



GOPS2017
Beijing



GOPS

2017 全球运维大会



指导单位:  信息中心联盟
Beijing Center Alliance

主办单位:  医疗运维社区
Health IT Community

 开放运维联盟
Open Ops Alliance

大会时间: 7月26-30日

大会地点: 北京朝阳悠唐皇冠假日酒店



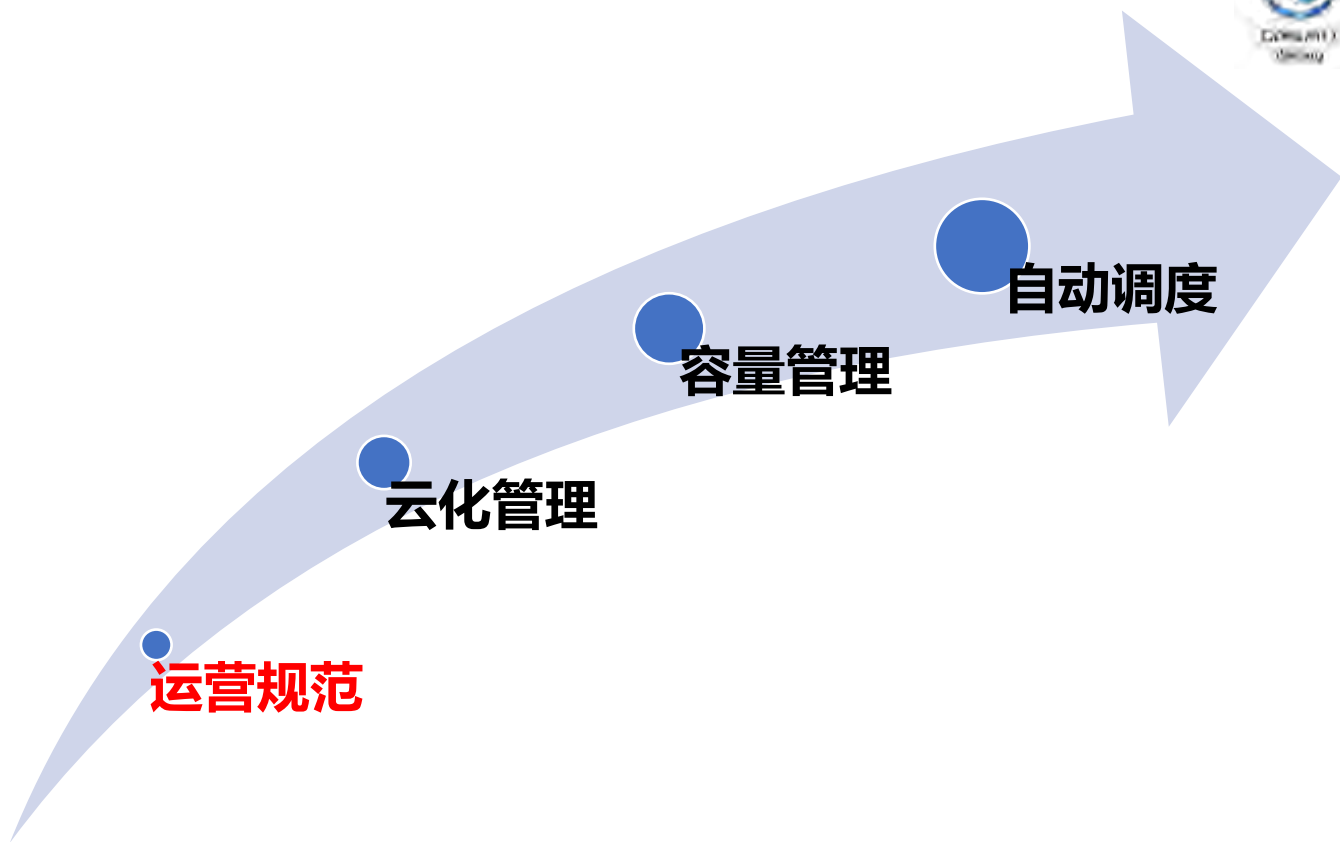
微信高效运维之路

吴磊



关于弹性伸缩的自动化

效率 → 成本





配置文件规范

- 目录结构标准
- 跨服务的相同配置项管理
- 同一服务内不同实例的差异配置项管理
- 开发/测试/现网的差异配置项管理

同一服务下同一版本的多个实例，在所有环境下配置文件的md5都严格一致

名字服务规范



接入层

类LVS实现

逻辑层

类etcd实现

存储层

路由配置
自动化

服务伸缩是运维工程，独立于研发的变更发布



数据存储规范

- 接入层 不带数据
- 逻辑层 带短周期cache、带静态数据、禁止动态数据落地
- 存储层 带长周期cache、基于paxos协议的数据落地

接入层和逻辑层的服务伸缩，无需考虑数据迁移和cache命中率

运营规范小结

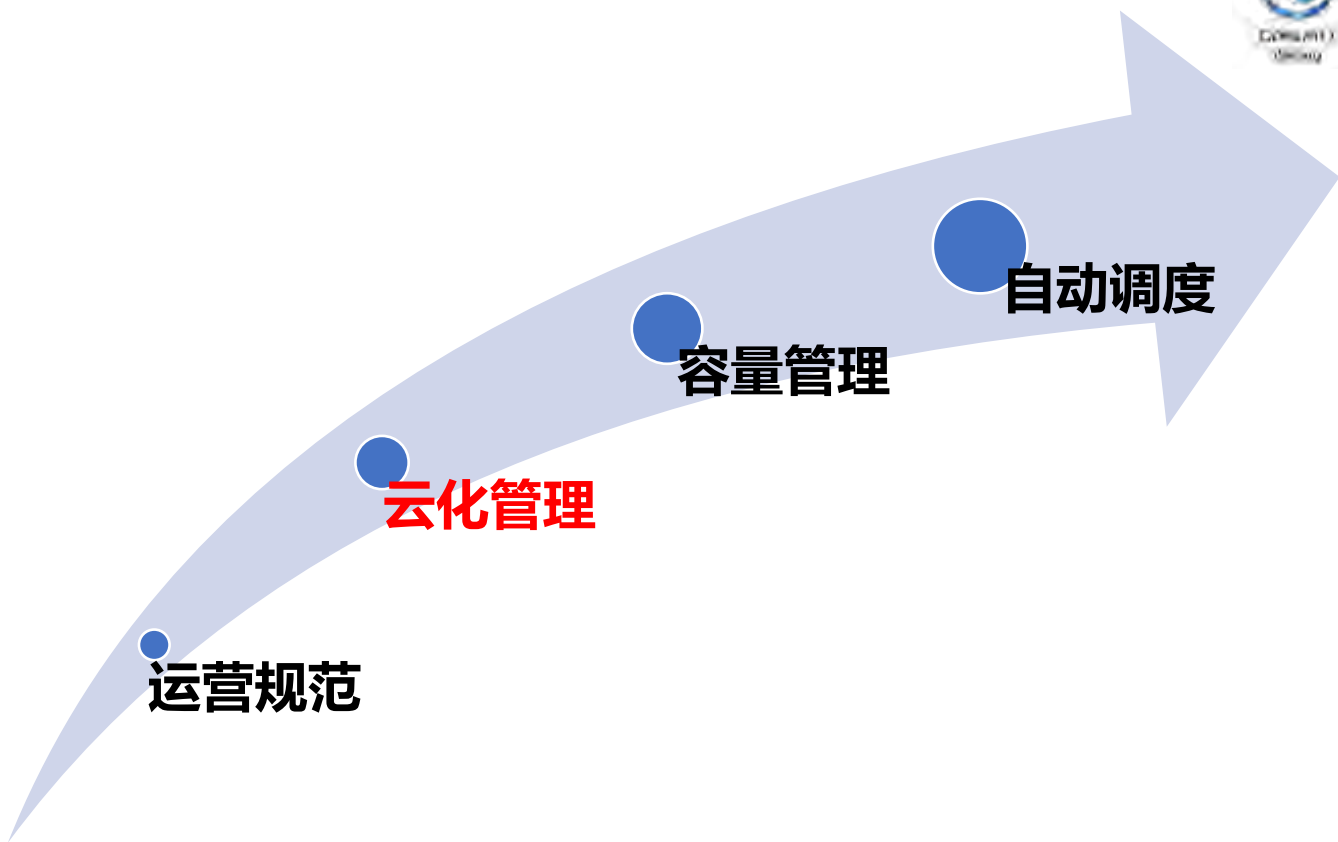


阶段目标

- 服务可运维

落实措施

- 变更系统拦截
- 全网扫描不规范的服务





为什么上云

- 微服务总数近5k
- 同物理机上多服务的资源抢占问题



哪部分上云

接入层

独占物理机 容量充足
变更少

逻辑层

混合部署 容量不可控
变更频繁

存储层

独占物理机 容量可控
变更少



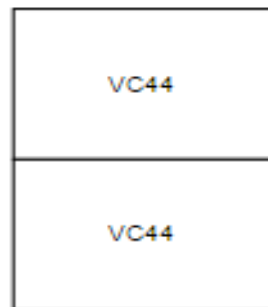
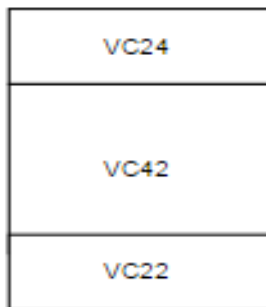
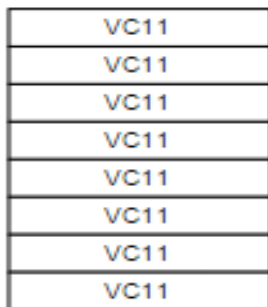
基于Cgroup的云化

- 虚拟机定制

VC11 = 1个cpu核 + 1G内存

VC24 = 2个cpu核 + 4G内存

- 物理机分片





线上未启用Docker

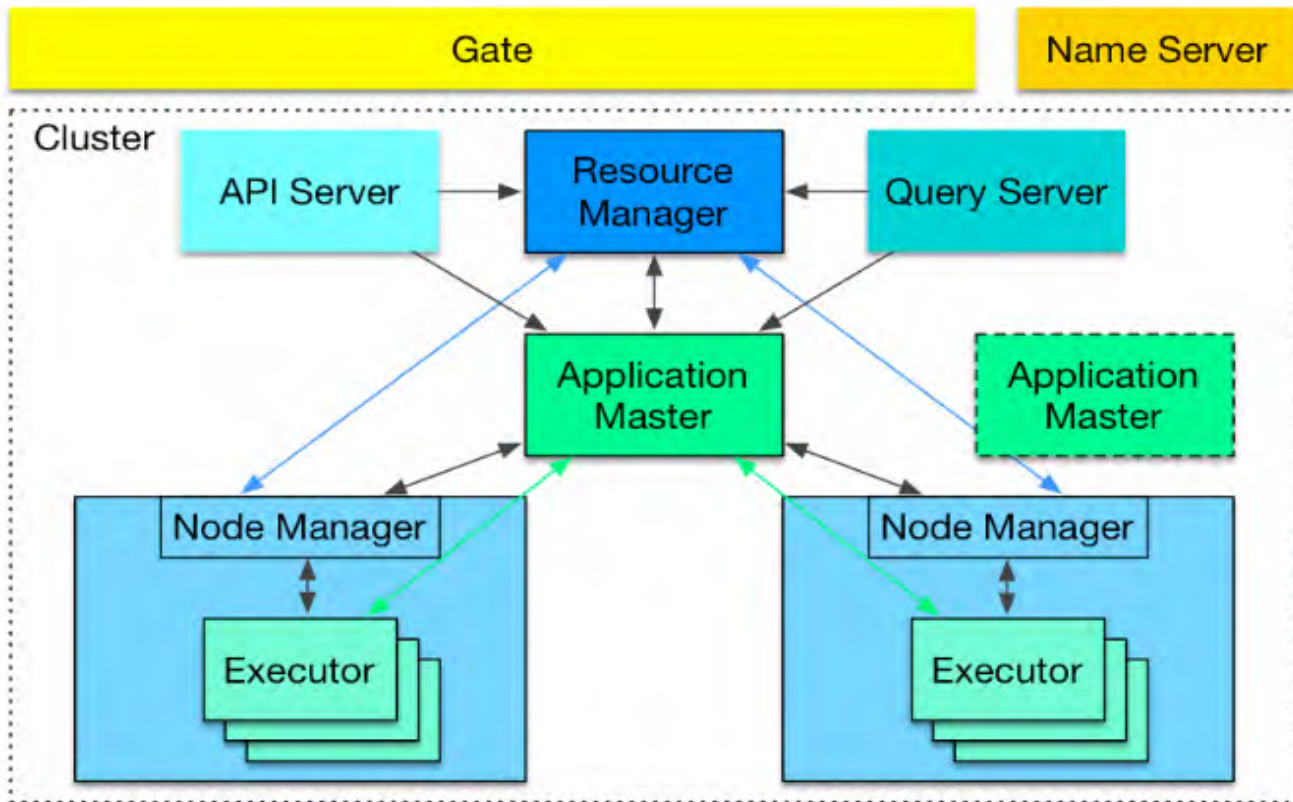
- svrkit框架100%覆盖全网，已实现标准化规范化
- 框架本身大量依赖IPC交互
- 自研非侵入式 vs Docker侵入式



私有云调度系统

- 基于svrkit框架自研
- 参考borg/yarn/k8s/mesos等主流调度系统的优点
- 覆盖80%的微服务

私有云调度系统架构



云化管理小结

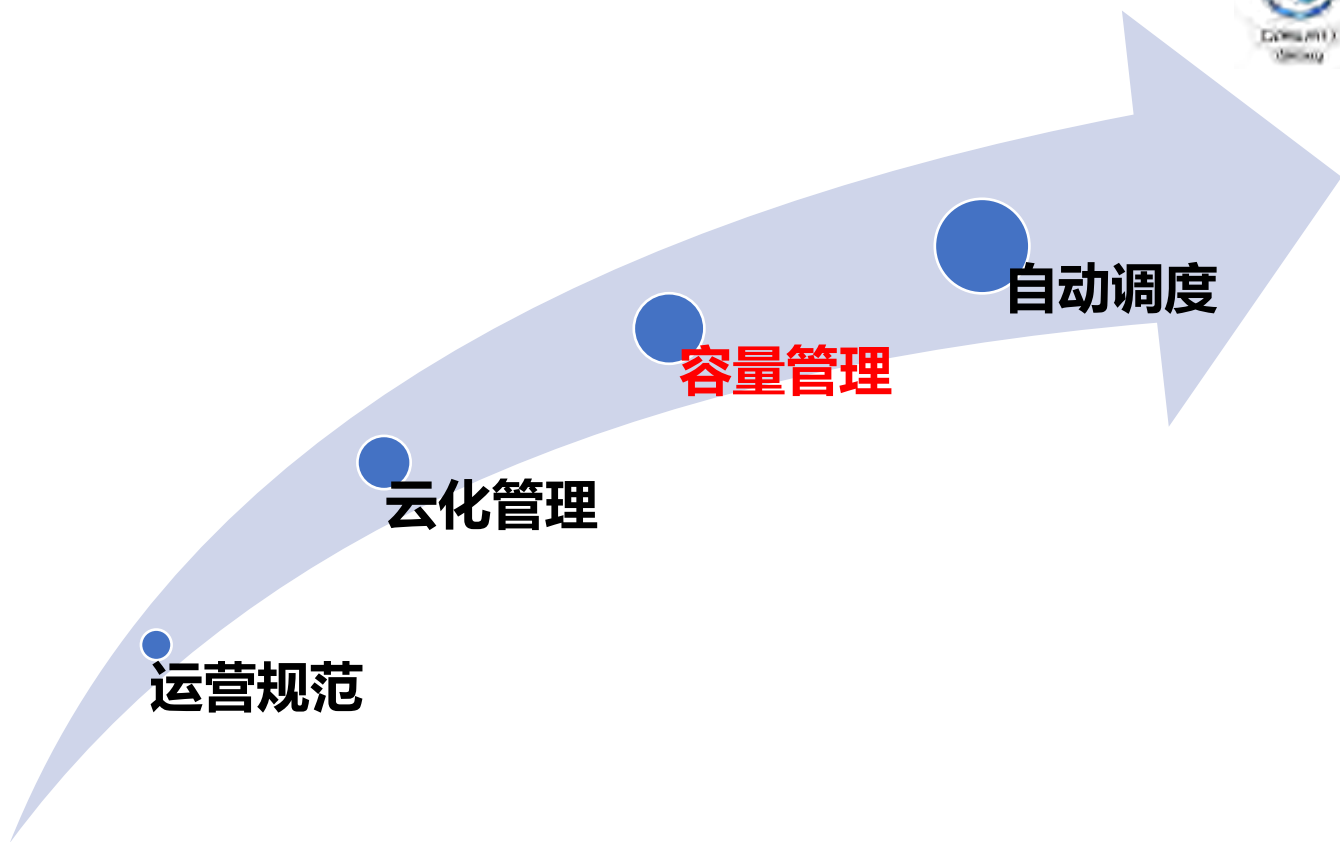


阶段目标

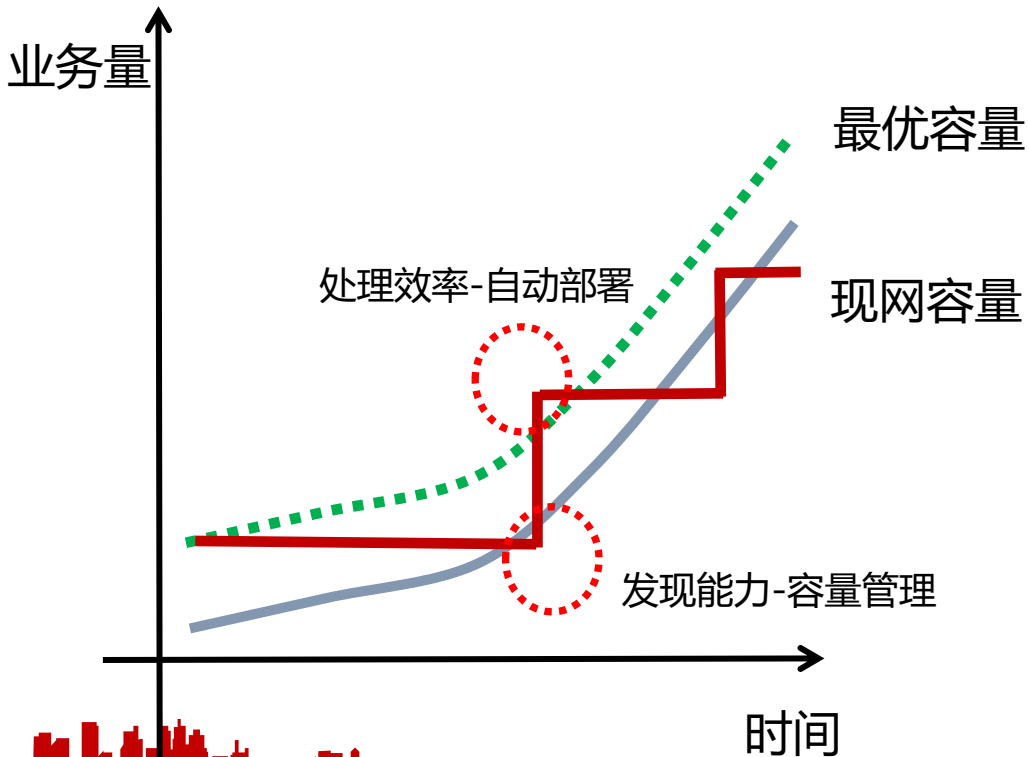
- 服务间资源隔离
- 服务伸缩页面化操作

落实措施

- 部署系统拦截未上云业务
- 主动改造核心业务



如何支撑业务发展



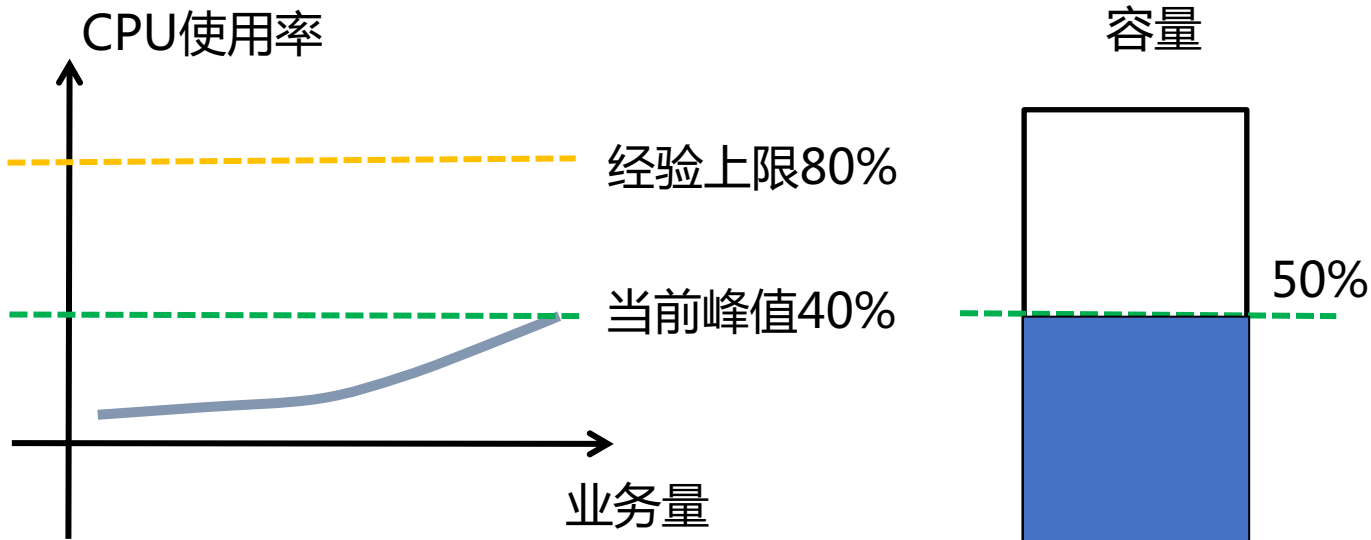
使用硬件指标评估容量



- 运算型 => CPU
- 存储型 => 磁盘空间
- 网络型 => 网卡流量



使用CPU评估容量



服务容量 = 当前峰值 / 经验CPU上限

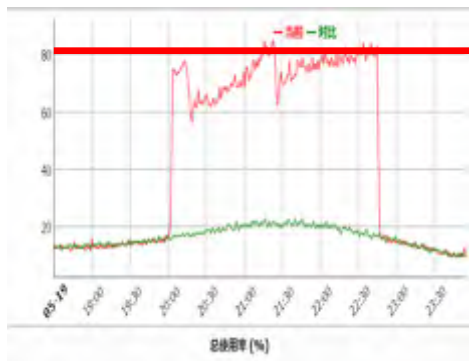
硬件指标是否可靠



真实案例

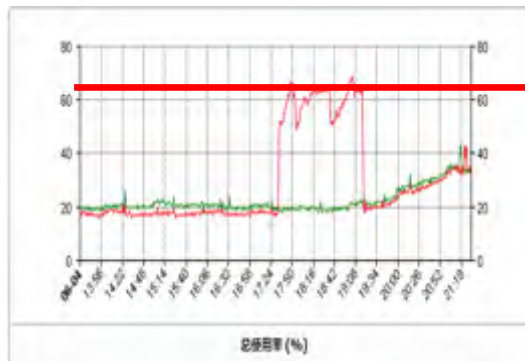


服务A



80%

服务B



60%

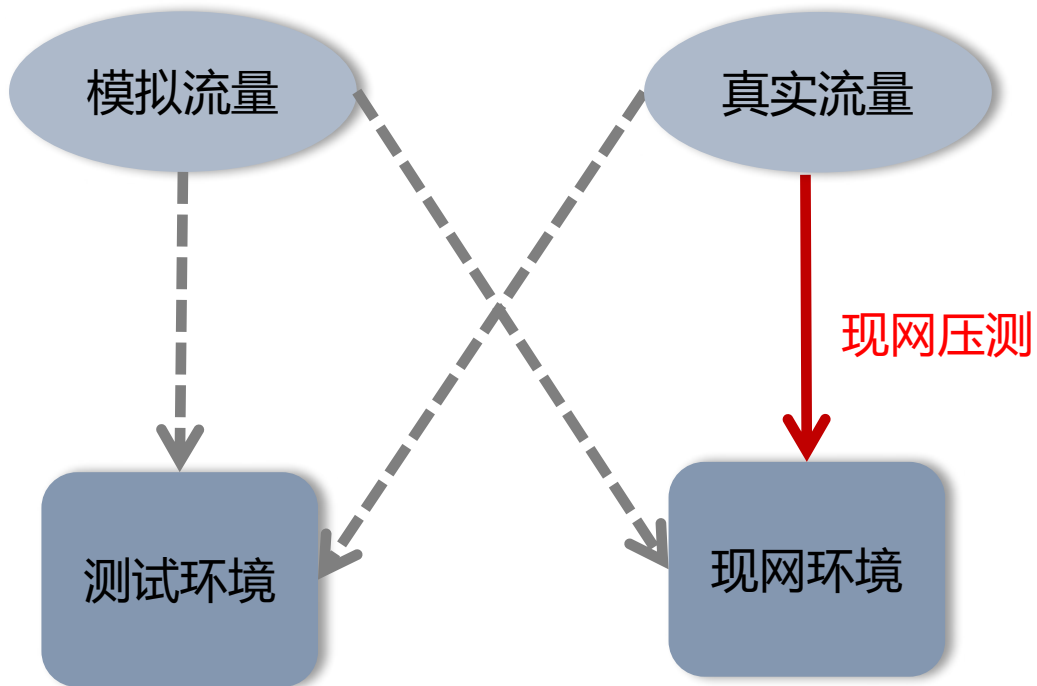


硬件指标的局限性

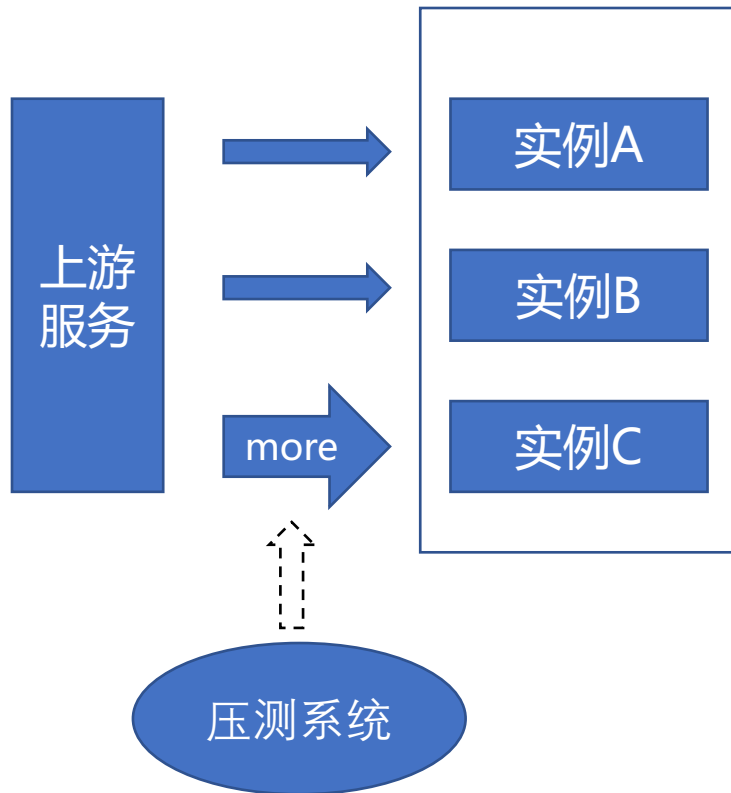
- 不同服务有不同硬件类型的上限
- 临界点的质量无法预测

通过压测，才能获得准确容量模型

压测方式



现网压测

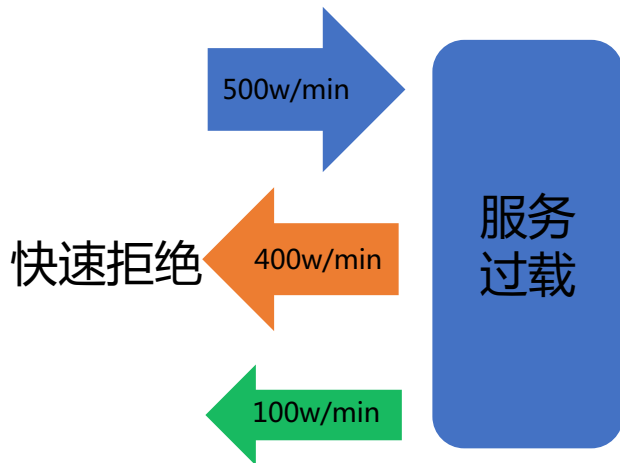
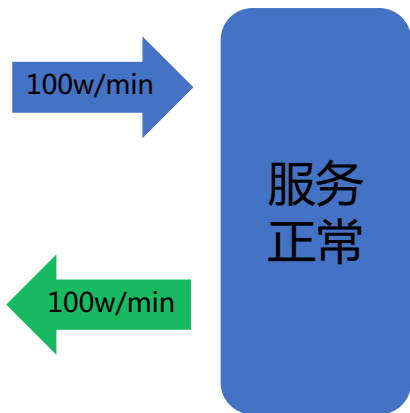




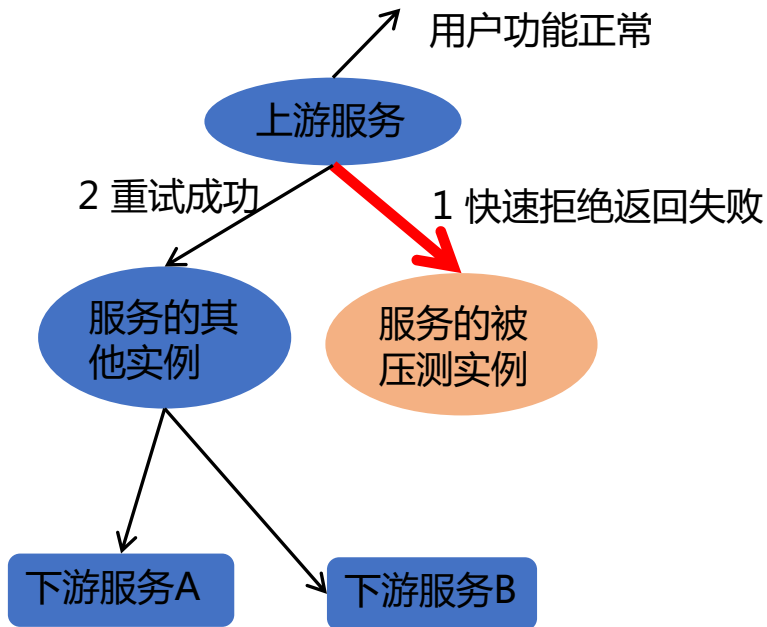
现网压测可能导致的异常

- 压测引发故障未能及时发现
- 压测带来底层问题未在上层业务表现
- 引发故障后未能快速恢复

服务的自我保护



上游服务的重试保护



立体化监控



业务监控

调用失败 调用重试

耗时 (svr client)

快速拒绝

硬件限制 (CPU 内存 网卡)

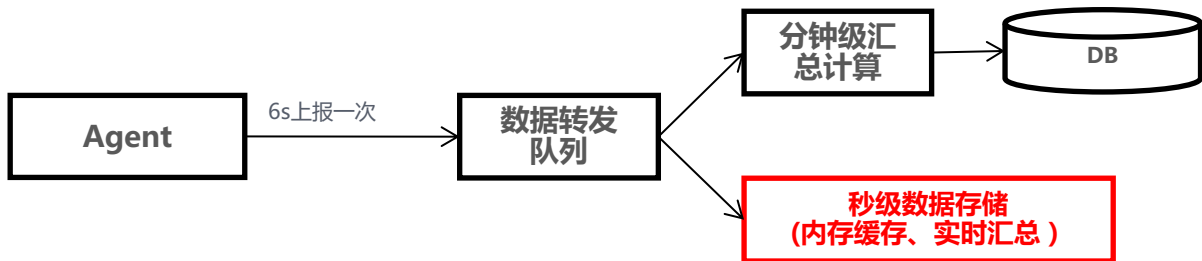
秒级监控



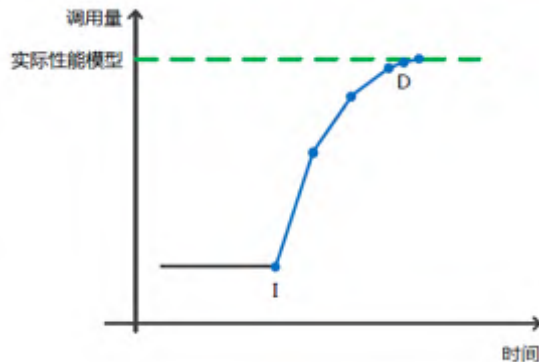
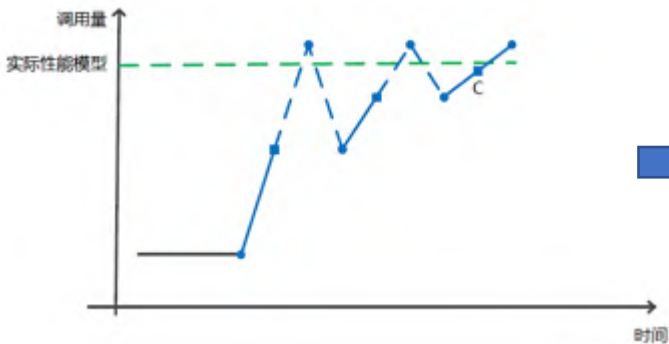
- 分钟级监控，入库延迟**3min**



- 秒级监控，入库延迟**10s**

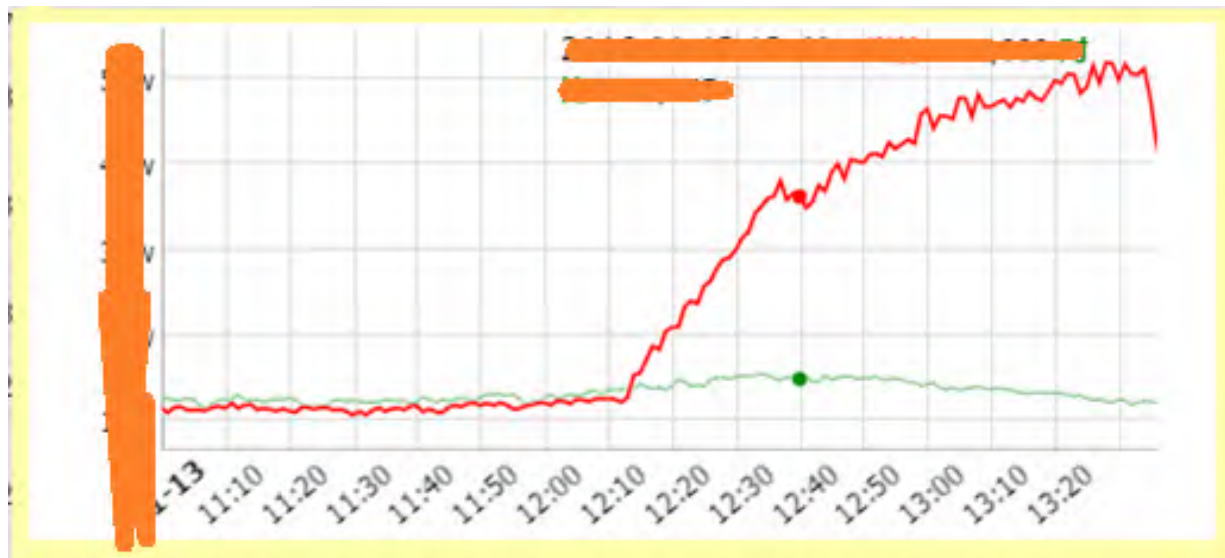


放量速率动态控制



$f(x)$ = x时刻获取请求入队延迟超过20ms的比例
 $P = f(D) - f(I)$, 实时计算P值并动态调整速率

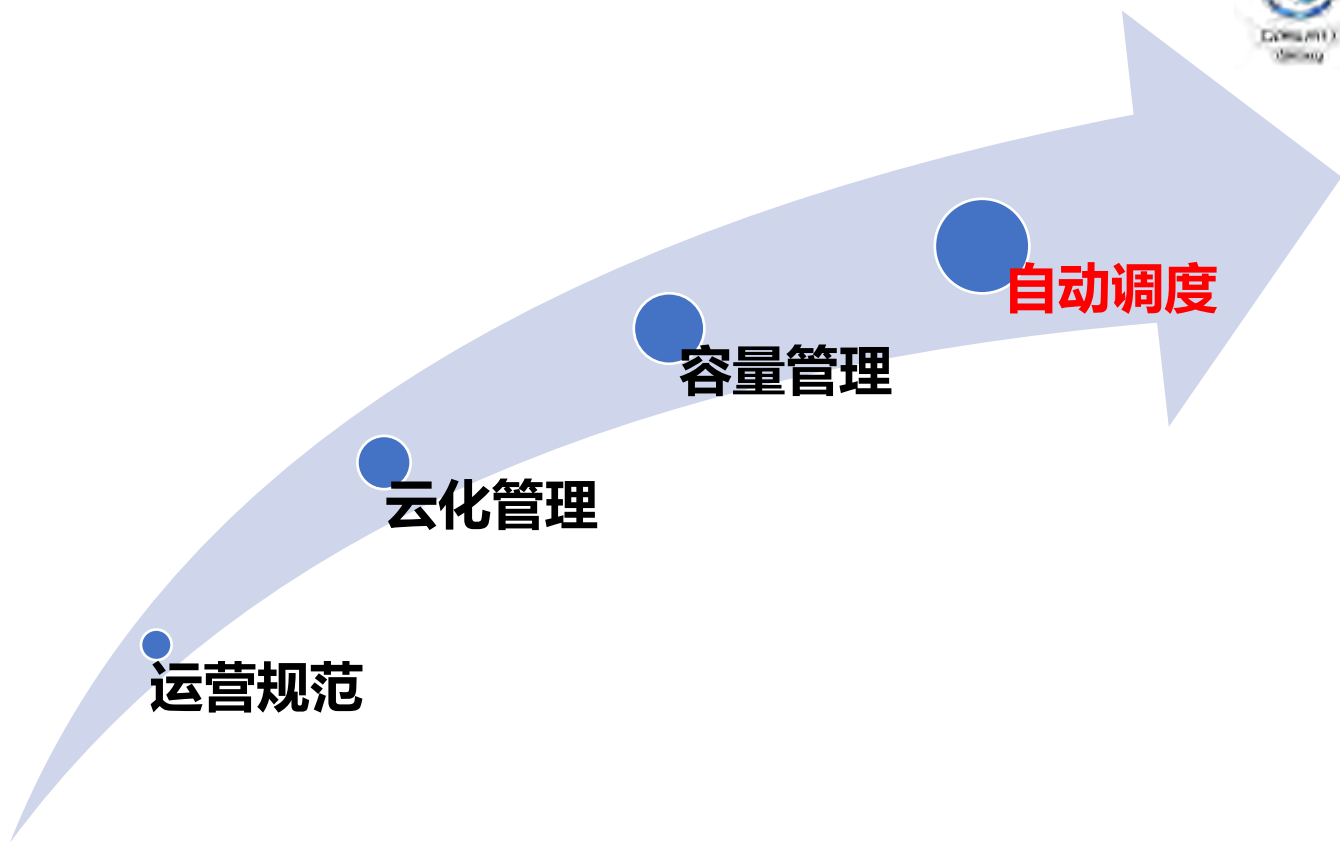
现网压测效果



容量管理小结



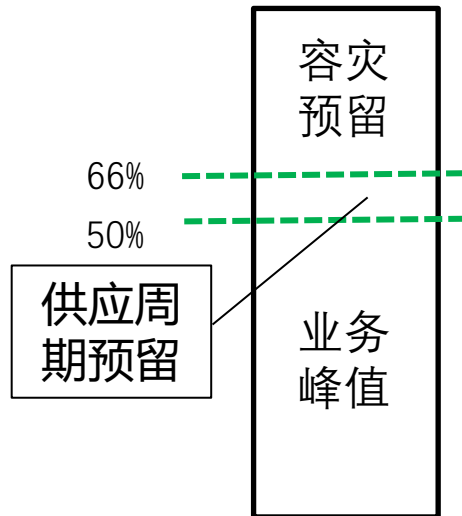
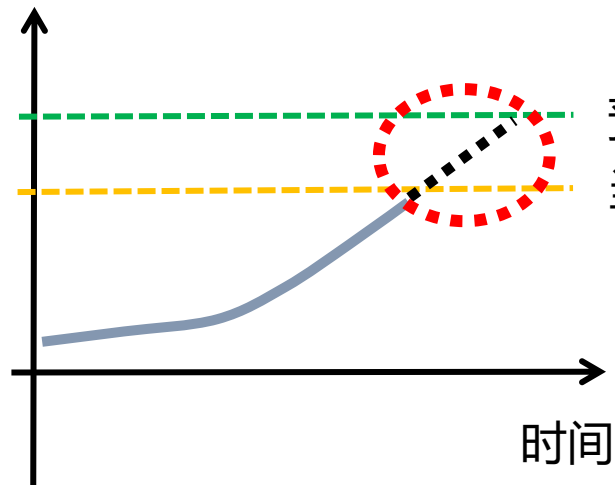
- 服务的资源需求可准确量化
- 可确认服务的最优机型



业务增长的自动扩容



业务量





异常情况的自动扩容

- 业务突增
- 程序性能下降



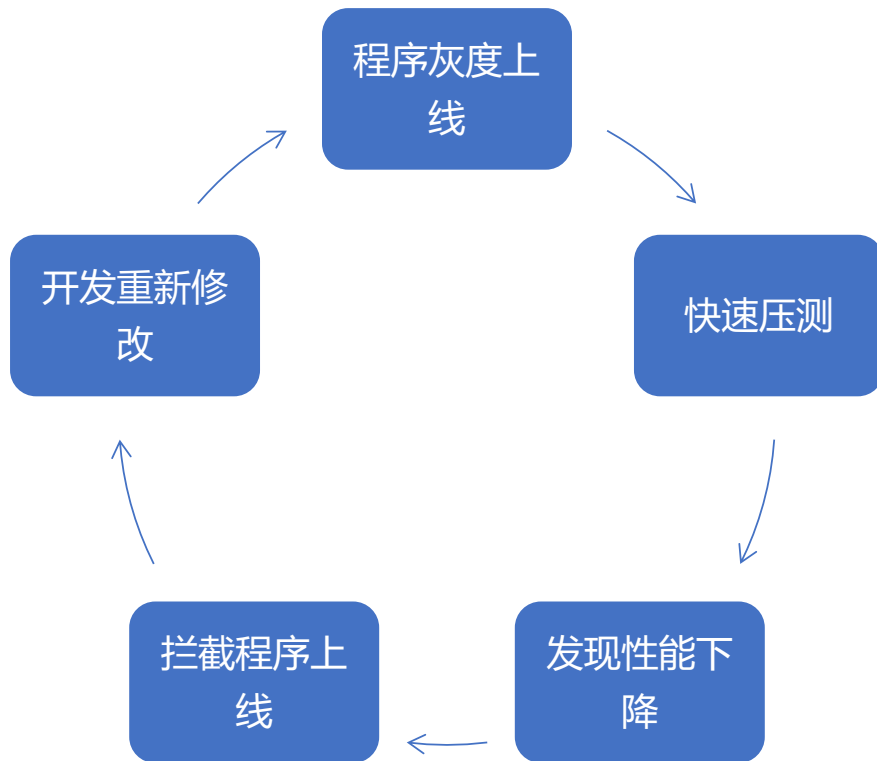
程序性能下降的合理性评估



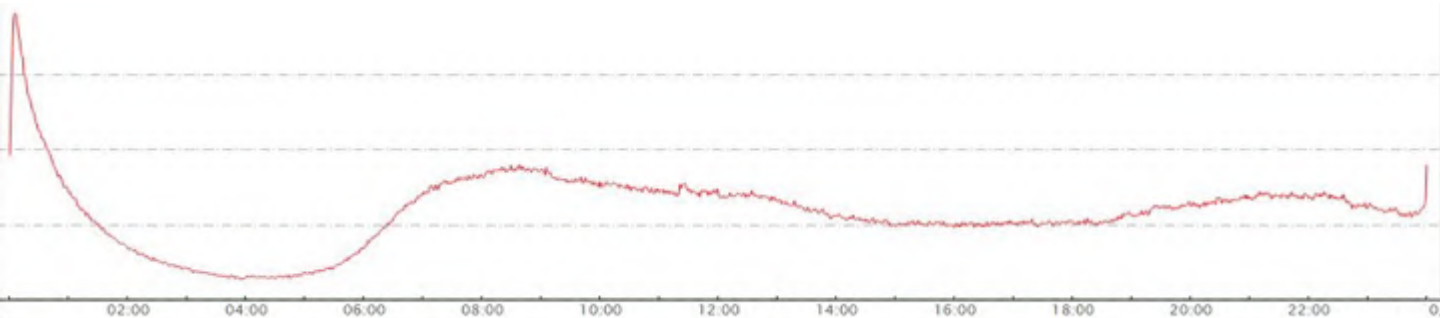
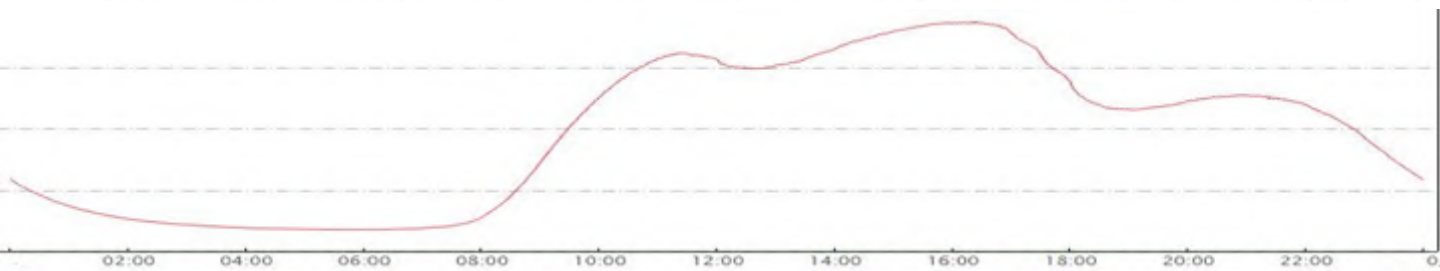
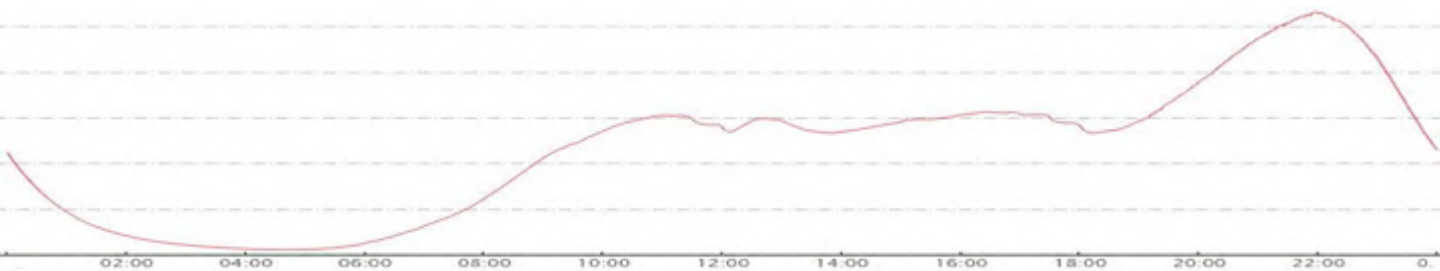
mmsynclogicsvr, shenzhen, C1(4.2.0)



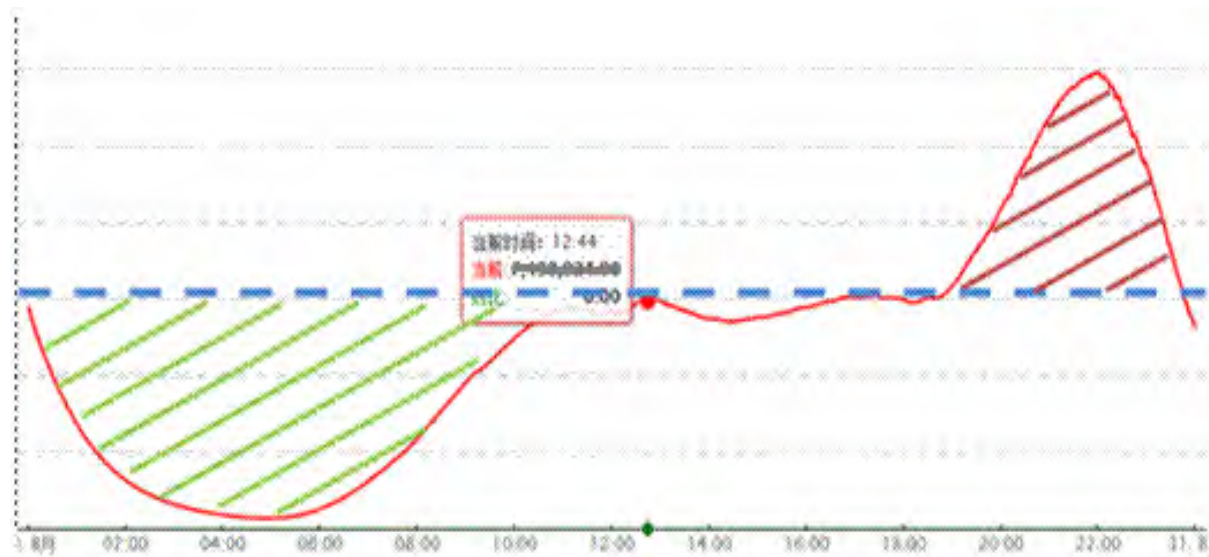
性能管理闭环



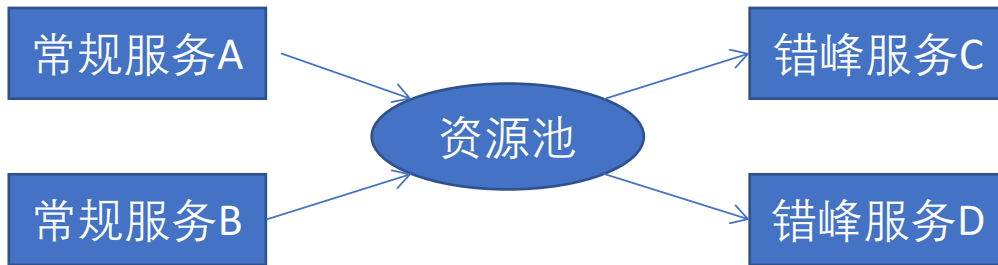
几种业务形态



削峰填谷



在线业务削峰



常规服务：高峰期后释放资源

错峰服务：获取资源



离线计算填谷

- 离线任务的运行时段
 - 01:00 ~ 08:00 不限任务
 - 08:00 ~ 20:00 任务入队控制
- 离线任务的资源占用
cpu.shares + memory.limit_in_bytes + blkio
- 离线任务的优先级最低



自动调度小结

- 全盘调度在线服务，充分利用资源
- 离线任务无需单独申请计算资源



高效运维社区
GreatOps Community



会议

培训

咨询

- 8月18日 DevOpsDays 上海
- 全年 DevOps China 巡回沙龙
- 11月17日 DevOps金融上海

- EXIN DevOps Master 认证培训
- DevOps 企业内训
- DevOps 公开课
- 互联网运维培训
- 企业DevOps 实践咨询
- 企业运维咨询



商务经理：刘静女士
电话 / 微信：13021082989
邮箱：liujing@greatops.com



Thanks

高效运维社区
开放运维联盟

荣誉出品



想第一时间看到
高效运维社区公众号
的好文章吗？

请打开高效运维社区公众号，点击右上角小人，如右侧所示设置就好

