

ArchSummit全球架构师峰会北京站2015

# 新浪微博高可用服务保障体系演进

新浪微博高级技术经理 李庆丰

# Geekbang

极客邦科技

整合全球最优质学习资源, 帮助技术人和企业成长  
Growing Technicians, Growing Companies

**InfoQ**  
LEUE

专注中高端技术人员的技术媒体



**EGO** EXTRA GEEKS' ORGANIZATION  
NETWORKS

高端技术人员  
学习型社交网络



**StuQ**  
LEUE

实践驱动的  
IT职业学习和服务平台



**GiT** GEEKBANG  
INTERNATIONAL  
TRAINING  
极客邦培训

一线专家驱动的  
企业培训服务



旧金山 伦敦 北京 圣保罗 东京 纽约 上海  
San Francisco London Beijing Sao Paulo Tokyo New York Shanghai

# QCon

## 全球软件开发大会

2016年4月21-23日 | 北京·国际会议中心

主办方 **Geekbang** & **InfoQ**  
极客邦科技

**7折** 优惠 (截至12月27日)  
现在报名, 节省2040元/张, 团购享受更多优惠

[www.qconbeijing.com](http://www.qconbeijing.com)



扫描获取更多大会信息

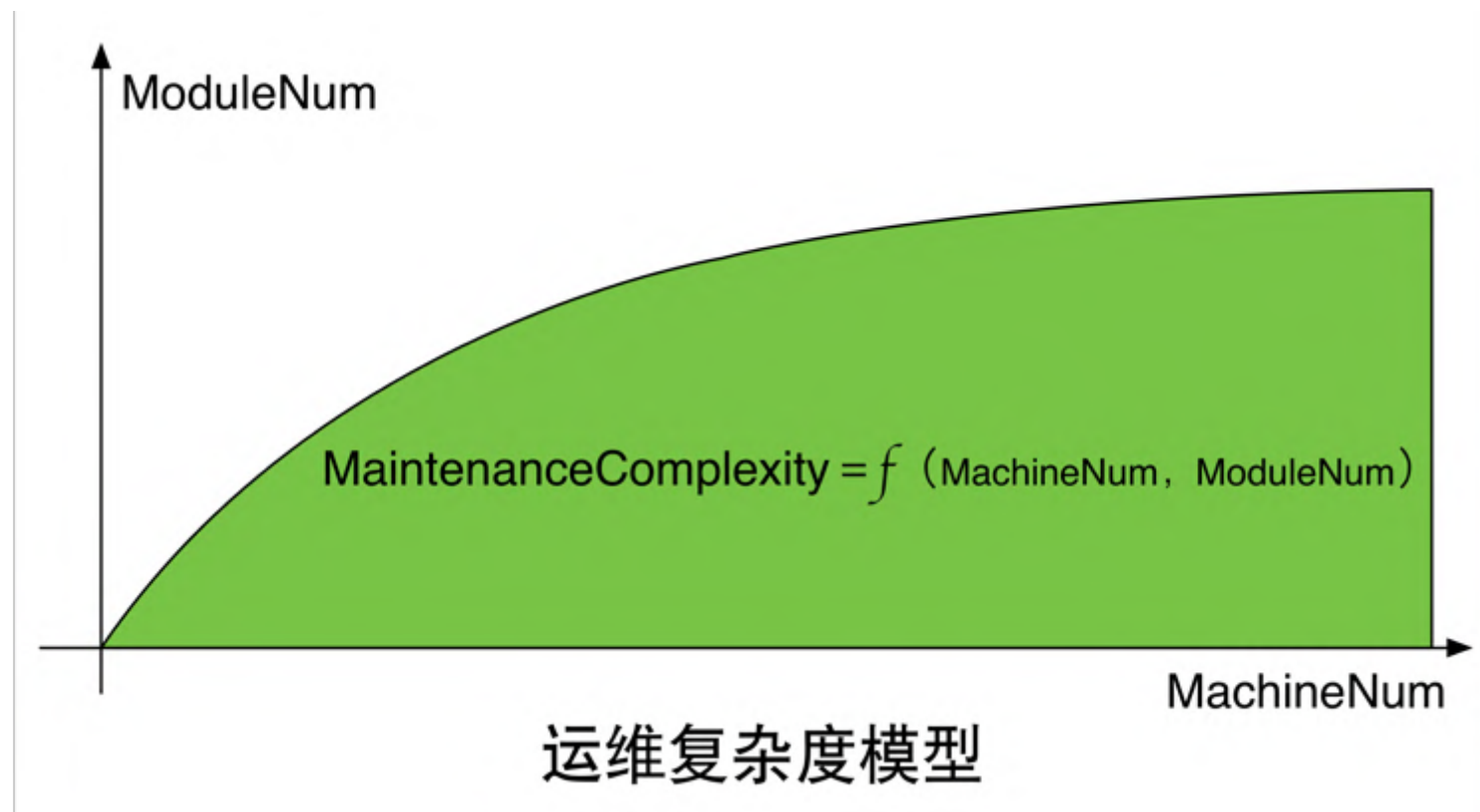
# 内容大纲

- **新浪微博服务保障面临的挑战**
- **伴随着业务发展，高可用服务保障体系的演进过程**
  - 起步阶段 - 纯手工
  - 快速成长 - 工具化、系统化
  - 稳定发展 - 异地容灾与多机房
  - 弹性调度时代 - 容器化及混合云
- **分享与探讨：如何保障高可用及高效运维**

# 服务保障的挑战



# 服务保障的挑战



运维复杂度与机器数量、服务模块数量正相关

# 服务保障的挑战



微博DAU

10000W+

6100W+

3600W+

起步

2015

2013

2012

2010

2009

功能模块数量

百个+

几十个

十几个

几个

机器数量

万台+

几千台

几百台

几十台

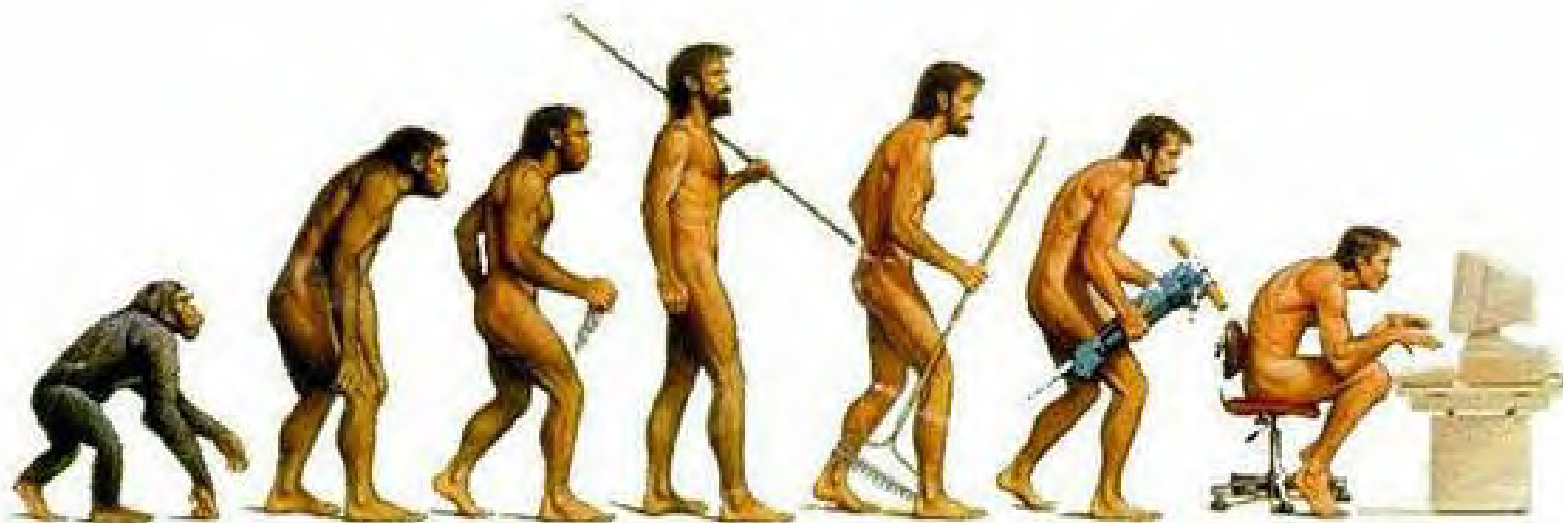
# 服务保障的挑战



不断演进的高可用保障体系



# 演进



# 演进：起步阶段



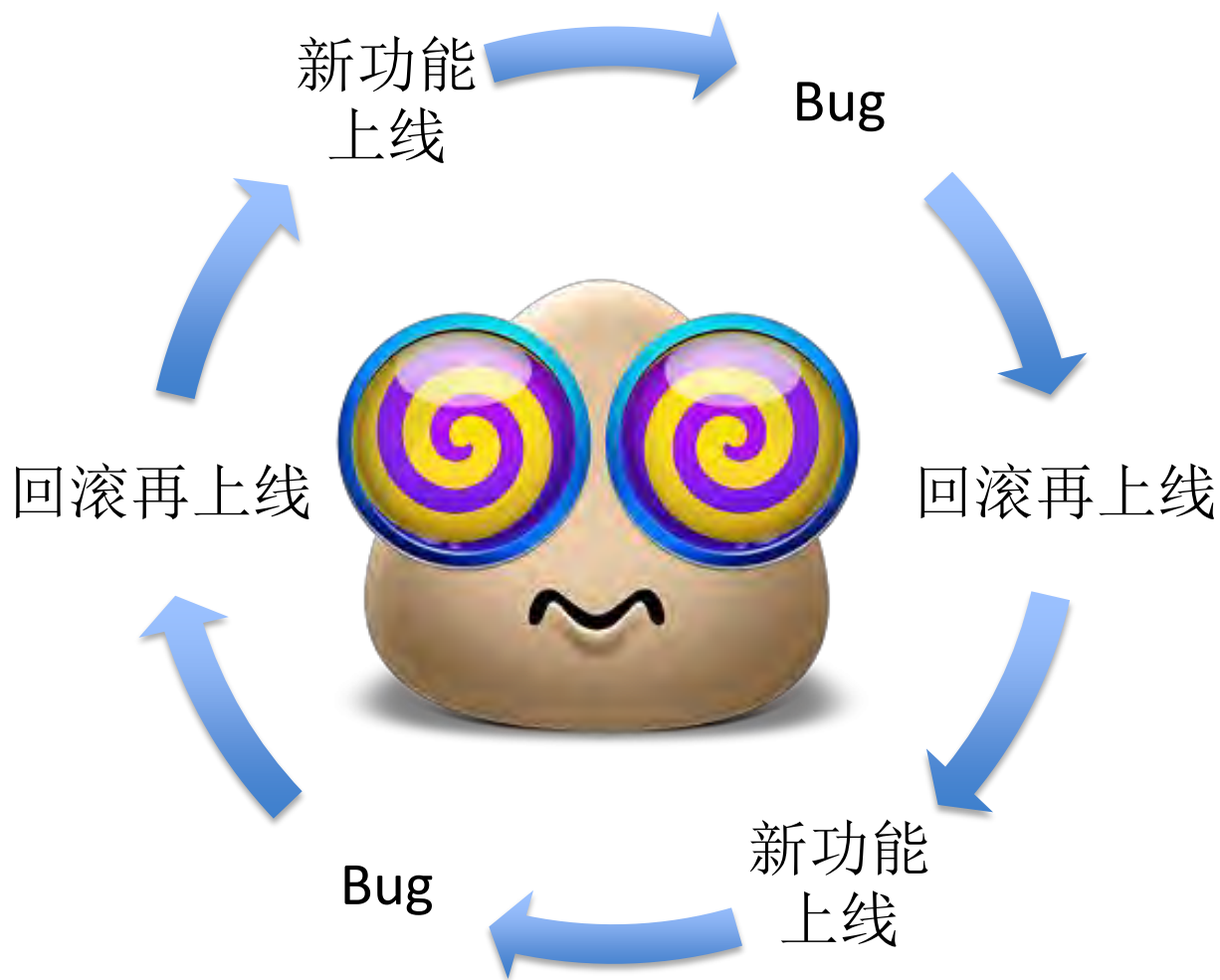
# 演进：起步阶段 - 纯手工



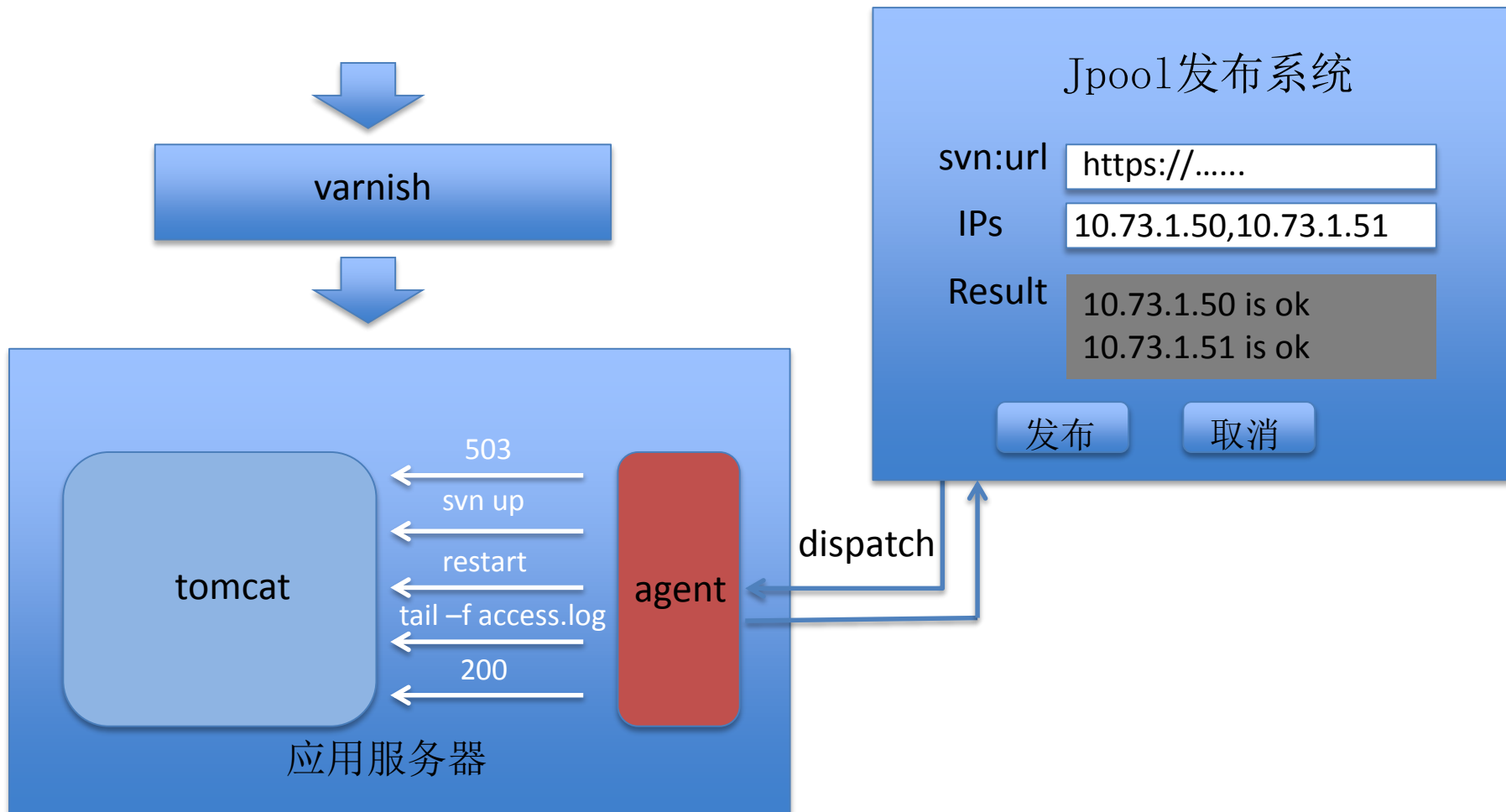
```
fingkideMacBook-Air:~ fingki$ ssh grep awk svn git_
```

CMD命令搞定一切

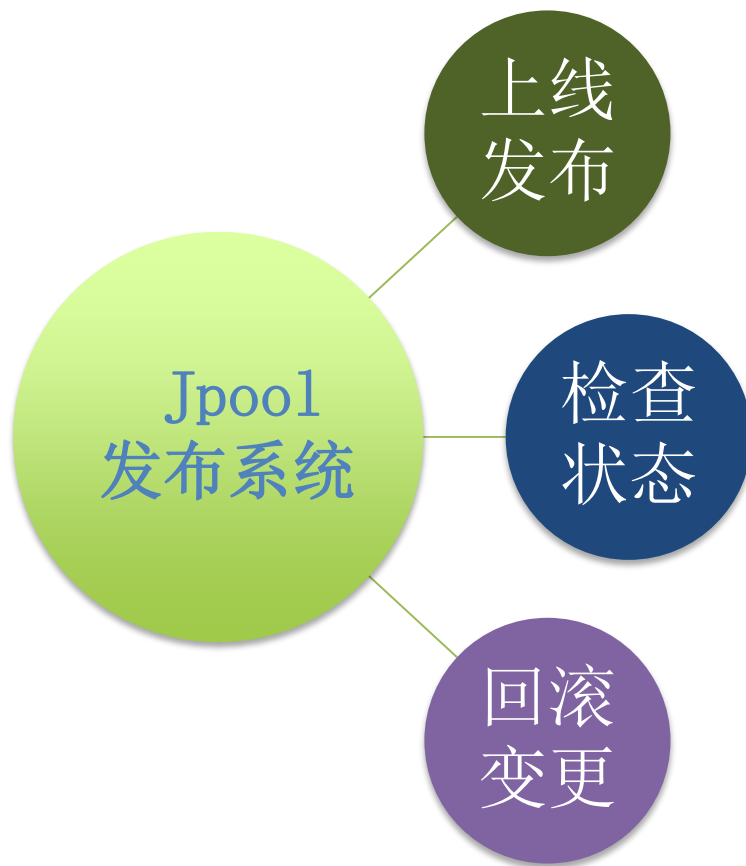
# 演进：起步阶段案例：10次上线的一天



# 演进：起步阶段案例 - 第一个工具系统



# 演进：起步阶段 - 第一个工具系统



# 演进：快速成长



# 演进：快速成长 - 工具化



微博运维工具系统1.0

用户:

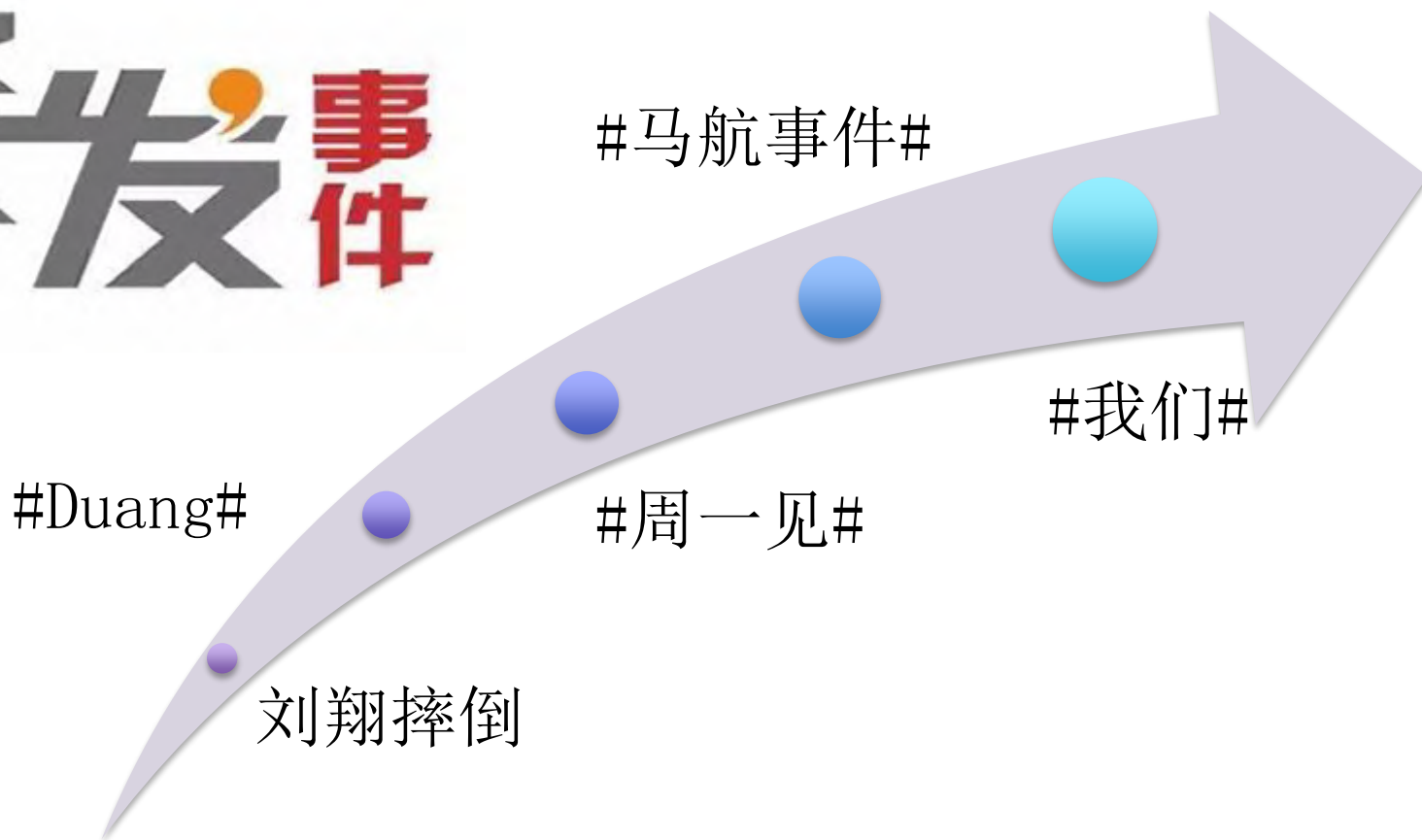
口令:

逐渐使用运维工具系统

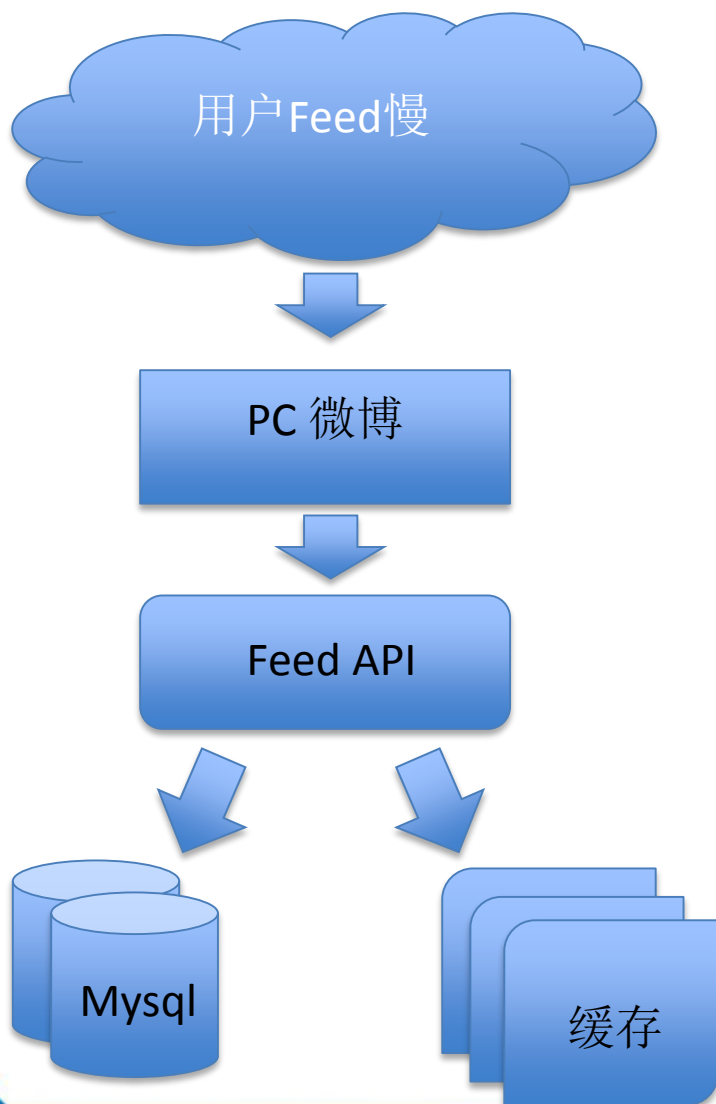


# 演进：快速成长 - 突发流量压力

突发事件



# 演进：快速成长案例 - 到底谁慢了



2012年刘翔意外摔倒事件引发微博短暂故障

# 演进：快速增长 - SLA让争执不再有

## What

- Service-Level Agreement
- 依赖双方的保障约定

## Why

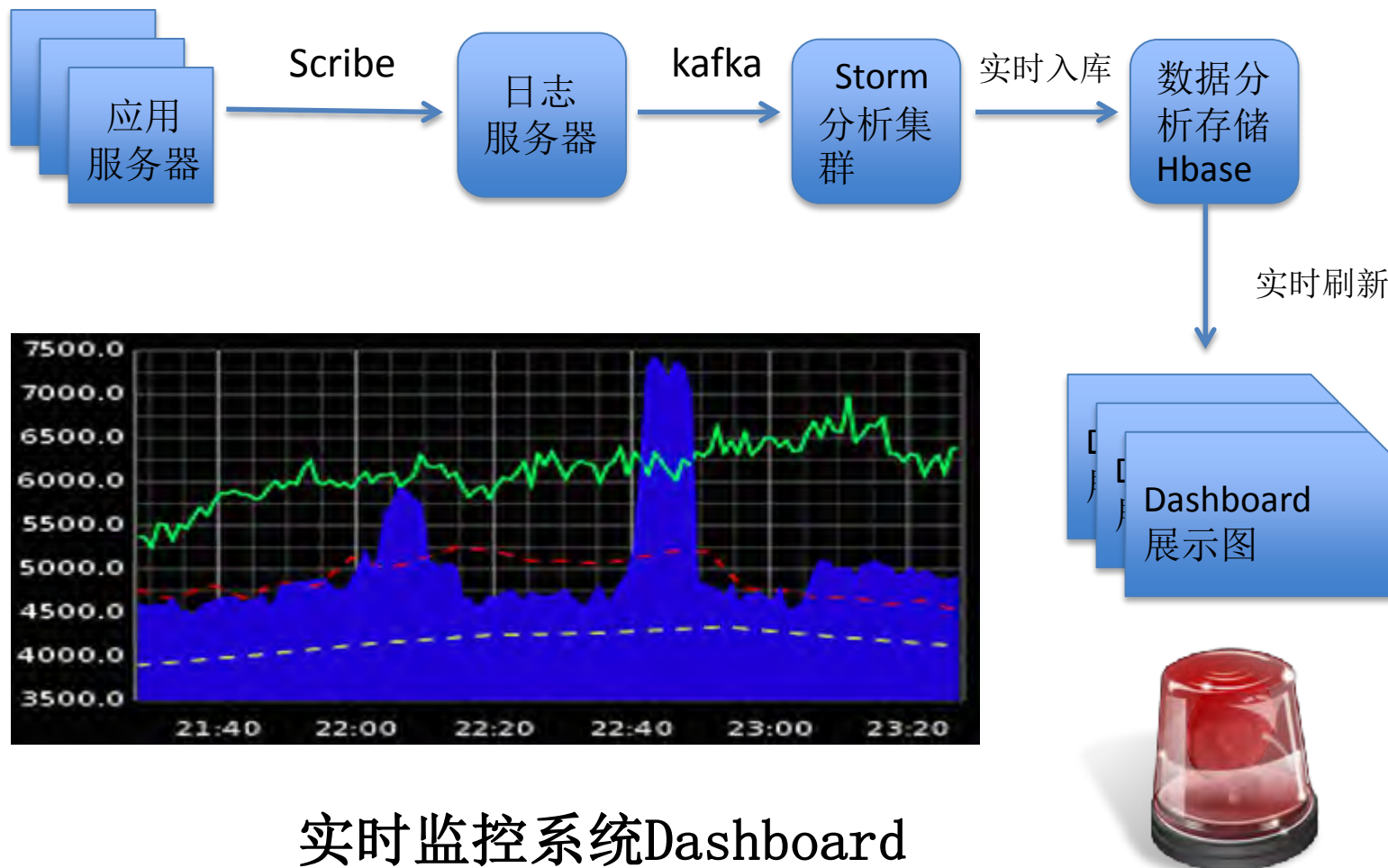
- 明确职责（Api, memcached, redis, mysql, network）
- 量化指标（one request）

## How

- 999规则
- 性能指标（如：99.9%<100ms）
- 容量指标（单机QPS，集群容量）



# 演进：快速增长 - 实时监控快速发现问题



# 演进：快速成长 - 日益完善的Jpool系统

## Jpool管理系统

任务管理

服务发布

Nginx管理

监控管理

降级系统

配置管理

Dispatch

Puppet master

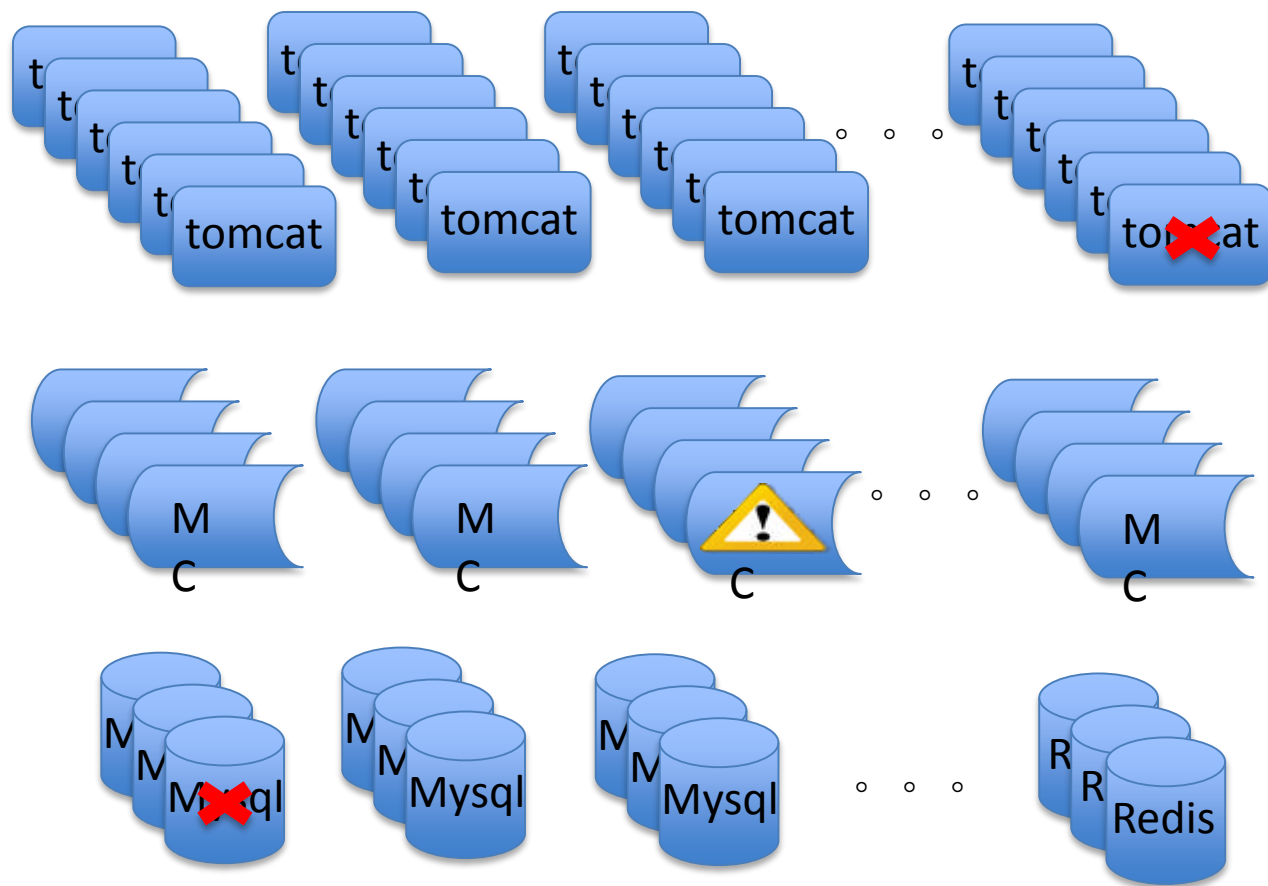
设备&集群管理

公司设备管理系统

The screenshot displays the Jpool management interface. At the top, there are radio buttons for 'OP' (check, off, on, stats). Below that, the 'REGEXP' field is set to '/2/status'. The 'SERVICE' section lists various configuration options, with '/2/statuses/friends\_timeline : true' selected. The '相关参数' (Related Parameters) section shows 'PORT' set to 880 and 'TYPE' set to 'api'. A dropdown menu is open over the 'TYPE' field, showing options: 'http', 'pullonoff', 'resource', 'hbaseregions', and 'graph'. At the bottom, the '执行结果' (Execution Results) section shows a successful execution of the service configuration on host 127.0.0.1.

```
OP:  check  off  on  stats
REGEXP: /2/status
SERVICE:
   /2/statuses/fo_me/ids : true
   /2/statuses/update : true
   /2/statuses/timeline_batch : true
   /2/statuses/tags/show_batch : true
   /2/statuses/latest_timeline : true
   /2/statuses/friends_timeline : true
   /2/statuses/last_time : true
   /2/statuses/close_friends_timeline_other : true
   /2/statuses/repost_by_me : true
相关参数: PORT:  880  881  19000, TYPE:  api
   pullonoff
   resource
   hbaseregions
   graph
执行结果:
state: success
err:
detail:
-----
ip: 10.75.0.180
state: success
detail:
switcher(/2/statuses/friends_timeline) enabled over 127.0.0.1 880 (type=api)
check 127.0.0.1 880 "/2/statuses/friends_timeline" (type=api) execute successful
```

# 演进：快速成长 - 机器多经常遇到服务器宕机



机器故障影响服务

收到报警

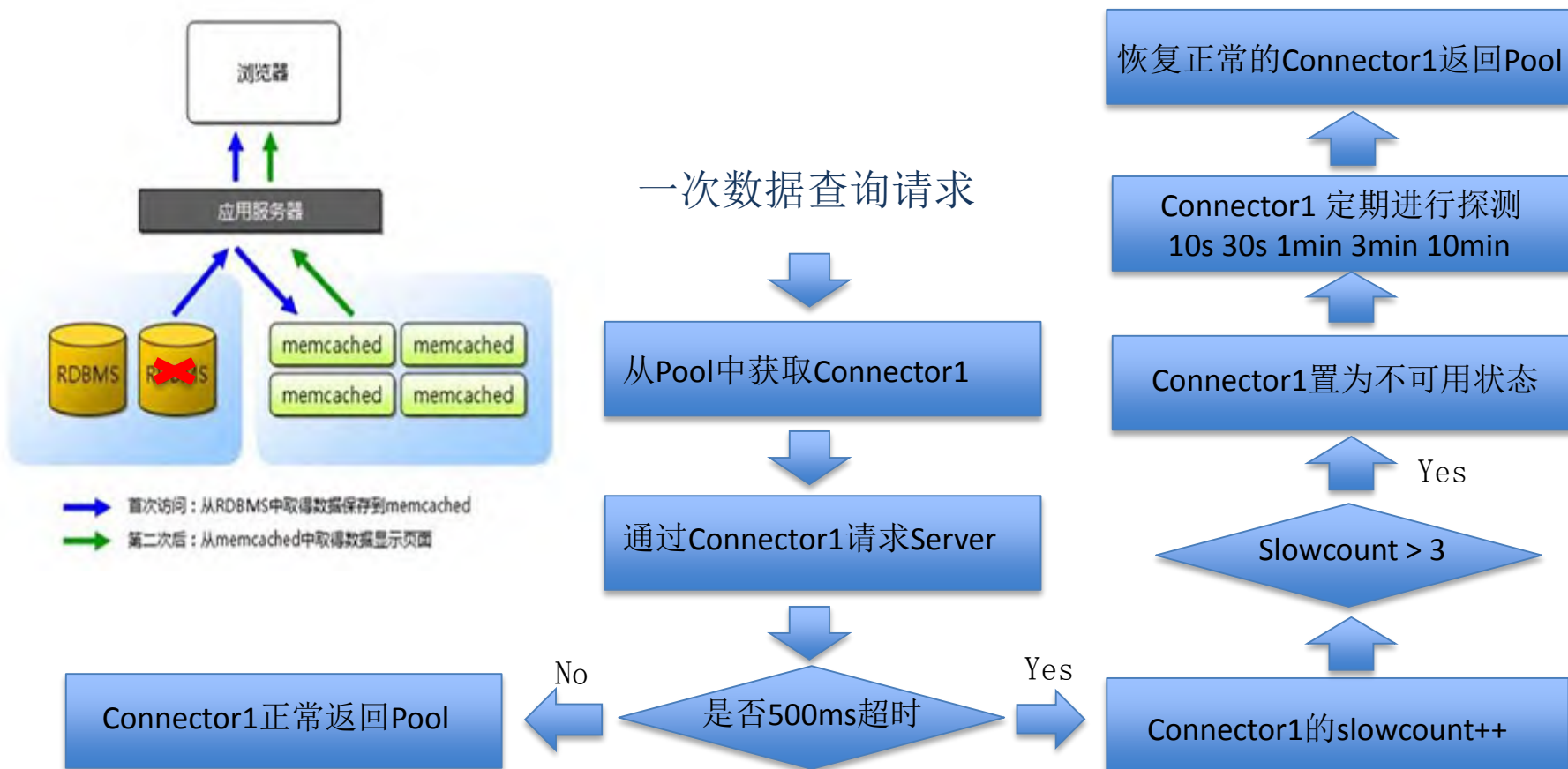
找到故障机器

从服务中摘除

服务恢复

核心资源问题处理不及时可能导致集群故障

# 演进：快速增长案例 - 针对Mysql自动降级



适当的自动降级和容错策略提高运维效率及可用性

# 演进：快速成长 - 工具化、系统化

建立内部SLA体系

建设实时的监控体系

建设服务依赖降级系统

推动架构容错能力





# 演进：稳定发展



# 演进：稳定发展 - 业务特点



相对完备运维保障系统

# 演进：稳定发展 - 互联网最大的威胁

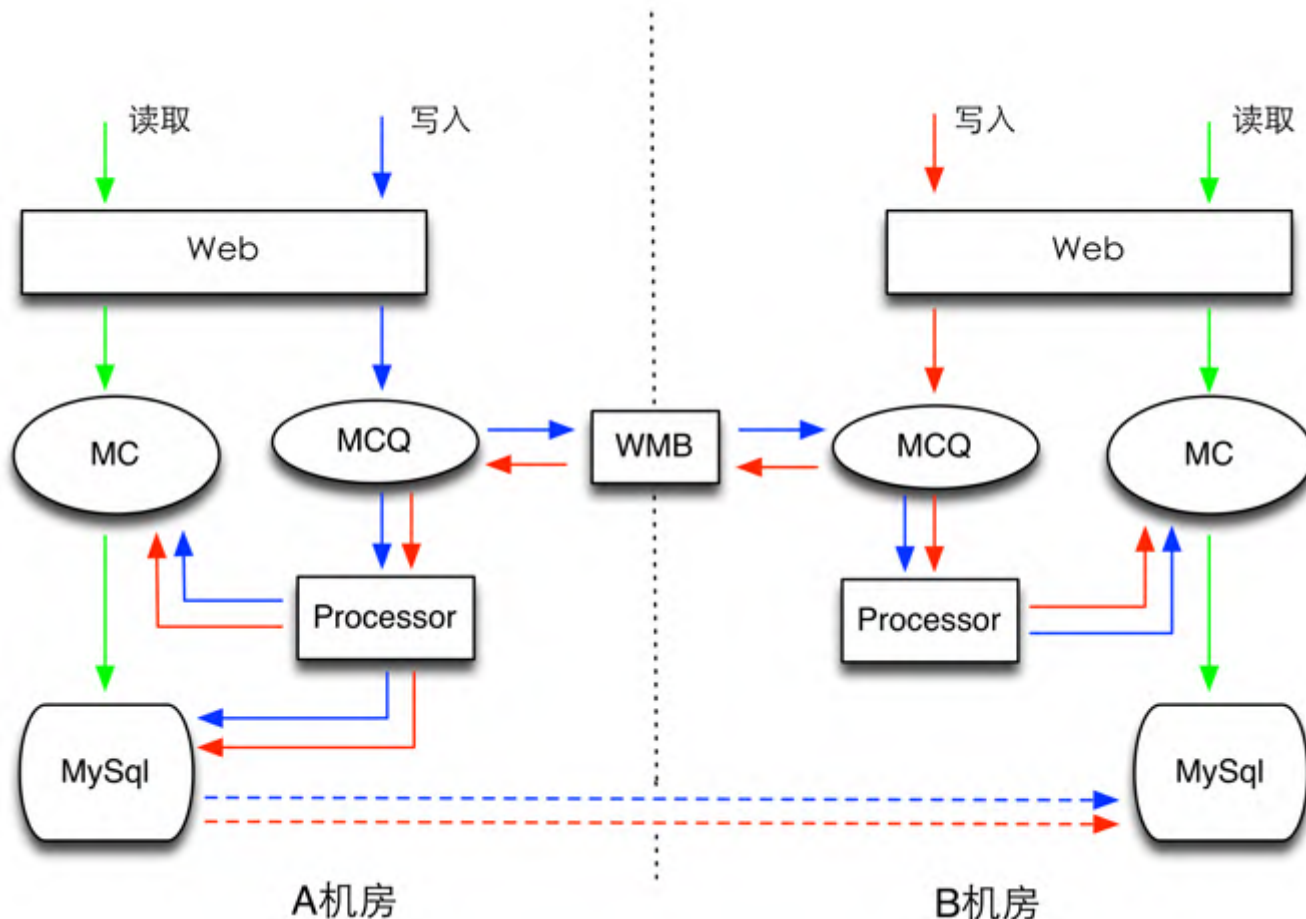


# 演进：稳定发展 - 异地容灾与多机房

多  
机  
房  
部  
署

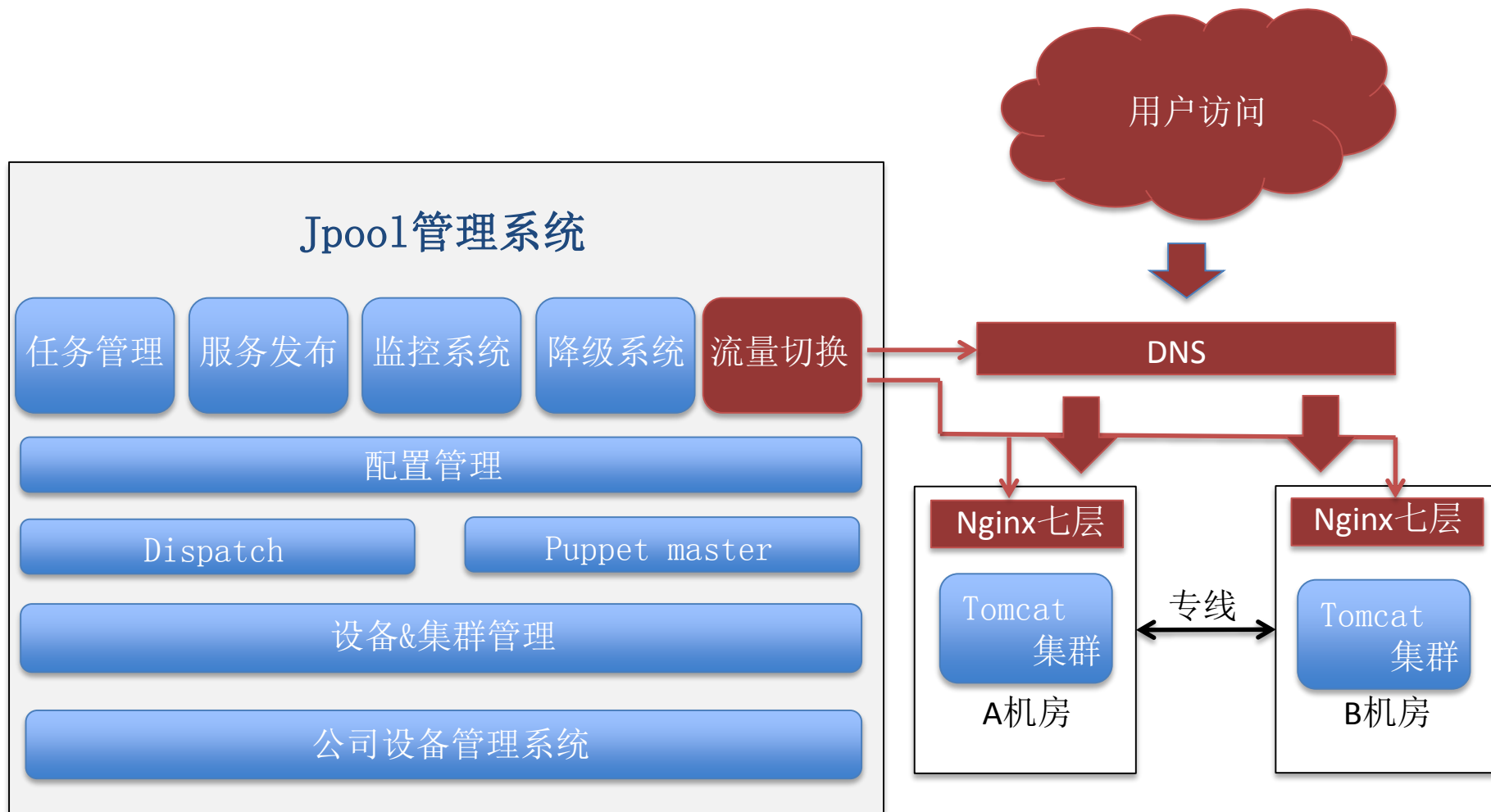


# 演进：稳定发展案例 - 微博多机房架构



微博多机房架构

# 演进：稳定发展案例 - 流量切换能力



# 演进：弹性调度时代

时间：2015

设备：万+台

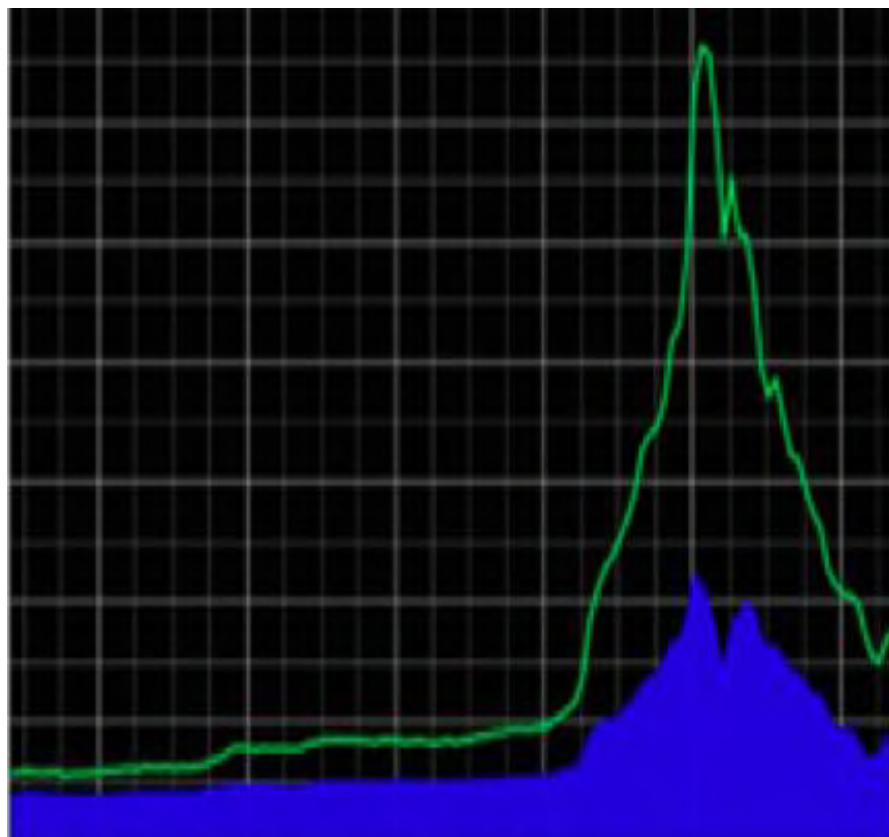
服务：百十个

业务：百亿级 / 天

运维：容器化部署 扩缩容<5min/百台



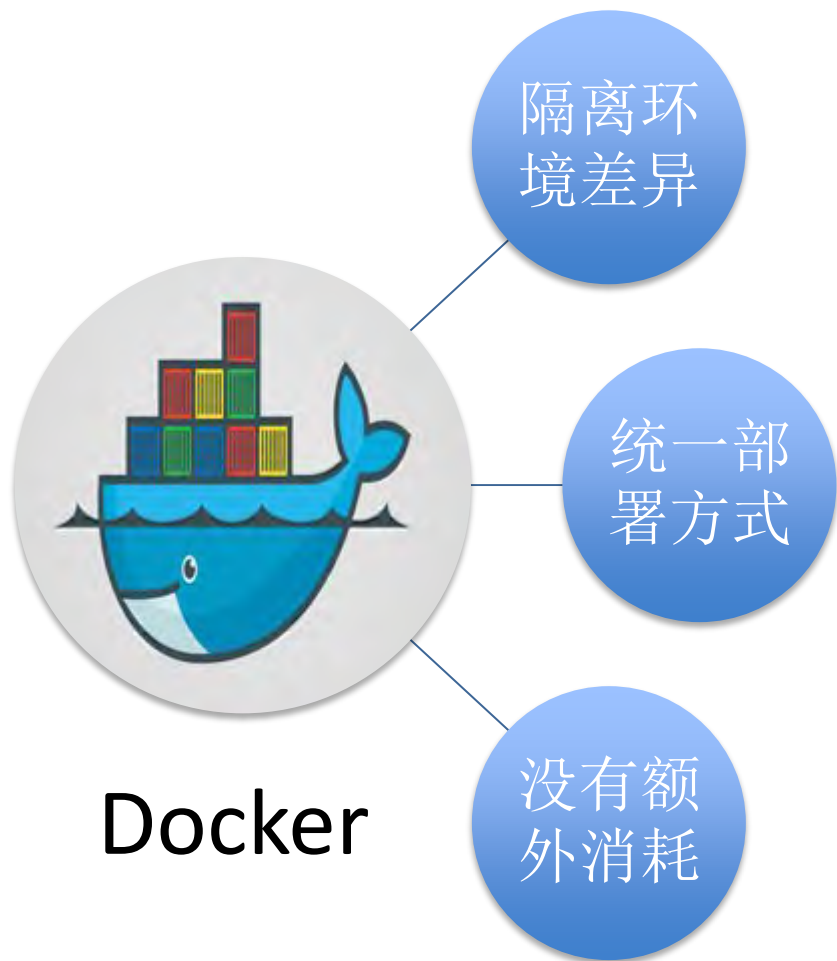
# 演进：弹性调度时代 - 用更低的成本保障服务



春节核心业务峰值



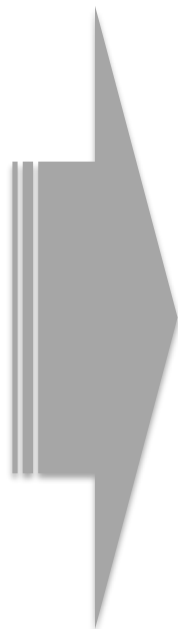
# 演进：弹性调度时代 - 服务容器化



# 演进：弹性调度时代 - 提高服务器上线效率

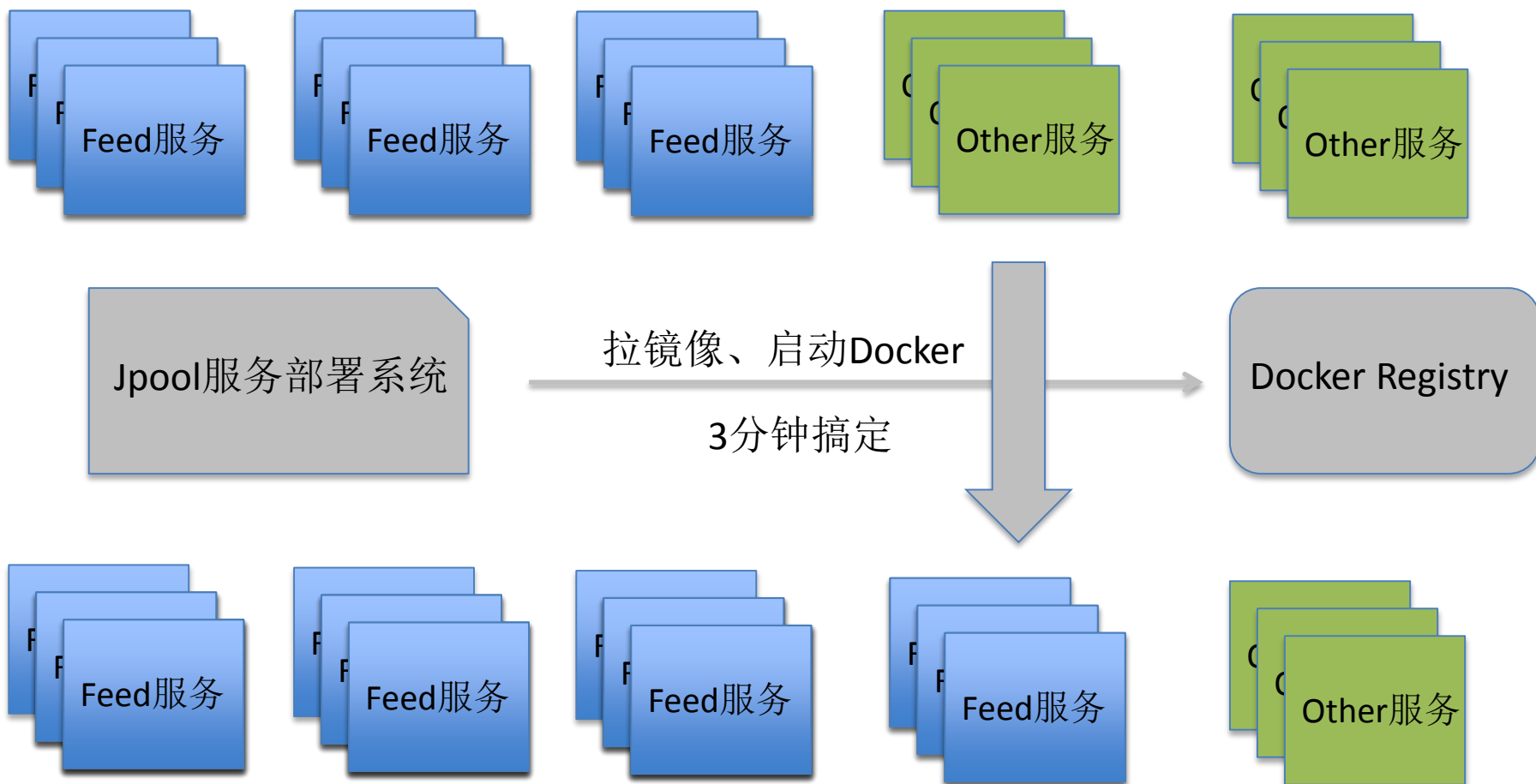


传统服务器上线（1天/百台）



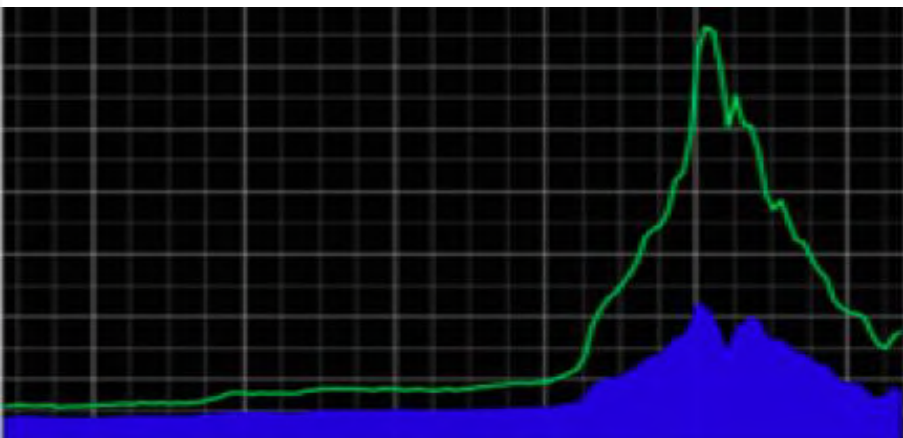
容器化服务器上线（5min/百台）

# 演进：弹性调度时代案例 - 2015春节抗峰

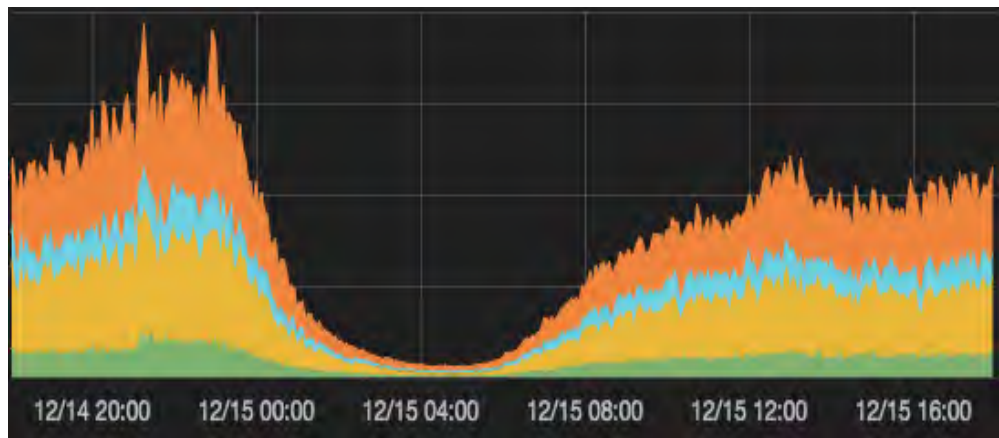


Docker集群助力微博2015春晚峰值挑战

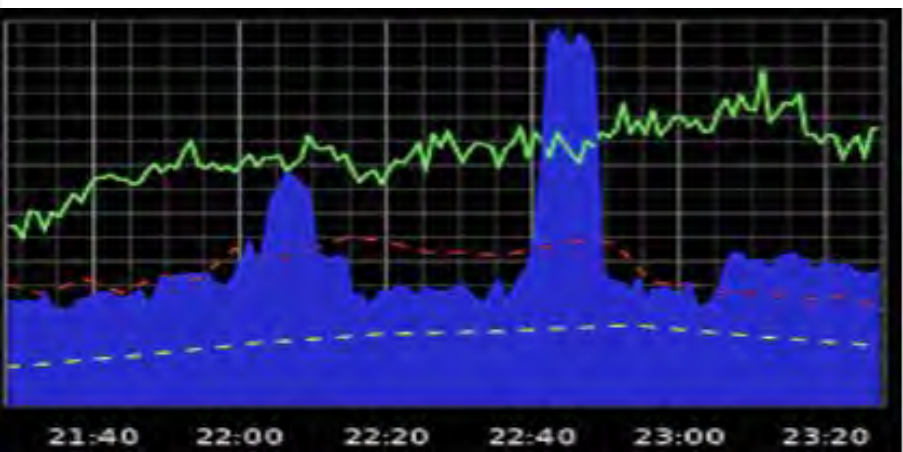
# 演进：弹性调度时代 - 成本再降低



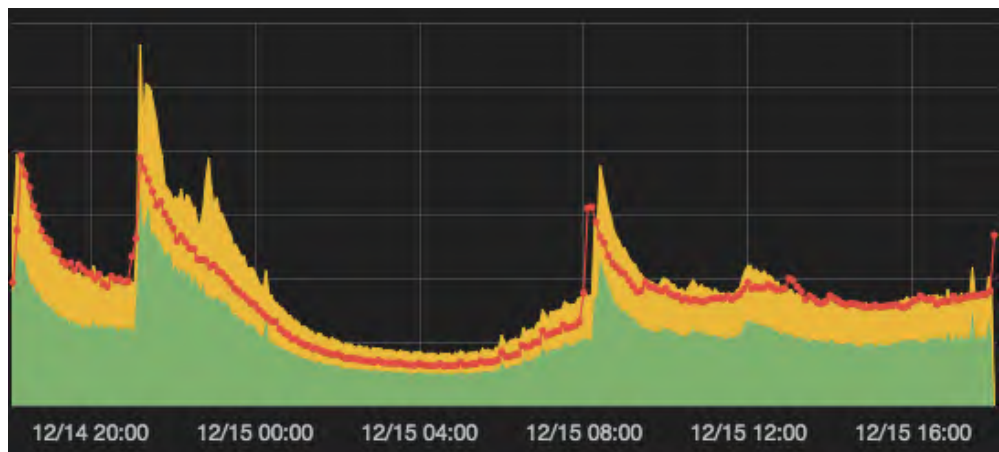
元旦春节峰值



日常业务峰值

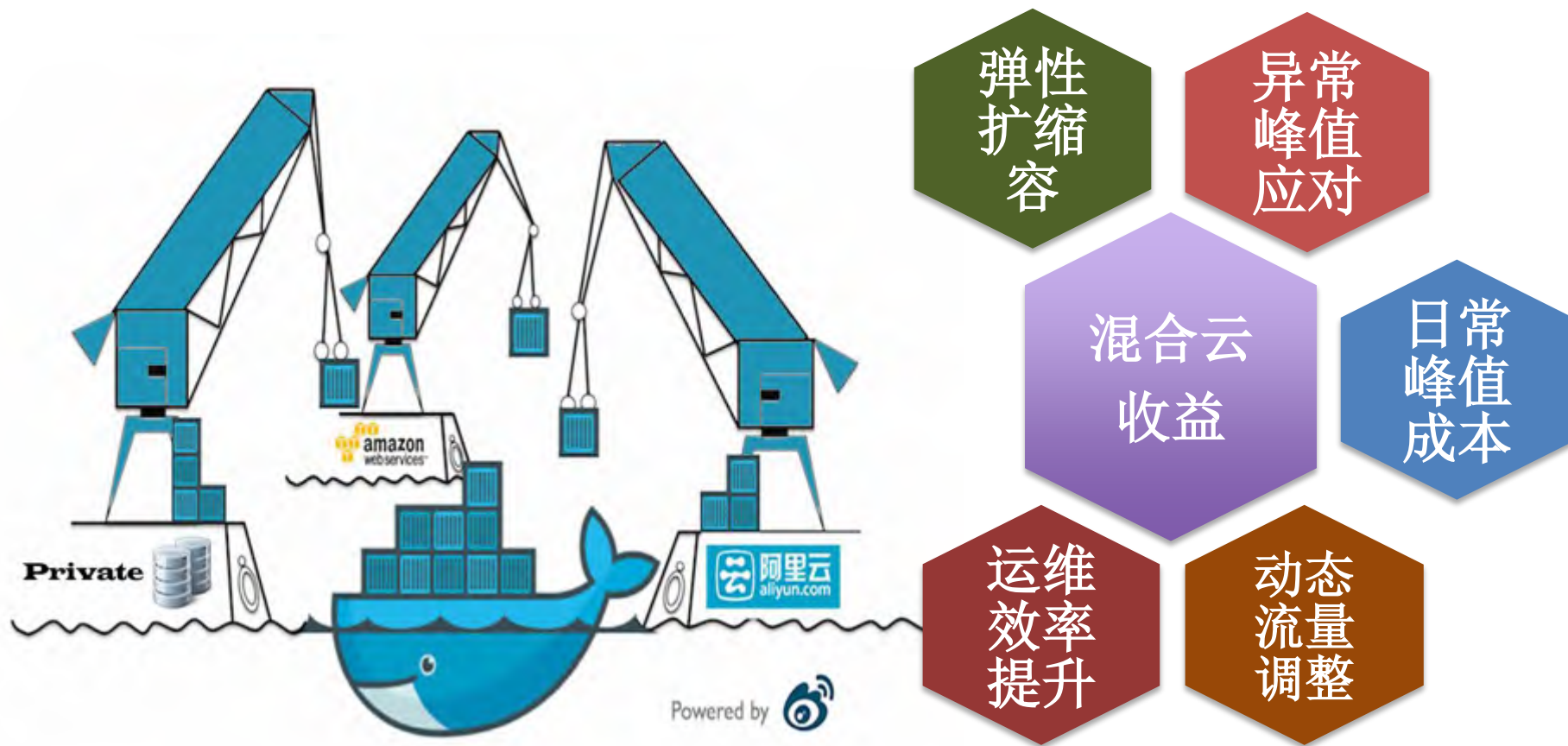


异常事件峰值



运营业务峰值

# 演进：弹性调度时代 - 公有云与混合云



公有云的成熟让混合云方案成为可能

# 演进：弹性调度时代 - 混合云可行性

## 容器技术成熟度

- Docker开源且被广泛接纳
- Swarm, Kubernetes, Mesos

## 业界经验和案例

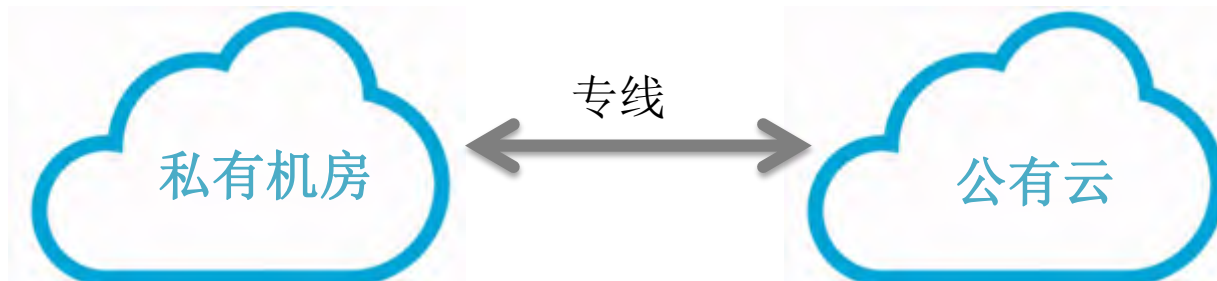
- Google10年前已经在做
- 百度, 京东, 12306
- DaoCloud, 希云, 云雀。。。

## 公有云的发展

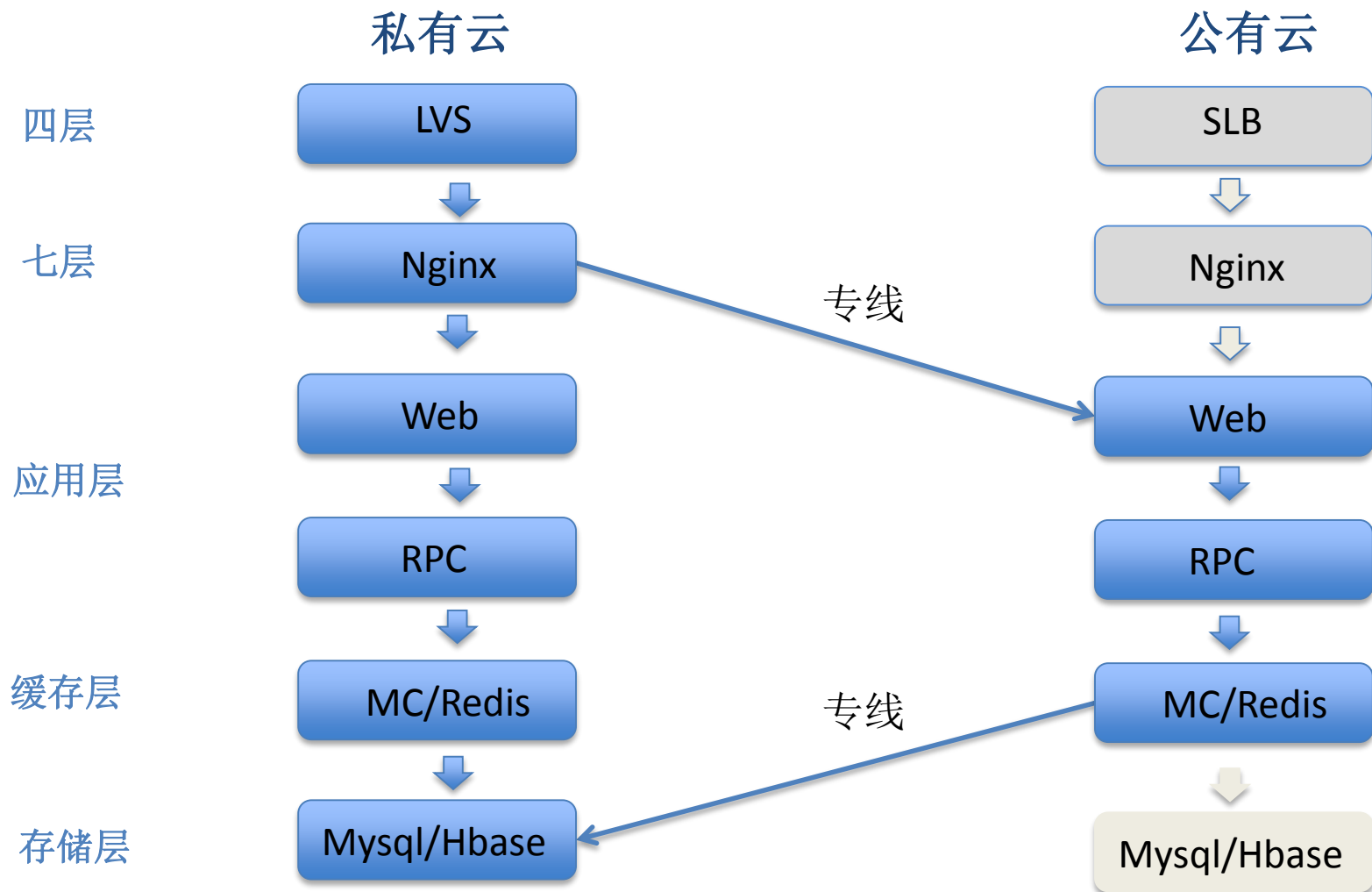
- 公有云设施日趋完善
- AWS, 阿里云



# 演进：弹性调度时代案例 - 微博混合云方案



# 演进：弹性调度时代案例 - 微博混合云部署

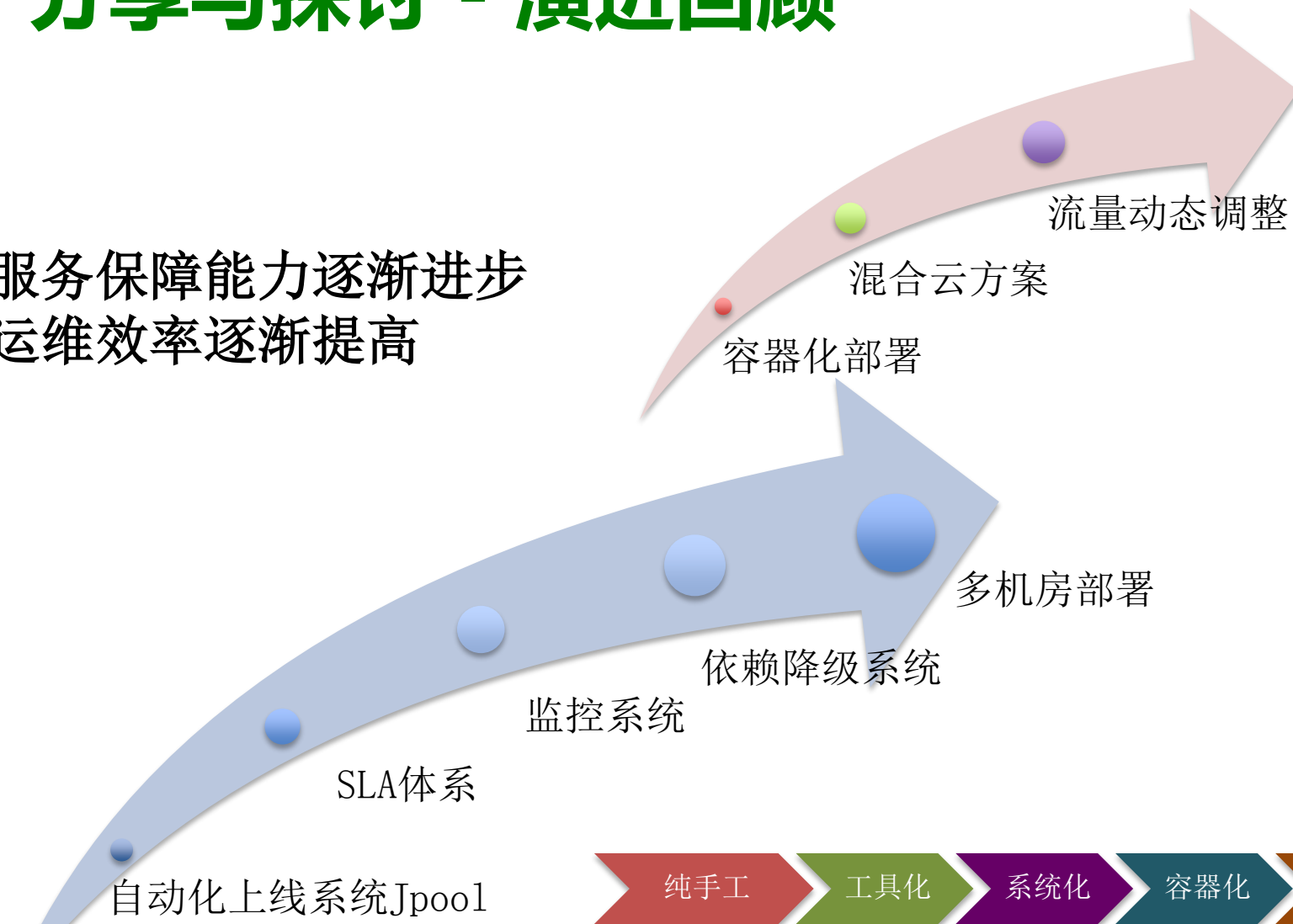


混合云方案即将助力微博2016春晚峰值挑战



# 分享与探讨 - 演进回顾

服务保障能力逐渐进步  
运维效率逐渐提高



# 分享与探讨 - 如何保障高可用及高效运维



## 服务保障四原则

# 分享与探讨 - 如何保障高可用及高效运维





长按二维码或搜索微信号 *【ArchNotes】* 添加高可用架构

---

# Thanks!

欢迎后续交流探讨，微博 @Fingki\_Li

