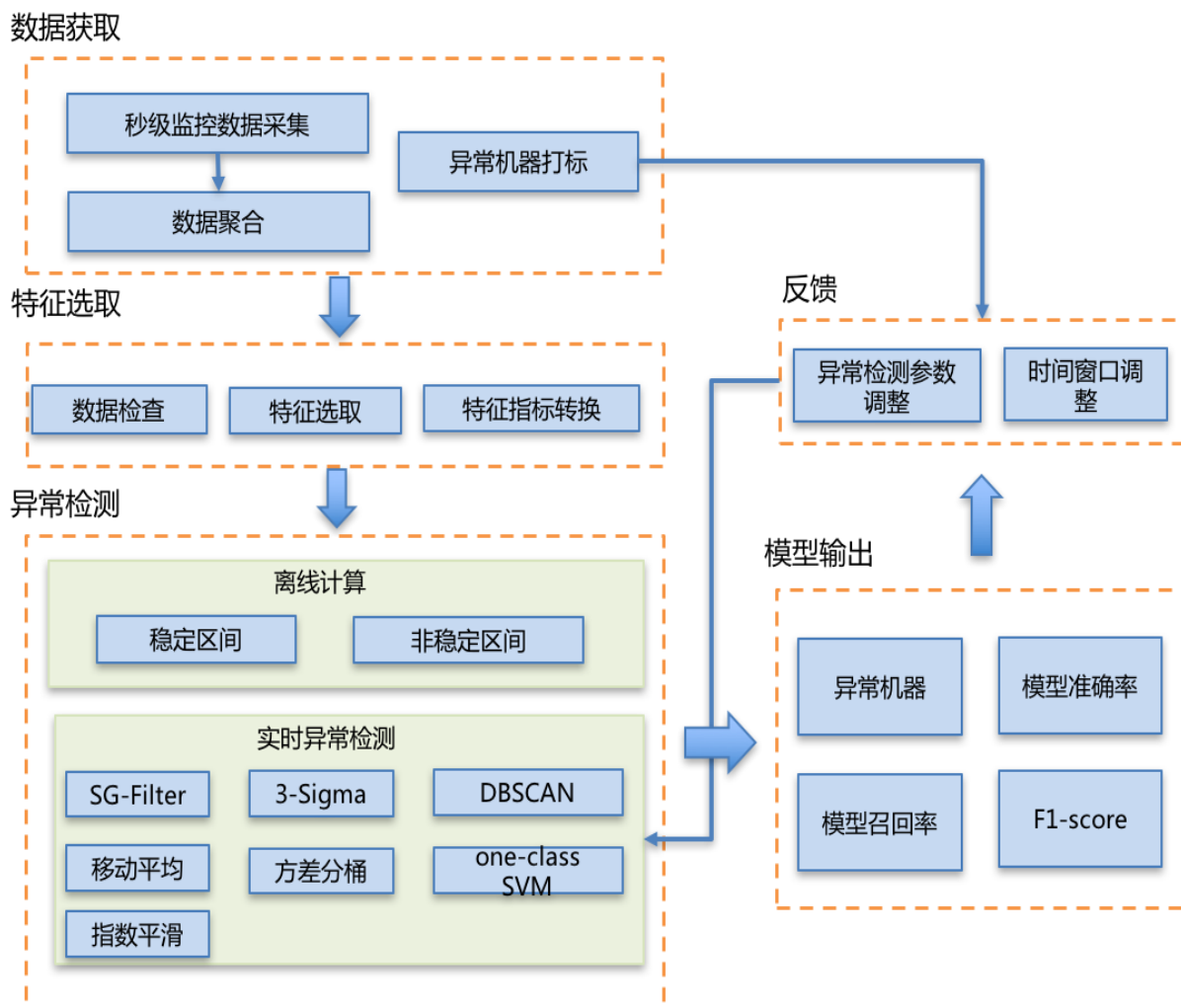
异常检测 - 离群检测框架

➤ **特征选取、处理:** 指标关联性处理, 指标转换

➤ **稳定区间、非稳定区间:** 针对不同的区间, 用不同的离群点检测算法来做处理。

➤ **多种离群算法综合刷选:** 离群是相对的, 多种离群点检测算法综合打分

➤ **监督学习:** 根据输出准确率、召回率、以及离线打标错误机器, 动态调整部分算法参数



异常检测 - 时序处理

➤ 波形特征提取:

在每个时刻，每台机器，由它四类指标前5分钟的波形

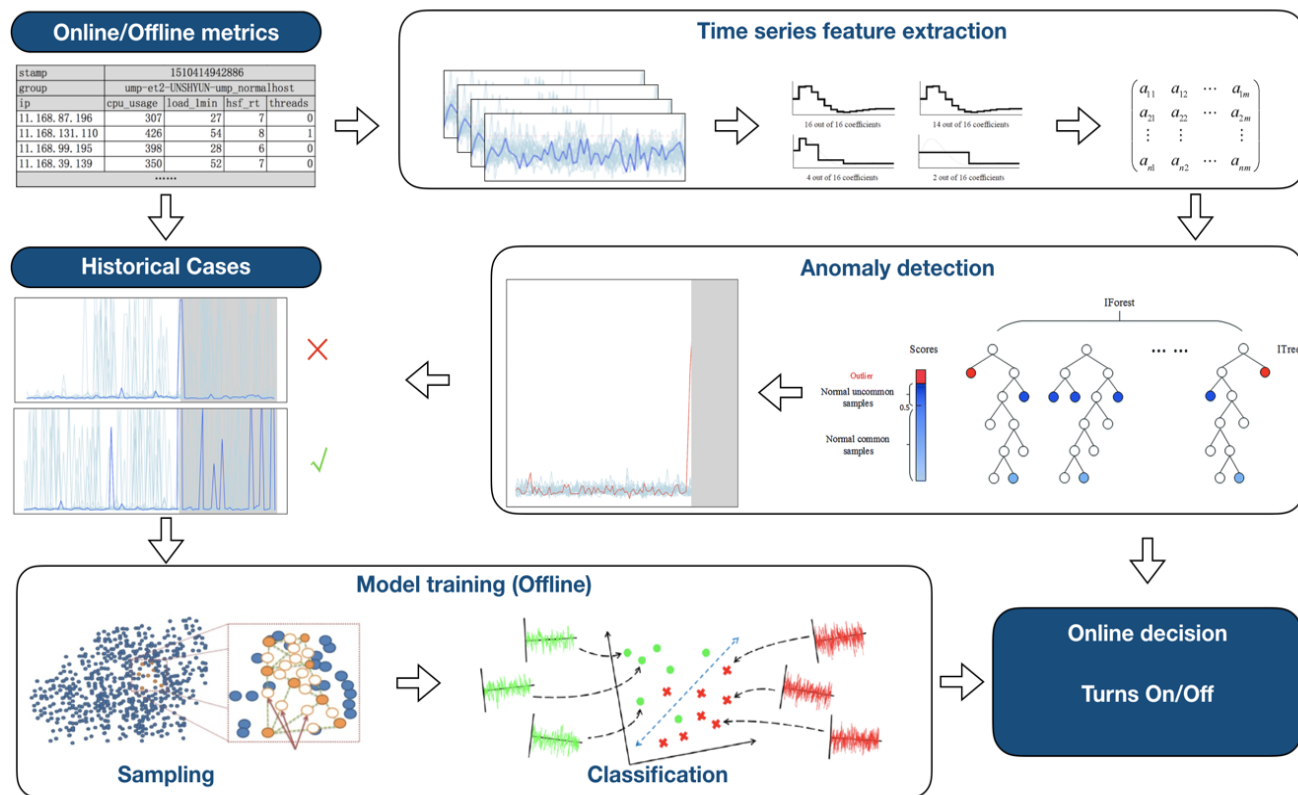
➤ 时序异常检测:

指标表现、持续时长，对比异常机器数据库

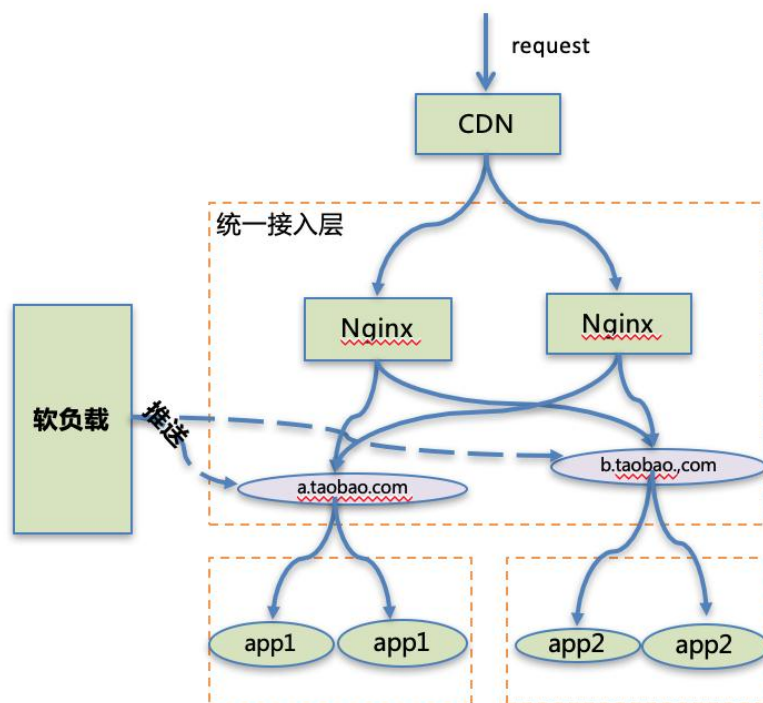
➤ 机器学习:

从历史绝对异常点上构建出来的机器学习模型，来判别它是否是绝对异常。

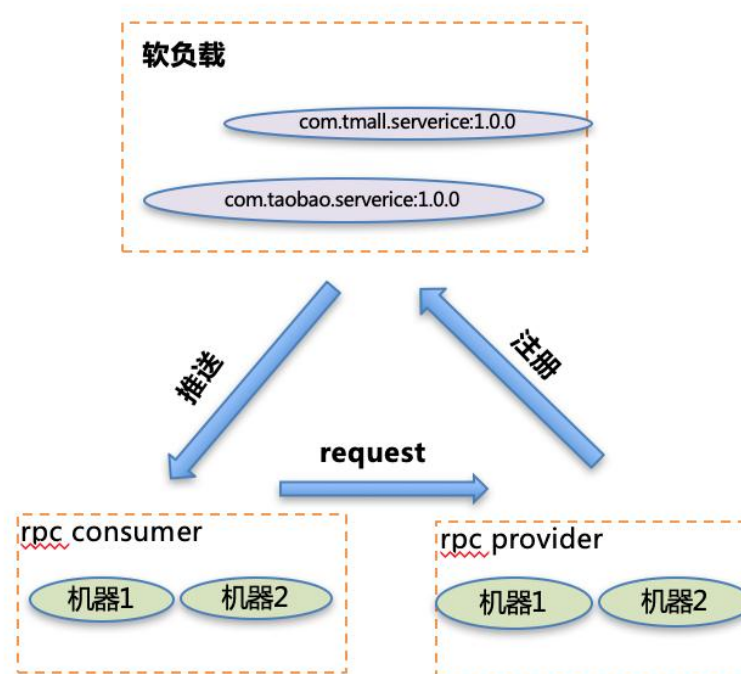
Anomaly Detection and Forecasting



HTTP流量调度



RPC流量调度



异常机器流量调度 - 总结



局部异常自愈能力

- 日均调度异常机器次数6000+, 日常屏蔽异常机器时长6600分钟

微观问题发现能力

- 异常检测累计为业务方发现系统性问题400多例

- ✓ 防刷单、流量清洗
- ✓ 单元切流
- ✓ 限流降级
- ✓ 按业务属性维度流量管控
- ✓ 异常机器流量调度