



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  
OOPSA Open OPS Alliance  高效运维社区  
Great OPS Community

指导单位：  数据中心联盟  
Data Center Alliance



# 支撑2亿用户并发 大规模Web Server集群的高效运维

周小军 腾讯社交网络运营部



# 提纲

- 业务场景
- 接入架构
- 运维的挑战
- 高效运维的实践
- 未来愿景



# 个人简介

- ❖ 周小军@腾讯社交网络运营中心
- ❖ 十几年IT架构和运维经验
- ❖ 擅长数据中心、IT架构&运维、云计算



# 业务场景

- QQ在线 2.59 亿
- 手Q在线 2.49 亿
- 刷一刷抢红包用户数 3.08 亿
- QQ日消息发送总量 200 亿条
- Qzone在线 3 千万
- 相册日上传图片 6 亿张

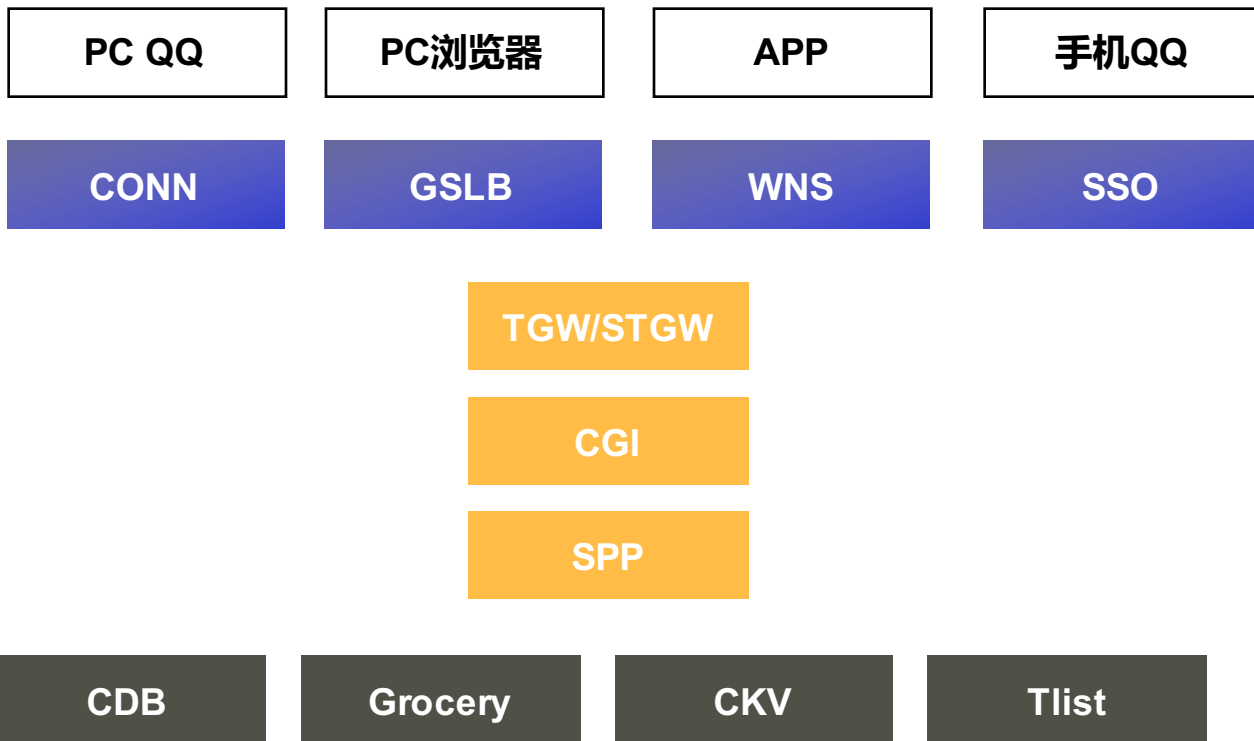


# 业务场景

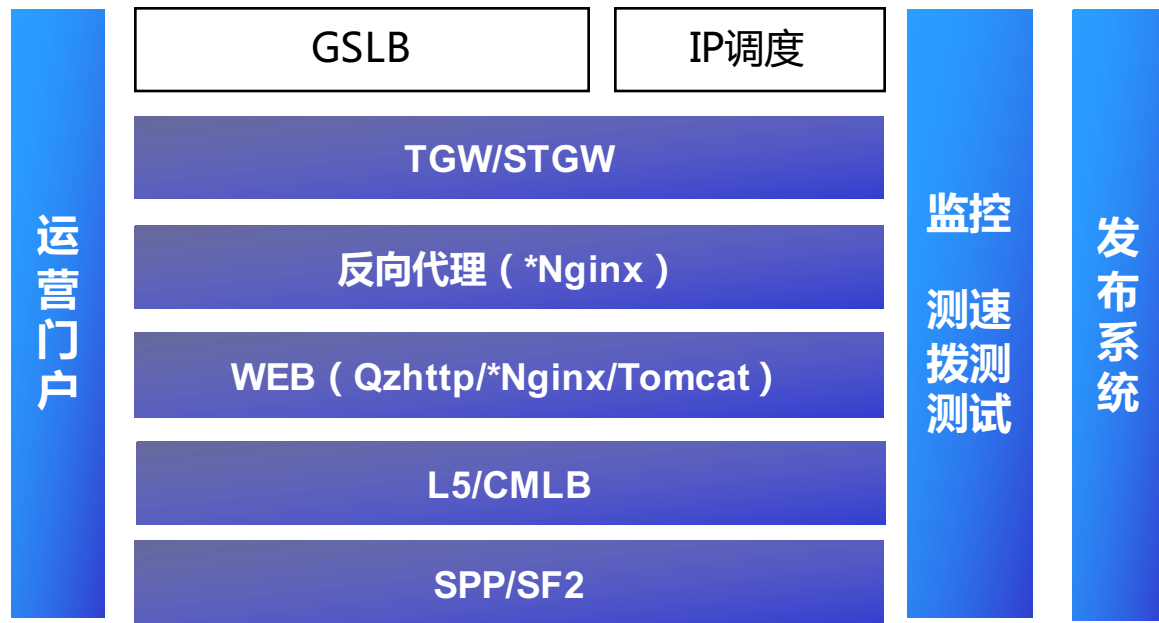
- 挑战：海量并发、延迟、服务可持续性
- 服务器规模 < 1万台
- 服务可用性 = 99.99%



# 三层架构



# 接入层服务架构



# Web架构的变迁

## 四个阶段

### 1) 使用域名来隔离业务

- 简单
- DNS解析太多

### 2) 业务统一域名

- DNS解析快
- 部署复杂，应用耦合

### 3) 添加反向代理

- 后端应用分离
- 增加了一层架构

qzone.qq.com



TGW

Nginx



应用2

应用3

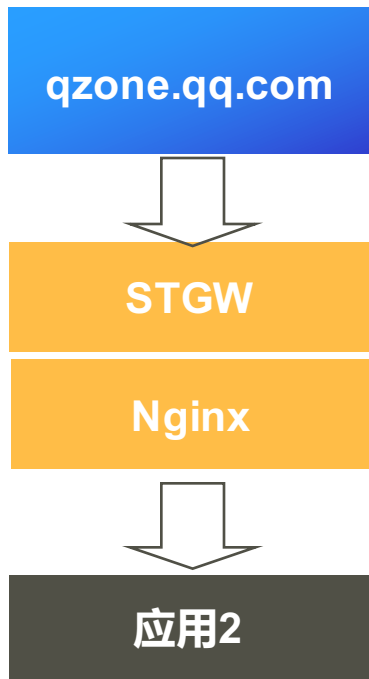
应用1



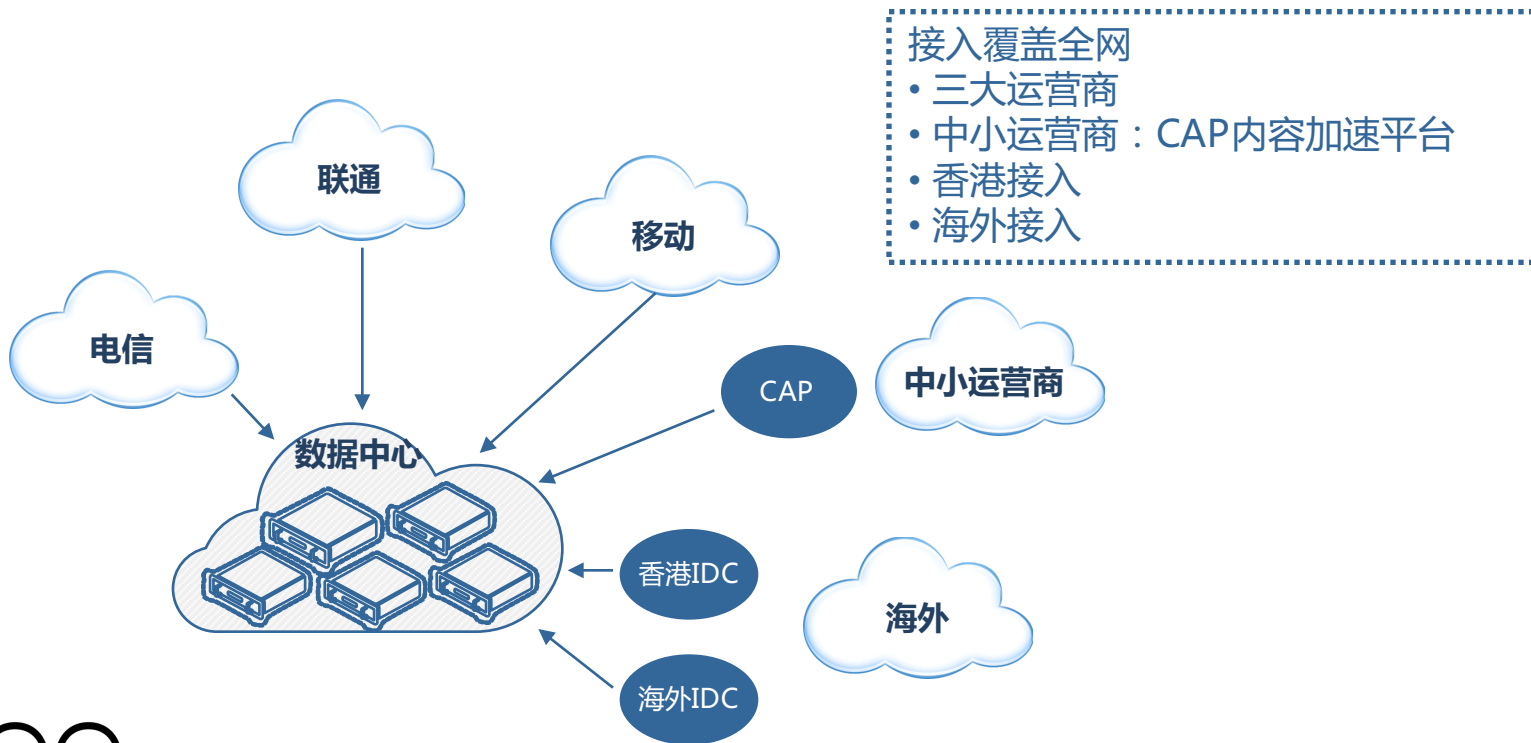
# Web架构的变迁

## HTTPS阶段

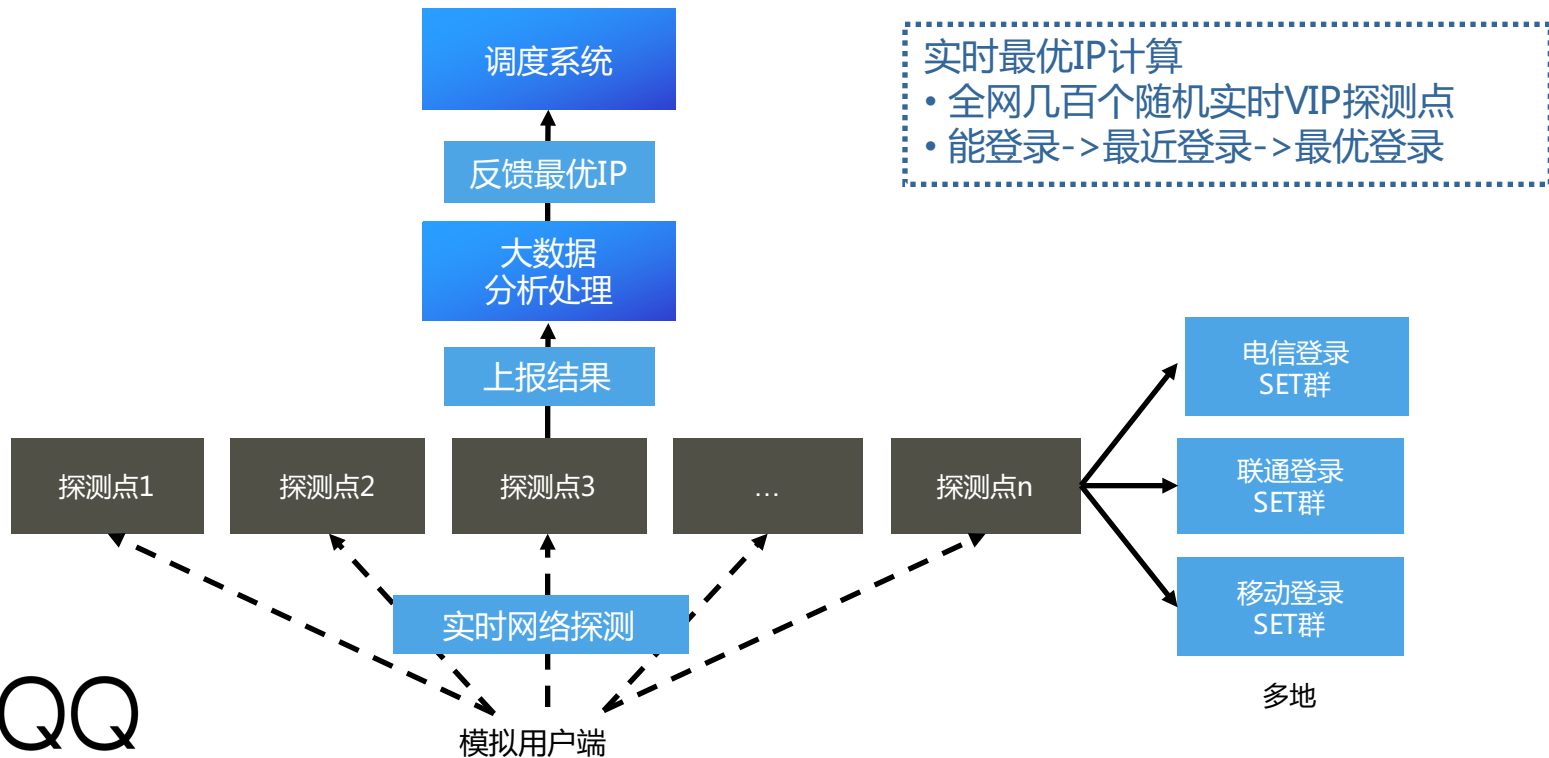
- 证书存储在STGW
- STGW使用硬件加速卡
- 保持现有业务架构
- 第一次请求延迟较HTTP增加



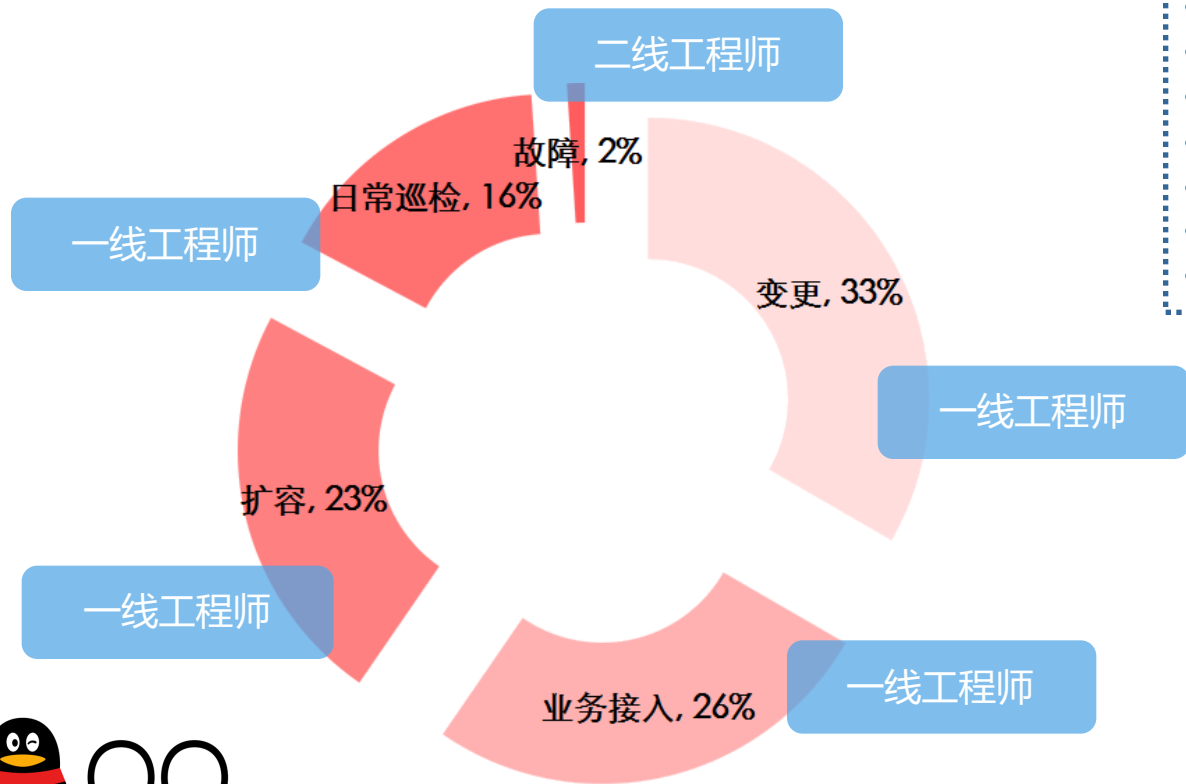
# 全网接入架构



# 移动终端接入架构



# 接入运维的挑战-一线事务多



- 设备故障处理
- 业务报障处理
- 业务接入 ( 域名申请到上线 )
- 日常扩缩容
- 版本升级
- 业务规则变更
- 其它





# 接入运维的挑战-坑多

标准化程度不够

组件多

路由复杂

业务部署不规范

自动化程度不够

人工运维易犯错误

拥抱架构变化

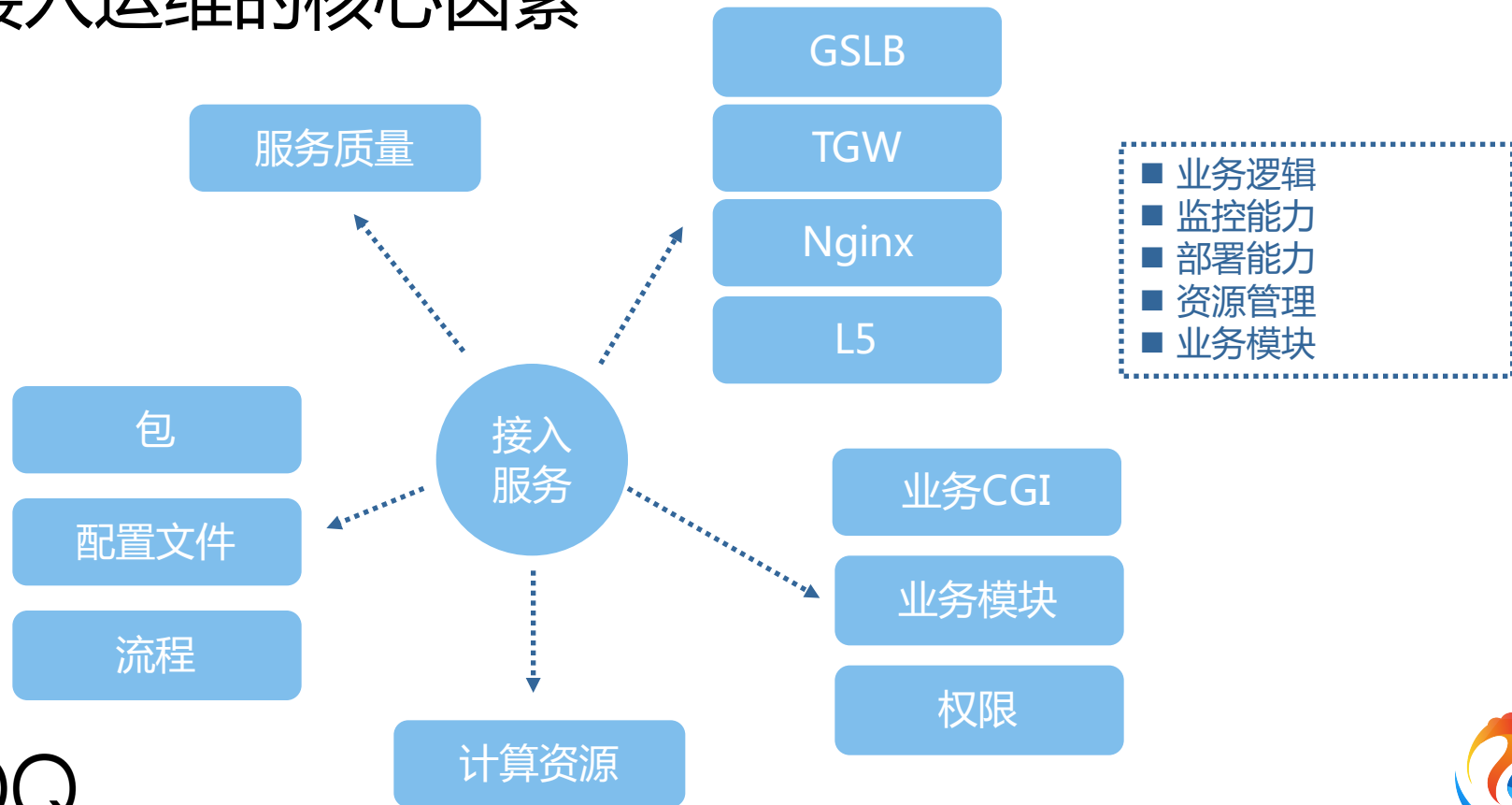
新旧架构迁移

监控覆盖率不高

标准化  
路由化  
流程化  
自动化  
监控率



# 接入运维的核心因素



# 高效运维-织云自动化

每个模块下的设备提供的服务都完全一致，每个模块可以看作是提供服务的最小单元。

一个模块下的服务抽象成以下几个资源配置：  
软件包、配置、文件、权限、脚本。

织云 首页 接入层 逻辑层 数据层 基础服务

[N][增值渠道] > qq.com[WEB] 虚拟机(CAE)

共计 9 个配置 + 添加

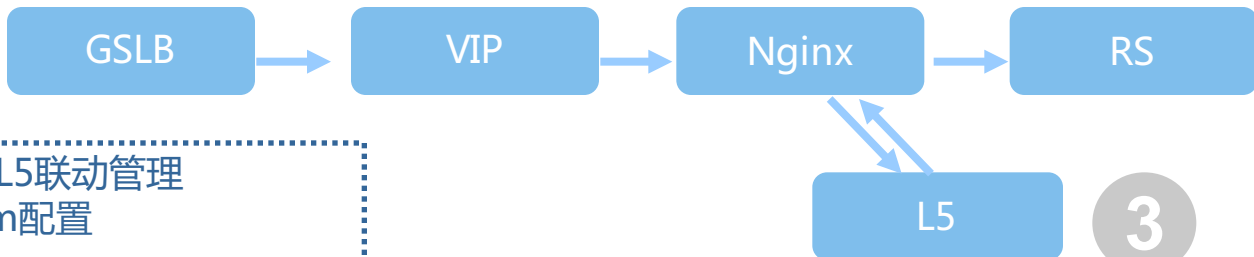
模块资源共 1 次变更 撤销全部 撤销上一步 保存到CMDB

配置 9

名称	操作	文件ID	描述
cardentry.conf	编辑   下发	211428	
config.xml	编辑   下发	211427	
LinkForward.conf	编辑   下发	241815	跳转链接用配
AppReportCgi.conf	编辑   下发	253258	
nginx.conf	编辑   下发	253261	nginx配置发
report.conf	编辑   下发	253390	
SchoolAccessImage.conf	编辑   下发	253409	
config.json	编辑   下发	253429	
clear.conf	编辑   下发	254585	[N][增值渠道



# 高效运维-路由策略



1

Nginx与L5联动管理  
upstream配置

upstream的格式：  
upstream music.qq.com\_pc  
{  
  l5id 201607:65535;  
}

2

自动生成Nginx upstream格式的ip-port：

```
server 10.10.1.18:80 weight=100;  
server 10.10.1.19:80 weight=100;  
server 10.10.1.20:80 weight=100;  
server 10.10.1.21:80 weight=100;
```

3

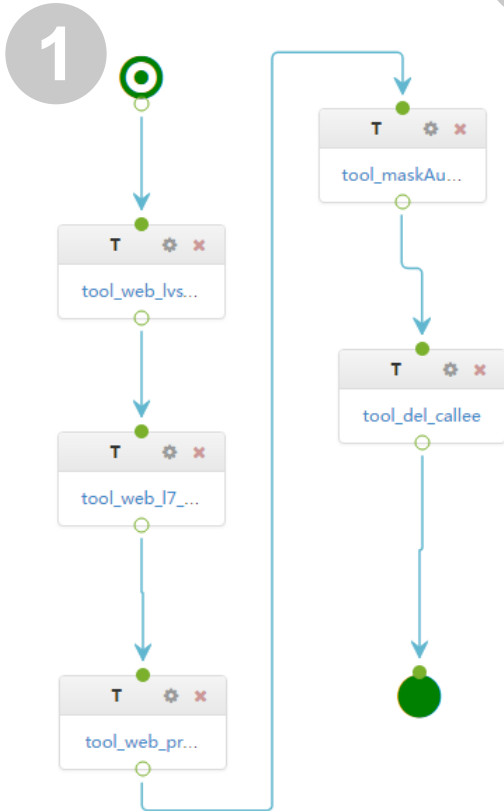
自动检测，检测L5有变更后重新生成upstream，并随机sleep来reload nginx服务。



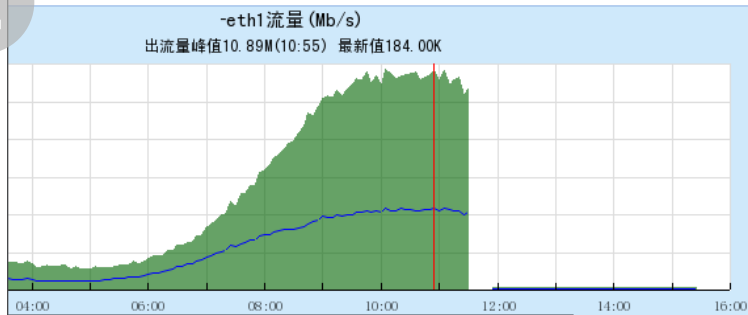
# 高效运维-路由策略

## 下线三步骤

- 1) 删除路由
- 2) 流量观察
- 3) 隔离操作



2



3

隔离信息

说明: 隔离后的设备会保留在当前模块, 状态改为部门回收中. 隔离到期后, 设备会自动下线并移除当前模块. 当设备还在隔离期时, 可以将其回收. 到现网故障机的设备将直接下线回收, 不可回收.

隔离IP(9):  
10.10.10.10  
10.10.10.10  
10.10.10.10  
10.10.10.10  
10.10.10.10  
10.10.10.10  
10.10.10.10  
10.10.10.10  
10.10.10.10

是否隔离:  否  是

隔离方式:  普通隔离(7天)  快速隔离(2天)

隔离原因:

是否是故障机:  是  否

业务下线  
 故障  
 其它

直接下线回收, 不可回收

关闭 启动维护流程



# 高效运维-包发布系统

包管理 / 查看程序包 [other] dbc\_raid\_tool - 1.0.1

基本信息 文件管理 进程 调度 占位符&监控 安装相关 启动相关 init.xml

- \* 类型和框架
  - 后台server包
  - SPP插件包
  - 脚本包
- \* 所属业务 其他
- \* 包名 dbc\_raid\_tool
- \* 程序包用户 root
- \* 版本 1.0.1
- 包依赖
- \* 作者 millerzhou
- \* 版本说明 在TS6上创建软RAID 10的版本包

问题反馈

启用桌面通知 + 创建新版本 导入安装记录 过滤IP ★ 已收藏

1台设备的 wnginx 正在热重启 0 / 1 详细

升级 启动 停止 重启 热重启 卸载 回滚

安装时间	其它 ...	当前操作
2016-05-11	[老业务接入][代理]	热重启中



像管理软件一样管理包



# 高效运维-配置文件管理

配置管理 / 与我相关

按文件名 ▾ 输入搜索词 搜索

共 14 条记录

文件ID	文件名	文件描述	最新版本	创建人	最近下发	最近更新	操作
254699	<a href="#">clear.conf</a>	[N][Qzone]-	0.0.0.1	millerzhou	millerzhou	2016-07-10 06:12:16	<a href="#">编辑</a> <a href="#">下发</a> <a href="#">下发记录</a> <a href="#">版本记录</a> <a href="#">修改描述</a> <a href="#">邮件抄送</a> <a href="#">删除</a>
254585	<a href="#">clear.conf</a>	[N][增值渠道	0.0.0.1	millerzhou	millerzhou	2016-07-04 11:13:09	<a href="#">编辑</a> <a href="#">下发</a> <a href="#">下发记录</a> <a href="#">版本记录</a> <a href="#">修改描述</a> <a href="#">邮件抄送</a> <a href="#">删除</a>
254438	<a href="#">clear.conf</a>	[N][兴趣部落	0.0.0.1	millerzhou	millerzhou	2016-06-27 08:45:46	<a href="#">编辑</a> <a href="#">下发</a> <a href="#">下发记录</a> <a href="#">版本记录</a> <a href="#">修改描述</a> <a href="#">邮件抄送</a> <a href="#">删除</a>
254314	<a href="#">rsync_access.sh</a>	[增值图片we	0.0.0.1	millerzhou	millerzhou	2016-06-17 19:53:33	<a href="#">编辑</a> <a href="#">下发</a> <a href="#">下发记录</a> <a href="#">版本记录</a> <a href="#">修改描述</a> <a href="#">邮件抄送</a> <a href="#">删除</a>



# 高效运维-配置文件管理

配置管理 / 编辑|下发

编辑: 文件名: nginx.conf; fileId: 253261; 文件描述: tnginx配置发布

历史版本: (只读) 0.0.0.3

```
user nobody;
worker_processes 8;

pid          log/nginx.pid;

events {
    use epoll;
    worker_connections 100000;
}
worker_rlimit_nofile 100000;

http {
    include      mime.types;
    default_type application/octet-stream;
    server_tokens off;

    server_names_hash_bucket_size 128;
    client_header_buffer_size 32k;
```

当前编辑: 0.0.0.1

diff 本地上传 生成版本

```
user nobody;
worker_processes 8;

pid          log/nginx.pid;

events {
    use epoll;
    worker_connections 100000;
}
worker_rlimit_nofile 100000;

http {
    include      mime.types;
    default_type application/octet-stream;
    server_tokens off;

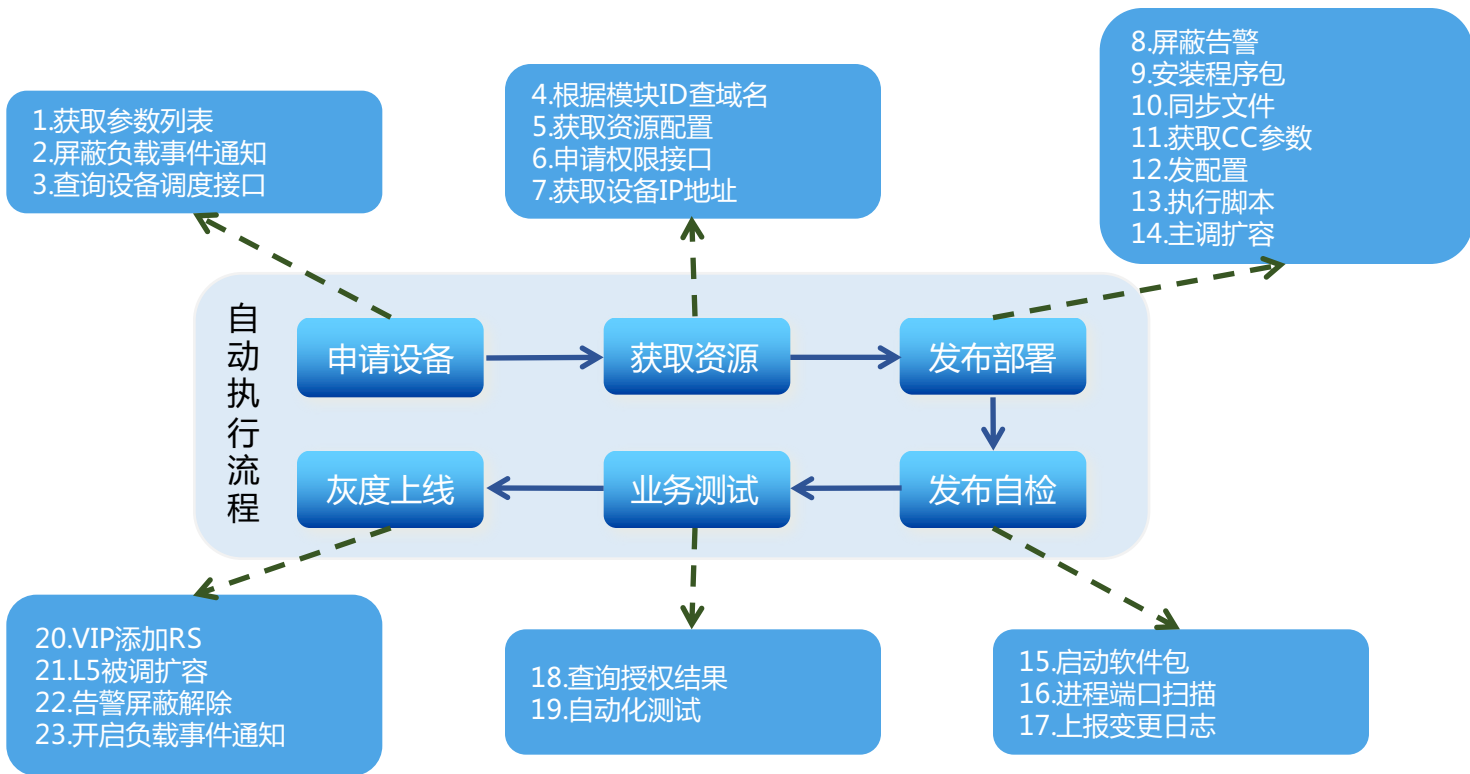
    server_names_hash_bucket_size 128;
    client_header_buffer_size 32k;
```

下发管理: 当前选中版本: 0.0.0.1; 版本描述: 版本MD5值: 43e3c ; 编码: UTF-8

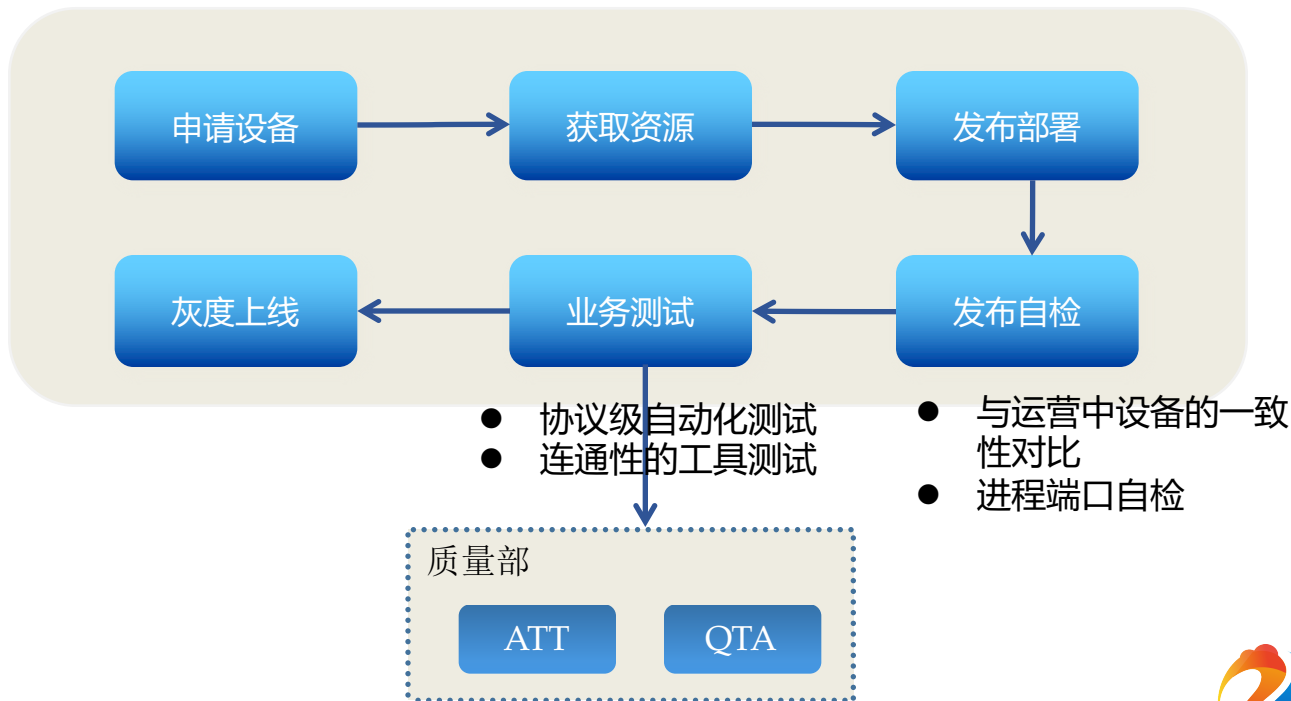
下发结果:



# 高效运维-一键扩容



# 高效运维-自动扩容流程



# 高效运维-流程剖析



- ✓ 检查设备连通性
- ✓ 告警屏蔽
- ✓ 获取资源配置
- ✓ 程序包安装
- ✓ 获取cc配置参数
- ✓ 下发配置文件
- ✓ 同步文件
- ✓ 执行脚本
- ✓ 申请权限
- ✓ L5主调扩容
- ✓ CMLB主调扩容
- ✓ 启动程序包
- ✓ 拷贝进程端口
- ✓ 进程端口扫描
- ✓ 获取端口信息
- ✓ 扫描端口连接性
- ✓ ATT自动化测试
- ▶ 上报变更日志



# 高效运维-流程剖析

工具管理 / 工具详情 / 工具ID : 244

\* 工具英文名  \* 工具功能

\* 工具类型  命令行参数  引用框架标准输入输出 \* 执行用户

\* 输入定义

ID	名称	类型
<input type="text" value="i_module_id"/>	<input type="text" value="模块id"/>	<input type="text" value="单行文本"/>
<input type="text" value="i_ip_list"/>	<input type="text" value="Ip列表"/>	<input type="text" value="IP列表"/>

\* 输出定义

ID	名称	类型
<input type="text" value="o_result"/>	<input type="text" value="执行结果"/>	<input type="text" value="文件列表[CC]"/>

\* 可重入  是  否 \* 命令类型  \* 工具分类

tool

+ 新增

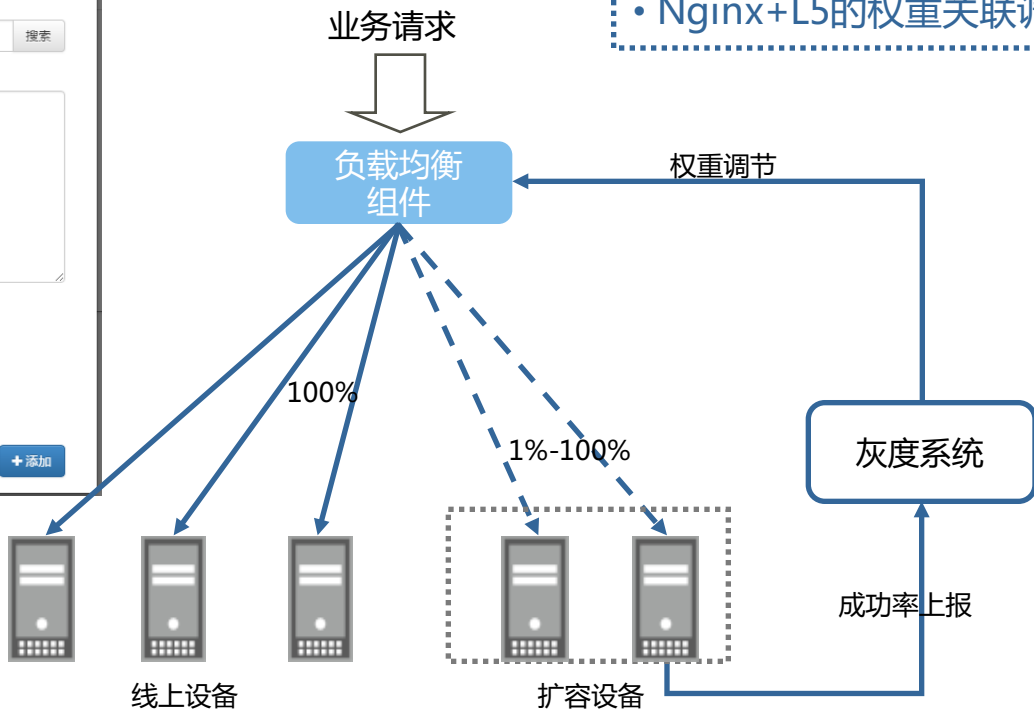
```
8 $obj = new WebProxyOffline();
9
10 $obj->offline();
11 $obj->getOpResult();
12
13 class WebProxyOffline{
14
15     private $magicTool;
16     ///操作结果
17     private $opResult;
18
19     private $ipList;
20
21     private $proxyUrl = 'http://proxy.i:
22
23     public function __construct() {
24
25         $this->magicTool = new MagicTool;
26         $tool_params = $this->magicTool
27
28         $this->ipList = $tool_params['i
```



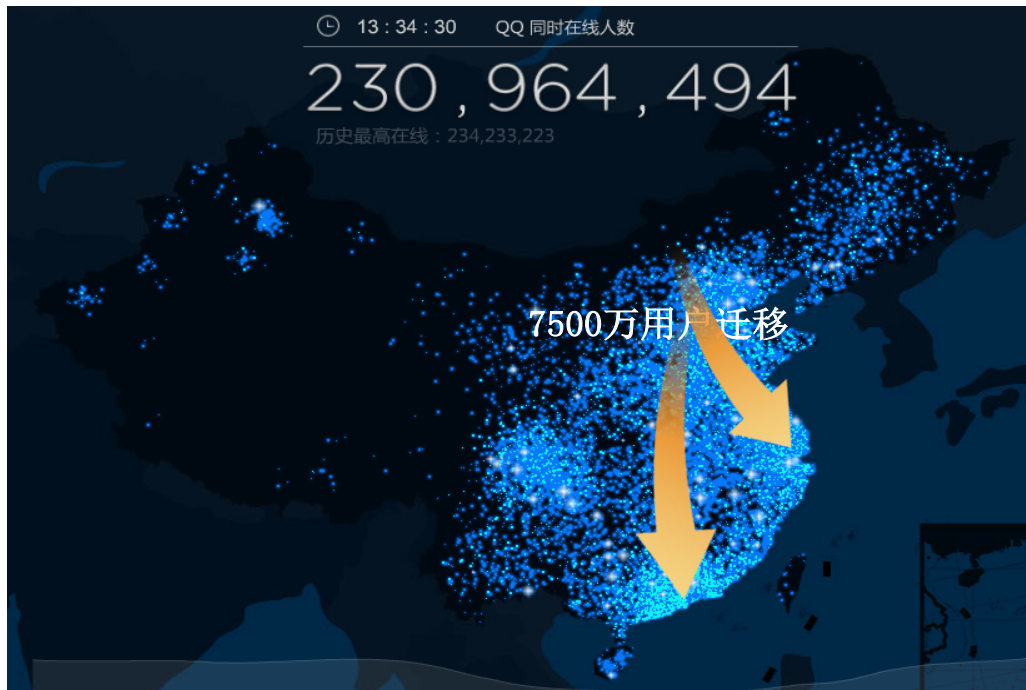
# 高效运维-如何做灰度

灰度的方法

- TGW上的权重调整
- L5上的权重调整
- Nginx+L5的权重关联调整



# 高效运维-快速调度能力



# 高效运维-快速调度能力

## ■ 变更GSLB

- 变更为异地TGW VIP
- 10分钟生效

## ■ 变更直连IP

- 下发新的VIP给移动终端
- 秒级生效

## ■ 变更L5内部服务

- 变更L5地址
- 秒级生效

国家:  运营商:

<input type="checkbox"/>	国家	运营商	地区
<input type="checkbox"/>	中国	电信	北京
<input type="checkbox"/>	中国	电信	天津
<input type="checkbox"/>	中国	电信	河北
<input checked="" type="checkbox"/>	中国	电信	山东
<input checked="" type="checkbox"/>	中国	电信	山西
<input checked="" type="checkbox"/>	中国	电信	陕西
<input checked="" type="checkbox"/>	中国	电信	内蒙古
<input checked="" type="checkbox"/>	中国	电信	宁夏
<input checked="" type="checkbox"/>	中国	电信	甘肃
<input checked="" type="checkbox"/>	中国	电信	青海
<input checked="" type="checkbox"/>	中国	电信	新疆
<input checked="" type="checkbox"/>	中国	电信	西藏
<input checked="" type="checkbox"/>	中国	电信	重庆
<input checked="" type="checkbox"/>	中国	电信	河南
<input type="checkbox"/>	中国	电信	湖南
<input type="checkbox"/>	中国	电信	福建
<input type="checkbox"/>	中国	电信	贵州
<input type="checkbox"/>	中国	电信	云南

### 批量修改RR

修改TTL

TTL:

修改指向

目标RR类型:

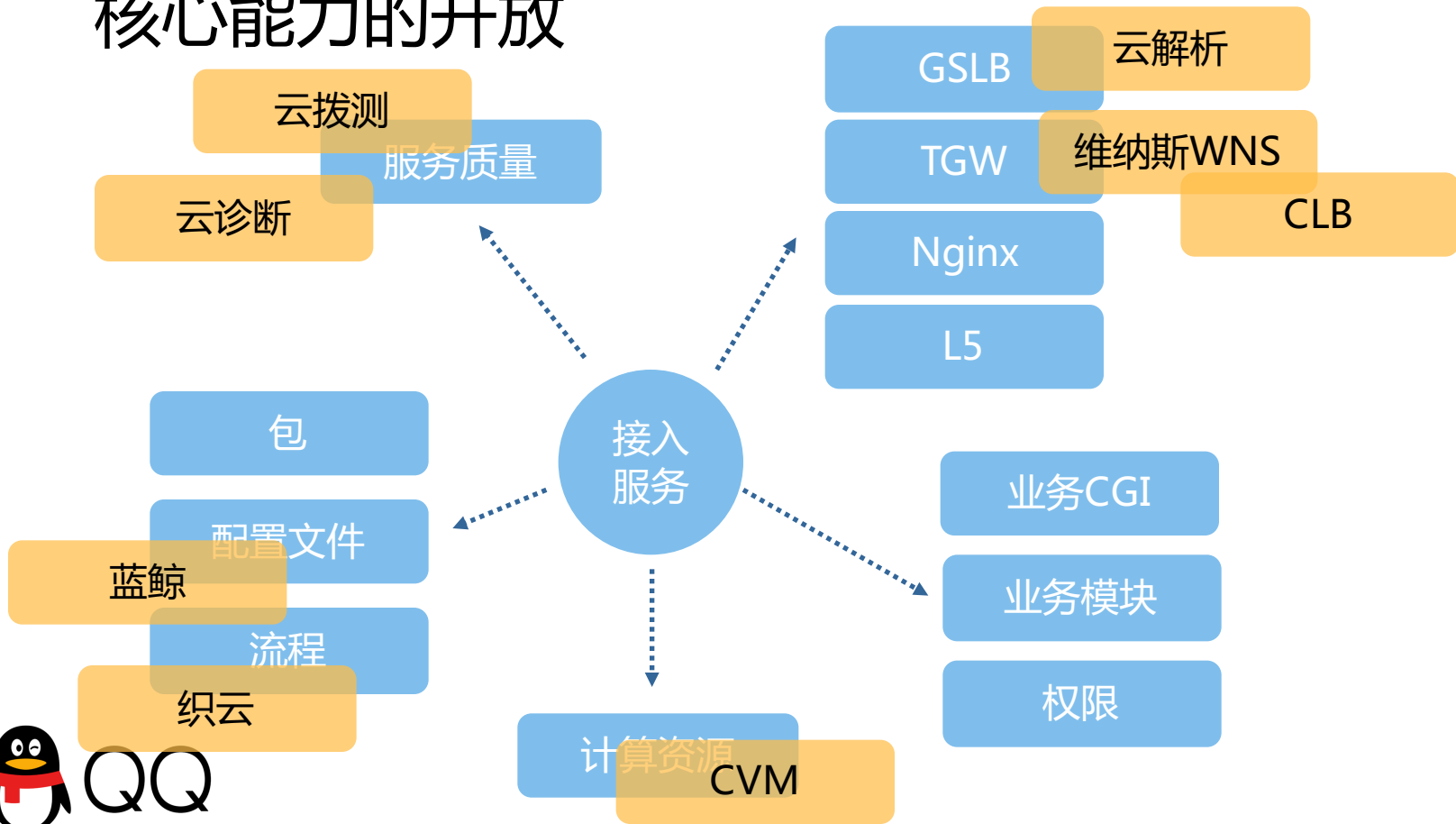
目标IDC:

- 成都 电信
- 上海\_ 电信\_有测速
- 广州\_ 电信\_有测速
- 香港\_ 香港\_无测速

深圳  
深圳  
深圳  
深圳  
深圳



# 核心能力的开放



# 接入运维工程师的价值和能力要求

业务变更

业务与架构能力

问题诊断能力

开源组件的改造能力

工具开发能力

运营平台构建能力

数据分析与决策

业务&架构

底层能力

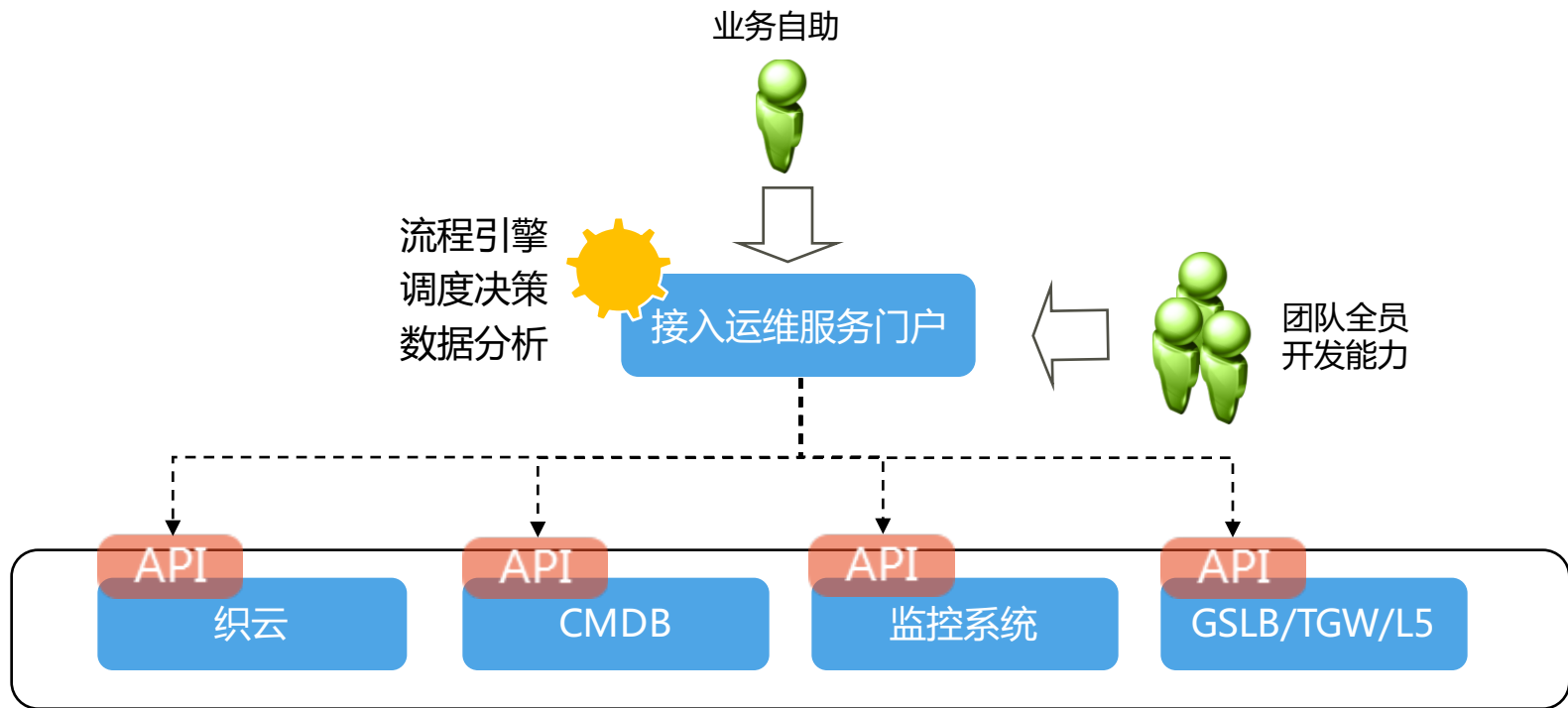
开发能力

运营意识

从防守型的运维到进取型的运营



# 接入运维门户



# 故障定位的挑战

连线 [服务] 08-23 19:25:08  
能让用户直接telnet ip +端口看看吗?

xu 08-23 19:25:53  
是在手机上 😞

连线 [服务] 08-23 19:29:03  
手机端确实不太好测试, 不过安卓端有挺多控制台程序 可以直接操作的

xu 08-23 19:32:44  
没有现成的工具? 我还要给用户装控制台程序? 🙄

连线 [服务] 08-23 19:33:38  
我这边没有呢

millerzhou(周小军) 08-23 19:49:01  
可以试下腾讯云诊断APP

xuai 08-23 19:51:05  
谢谢, 等用户回来我让她装个试试



使用云诊断能快速定位用户网络、域名、服务、运营商等问题



# 未来愿景

打通最后一公里--业务扩缩容自动化

服务化能力--实现Web as a Service

用户访问日志的数据挖掘与分析





# Thanks

高效运维社区  
开放运维联盟

荣誉出品





## 想第一时间看到高效运维公众号的好文章么？

请打开高效运维公众号，点击右上角小人，并如右侧所示设置即可：



# **GOPS2016 全球运维大会更多精彩**



## **GOPS2016** **全球运维大会•北京站**

2016年12月16日-17日  
北京国际会议中心

