



GOPS2018  
Shenzhen

# GOPS

全球运维大会 2018

2018.4.13-4.14

中国·广东·深圳·南山区 圣淘沙大酒店（翡翠店）





GOPS2018  
Shenzhen

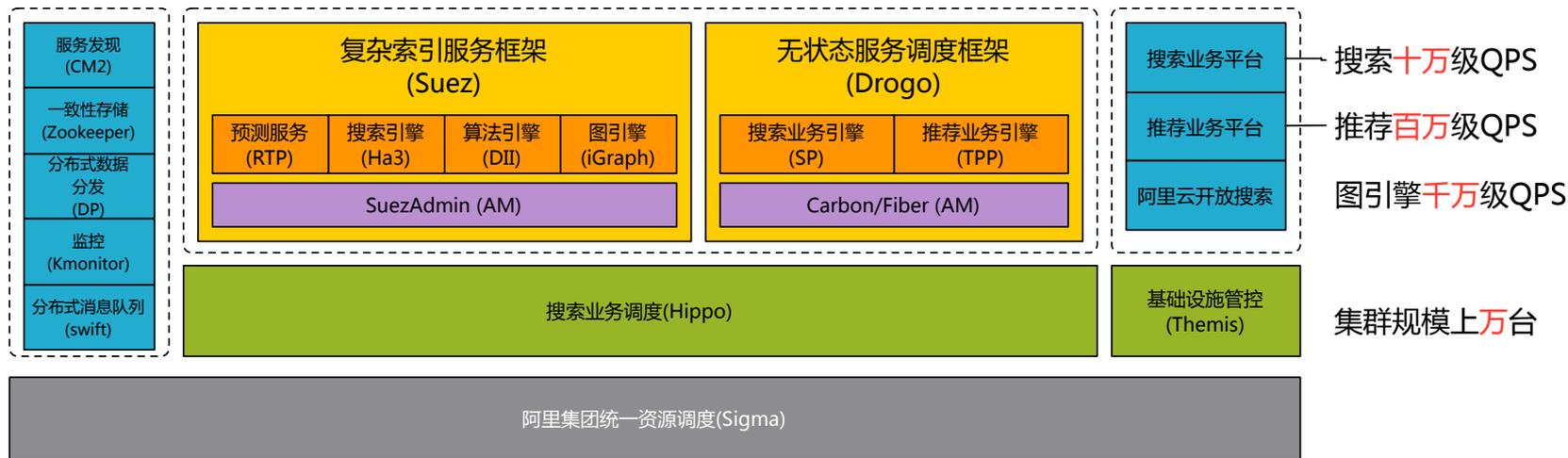
# 阿里巴巴搜索稳定性实践

李伟 高级技术专家

# 搜索在线服务体系



GOPS2018  
Shenzhen



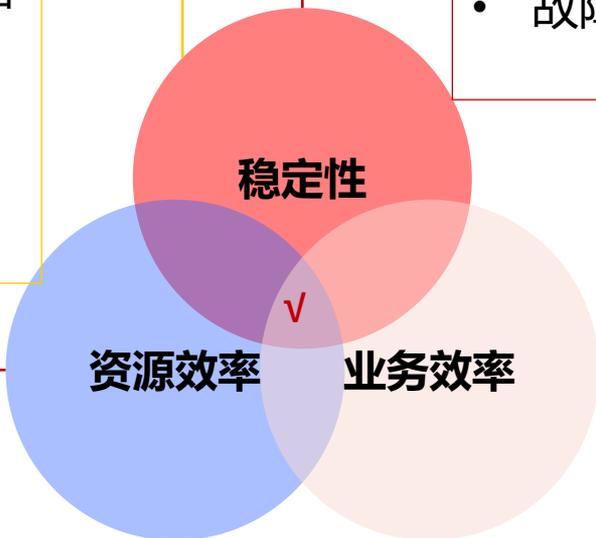


# 稳定性 VS. 效率

- 业务的快速迭代和高资源利用率带来的挑战：变化和错误成为常态

- **管理变化：**  
业务/性能/容量/故障
- **高可用设计**

- 可用性 > 99.99%
- 故障恢复时间 < 5min



- CPU利用率 > 50%

- 快速迭代，单业务发布频率 > 1/d

# 管理变化：Devops



GOPS2018  
Shenzhen

开发只关注业务，  
管控保障变更风险



开发/测试

变更/回滚

扩/缩容

监控/报警

Quota/成本

管控

灰度

容灾/管控

可用性保障  
(集群间)

目标驱动  
(集群间)

渐进一致性  
(集群间)

业务  
调度

Rolling

节点Failover

数据分发

服务注册

可用性保障  
(集群内)

目标驱动  
(集群内)

渐进一致性  
(集群内)

接口幂等

资源  
调度

资源分配和管理

故障机检测和维修

内核管理

混部

隔离

AIOPS

弹性伸缩

自动化压测

智能预测和异常检测

智能诊断和容量评估

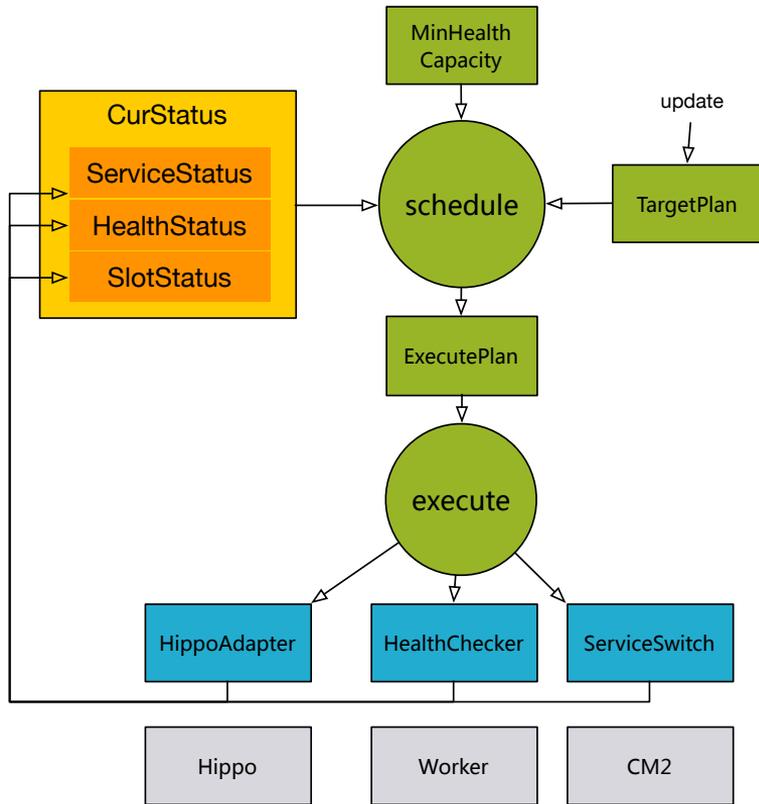
故障自动切流

自动化容灾演练



# 管理变化：业务调度和管控

- 自动部署
- 服务挂载
- 可用性保障
- 平滑扩缩容
- 异常恢复
  - 机器故障
  - 应用错误
  - 网络分割
- Rolling升级/回滚

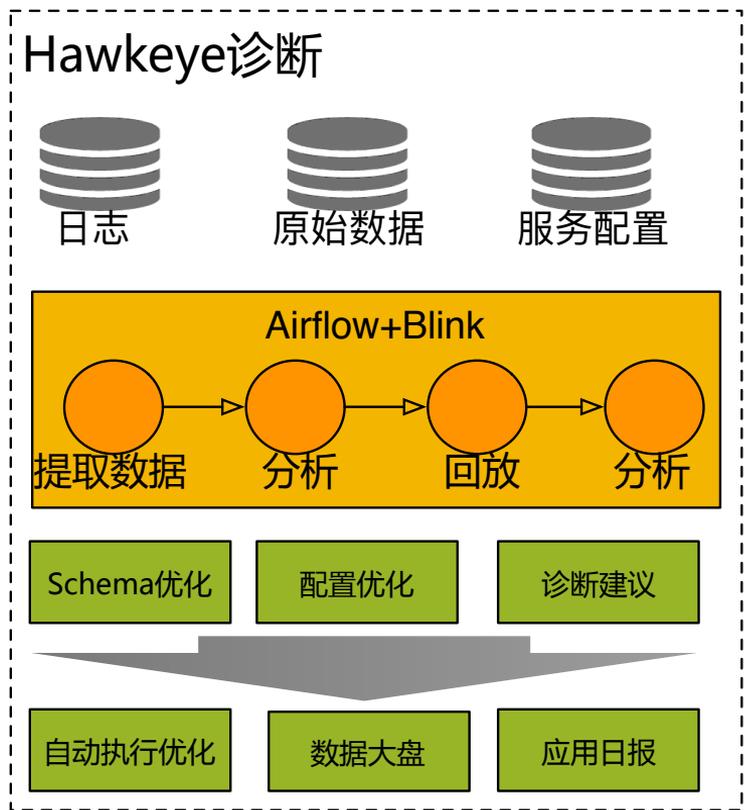


- 渐进最终一致性
- 接口幂等
- 异步调度
- 无状态



# 管理变化：智能诊断

- Schema优化
  - 索引优化提升性能 (倒排索引/bitmap)
  - 字段压缩节省内存
  - 冷数据优化节省内存
- 配置优化
  - QueryCache提升性能
  - 合理实时内存配置避免OOM
  - 最优Segment数提升性能
- 诊断建议
  - 检索截断，降低精排doc数，插件优化
- 部分应用诊断优化后性能最大提升8倍





# 管理变化：容量评估

- 优化目标  $min(price) = f(R_{used}, R_{target}, Price_R, QPS_{current}, QPS_{test})$

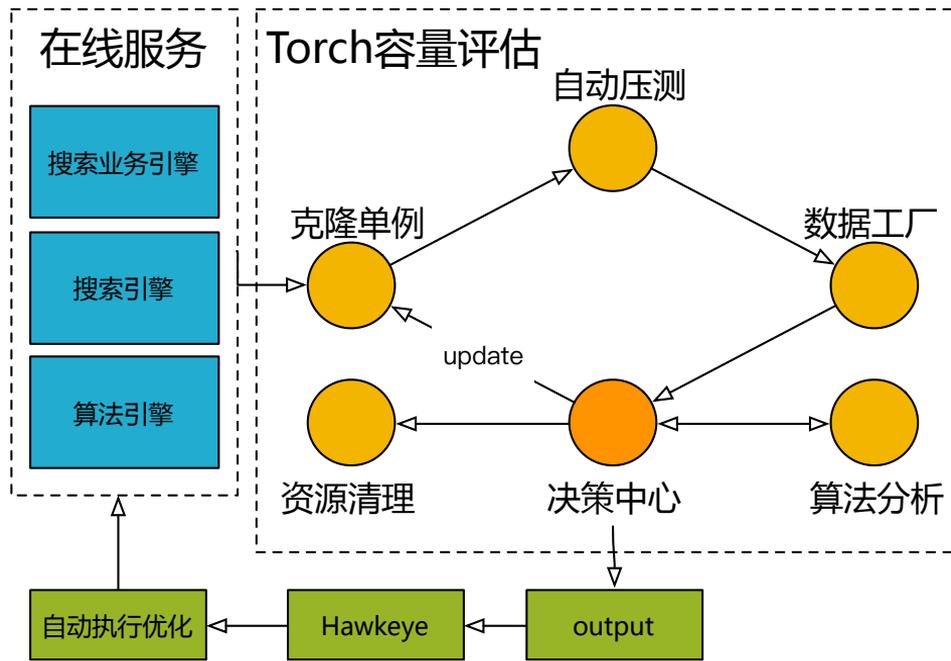
优化replica/sharding数平衡CPU/Mem成本和容量

优化build并发数平衡全量索引耗时和成本

调控各项资源单价促进平台资源优化

- 效果

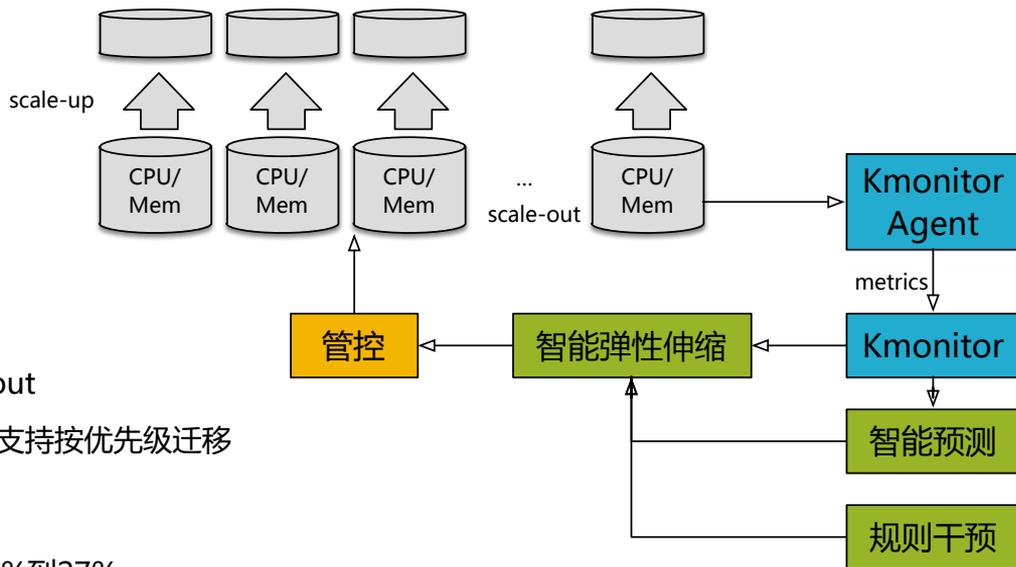
- 长尾应用自动进行资源优化，磁盘和内存使用率60%~70%，CPU使用率30%~40%，日均节省成本6w+元。





# 管理变化：弹性伸缩1

- 采集数据
- 决策
  - 响应式 & 预测
- 执行
  - scale-out vs. scale-up
  - FaaS解决scale-out秒级部署
  - 带数据服务通过存储计算分离优化scale-out
  - scale-up：运行时动态调整线程池，调度支持按优先级迁移
- 效果
  - 日常集群平均CPU利用率提 3.37 倍，从8%到27%
  - 双11 促期间，P1场景利用率保持在 30% ，其余保持在50%。99%的P1场景降级率低于2%，超时率低于0.6%





# 管理变化：弹性伸缩2

$N_i = f(\text{CPU}, \text{Load}, \text{异常QPS}, \text{真实QPS}, \text{预测QPS}, \text{分级压测QPS}, C_i)$

其中  $C_i$  为场景当前机器

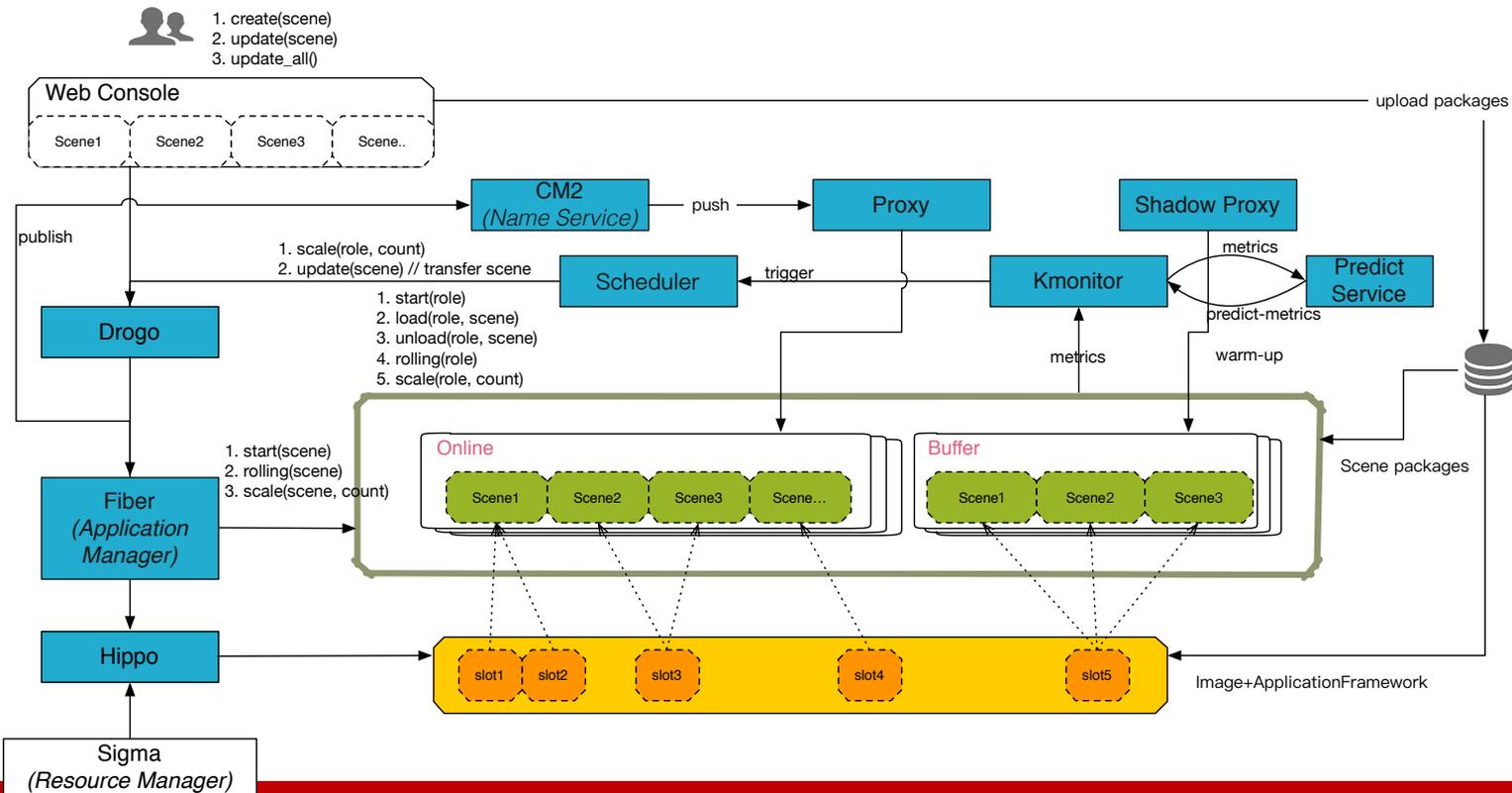
$$\min \sum w_i * (\text{CPU}_i - \text{OptimalCPU})$$

$$\text{其中 } \text{CPU}_i = \frac{\max(\text{当前QPS}, \text{预测QPS})}{\text{压测QPS} * (C_i + N_i)}$$

- 定义业务优先级(P1~P4)+服务等级(LU~L3)
- 算法实时预测QPS
- Daily-run压测获取各服务等级QPS
- Buffer提前预热加快扩容速度
- 自动降级应对突发流量，自动限流(过载保护)保证最大服务能力
- 动态负载均衡解决扩容预热不充分导致的服务能力不足



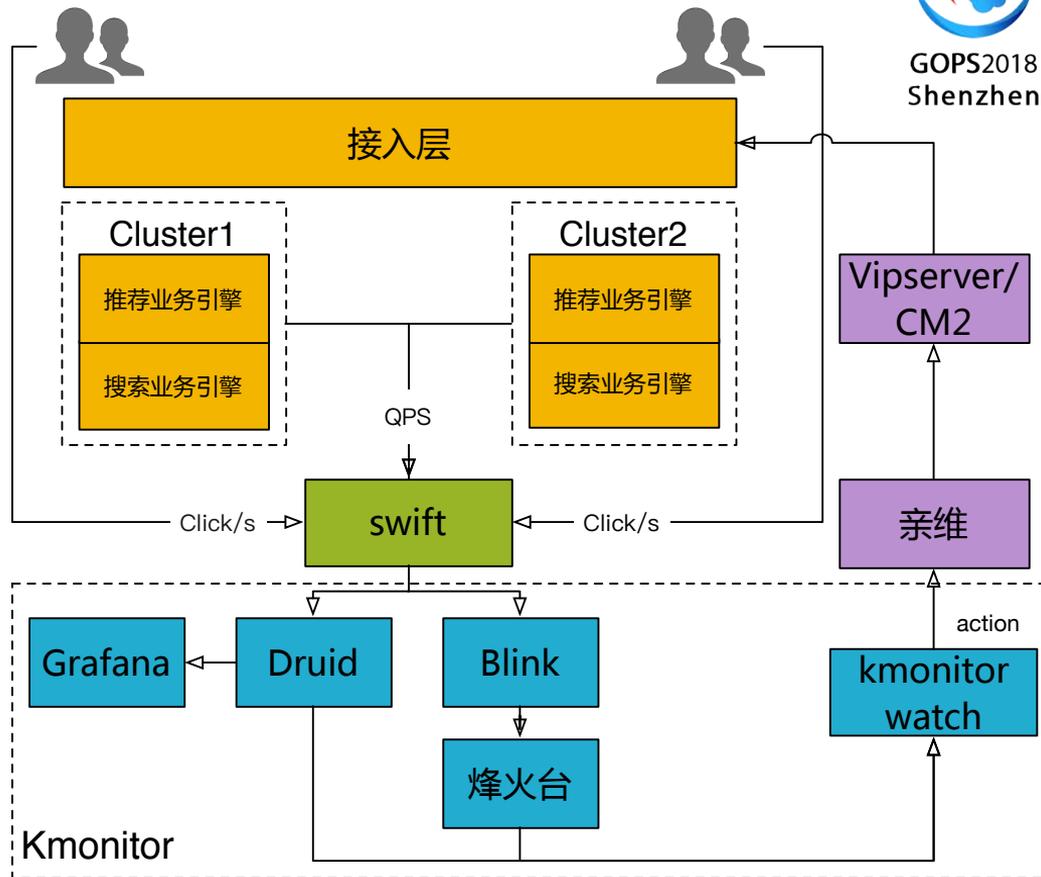
# 管理变化：无状态服务秒级扩容





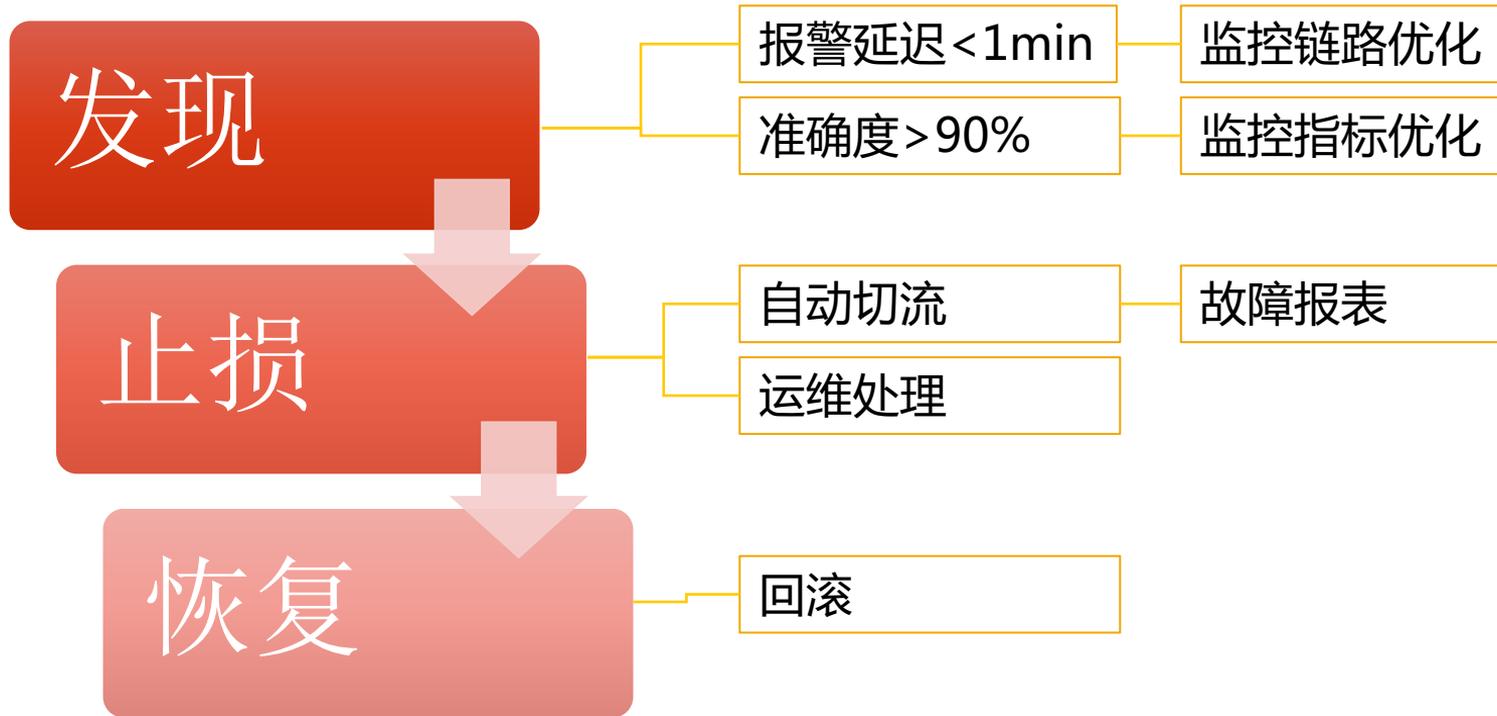
# 管理变化：监控

- 监控链路优化，1分钟报警
- 监控项优化
  - 对P1/P2和P3/P4故障定制不同监控策略
  - 业务效果监控提升准确率
- 监控自动执行action
  - 报警
  - 切流
  - ...





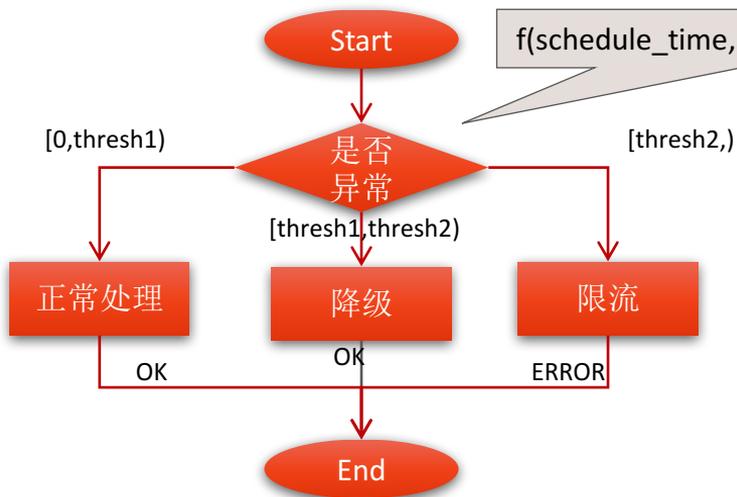
# 管理变化：故障自动切流和快速恢复







# 高可用设计：自动降级和限流



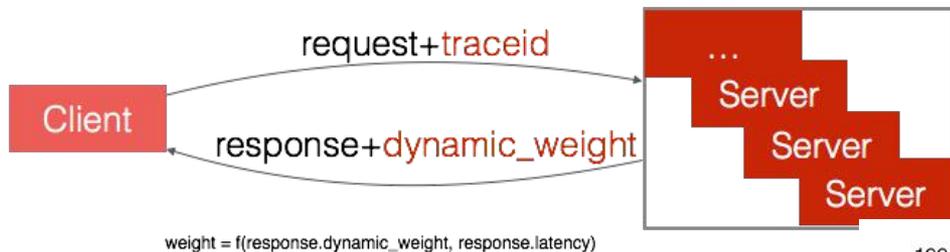
降级判断：  
异常率、Load

自动降级策略：  
L1: 减少Search和Rank数量  
L2: 降低算法模型复杂度  
L3: 取消Rank  
L4: 静态Cache

- 超出正常服务能力后降级提升服务能力
- 超出最大服务能力后限流确保不雪崩

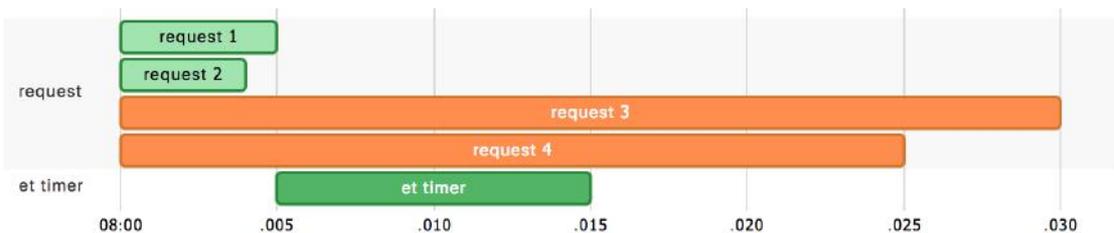
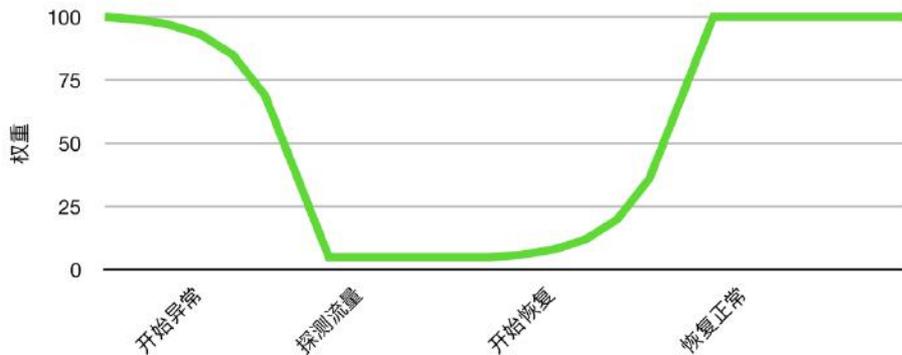


# 高可用设计：动态负载均衡&熔断



结合服务端返回的动态权重和响应时间，均衡服务端各节点负载

连续超时或不可用节点逐步降低权重直至0。  
权重为0的节点通过copy流量探测实现恢复。



通过提前中止(EarlyTerminate)  
解决并发查询的短板效应

# 线上效果



GOPS2018  
Shenzhen



# 总结



GOPS2018  
Shenzhen

## 稳定性风险

业务变更

用户行为变化

自然因素/变更  
导致异常

## 管理变化

Devops

管控&容灾

智能诊断

容量评估

故障快速恢复

弹性伸缩

## 高可用设计

自动降级

自动限流

动态负载均衡

熔断

- 搜索和推荐两大平台支持阿里集团大部分流量
- 高性能分布式搜索和推荐引擎以及图检索引擎，支持千万级QPS
- 大规模集群的智能化调度
- AIOPS无人值守运维
- 国际化支持东南亚、印度、俄罗斯业务





GOPS2018  
Shenzhen



# Thanks

高效运维社区  
开放运维联盟

荣誉出品



GOPS2018  
Shenzhen

想第一时间看到高效运维社区  
的新动态吗？

