智能弹性容量管理

by 张娟（希宁）

ABOUT ME

2011年加入阿里

2016年~至今，从事集团弹性资源管理

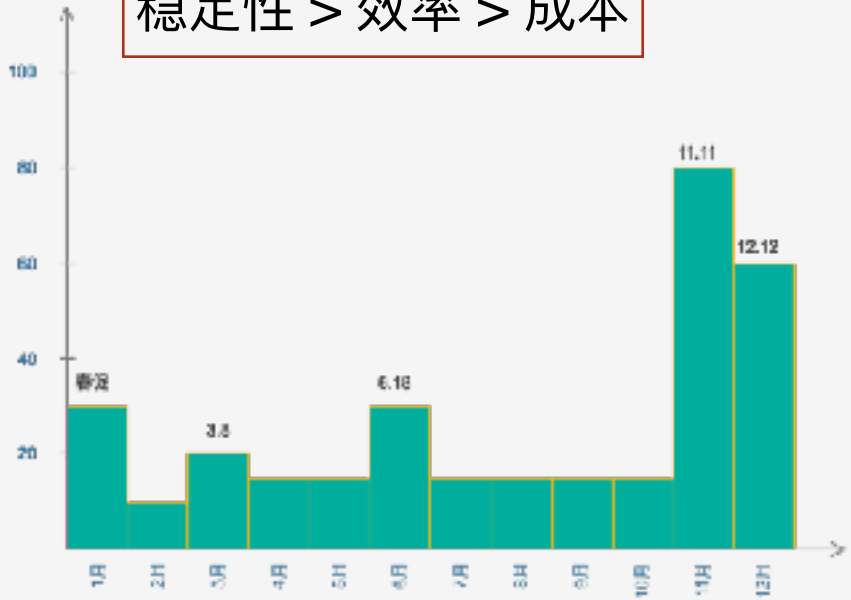
Contents

01

背景

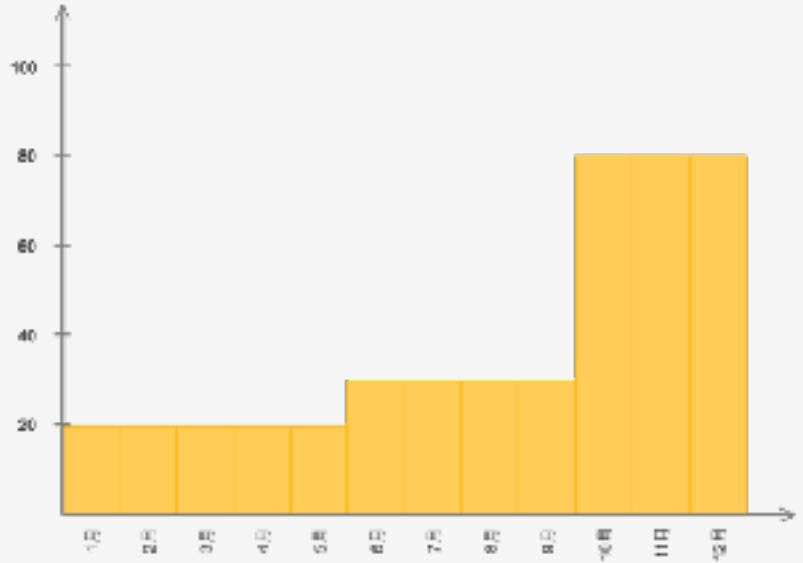
成本运维的困境

稳定性 > 效率 > 成本



月度流量趋势图

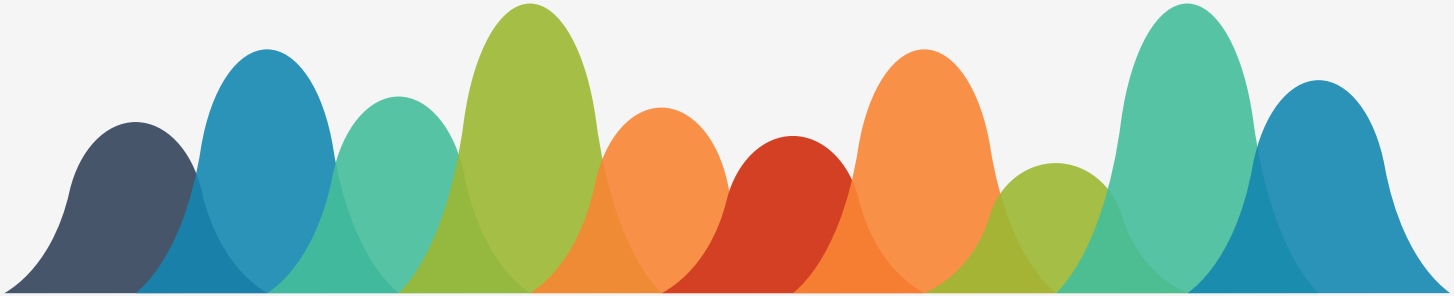
月度资源保有量趋势图



成本运维的困境

传统做法

- 经验预估
- 等比预估
- 压测到目标量级



容量规划的意义

- 用更科学手段做资源运营。

预算规划 1

2 软件迭代容量评估

峰值流量应对 3

4 新IDC引流容量初始化

02

智能弹性容量管理

概述

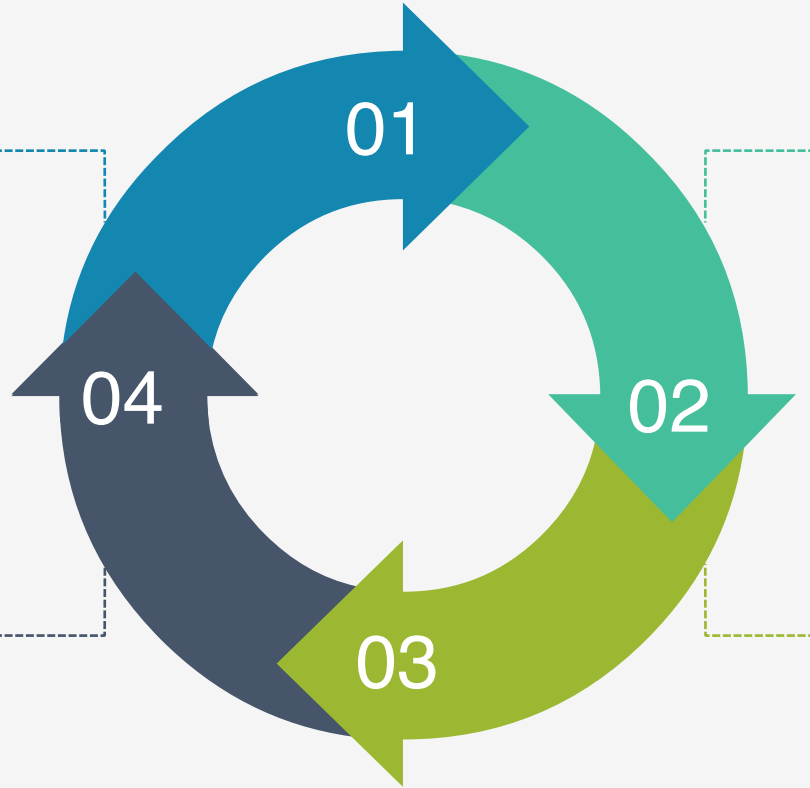
What “容量规划”+“弹性伸缩”+“风险评估”

How “智能决策” + “自动执行”

Target “最优利用” + “容量自治”

Who “Dev+Ops” + “业务”

智能容量管理反馈环



需求管理

容量预测

数据分析

弹性伸缩

容量预测公式

抽象成简化公式：

目标机器数 = 预测流量 / 应用单机能力（预测）

流量预测

自然态流量预测	利用集群流量时序特征回归
非自然态流量预测	全链路流量模型，基于业务目标的各应用集群流量预测，线性回归

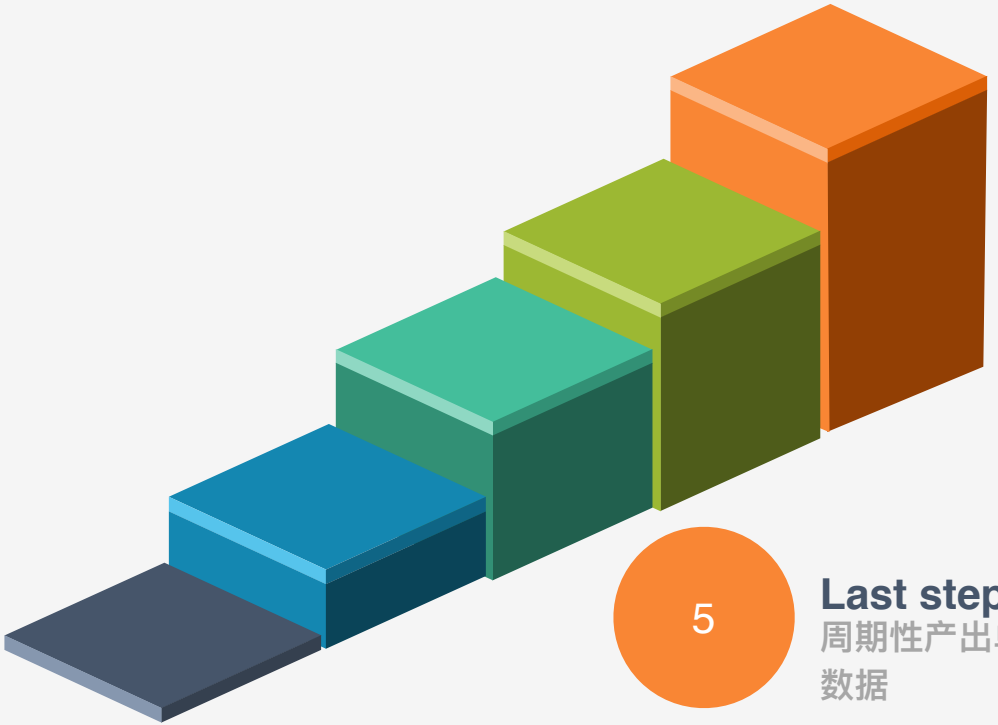
单机能力预测

First step
应用特征分析
1

Second step
提取关键性能指标，建立算法模型
2

Third Step
通过日常样本数据回归预测
3

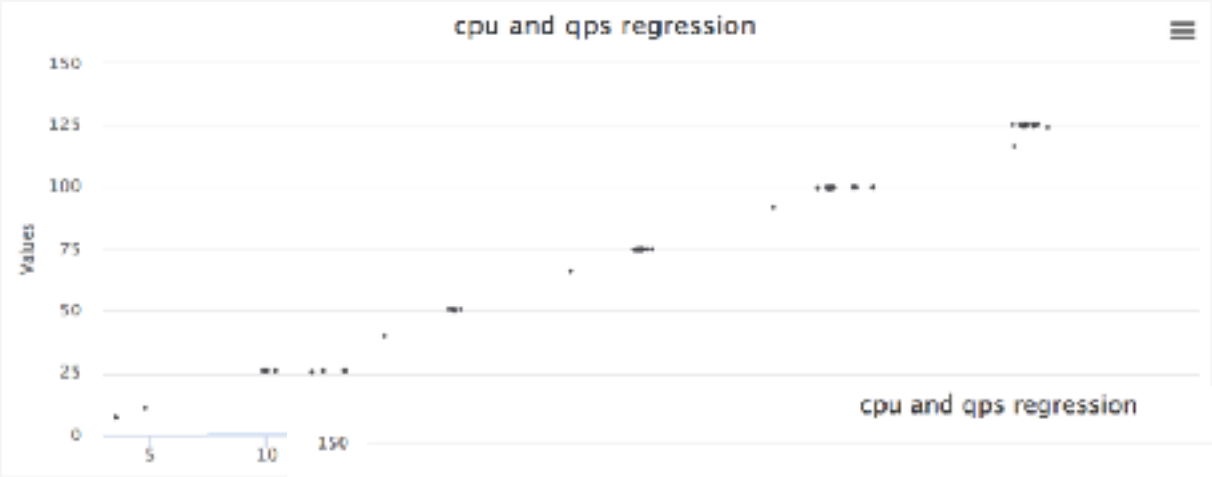
Fourth step
压测验证，效果反馈
4



5

Last step
周期性产出单机性能数据

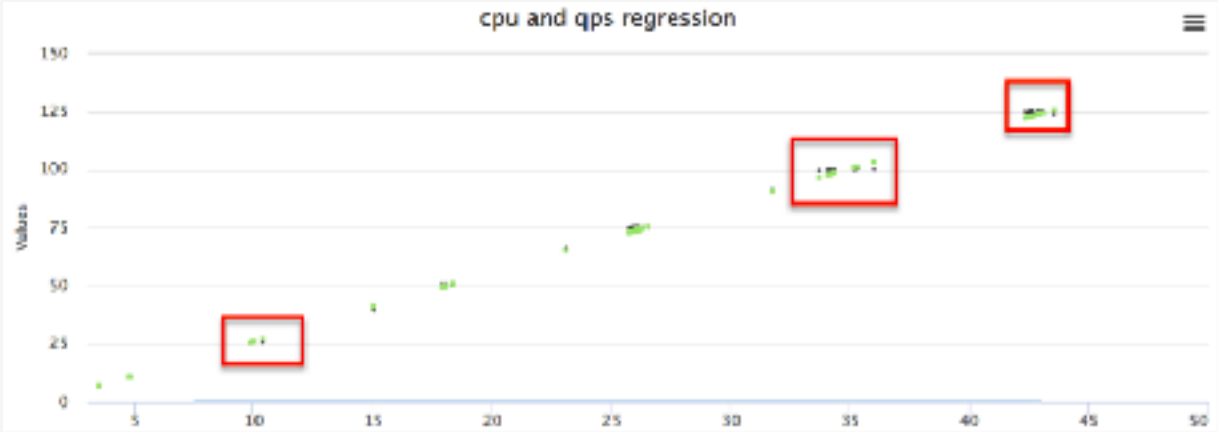
线性回归模型



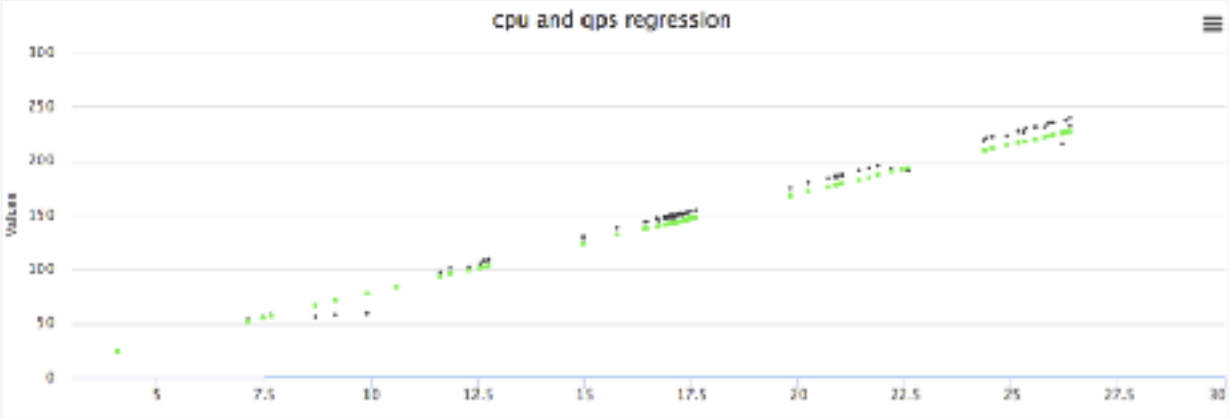
“.” 为真实样本点

“.” 为拟合后的点

模型校正



残差降噪



局部加权

效果评估

理论评估

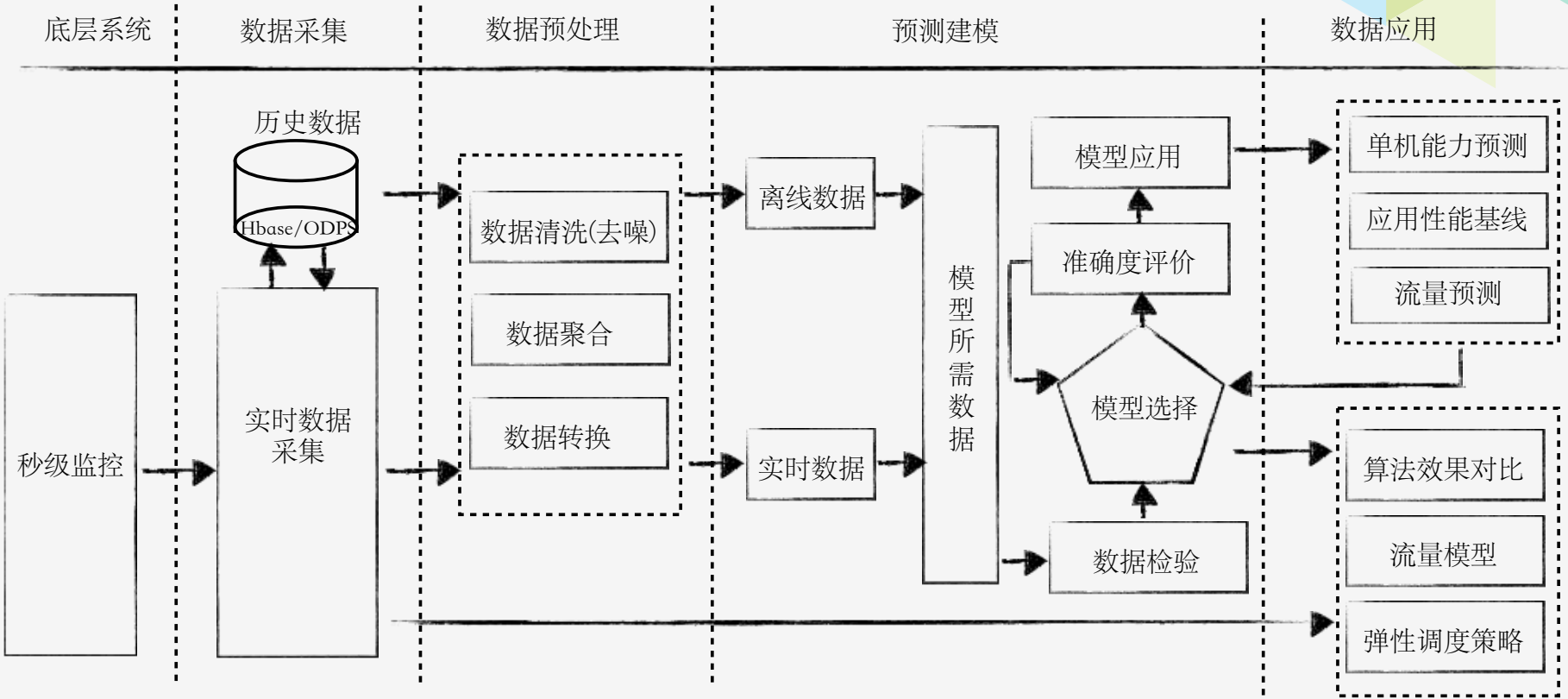
QPS {CPU(Max) / 2 } 预测 QPS {CPU(Max) }

1-abs(qps真实 –qps预测)/qps真实

实际评估

压测验证

数据处理框架


底层系统
数据采集
数据预处理
预测建模
数据应用
历史数据
Hbase/ODPS
秒级监控
实时数据采集
数据清洗(去噪)
数据聚合
数据转换
离线数据
实时数据
模型所需数据
模型应用
准确度评价
模型选择
数据检验
单机能力预测
应用性能基线
流量预测
算法效果对比
流量模型
弹性调度策略

APM应用性能管理

维护应用性能基线。

04

具体实践

实践场景

01 日常弹性

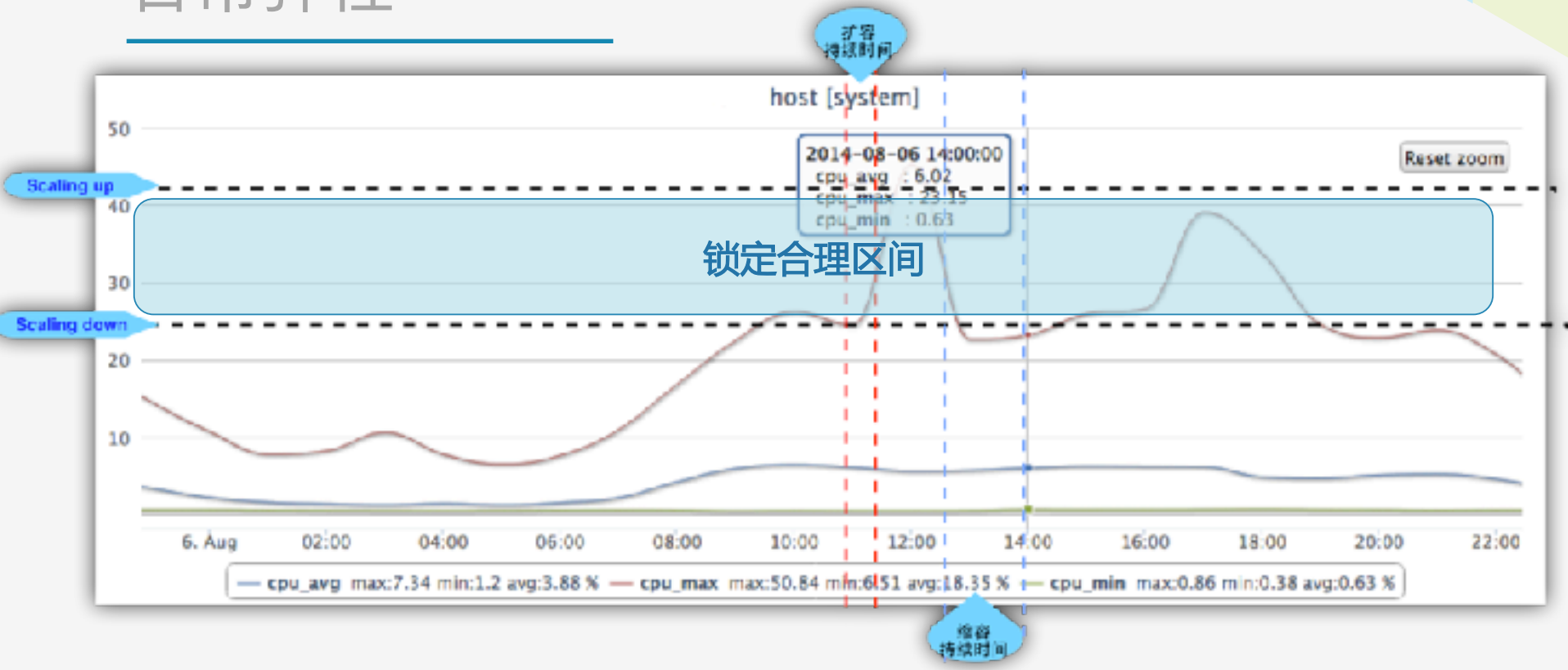
02 分时复用

03 边压边弹

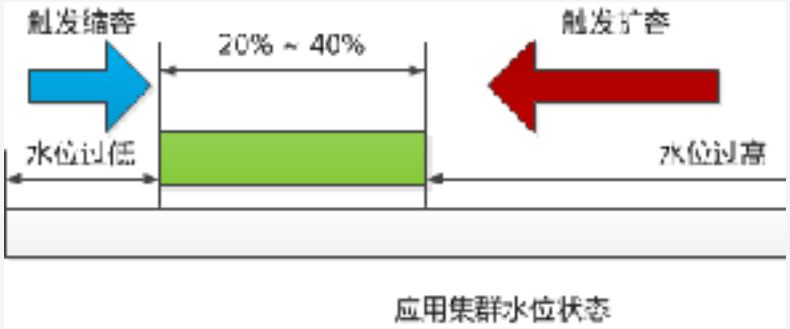
04 IDC引流评估

05 其他

日常弹性

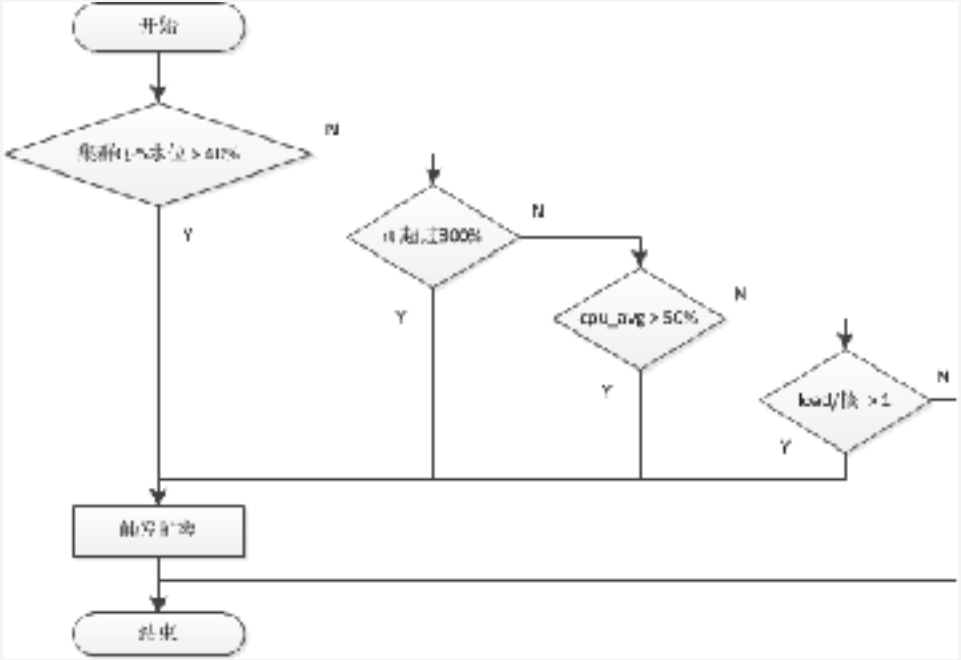

扩容
持续时间
host [system]
50
40
30
20
10
Scaling up
Scaling down
2014-08-06 14:00:00
cpu_avg : 6.02
cpu_max : 23.15
cpu_min : 0.63
Reset zoom
锁定合理区间
6. Aug
02:00
04:00
06:00
08:00
10:00
12:00
14:00
16:00
18:00
20:00
22:00
cpu_avg max:7.34 min:1.2 avg:3.88 %
cpu_max max:50.84 min:6.51 avg:18.35 %
cpu_min max:0.86 min:0.38 avg:0.63 %
缩容
持续时间

日常弹性



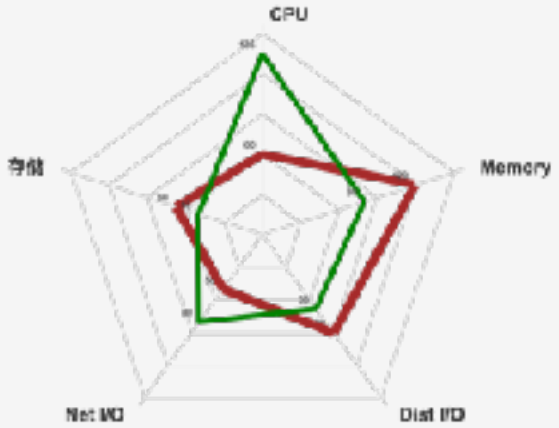
备注： 集群水位 = 集群QPS / （单机QPS极限值 * 机器数） 单机QPS极限能力

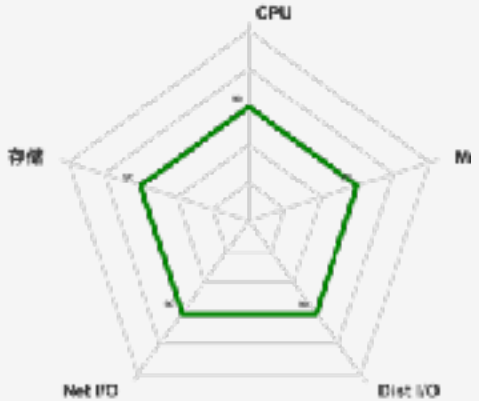
触发模式：手动、自动、定时

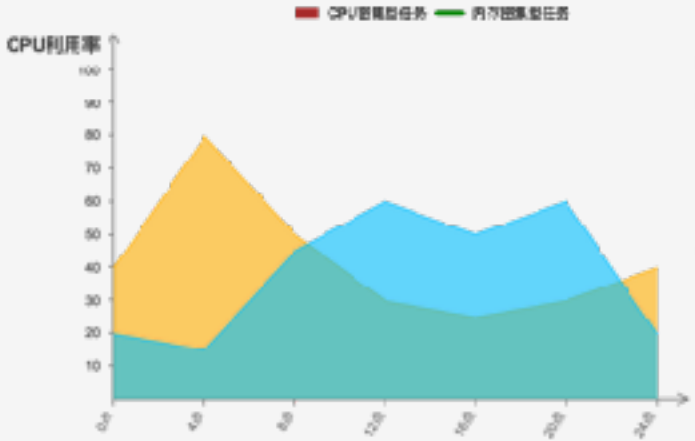


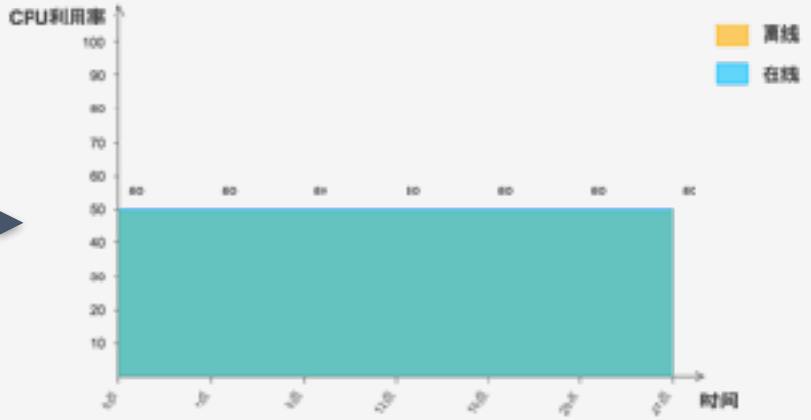
触发策略

分时复用背景


CPU
Memory
Dist I/O
Net I/O
存储


CPU
存储
Net I/O


CPU利用率
离线
在线


CPU利用率
离线
在线
时间

分时复用挑战

服务SLO保证 I

事件模型、QoS监控和熔断机制

资源边界最大化 II

精准权重调度和容量预测

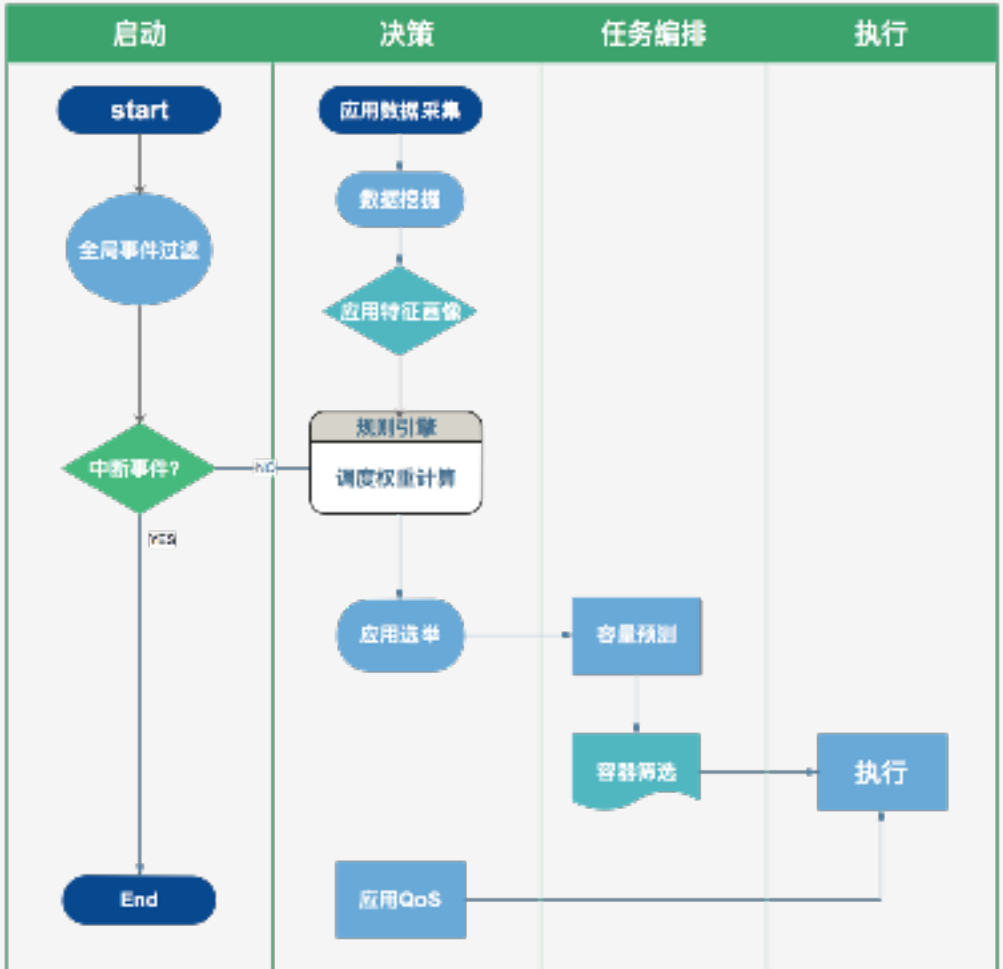
精细化分时调度 III

全时段精细化削峰填谷

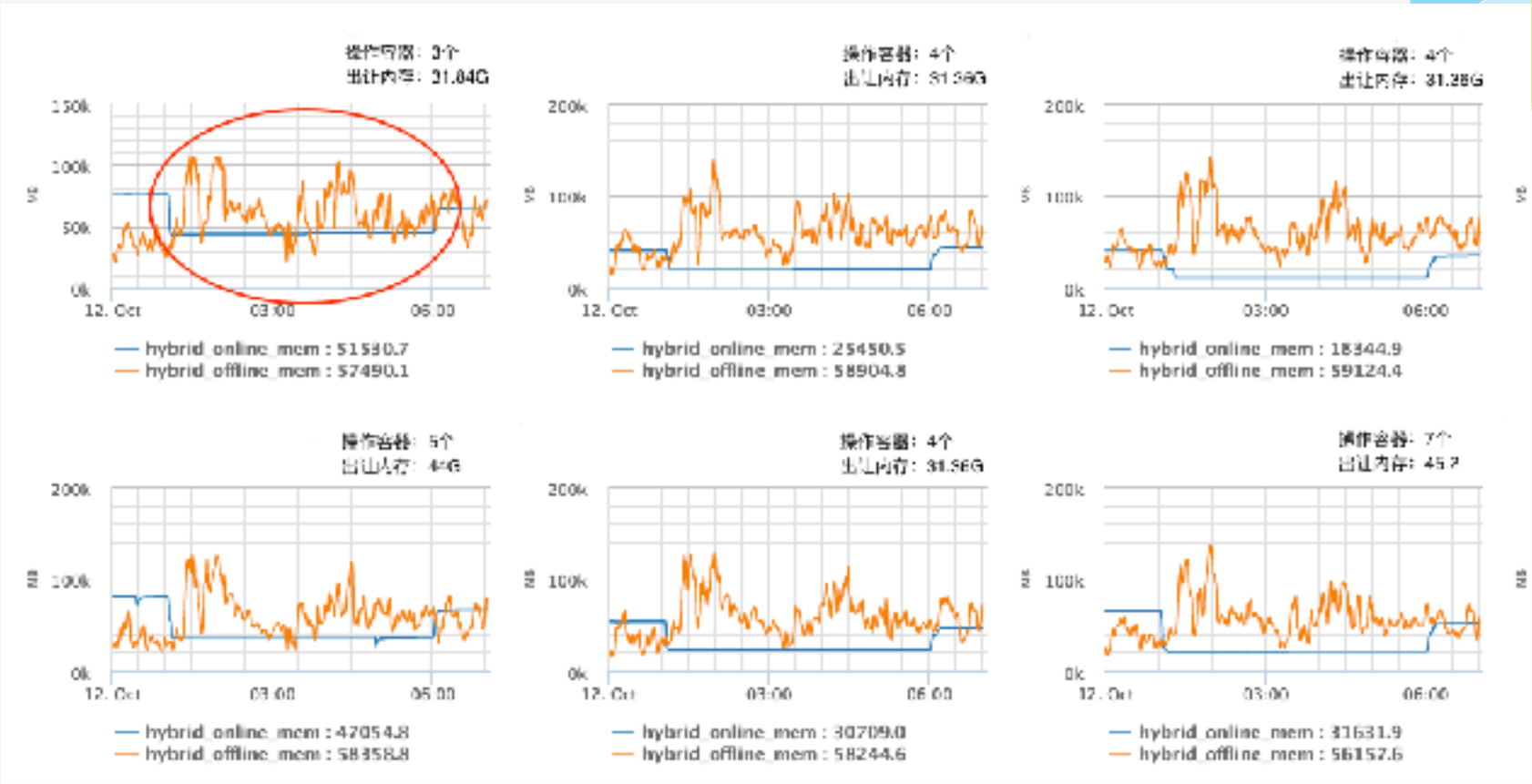
高效执行，快速恢复 IV

数据分析为前提，多种执行策略

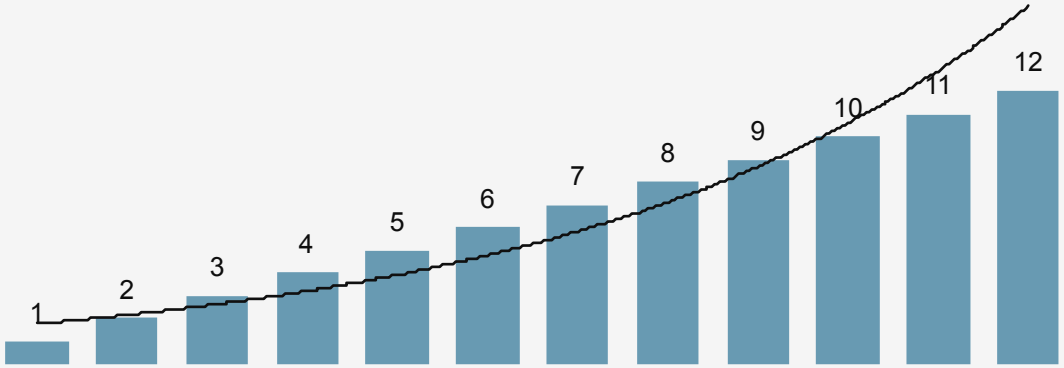
分时复用核心模块


启动
决策
任务编排
执行
start
全局事件过滤
中断事件?
End
应用数据采集
数据挖掘
应用特征画像
规则引擎
调度权重计算
应用选单
容量预测
容器筛选
执行
应用QoS

分时复用效果

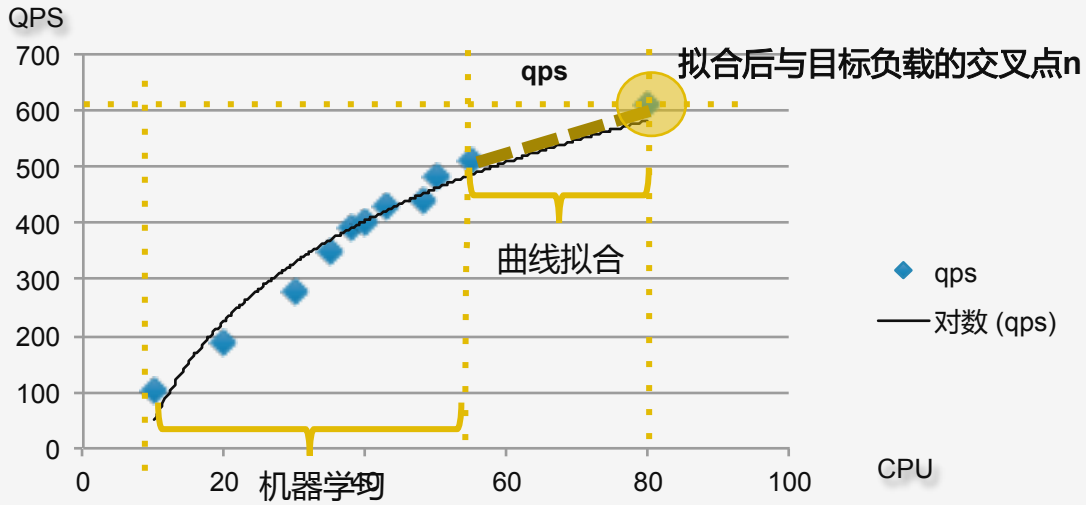

150k
100k
50k
0k
12. Oct
03:00
06:00
hybrid_online_mem : 51590.7
hybrid_offline_mem : 57490.1
200k
100k
0k
hybrid_online_mem : 25450.5
hybrid_offline_mem : 58904.8
200k
100k
0k
hybrid_online_mem : 18344.9
hybrid_offline_mem : 59124.4
200k
100k
0k
hybrid_online_mem : 47054.8
hybrid_offline_mem : 58358.8
200k
100k
0k
hybrid_online_mem : 30709.0
hybrid_offline_mem : 58244.6
200k
100k
0k
hybrid_online_mem : 31631.9
hybrid_offline_mem : 56157.6

边压边弹

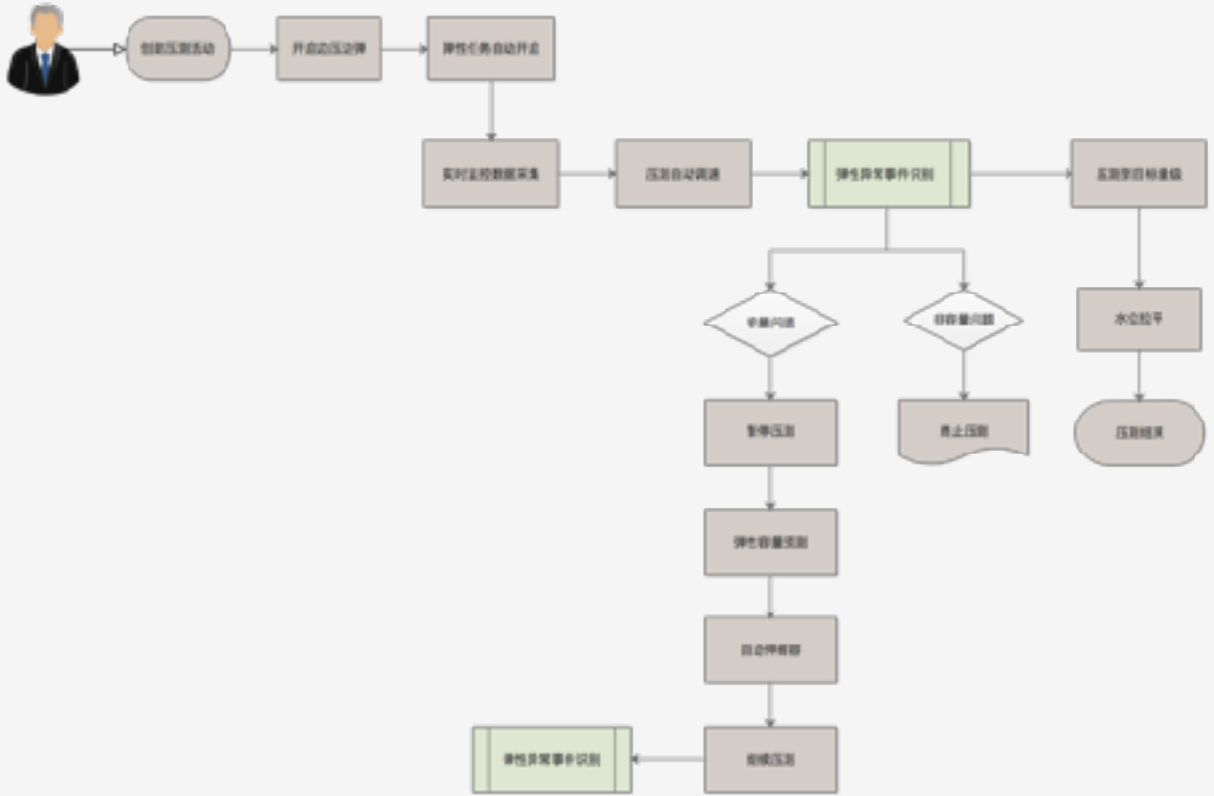


压测交易量增长趋势

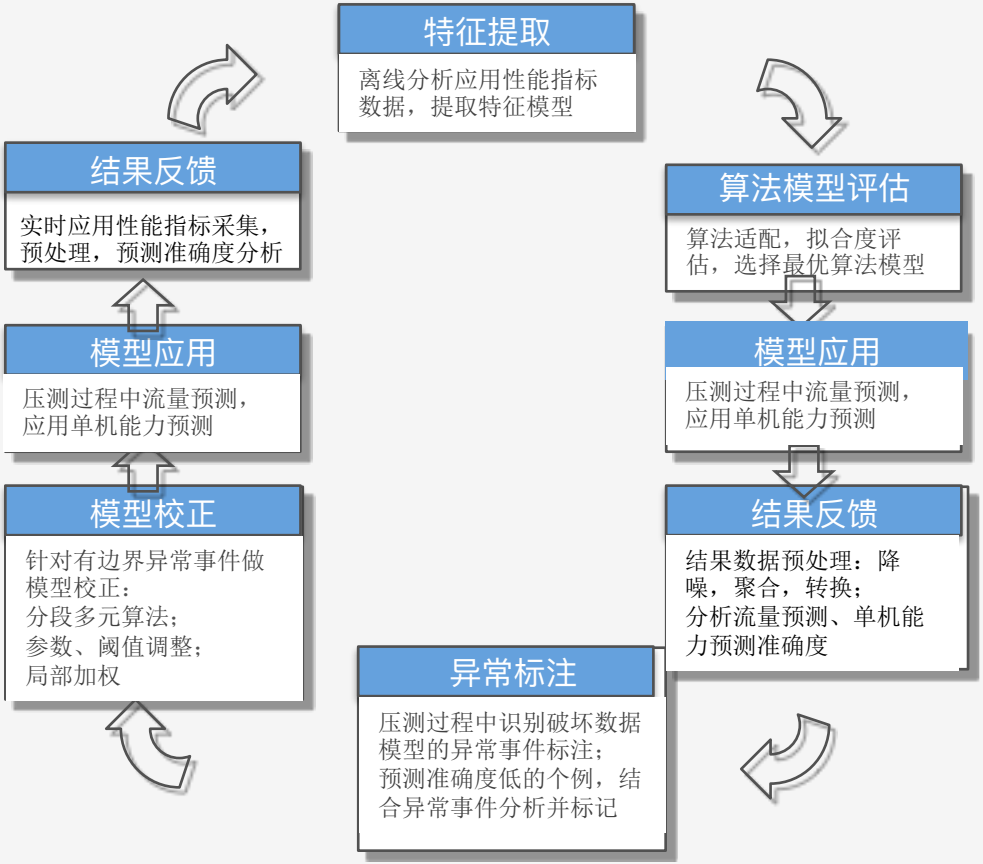
负载预测,容量评估



边压边弹执行流程



边压边弹智能决策


特征提取
离线分析应用性能指标数据，提取特征模型
算法模型评估
算法适配，拟合度评估，选择最优算法模型
模型应用
压测过程中流量预测，应用单机能力预测
结果反馈
结果数据预处理：降噪，聚合，转换；
分析流量预测、单机能力预测准确度
异常标注
压测过程中识别破坏数据模型的异常事件标注；
预测准确度低的个例，结合异常事件分析并标记
模型校正
针对有边界异常事件做模型校正：
分段多元算法；
参数、阈值调整；
局部加权
模型应用
压测过程中流量预测，应用单机能力预测
结果反馈
实时应用性能指标采集，预处理，预测准确度分析

边压边弹目标

压测无人值守

05

未来展望

- 智能化容量自治。
- 整体集群资源0冗余。

F&Q

Please Join Us!

Thanks!