



GOPS2018  
Shenzhen

# GOPS

全球运维大会 2018

2018.4.13-4.14

中国·广东·深圳·南山区 圣淘沙大酒店（翡翠店）





GOPS2018  
Shenzhen

# 万台服务器一人挑的实战技巧

张黎明 组件运维团队负责人，8年运维



GOPS2018  
Shenzhen

# 海量服务运维过程中的挑战

1.8万个域名、3000个业务模块、4万台设备，单人运维设备超2万

- 上万域名如何保证就近接入，如何应对运营商出口网络故障
- 上万域名的https证书如何高效统一维护
- 服务器宕机如何保证业务无损和故障自愈
- 如何维持现网服务的整齐划一，实现自动化能力
- 如何有效应对大型活动事件几百模块、上万设备的快速缩扩容



GOPS2018  
Shenzhen

# 目录

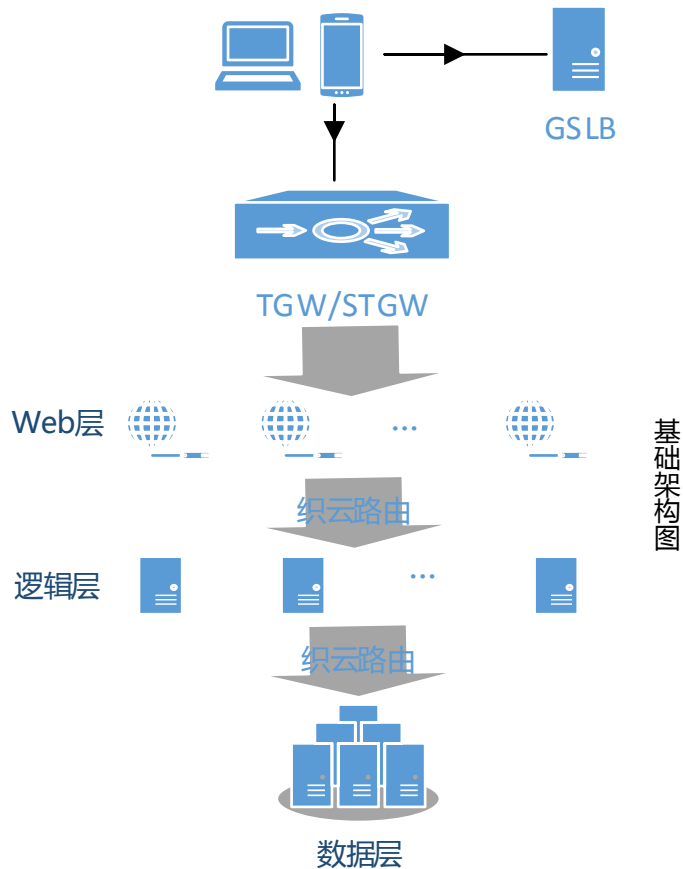
- ➔ **1** 海量服务的基础架构
- 2** 运维实践中总结的几个原则
- 3** 支撑大型活动事件的实战技巧



# 海量服务的基础架构

1. 名字服务实现没有调不走的流量
2. 容错做到没有不能宕机的设备
3. 统一框架提升研发运维效率

牛顿说：“如果我看得比别人更远些,那是因为我站在巨人的肩膀上。”  
我说：“我们只是站在巨人的肩膀之上，踮起脚尖，做了一次身高一米八的眺望”





# 巨人的肩膀：GSLB

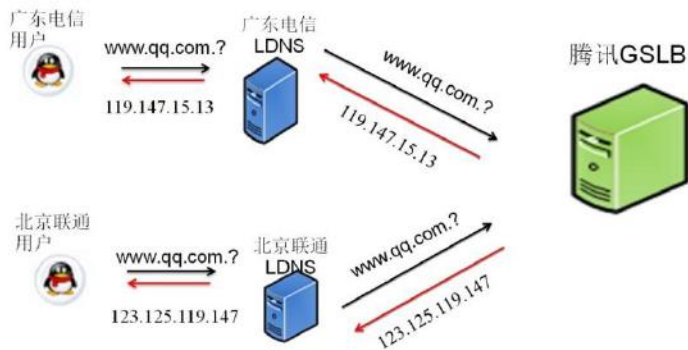
GSLB 是英文Global Server Load Balance的缩写，GSLB 是腾讯的自研DNS服务，通过识别LocalDNS出口IP的国家+省份+ISP属性，然后给对应请求返回相应的IP，实现就近接入

域名接入面临的问题：

- 1、跨运营商通信网络差，要尽量避免
- 2、运营商出口网络故障要能快速切换

GSLB强大的基础数据：

- 1、全面而精准的IP地址库
- 2、拥有各IDC机房对各地用户覆盖质量的实时数据



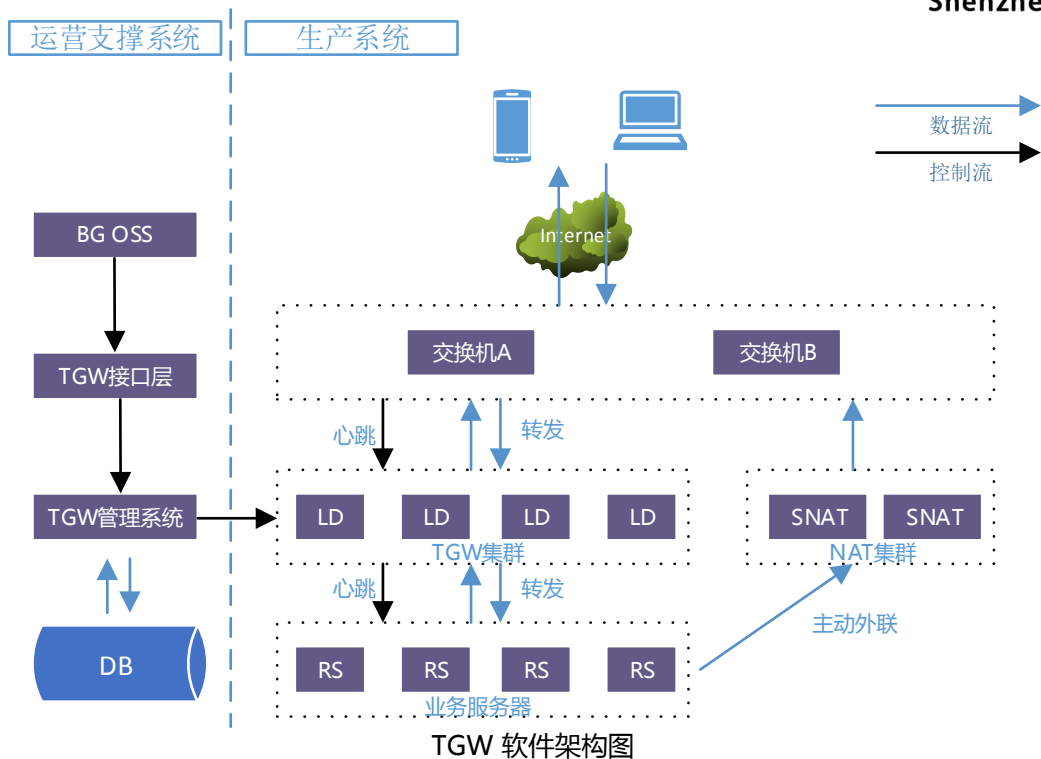
用户GSLB查询过程



# 巨人的肩膀：TGW

TGW，全称Tencent Gateway - 腾讯网关系统。具备以下功能特点：

- 负载均衡
- 支持多运营商接入
- 外网IP收敛
- 支持4/7层，业务接入透明



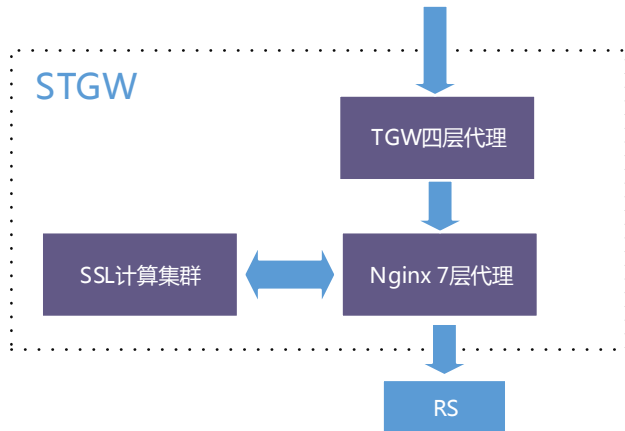


# 巨人的肩膀：STGW

STGW是在TGW基础之上构建的支持https的网系统，“stgw=tgw+https加解密集群”

具备以下特点：

- ✓ 证书统一集中管理，完全托管
- ✓ SSL加速，计算资源共享
- ✓ 多协议接入及适配
- ✓ 安全防攻击



STGW架构图

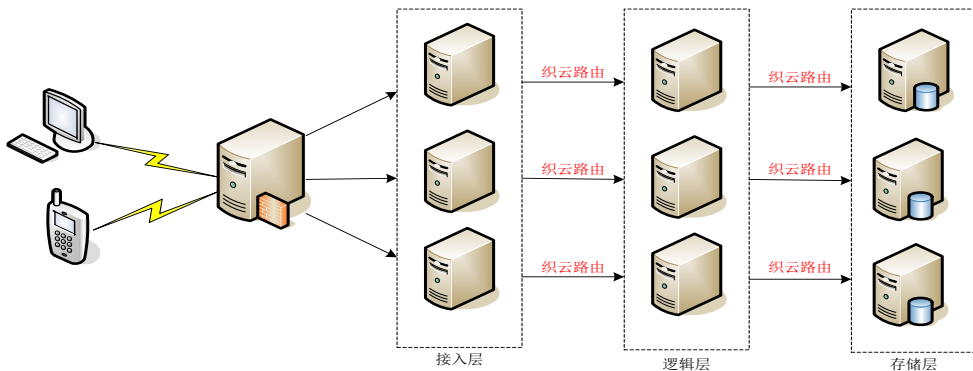




# 巨人的肩膀：织云路由

织云路由是应用层实现的一套兼具负载均衡和自动容错的**内网**名字服务系统，能有效解决以下问题：

- 缺乏一套稳定的内网服务注册，服务发现方案，调用方写死ip导致无法快速切换
- 业务多地部署缺乏一套就近访问,异地自动容灾切换的解决方案
- 同一个服务下的各个服务器负载不均衡，不能较好的做到能者多劳，物尽其用
- 单机故障会导致服务可用性受影响，需要运维人员半夜进行紧急变更
- 业务压测，灰度时需要引流，限流
- 业务需要过载保护，防雪崩

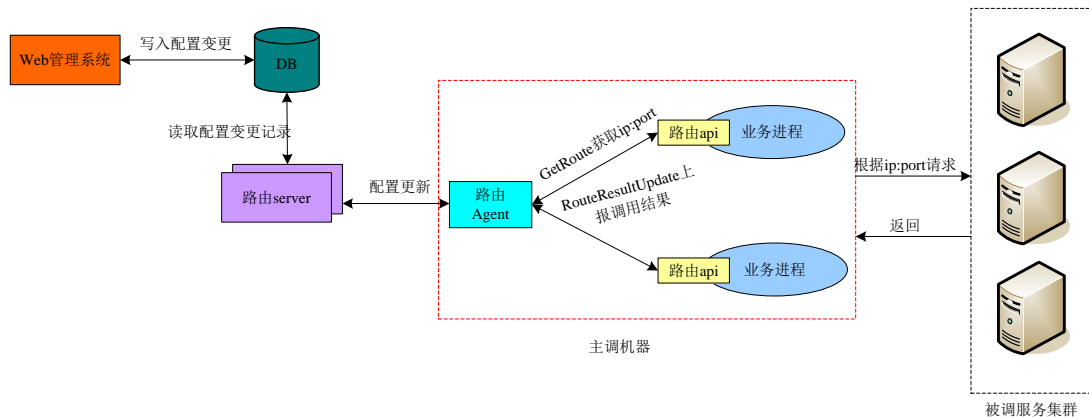




# 巨人的肩膀：织云路由

织云路由功能：

- 名字服务
- 负载均衡
- 过载保护
- 有状态路由
- 就近访问,异地容灾
- 服务可用性和返回码监控



目前织云路由已支持私有化交付，可以方便的对接客户的CMDB和OA认证系统，也可以和织云一整套自动化运维解决方案打包交付，[有意向的客户可以联系contactus\\_zy@tencent.com](mailto:contactus_zy@tencent.com)



# 织云路由与其他负载均衡组件对比

对比功能项	织云路由	LVS/F5	DNS
就近访问寻址	支持	LVS不支持/F5 GTM支持	支持，根据不用国家、省份、运营商进行调度
故障风险	只提供寻址，不转发业务数据；有缓存机制，故障对业务几乎无影响	需集中转发业务数据；故障对业务是致命的（通过规范标准的集群容灾规避整集群故障）	有缓存机制；故障对业务有影响
负载均衡	基于业务应用层反馈，可靠性更高；支持权重，动态自适应调节；	负载均衡算法比较简单。	基于简单的轮询机制
容错能力	支持RS自动容错，恢复	支持自动容错；	无
过载保护	有	无	无
监控统计	基于API上报，对服务的请求量、成功率和返回码进行统计，实现服务质量的精细化管理	无法提供业务纬度的精细监控	无法提供业务纬度的精细监控
部署维护	主调机器安装agent即可	需要专用设备，或对系统进行特殊配置	系统缺省提供
使用门槛	低，需业务代码调用API	无，VIP方式，对业务透明	低，需业务调用域名查询API

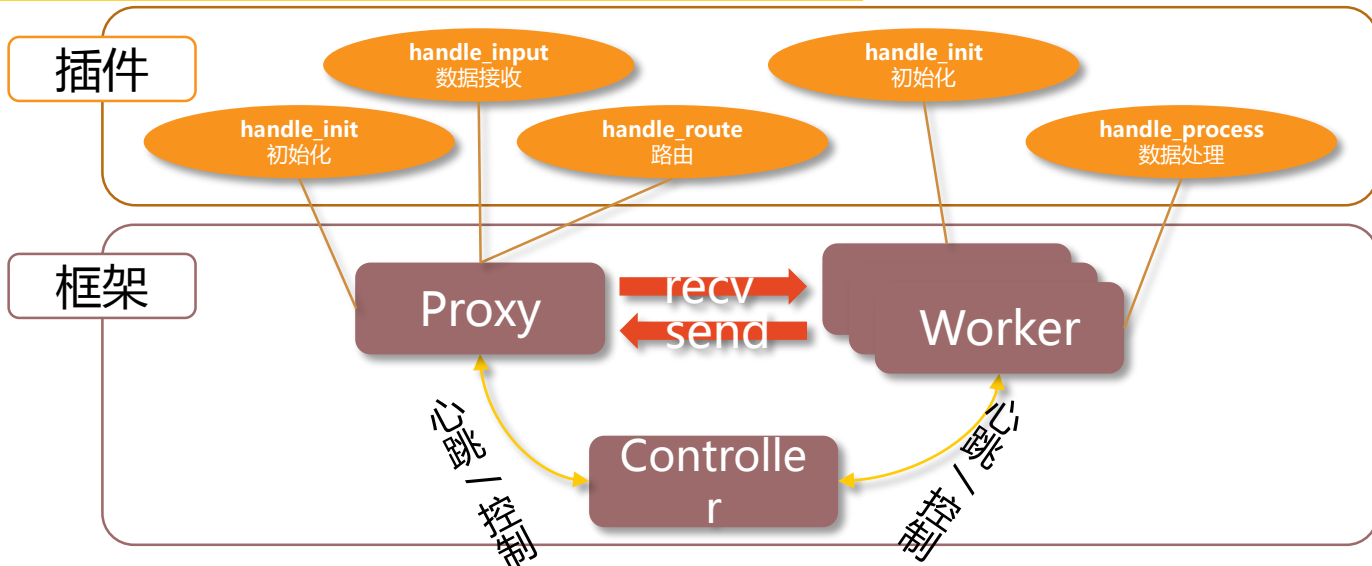


# 巨人的肩膀：逻辑层统一框架SPP

Spp是SNG使用最广泛的健壮、高性能通用逻辑框架，开发可以只实现几个回调函数就搭建起自己的逻辑server。

统一框架对于运维的意义：

- 跨业务实现维护策略统一，是后续自动化运维的基础
- 网络框架和业务逻辑 SO 分离管理，运维具备快速升级框架的能力
- 运维专业度提升，具备比开发更强的troubleshoot能力
- Controller对proxy和worker具备监控和起停的能力，服务稳定性提升





GOPS2018  
Shenzhen

# 目录

**1** 海量服务的基础架构

**2** 运维实践中总结的几个原则

**3** 支撑大型活动事件的实战技巧



# 名字服务原则

名字服务带来的流量快速切换和自动容错对于运维意义重大：

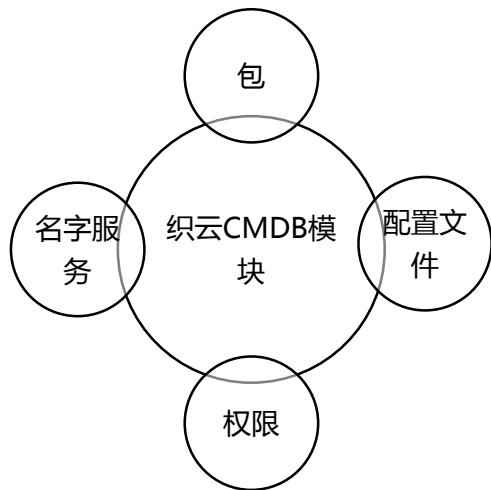
- 每年上万台设备，上千个VIP的裁撤
- 调度演习、发布、灰度和缩扩容而产生流量调度是常态
- 对于近十万台设备，每天几十台设备宕机成为常态
- 自动化运维的基础之一

名字服务如何在研发侧落地执行：

- 联合QA建立质量考核体系，对业务研发进行考核
- 后台框支持RPC，对业务开发封装路由寻址的过程
- 织云路由在内网支持DNS协议



# 一致性原则



一致性包含两个方面：

- 织云CMDB录入与现网机器上的是否一致
- 同模块下的各个IP资源是否一致

## 模块资源一致性

业务包：  
回收非公共包安装权限

配置文件：  
支持key-value；定期和中心对比，开启强一致

权限：  
推动常用权限管理维度模块化

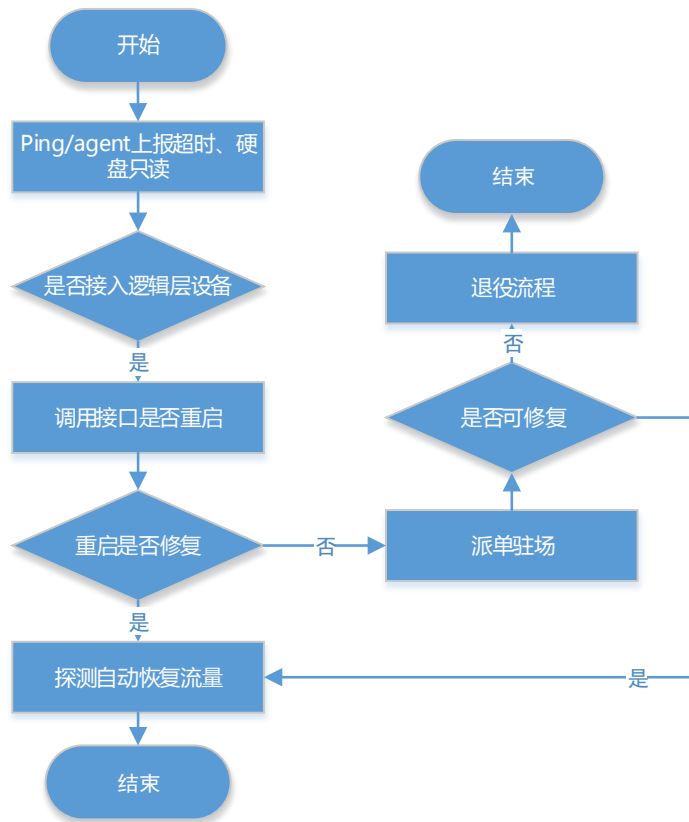
名字服务：  
梳理并推动业务模块的名字服务接入一致性；不一致对操作人派发工单



# 无数据原则

接入层和逻辑层设备无数据原则是告警自动化处理的重要基础

- 无cache内存型服务，对于重启服务恢复故障意义重大
- 设备要求简化，无raid，硬盘故障之后直接换盘重装
- 磁盘在保存尽可能多日志的前提下实现自动化清理



单机故障自动处理流程，简单才会可靠

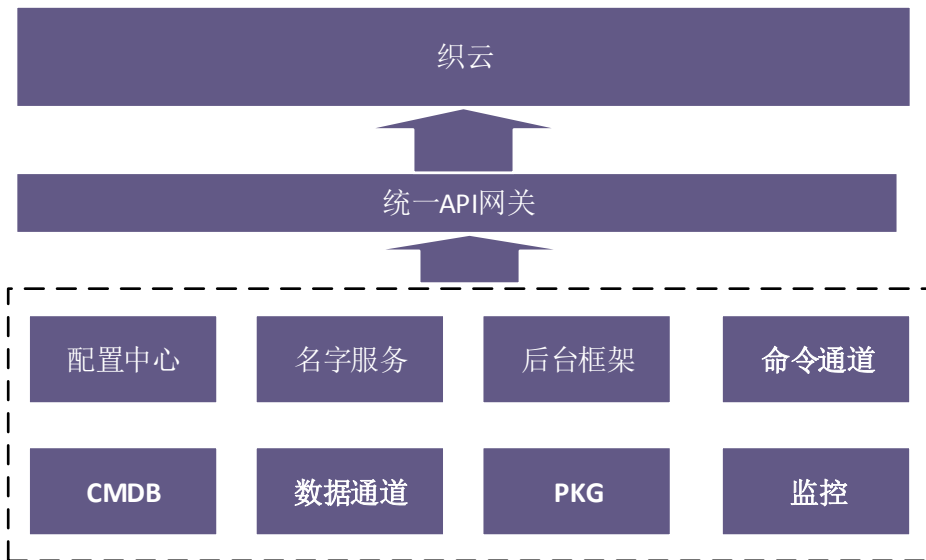




# 统一原则

- 统一后台服务框架SPP
- 统一名字服务织云路由
- 统一配置中心
- 统一数据上报通道
- 统一包发布系统PKG

统一原则带来的维护对象的减少和对上层系统呈现出来的简单可以大大提升上层自动化系统可靠性，并降低其复杂度





GOPS2018  
Shenzhen

# 目录

1 海量服务的基础架构

2 运维实践中总结的几个原则

➔ 3 支撑大型活动事件的实战技巧

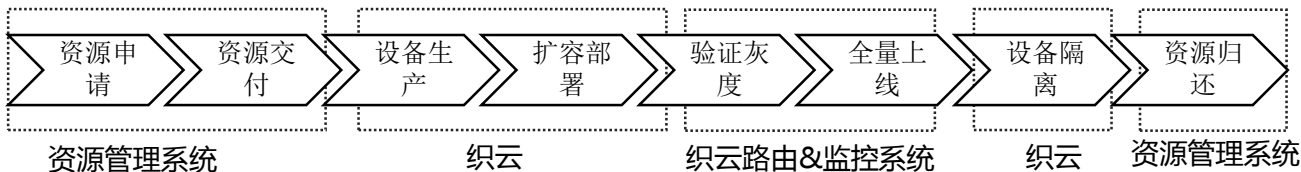


# 大型活动的挑战

- 扩容设备量大、模块多、扩容状态跟进难、设备采购交付周期长，留给扩容部署的时间紧

刚刚过去的2018年春节和红包活动，组件运维团队扩容641次，涉及535个模块，15701台设备，扩容操作要求在设备交付后的一周内完成

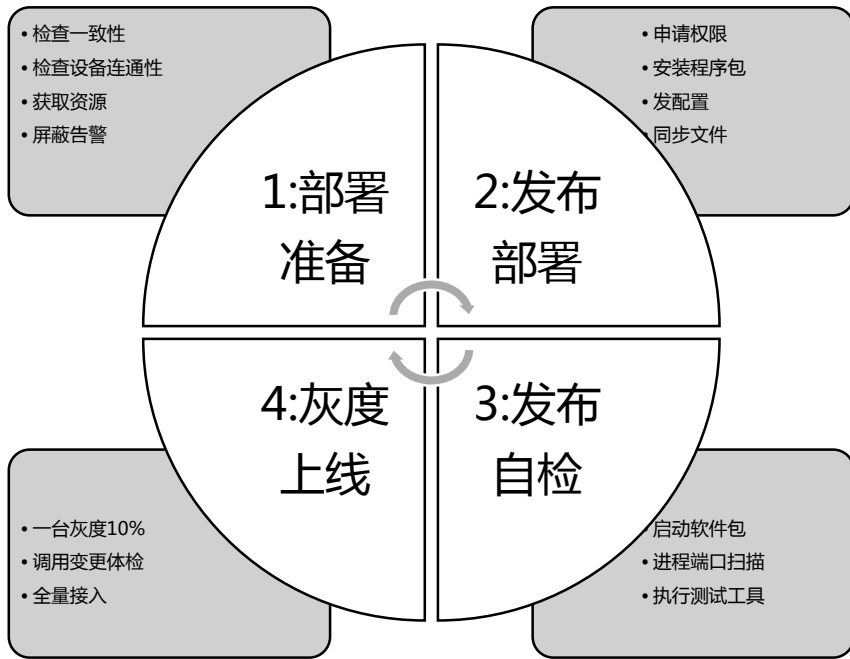
- 步骤繁多，各个系统之间跳转，沟通和单独操作消耗量巨大





GOPS2018  
Shenzhen

# 织云自动部署流程



✓ 获取参数列表

✓ 检查一致性上报

✓ 检查设备连通性

✓ 获取资源配置

✓ 申请权限

✓ 屏蔽告警

✓ 安装程序包

✓ 获取cc参数

✓ 发配置

✓ 同步文件

✓ 执行脚本

✓ L5主调扩容

✓ cmlb主调扩容

✓ 启动软件包

✓ 进程端口扫描

✓ 告警屏蔽解除

✓ 执行测试工具

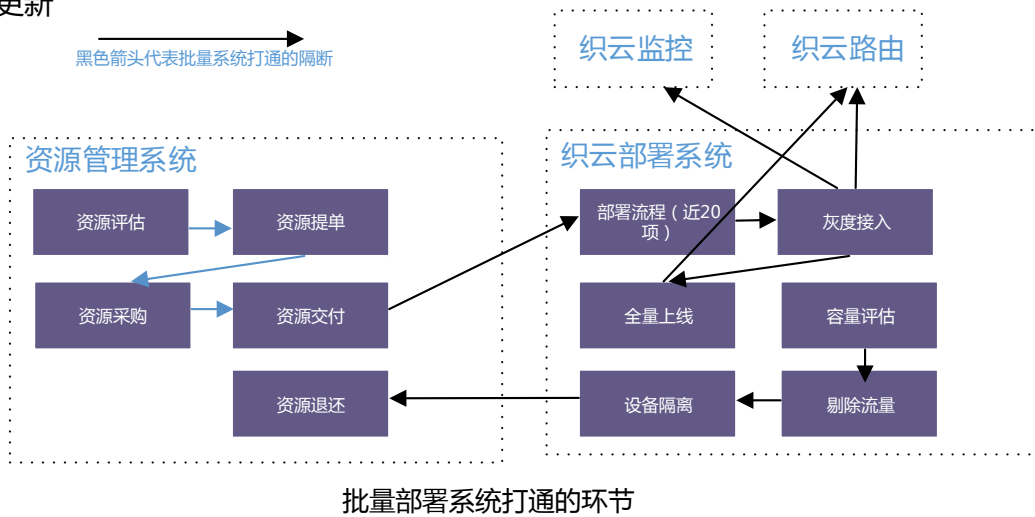
织云扩容流程截图



# 基于织云之上的批量部署系统

## 批量部署系统：

1. **集中展示和管理**：各个步骤封装成原子接口，在统一界面操作
2. **对接各个系统**：打通资源系统、织云和监控系统，扩容流程自动发起，自动上线
3. **规模效应**：批量操作上百个业务模块
4. **状态自动跟进**：状态自动更新



# 18年春节数据



GOPS2018  
Shenzhen

一周内组件层共扩容641次，涉及535个模块，15701台设备  
其中，单人最高扩容310次，涉及272个模块，9757台设备  
组件层总体成功自动上线522次，涉及414个模块，10599台设备

春节期间组件团队共41556台设备，其中单人最多名下21069台



GOPS2018  
Shenzhen



 腾讯织云  
CLOUD OPERATIONS CONSOLE



高效运维社区  
开放运维联盟

# Thanks

腾讯运维体系专场  
荣誉出品



GOPS2018  
Shenzhen

想第一时间看到高效运维社区的  
最新动态吗？

