

基于空闲资源的 弹性计算实践

陈杰

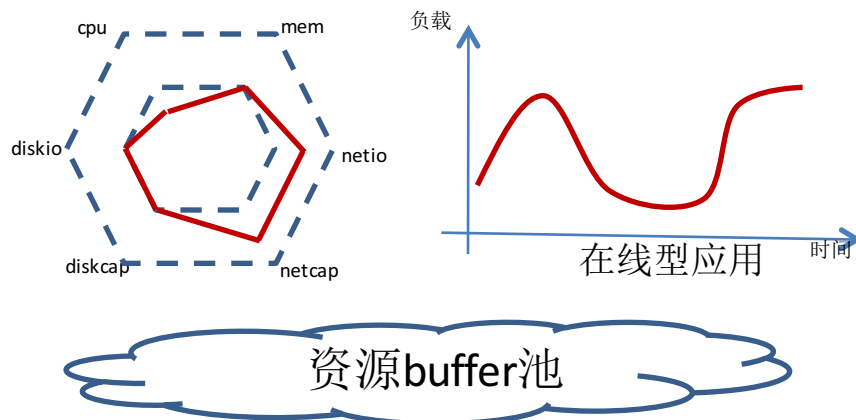
腾讯-架构平台部

2017-6-10

大纲

- 项目背景
- 关键挑战
- 技术架构
- 经验分享

项目背景



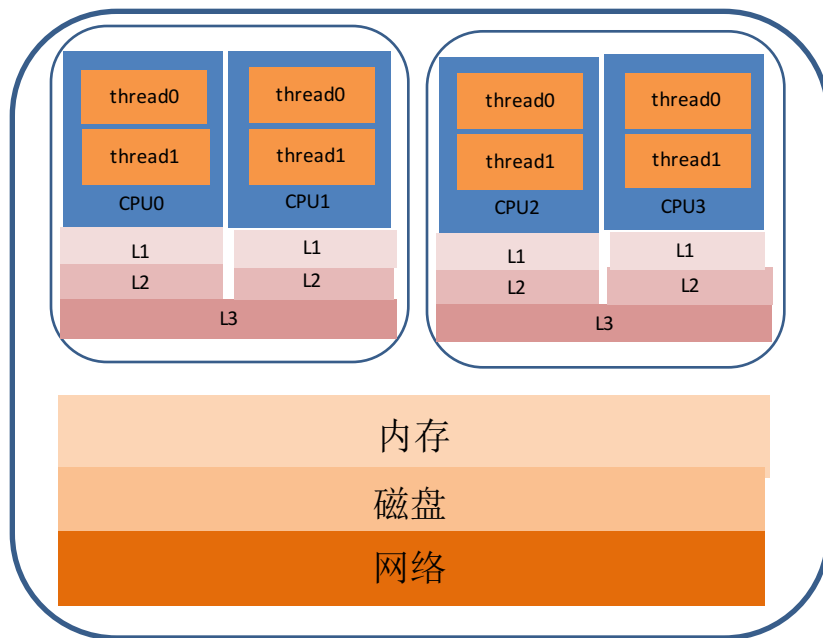
图片，视频，AI等计算成本
逐渐成为不可承受之重



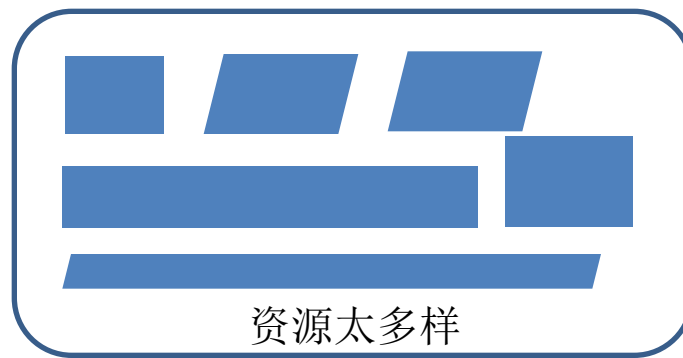
资源利用并不充分，可挖掘
空闲资源多

用闲置资源满足不断增长的计算需求

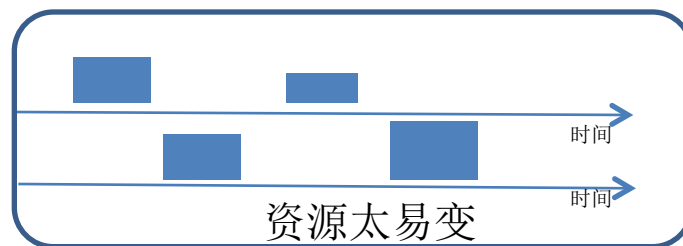
关键挑战



共享资源易冲突



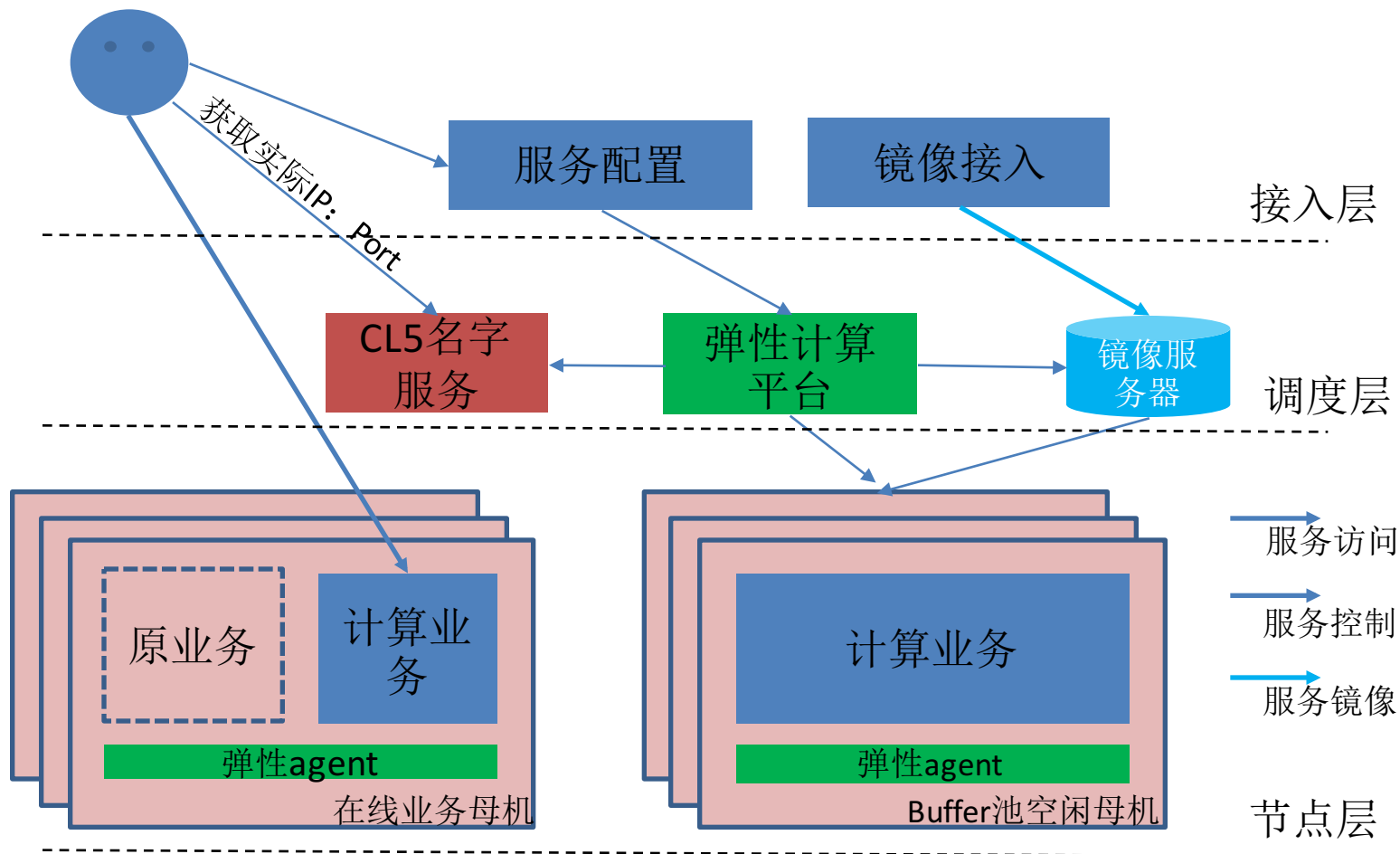
资源太多样



资源太易变

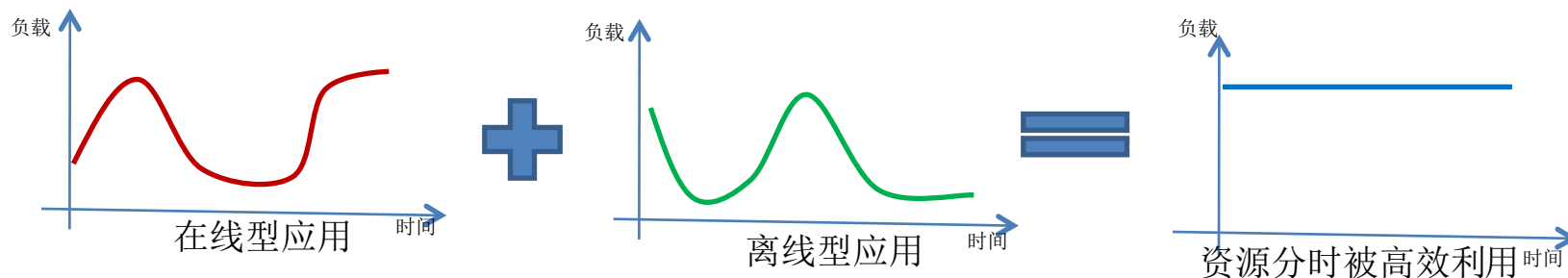
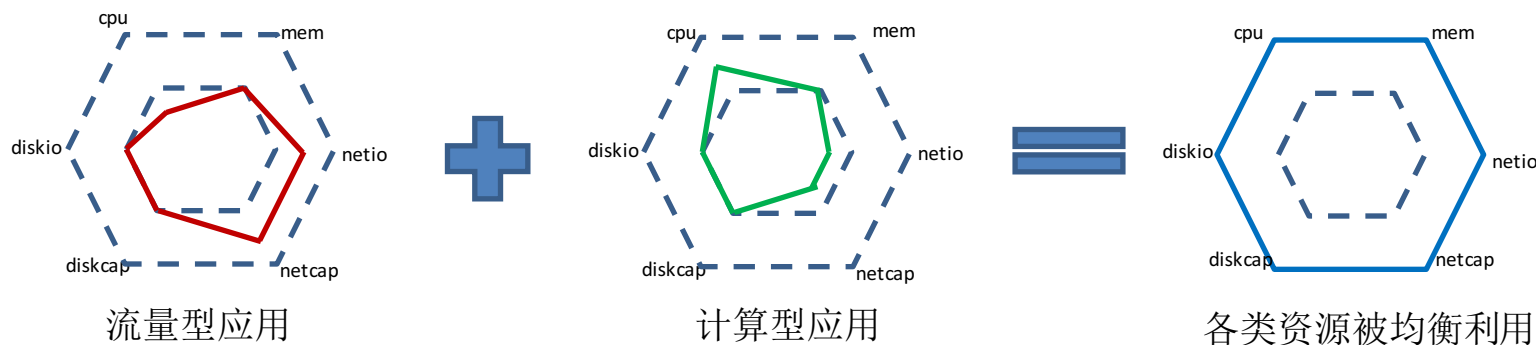
资源难以被用好

技术架构



技术架构-怎么做好在线业务的质量保障？

保容量：依据业务资源模型做好混搭与隔离



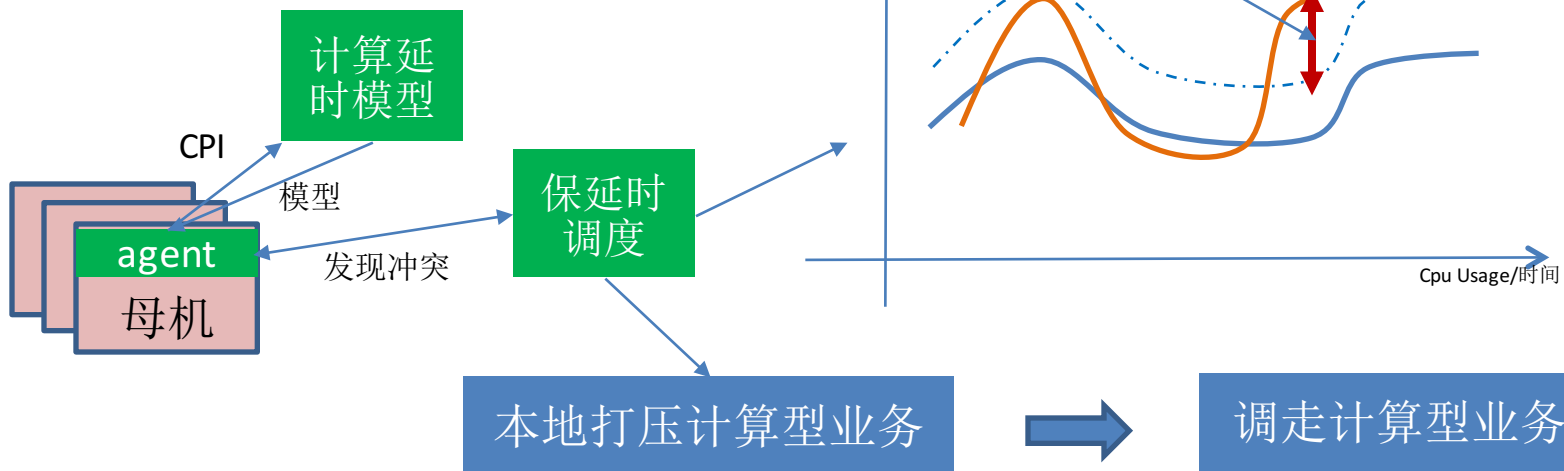
挑战：怎么建立合理的资源模型？

技术架构-怎么做好在线业务的质量保障？

保延时：及时识别计算延时下降并调度

L1	L2	L3	MEM
2ns	6ns	15ns	80ns

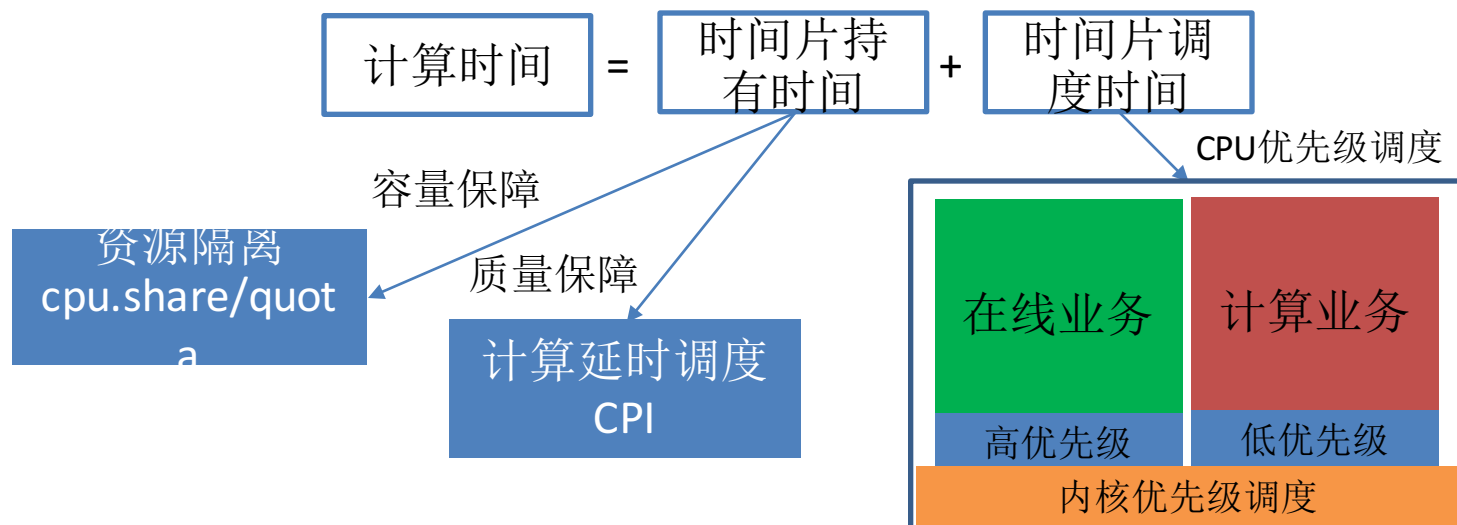
CPI = 指令周期数/发送指令数



挑战：怎么应对误报及瞬时毛刺？

技术架构-怎么做好在线业务的质量保障？

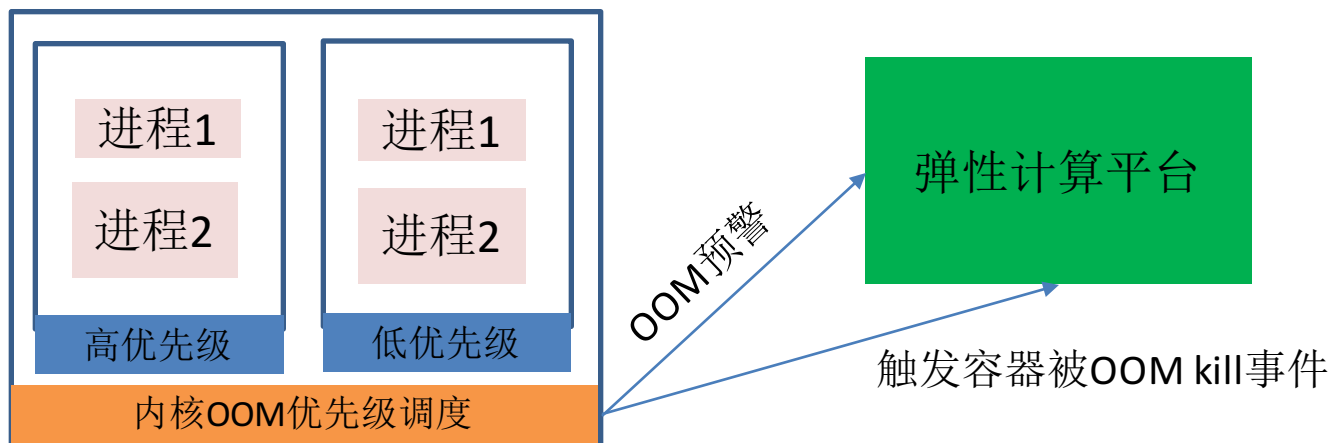
消毛刺：支持CPU组业务优先级调度



4608: 3	无共享	共享优化前	共享优化后	优化效果
平均延时ms	1.59	1.65	1.62	1.8%
最大延时ms	23.85	29.67	23.79	19.8%
毛刺数	19768	31576	21282	32.6%

技术架构-怎么做好在线业务的质量保障？

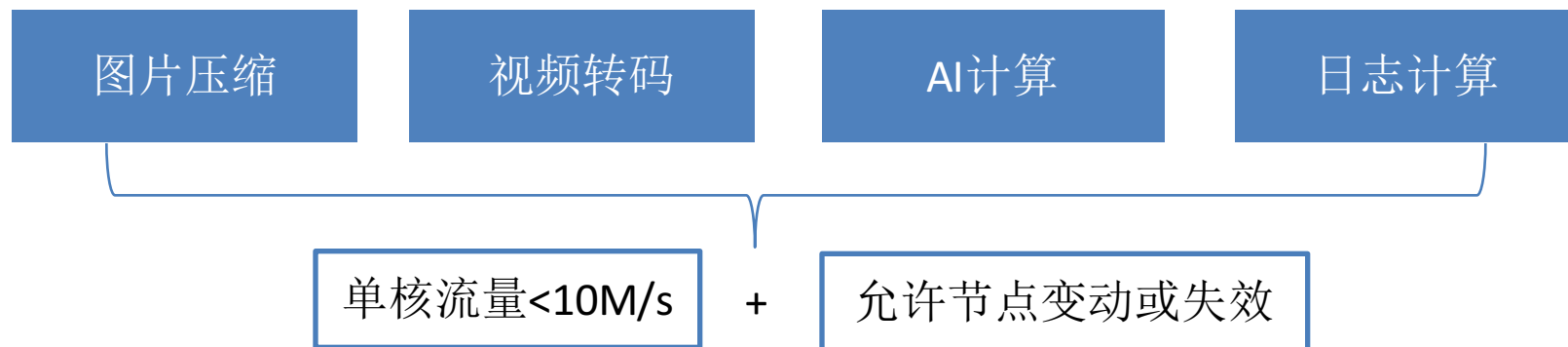
消异常：OOM预测及优先级调度



OOM场景	默认行为	优化后
内存不足预警	无动作	平台调走低优先级容器
容器内存不足	Kill 存活短内存高oom score高的进程	Kill低优先级容器，并触发事件通知
母机内存不足		

技术架构-怎么用好弹性资源？

筛选业务，提供场景式计算服务



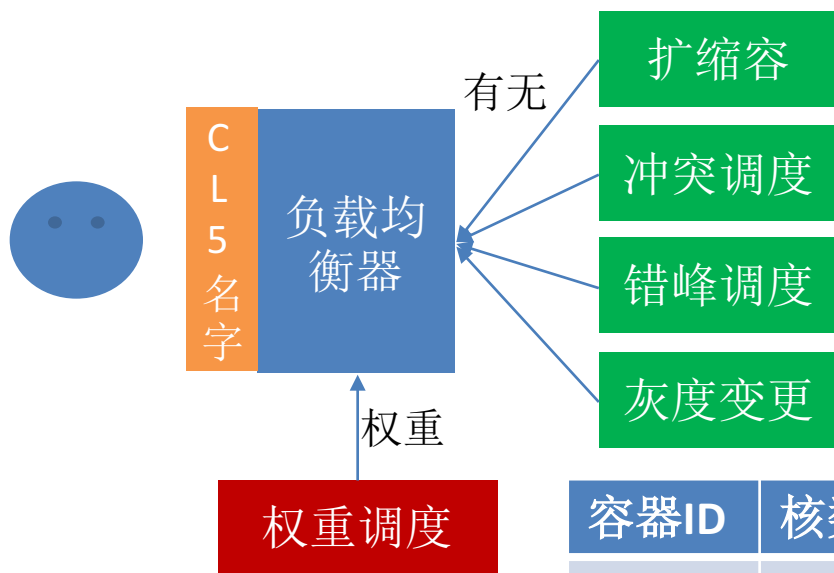
计算优先级	场景举例	资源类型
在线计算型-关键路径	图片上传压缩	资源规格固定，调度少
在线计算型-非关键路径	图片下载压缩/视频转码	资源规格固定，调度较多
离线计算型	AI/日志计算	资源规格不固定，调度多

挑战：怎么和原有业务计算框架集成？



技术架构-怎么用好弹性资源？

对业务封装资源的易变和多样性



CPU	性能
Intel E3-1230v2	1.00
Intel X3440	0.57
Intel E5-2680v4	0.92
..	..

容器ID	核数	性能	quota	单核权重	总权重
1	4	1.00	90%	0.90	3.6
2	6	0.92	60%	0.55	3.3
3	8	0.57	70%	0.40	3.2

挑战：怎么通过调度实现负载均衡？



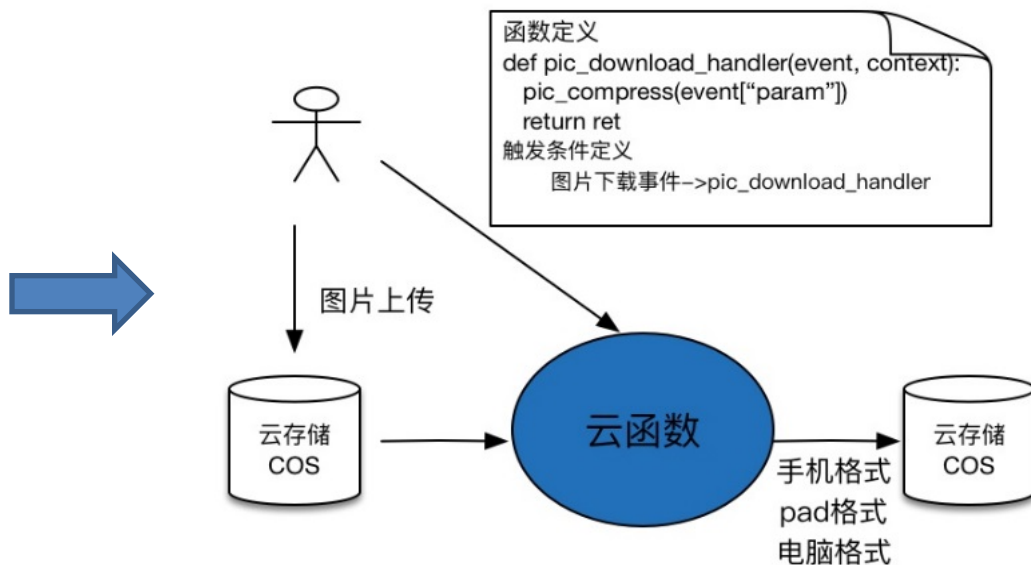
技术架构-怎么用好弹性资源？

以云函数方式进一步解放生产力

用户依然需要
关心任务调度

用户依然需要
关心资源细节，
比如端口

用户依然需要
编写大量后台
代码



欢迎试用腾讯云-无服务器云函数

经验分享-想清楚提供机制还是提供策略

故事1:

A业务利用率阈值设置不合理，高峰期保留大量资源没充分利用

B业务设置很合理，高峰期确没资源扩容了

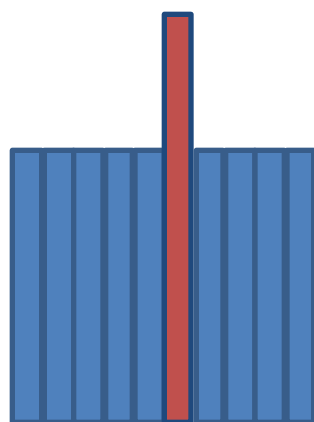
故事2:

平台默认打开自动扩缩容，自动调配资源；

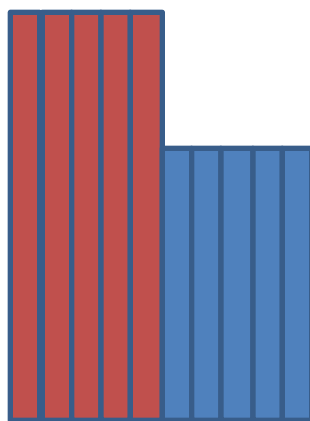
A业务对自动扩缩容机制不知情，发起了版本变更

1. 业务流程要闭环
2. 未闭环前提供机制不提供策略

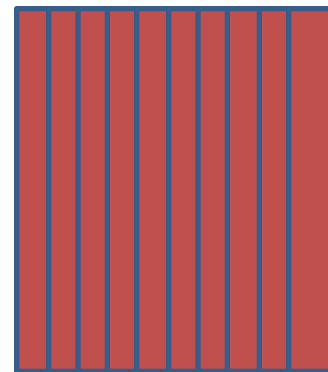
经验分享-自动扩缩容前负载先确保均衡



业务不均



资源不均



完美状态

1. 业务请求要先均匀：如大小图分离
2. 资源与业务权重需匹配：如根据整机负载调权重

经验分享-重视底层能力及风险

Loopback +
device mapper

无须格式化

XFS + Overlay

须格式化

Ceph..

网络穿越流量

1. Loop back + DM切忌生产环境使用
2. 规模越大越会遇到社区碰不到的问题
3. 热补丁机制能救急，应该成为平台基础能力

欢迎探讨，一起提升行业资源利用效率

