



GOPS 2016  
Shanghai



# 全球运维大会

2016

重新定义运维

上海站

会议时间： 9月23日-9月24日

会议地点： 上海·雅悦新天地大酒店

主办单位： 开放运维联盟



高效运维社区



指导单位： 数据中心联盟



服务器资源利用率低是常态么？  
怎么平衡稳定性和资源使用率？  
如何安全地提升利用率？

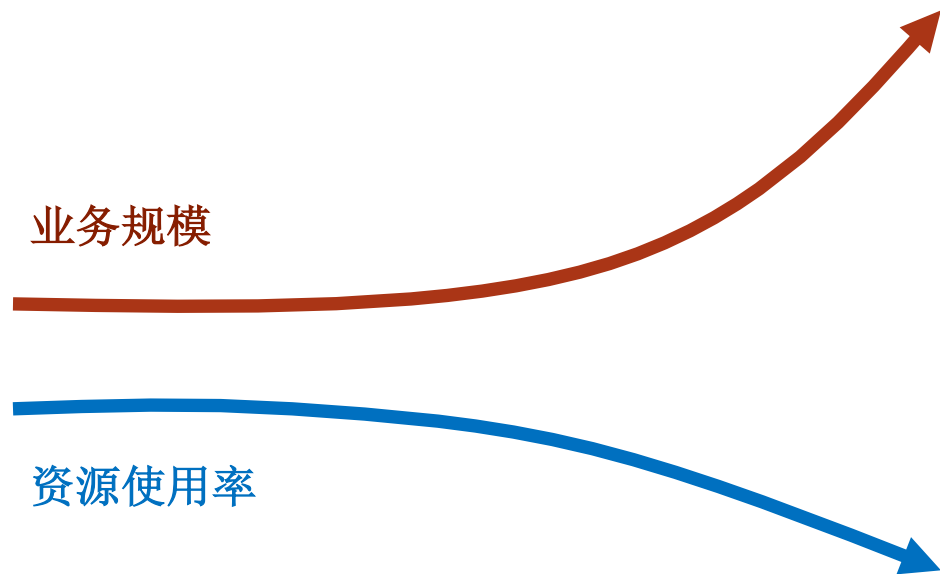
## 基于应用的自动化容量管理与评估

利用率达到多少才算合理？  
如何实现自动化弹性扩容？  
开发人员能得到什么好处？  
有必要限制资源使用么？  
需具备哪些前提？

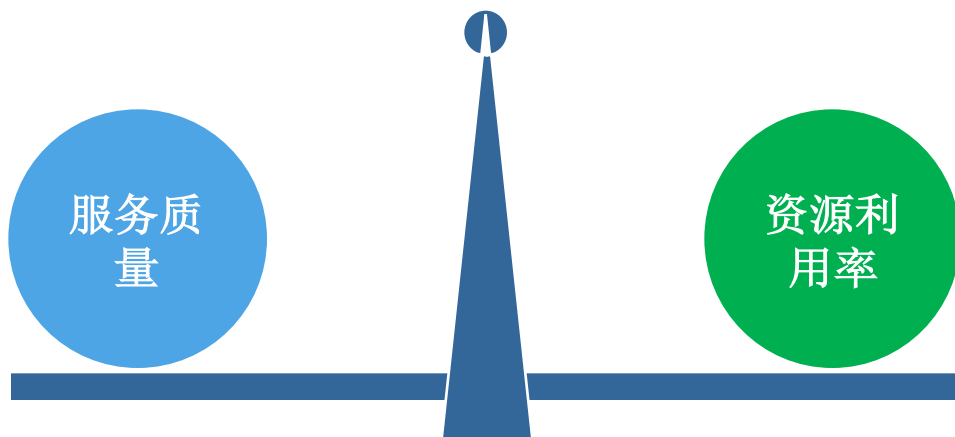
陈剑明·携程网站运营中心

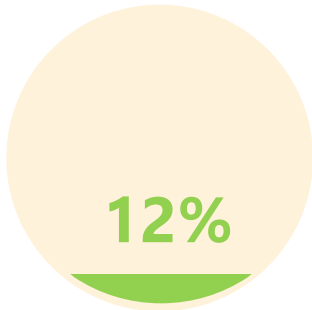


# 网站发展过程

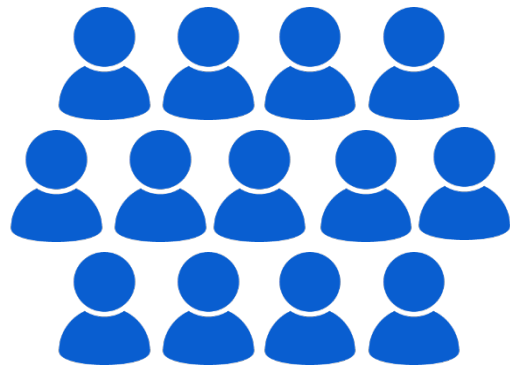


- 追求响应速度
- 夸大需求
- 资源到位时间长





↑ 1% ≈

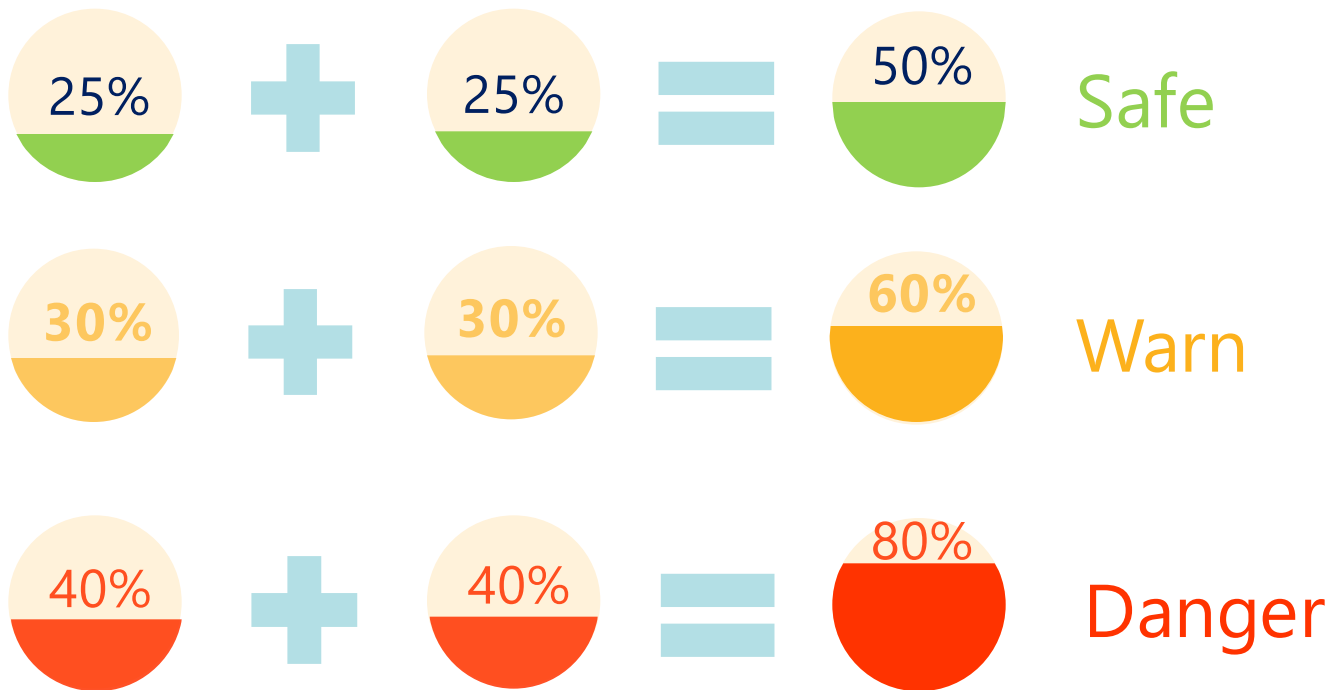


1人 = 1000人

# 度在哪里？

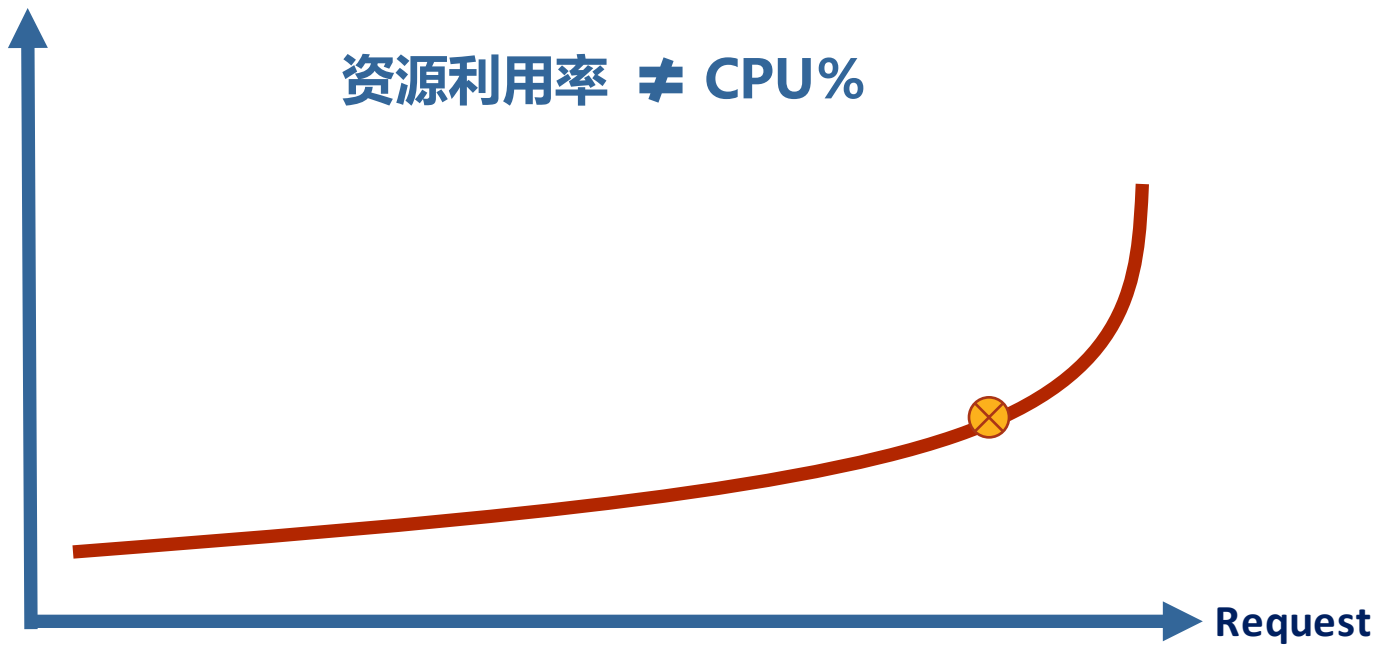
满足上层服务的**稳定性和可靠性**要求





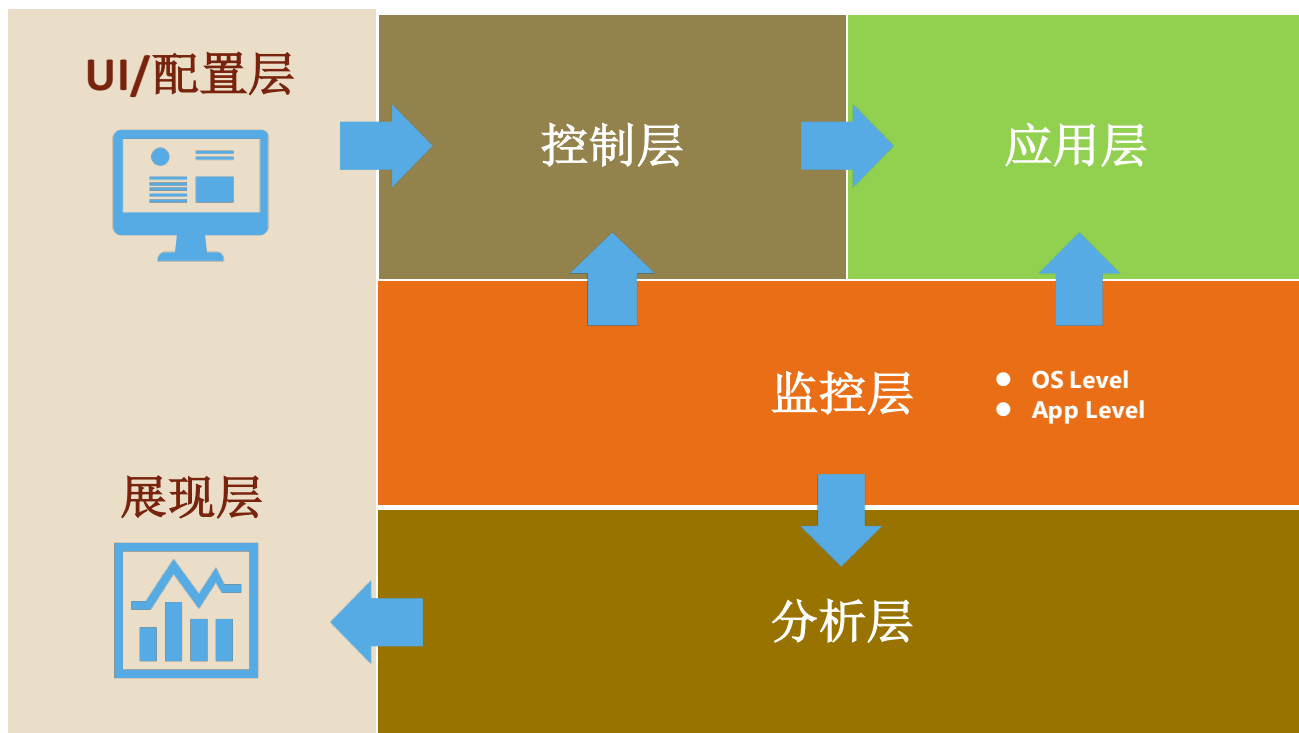
Response Time

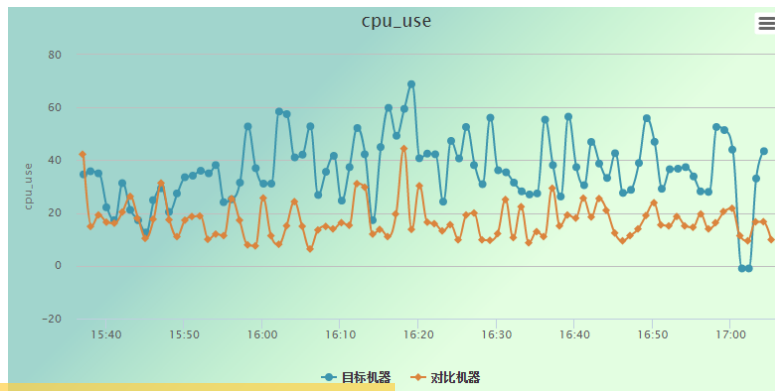
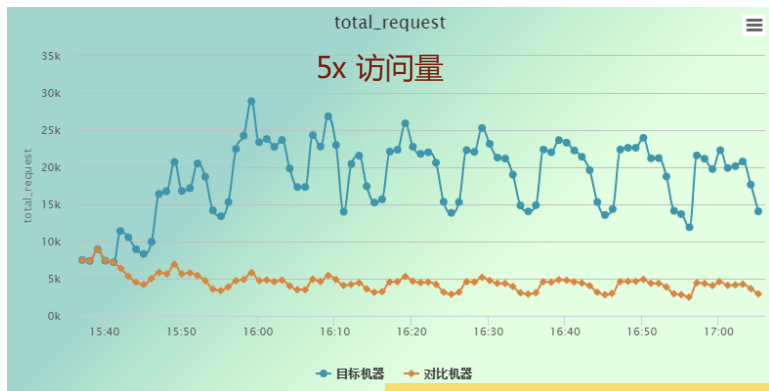
资源利用率  $\neq$  CPU%



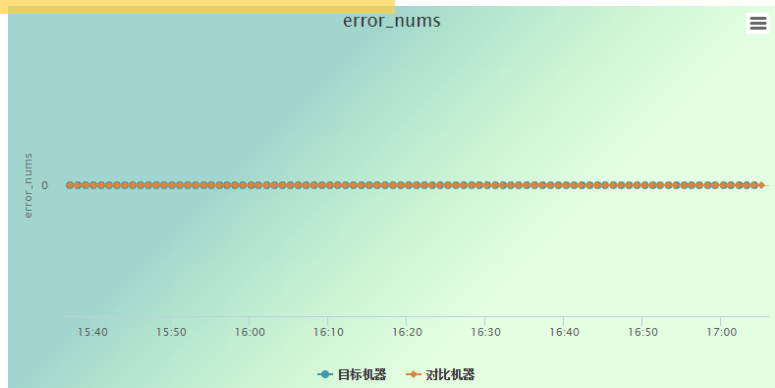
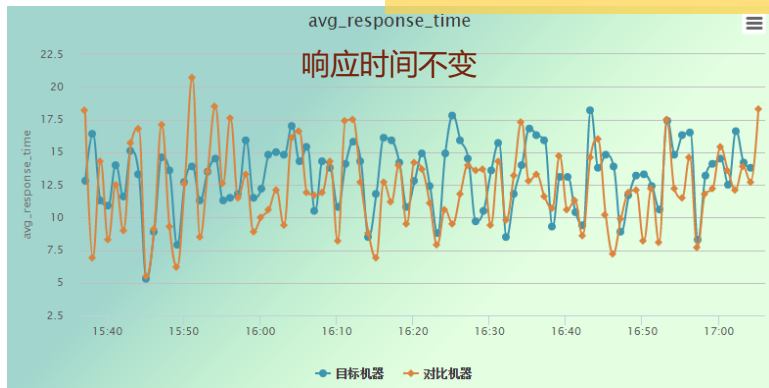


# 逻辑架构



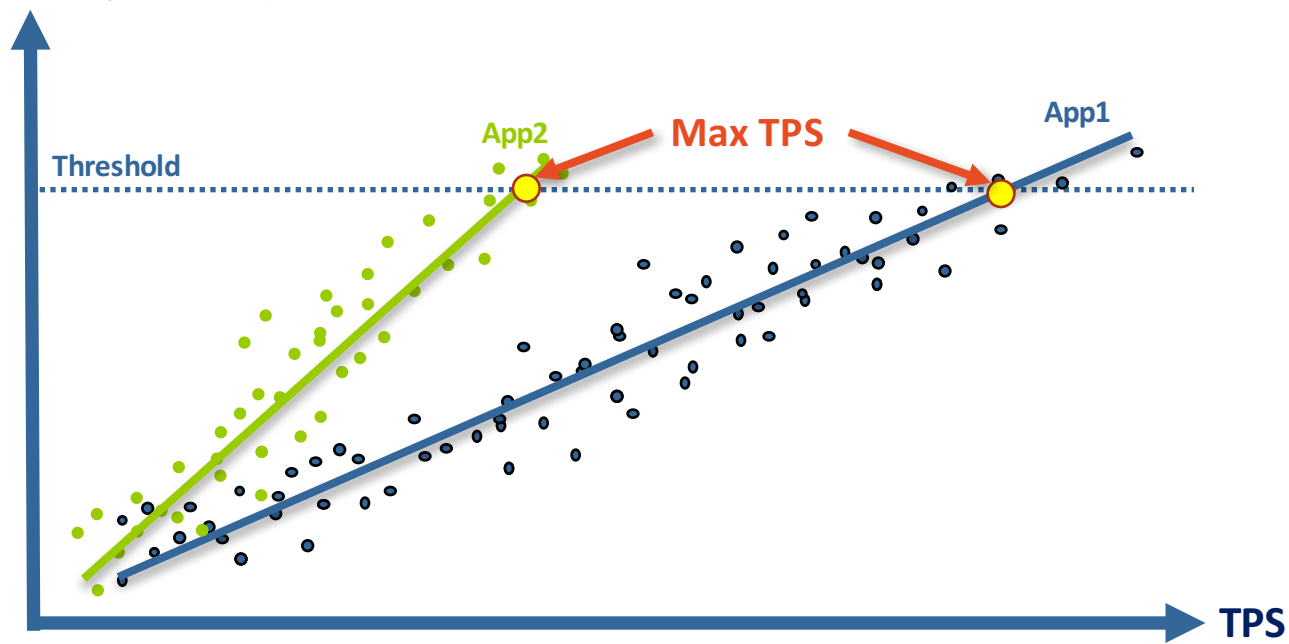


## 5倍流量下应用表现



# 算法机制

Performance Metrics  
(CPU/GC/Thd/Err/ResponseTime...)



# 对开发人员有什么帮助？

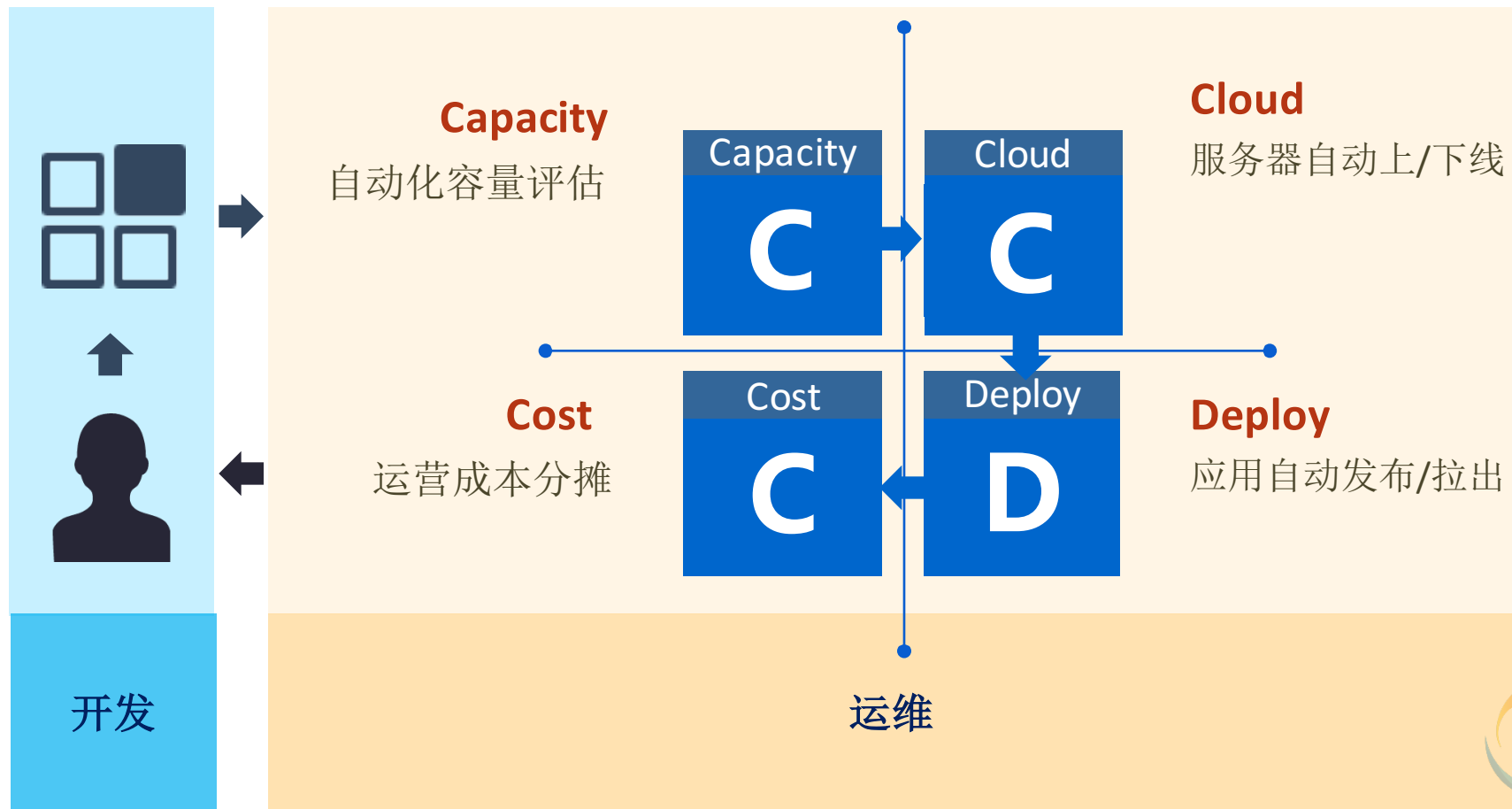
Q1：我的应用在**生产环境**上最大可以支持多少访问请求？

Q2：**生产环境**上了新版本后性能变好了还是差了？

Q3：业务说明天搞活动会带来**X倍**的流量，我的应用撑得住么？

Q4：我能掌握当前应用的容量使用情况么？







# Thanks

高效运维社区  
开放运维联盟

荣誉出品

