

# 阿里管控体系在双十一的实践



子矜

阿里巴巴中间件高可用架构团队

2016/4/22

# QCon

2016.10.20~22

上海·宝华万豪酒店

## 全球软件开发大会 2016

### [上海站]



购票热线: 010-64738142

会务咨询: [qcon@cn.infoq.com](mailto:qcon@cn.infoq.com)

赞助咨询: [sponsor@cn.infoq.com](mailto:sponsor@cn.infoq.com)

议题提交: [speakers@cn.infoq.com](mailto:speakers@cn.infoq.com)

在线咨询(QQ): 1173834688

团·购·享·受·更·多·优·惠

**7折** 优惠(截至06月21日)  
现在报名, 立省2040元/张

# 阿里管控体系 在双十一的实践



简介



阿里线上管控体系，应对挑战



关于我们



# 阿里中间件技术部

打造世界第一流的中间件







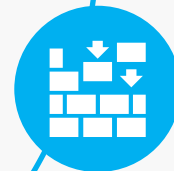
开关预案



阿里线上管控体系



限流



降级&流量调度

# 阿里管控体系 的最佳实践



简介



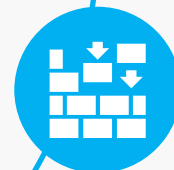
阿里线上管控体系，应对挑战



关于我们



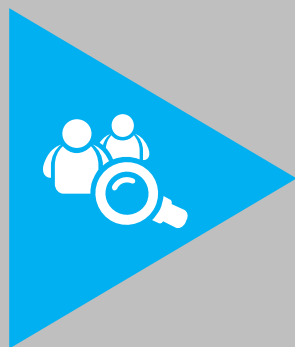
# 阿里线上管控体系



限流



# 场景一:零点零分小唐下单了



# 天猫 Tmall.com

男人的品味 始于足下



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。

4 秒后重试...

# 用户洪峰在双十一

天猫移动端销售金额  
突破1亿

**140000** 笔/秒

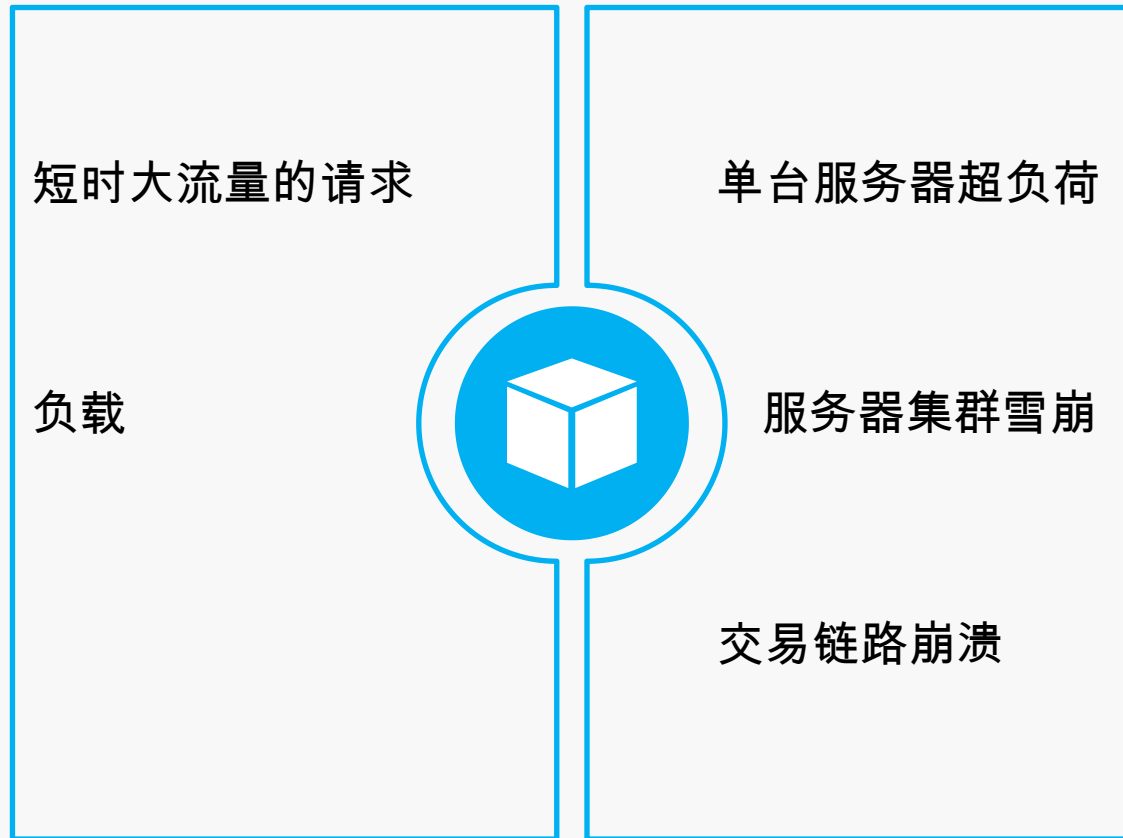
销售金额破百亿

**75** 秒

创建订单当天峰值

**38** 分钟

# 洪峰对系统意味着什么





# 限流的考虑因素

①

允许访问  
的速率

②

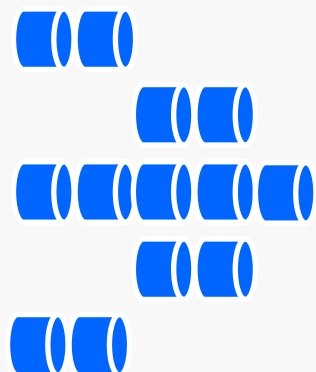
爆发量

③

爆发间隔  
时间

# 令牌桶限流

来的请求

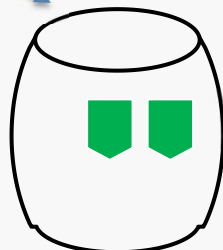


拦截器



按照每 $1/r$ 秒的速度向桶中存放1令牌

存储量为 $b$ 的令牌桶

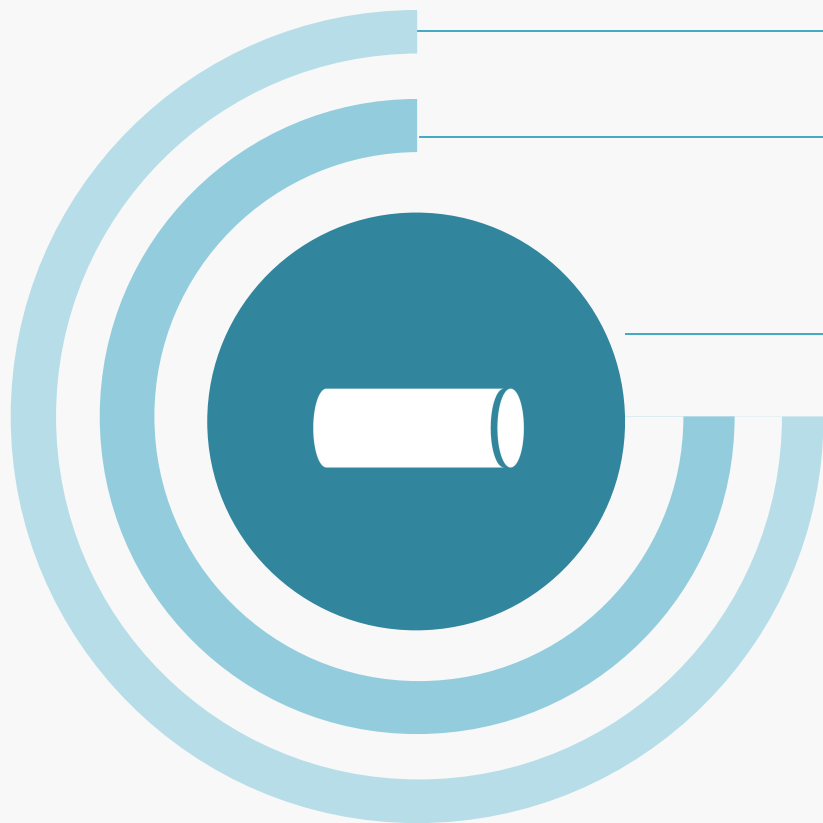


继续发送



丢弃

# 双十一零点:小唐下单



- Token rate:  $r$

- 允许通过的最大流量是  $Q_{max} = b + r$

- 持续流量:  $T_{max} = \begin{cases} \frac{b}{m-r}, & m < r, \text{ 否则洪} \\ \infty & \end{cases}$

峰不会到来

- 承诺突发量:  $L_{max} = T_{max} * r$

# 如何把Token Rate转换成通过率？

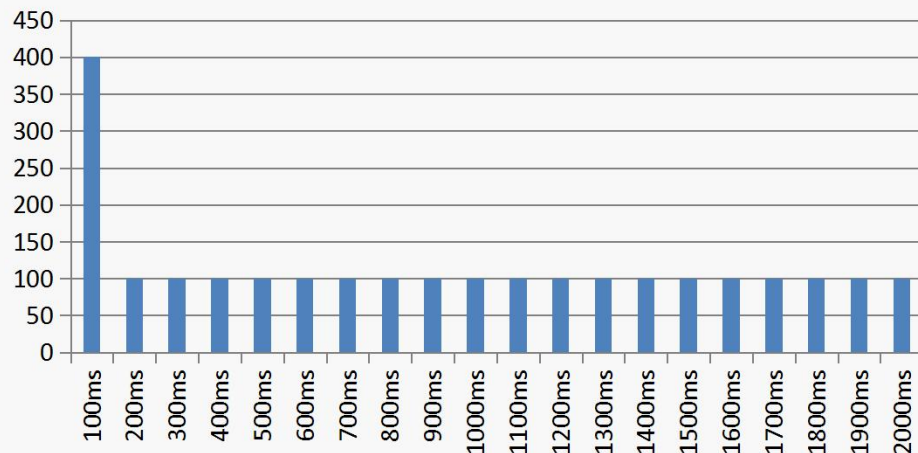
桶的大小是**300**个令牌

系统的通过率为**1000**qps，把一秒切成**10**个格子，每个格子的时间窗口为**100**ms，每个格子发放**1000/10**个令牌

洪峰场景：

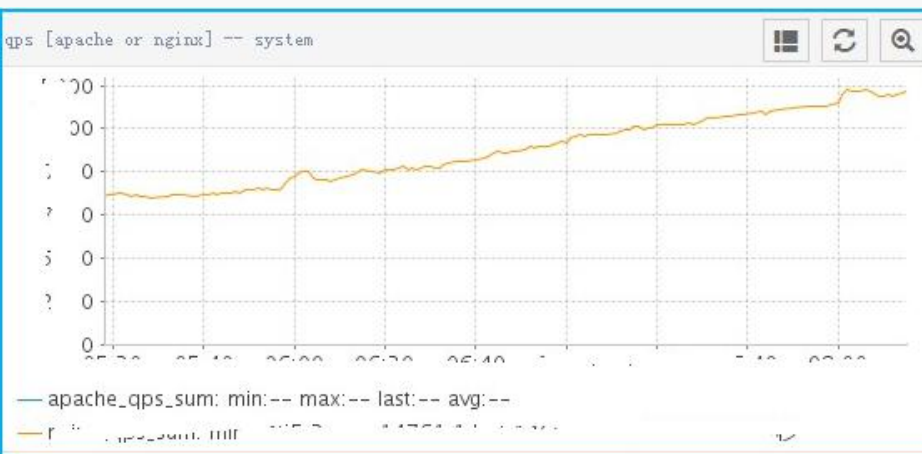
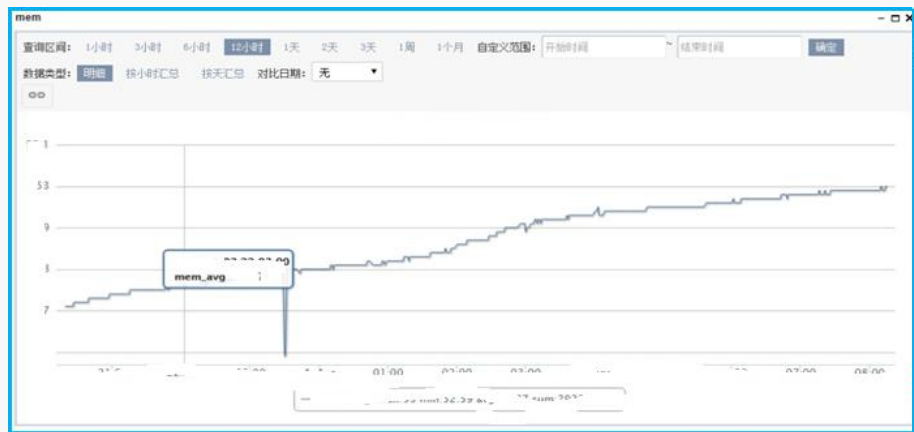
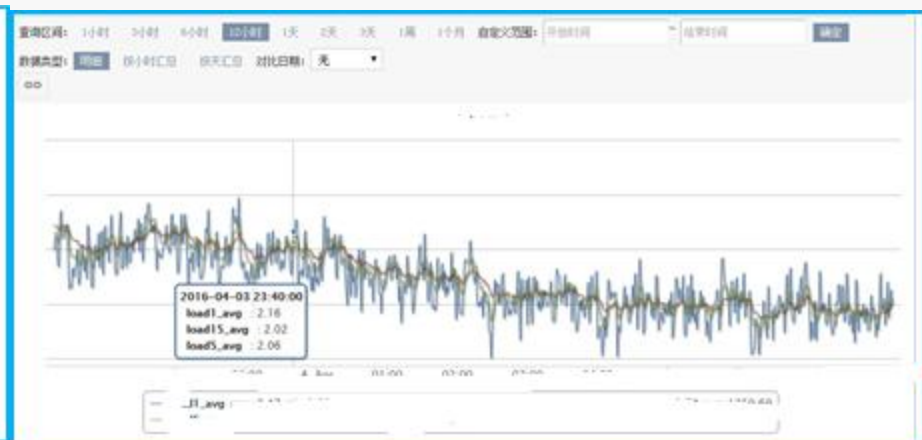
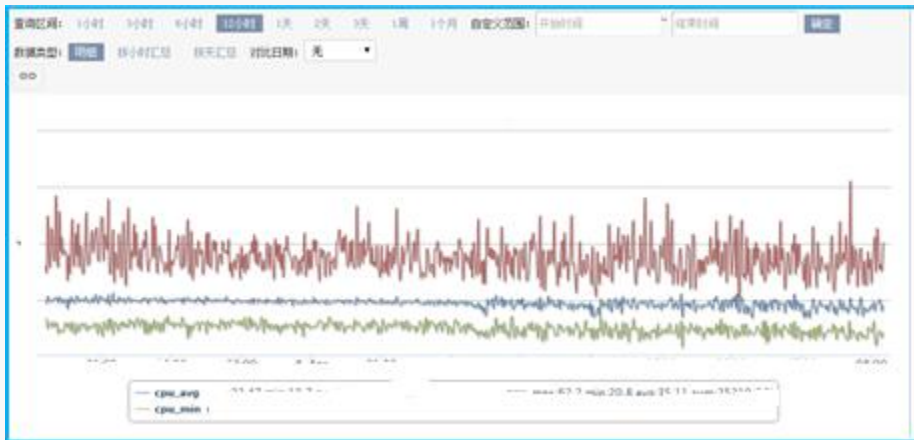
- 双十一0点之前，桶里放满了令牌
- 在双十一0点到0点10分，每秒的请求超过了10000

通过量

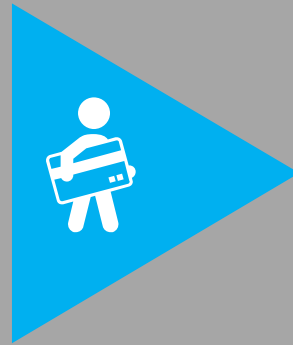
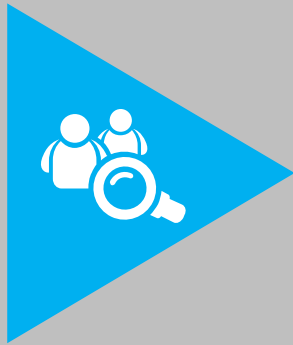




# 双十一系统表现



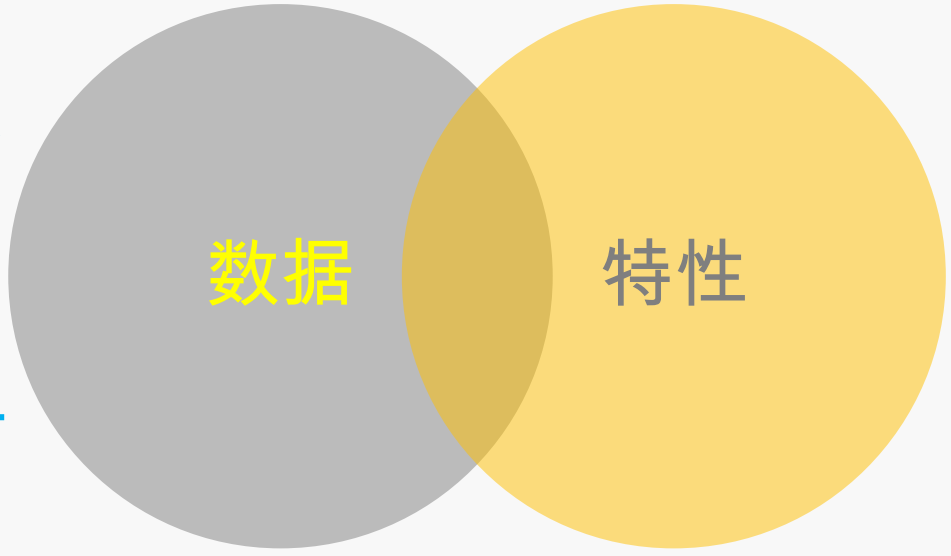
# 场景二：双十一零点零五分:小唐的 订单状态变为已经发货



# 系统间的回调洪峰

前5分钟包裹  
处理量是484万

前38分钟包裹  
处理量是3754万

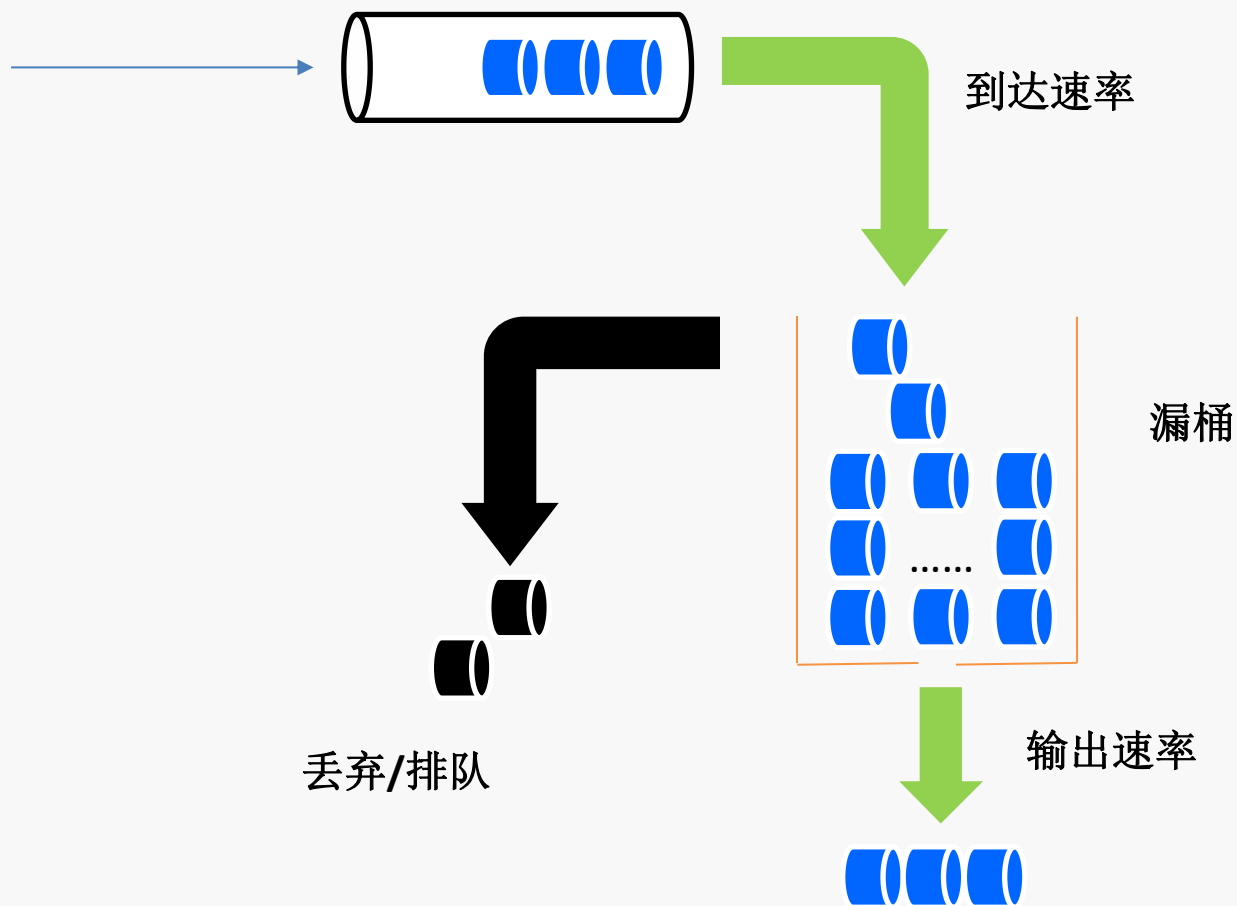


有时间间隔

调用量大

允许有延迟

# 漏桶算法





# 小唐的订单什么时候回调完成？



当回调洪峰到来时， $\tau$ 为桶的容量， $T$ 为请求到来的个数， $\sigma$ 为漏水的个数，最坏延迟时间：

$$\tau / (T - \sigma)$$

# 限流框架的要素

监控模块

实时监控  
收集数据  
反馈分析

限流决策

区分场景  
用户洪峰  
回调洪峰  
系统保护

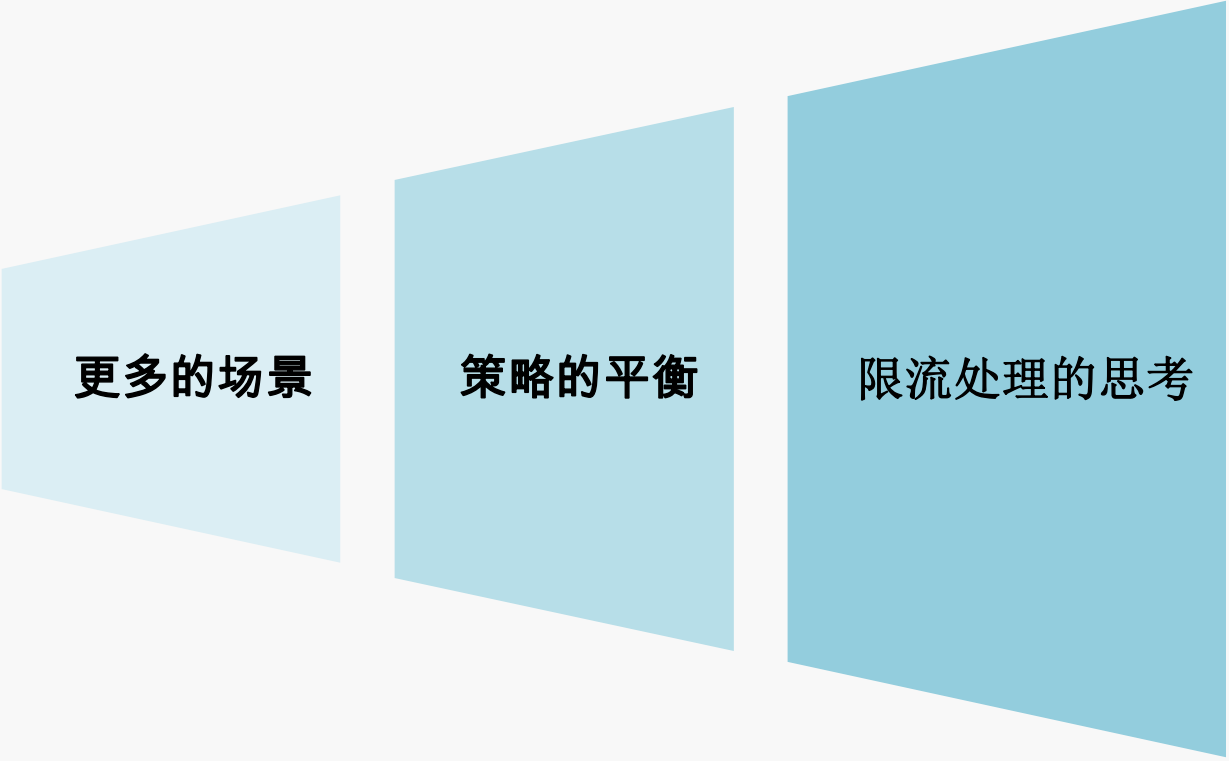
规则变更

动态调整  
令牌桶容量  
令牌产生速率

限流处理

丢弃请求  
等待队列

# 场景延伸



更多的场景

策略的平衡

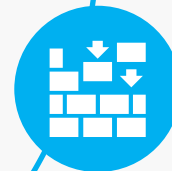
限流处理的思考



## 阿里线上管控体系



限流



降级&流量调度



# 系统的可用性



## 抱歉！页面无法访问.....

可能因为：

网址有错误 >请检查地址是否完整或存在多余字符

网址已失效 >可能页面已删除，活动已下线等

[了解更多原因>>](#)

[寻找答案](#)

或者逛逛：[淘宝首页](#) | [我的淘宝](#) | [搜宝贝/店铺](#)

# 系统的可用性



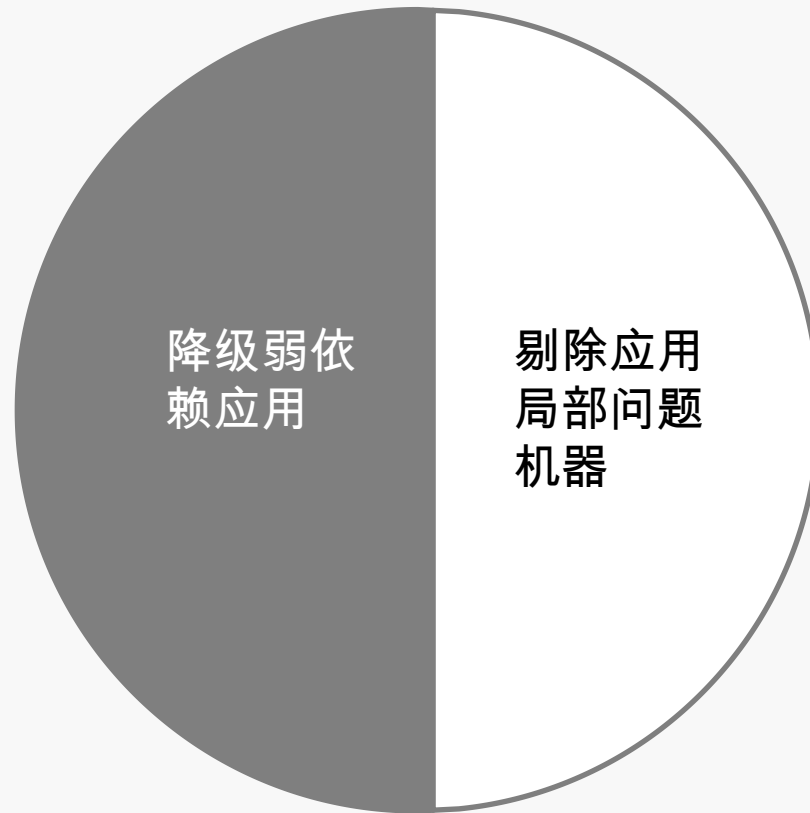
# 系统的可用性

买不到这个商品的可能性是： $\rho = \sigma_1 + (1 - \sigma_1)\sigma_2 + (1 - \sigma_1)(1 - \sigma_2)\sigma_3 + (1 - \sigma_1)(1 - \sigma_2)(1 - \sigma_3)\sigma_4 + \dots \prod_{i=1}^n (1 - \sigma_{i-1}) * \sigma_i$

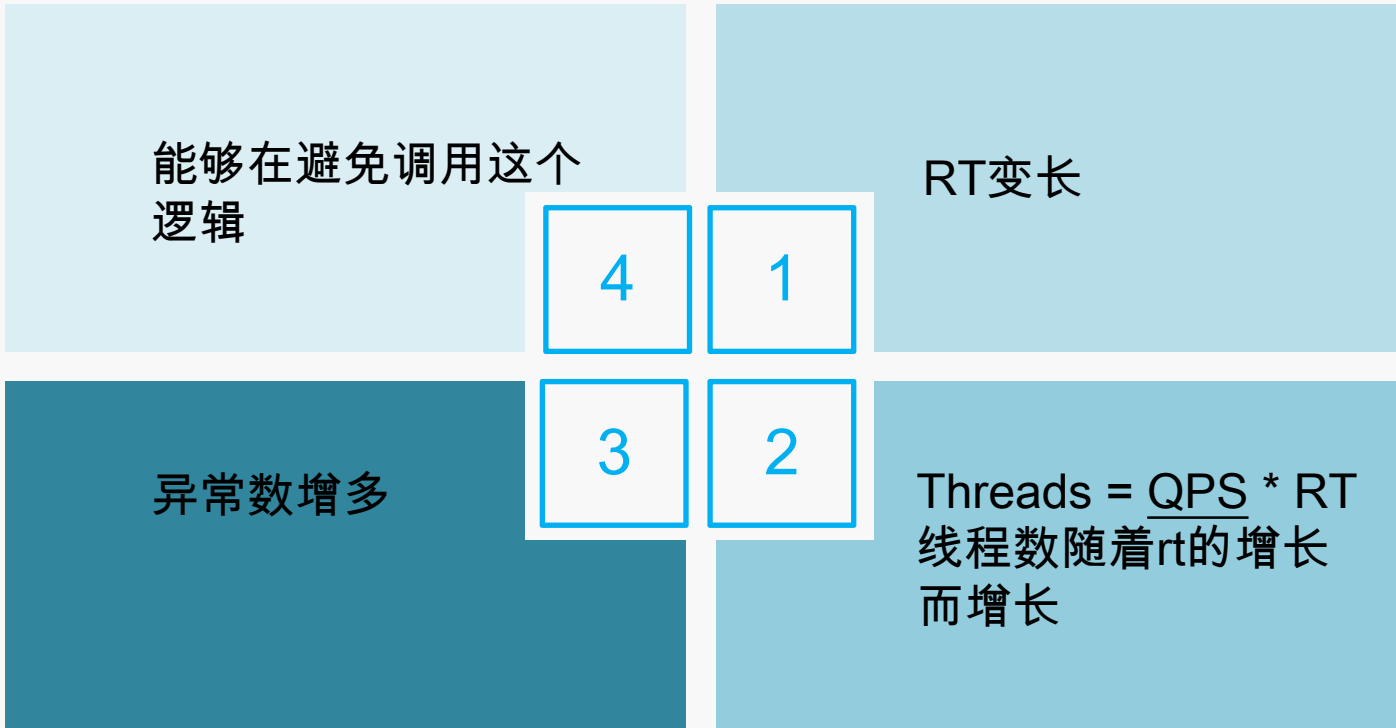
假设小唐下单经过了n个系统，每个系统的不可用性为 $\sigma(0 \leq \sigma \leq 1)$

N越大，小唐失败的概率就越高

# 提高可用性的方式



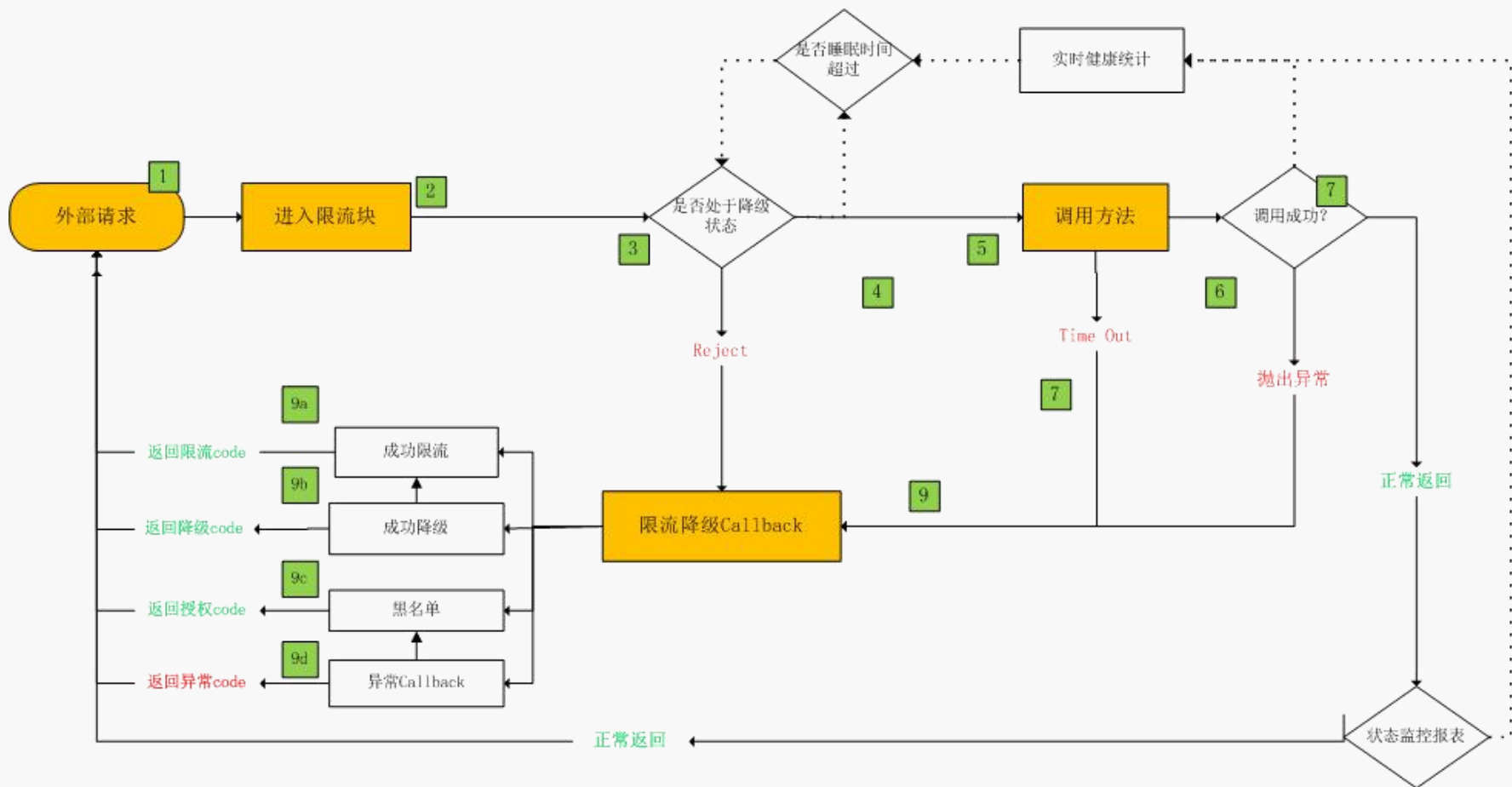
# 上游系统如何发现弱依赖应用不可用



# 弱依赖应用不可用

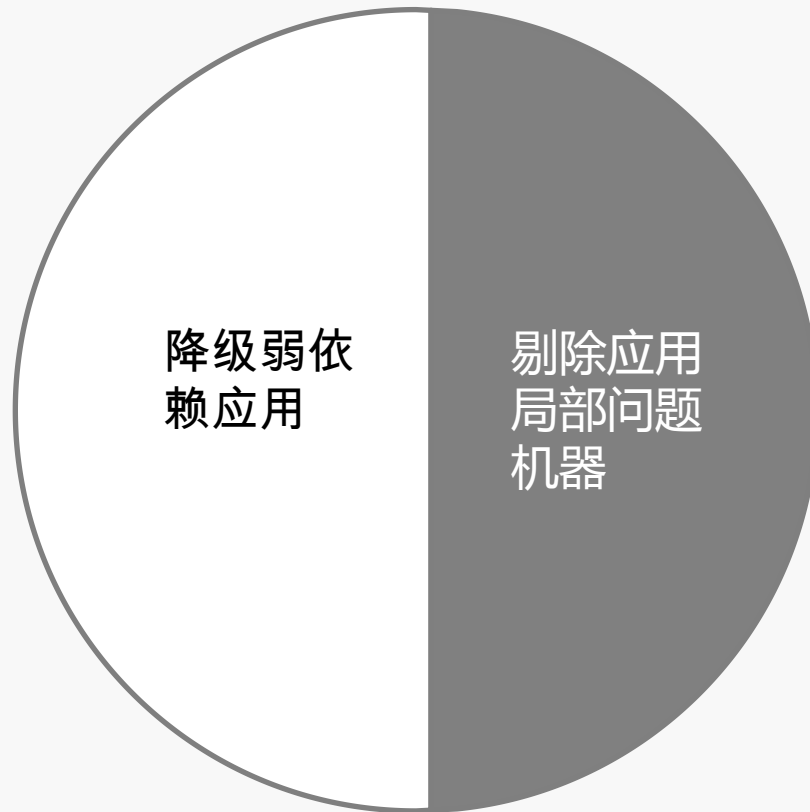


# 自动降级框架

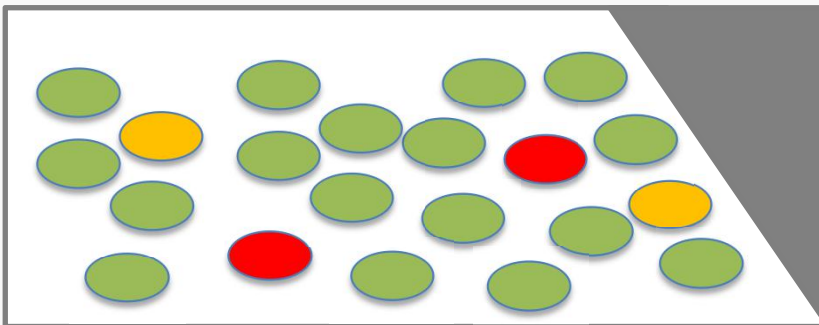




# 提高可用性的方式

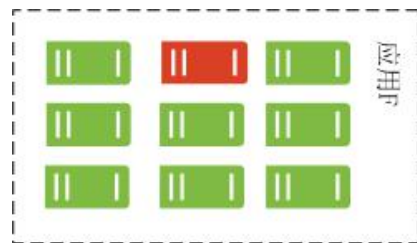


# 应用局部不可用

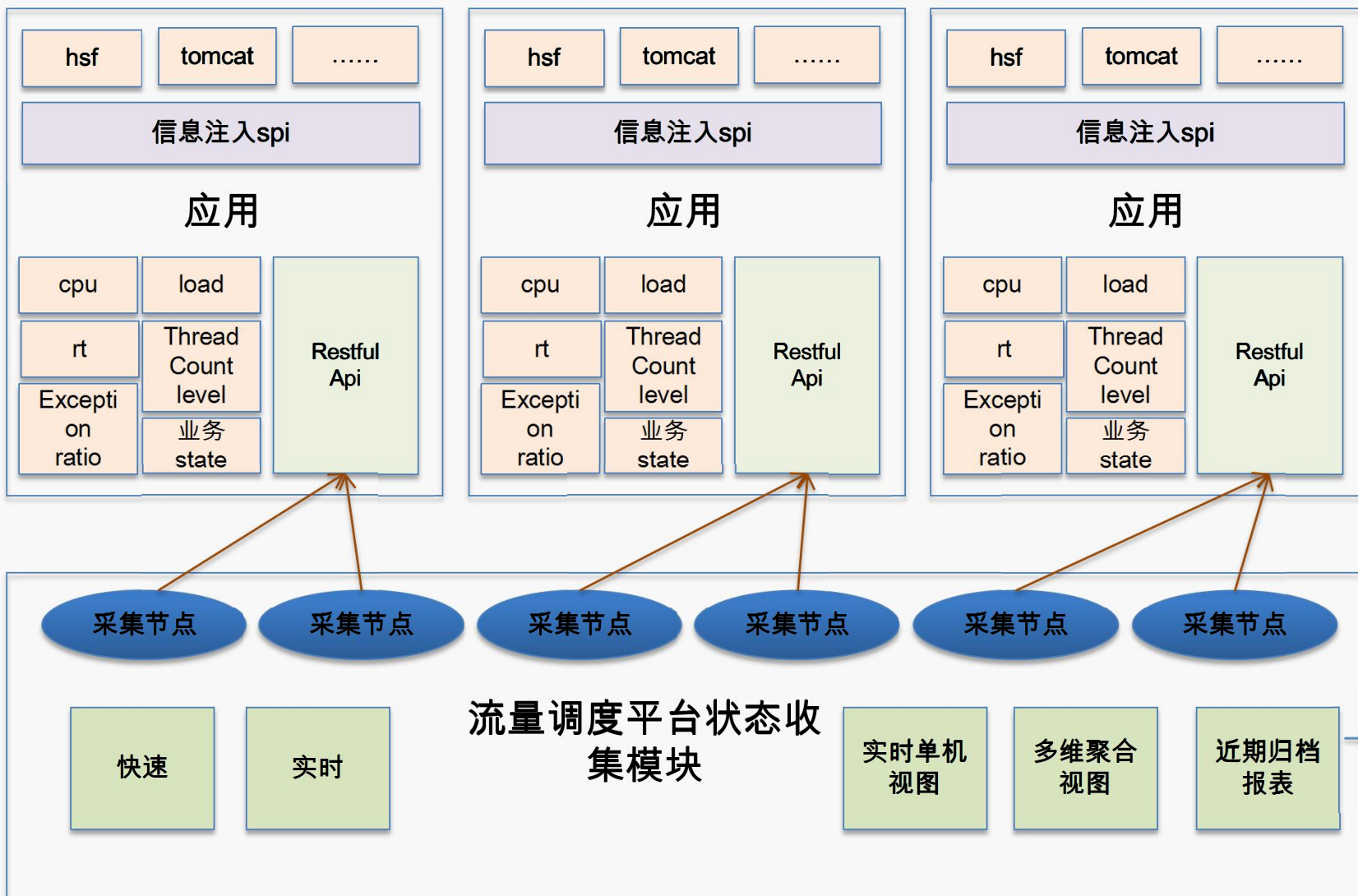


分布式环境同一个应用不同机器由于硬件、网络、超卖、程序自身等众多因素出现服务状态的差异

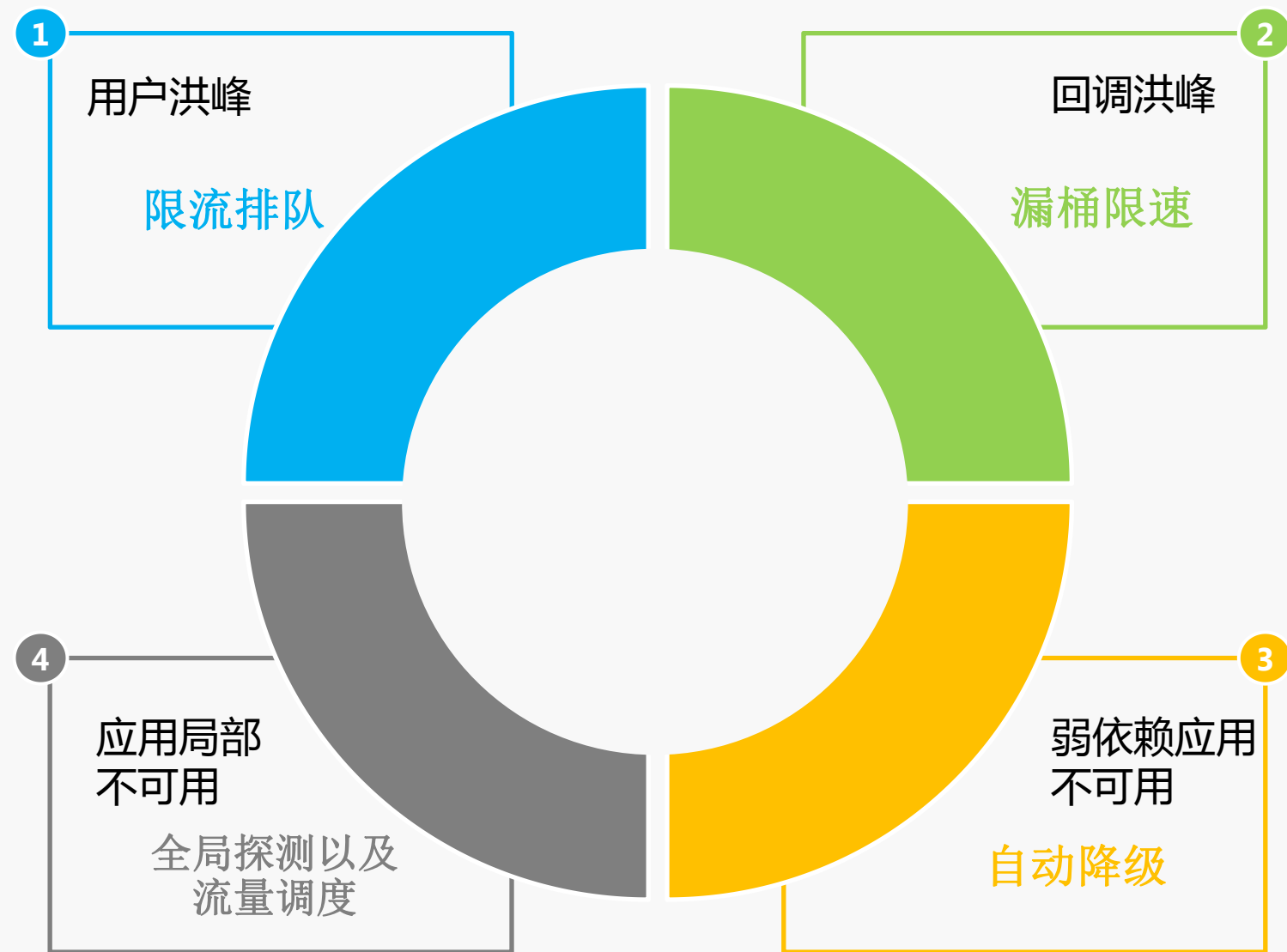
通过流量调度策略，使分布式服务具备自愈能力和自我隔离能力



# 全局探测与流量调度



# 小唐能够顺利下单

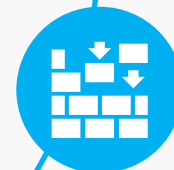




开关预案



阿里线上管控体系

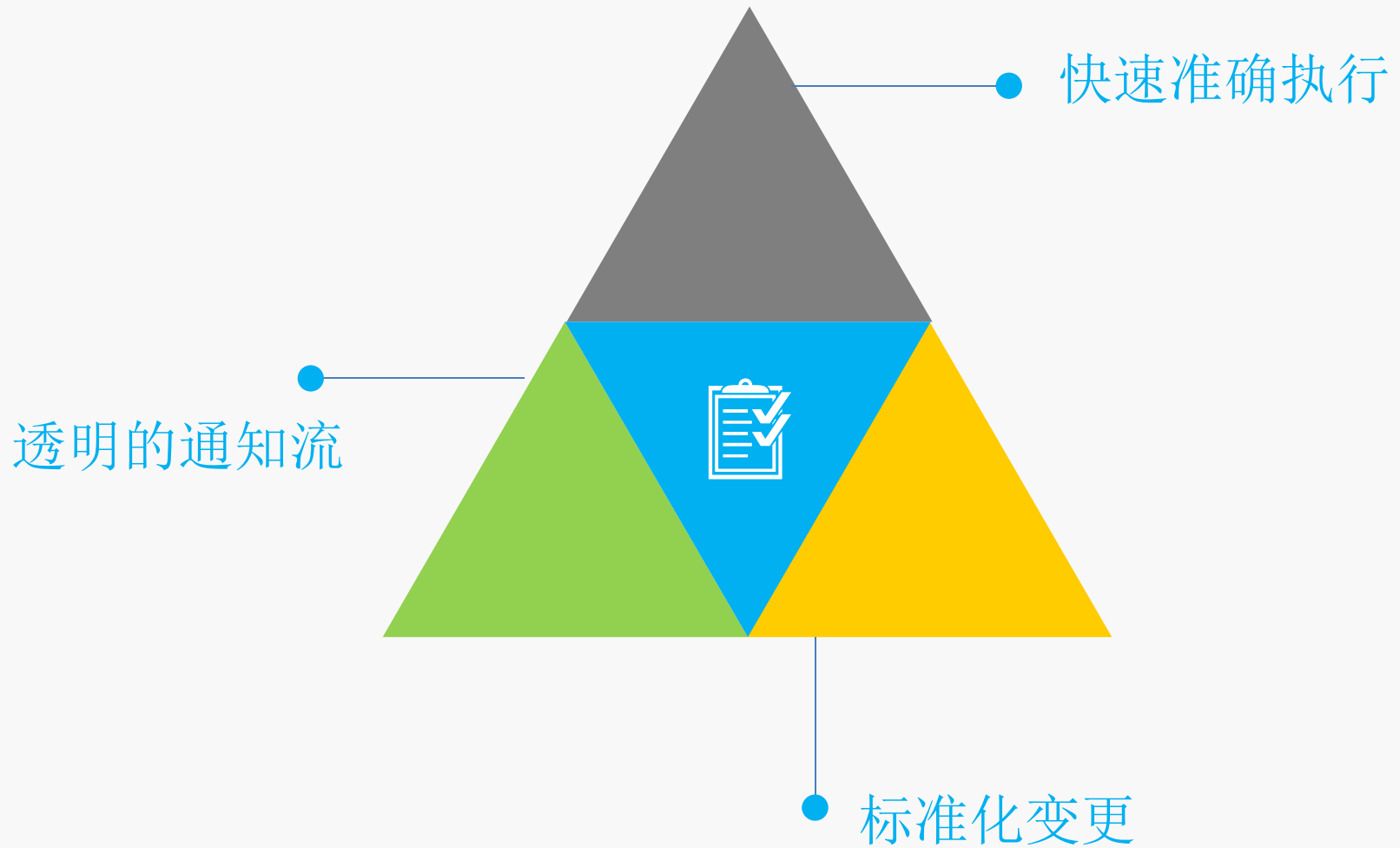


降级&流量调度



限流

# 大促当天



# 标准化变更

## 开关中心控制台

开关管理

开关报表

开关监控

操作权限

分机房执行

开关分组

定时开关

通知机制

使用  
简单

标准

规范

维护  
便捷  
高效

## 稳定性开关客户端

开关定义

使用开关

获取开关信息

更新开关

控制台交互

权限与日志

内存态与持久化

容灾机制

安全机制

## 定义开关

```
@AppSwitch(des = "String 类型开关", level = Level.p2)  
public static String stringSwitch = "string";
```

```
@AppSwitch(des = "Integer 类型开关", level = Level.p1)  
public static Integer integerSwitch = 2;
```

## 变更开关

级别	名称	描述	操作
p1	isOpenProfileDump	是否必须输出dump性能分析	  

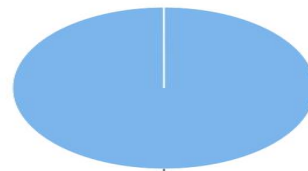
  

IP	HostName	Armory State	当前值	操作
100.69.171.20		false		
100.69.199.6		false		

## 校验开关

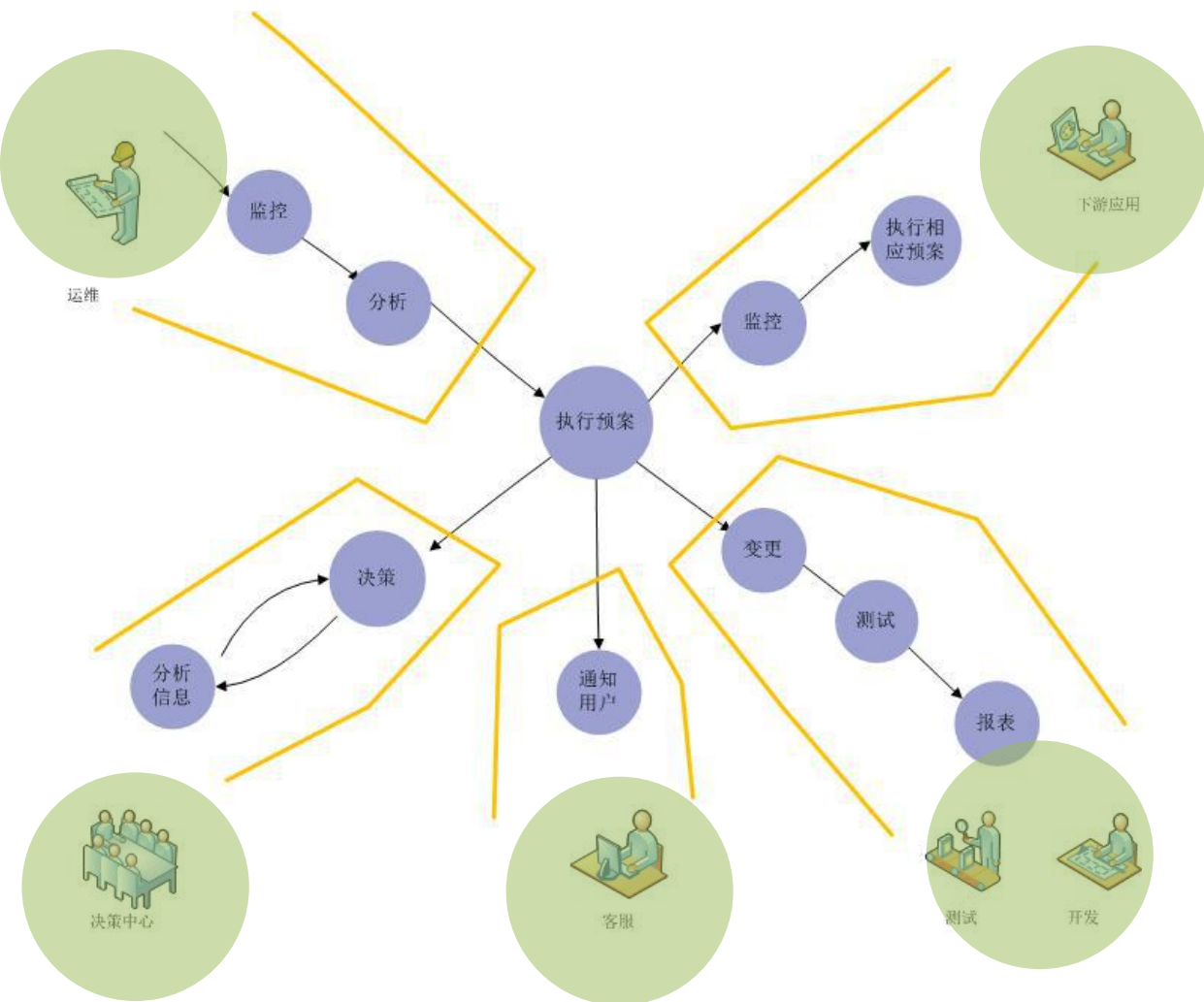
开关值分布统计 (isOpenProfileDump)

开关值集群分布统计  
数据采集时间: 2016-03-30 19:24:06





# 透明的通知流



隔离流程



权限控制



透明通知



角色视图



多维报表

# 阿里管控体系 的最佳实践



简介



阿里线上管控体系，应对挑战



总结，回说一下，关于我们



高可用架构团队

阿里双十一核心保障团队

想知道更多？

<http://jm.taobao.org/>

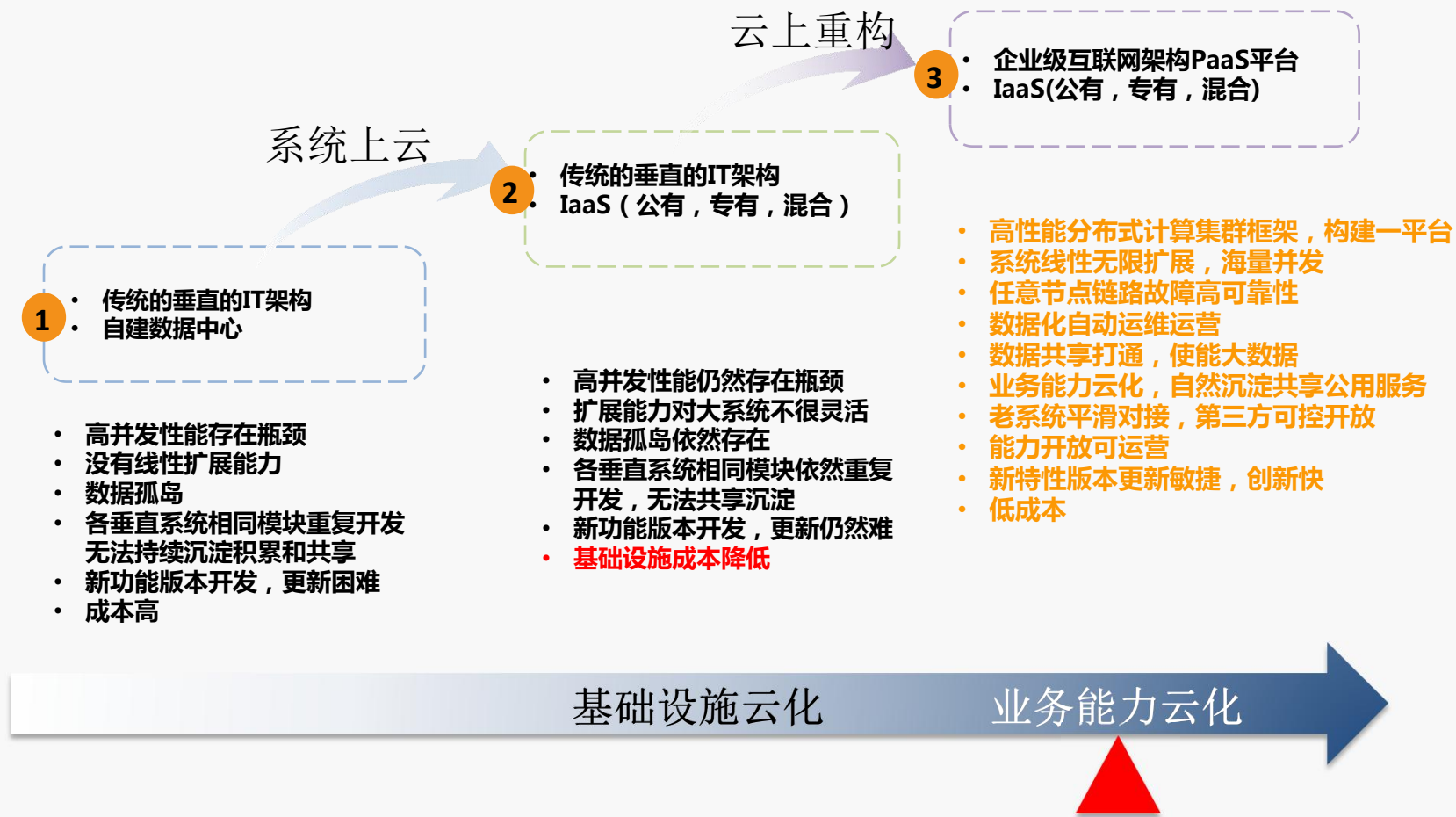
EDAS

弹性容量

一键建站

异地多活

# 企业级信息系统演进的历程



# Q&A

- 旺旺-子矜
- 钉钉 - 子矜

