

# Qzone后台基础架构之路

# 讲师介绍

讲师： 邓正伟（antideng）

部门： 腾讯/SNG/社交平台部

11年加入到腾讯QQ空间，从事Qzone后台基础架构的工作，负责Qzone的基础架构搭建，多地部署，柔性可用，直播等平台的研发工作；欢迎大家下载空间 app体验。

# 目录

## Qzone简介

## 整体架构

## 网关接入

### 接入层介绍

## 异地分布

### SET分布

### 同步中心

## 系统容灾

### 全网调度

# Qzone介绍

- 月活跃用户超6亿，日活跃用户超3亿
- 首页日访问量超40亿次
- 日新增UGC几亿条，日互动互访数上百亿次

空间的量级？

亿级，十亿级，甚至百亿级访问如何提供用户安全，快速，可靠，稳定的服务？

# 目录

## Qzone简介

## 整体架构

## 网关接入

### 接入层介绍

## 异地分布

### SET分布

### 同步中心

## 系统容灾

### 全网调度

# 整体架构



# 目录

Qzone简介

整体架构

网关接入

接入层介绍

异地分布

SET分布

同步中心

系统容灾

全网调度

# 为什么加一层网关接入

增加成本

加长链路



# 为什么加一层网关接入

## 雅虎承认发生大规模数据泄露 2亿账户信息被盗



2016年9月22日 - (原标题 雅虎承认被盗) 9月22日消息,据 投资者们称,雅虎最大痛点。笔者细数了一  
money 163.com/16/0922/

## 微信一度大面积崩溃 回应称光缆因施工被挖断

2013年1月20日 - 发生连接故障 北京、上海、广东、浙江等  
接受消息,也无法刷新朋友圈,微信官方团队称因光缆被市政



# 为什么加一层网关接入

**安全：**数据保护的大门，登录权限校验，IP授权，数据加密，监控上报。

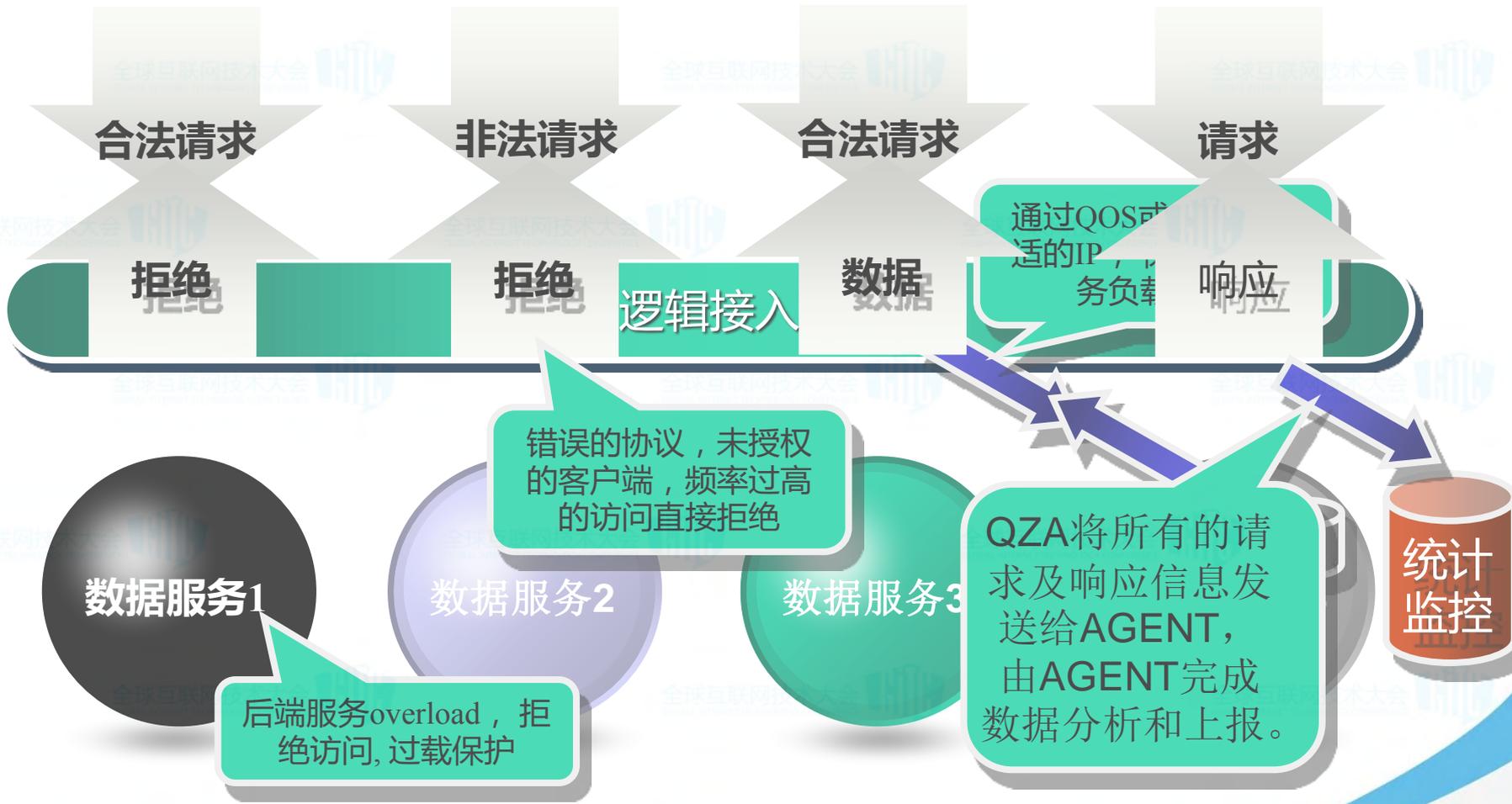
**标准：**团队之间协议统一，降低沟通成本；让业务只专注于逻辑实现。

**容错容灾：**拒绝非法访问，对后端服务过载保护。

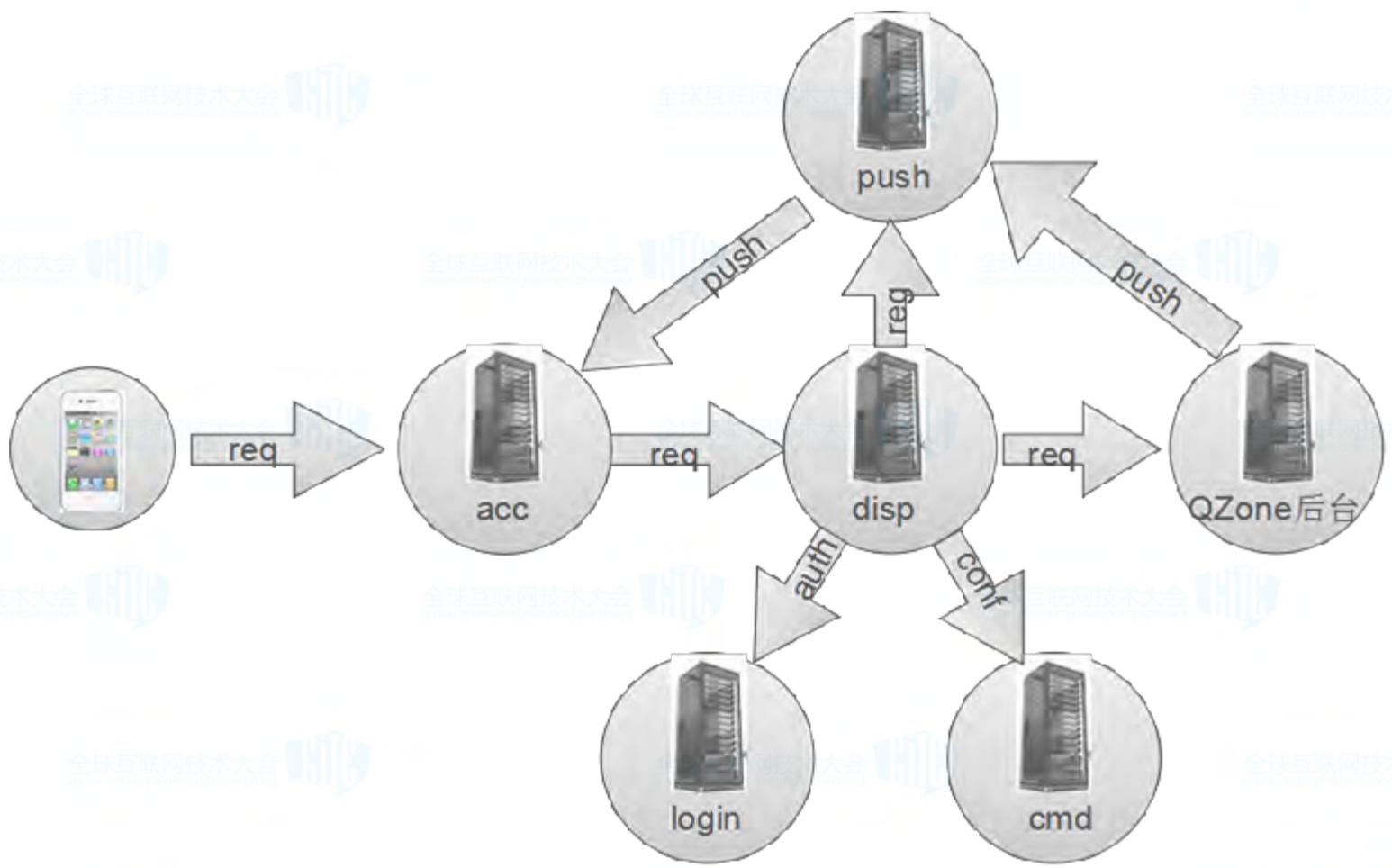
**全网调度：**从网关接入层进行调度/容灾，后端一步到位。

**访问监控：**对访问来源，访问量，服务成功率，服务延时，非法访问等进行有效监控告警。

# 逻辑接入—QZA+JCE



# 用户接入—WNS+JCE



# 目录

## Qzone简介

## 整体架构

## 网关接入

### 接入层介绍

## 异地分布

### SET分布

### 同步中心

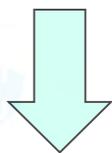
## 系统容灾

### 全网调度

# NO-SET模式

## NO-SET模式:

- 逻辑层: 负载均衡, 读写分离
- 数据层: 分库分表, cache



告警-扩容-告警...

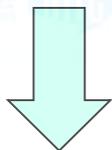
机房瓶颈(流量, 掉电)



# 同城多SET分布

## SET标准

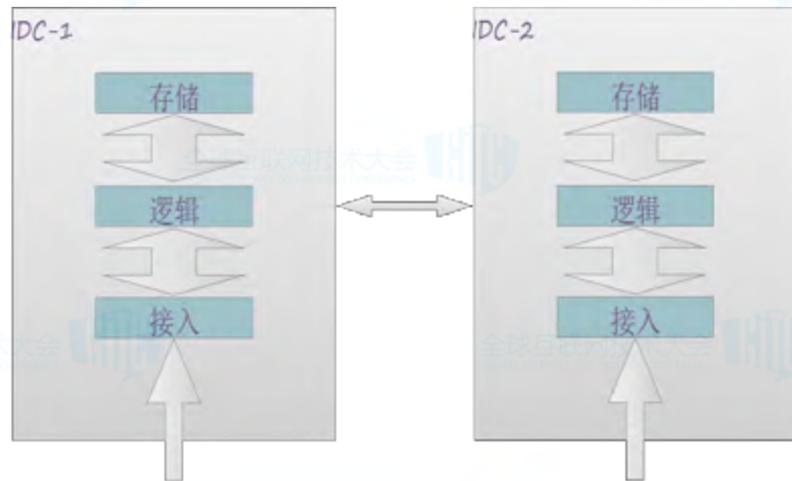
- 同构，功能一样
- 独立，物理隔离
- 大小标准（在线用户1KW）



机器不能满足业务增长需要（农场）

区域出口网络故障

北方的用户使用慢



# 异地多SET分布

## 异地多SET（深圳、上海，天津）

- 缓解机器机架瓶颈
- 就近接入
- 容灾调度（防主干光钎挖断、爆炸、掉电）



# 目录

## Qzone简介

## 整体架构

## 网关接入

### 接入层介绍

## 异地分布

### SET分布

### 同步中心

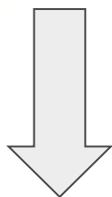
## 系统容灾

### 全网调度

# SET间数据同步

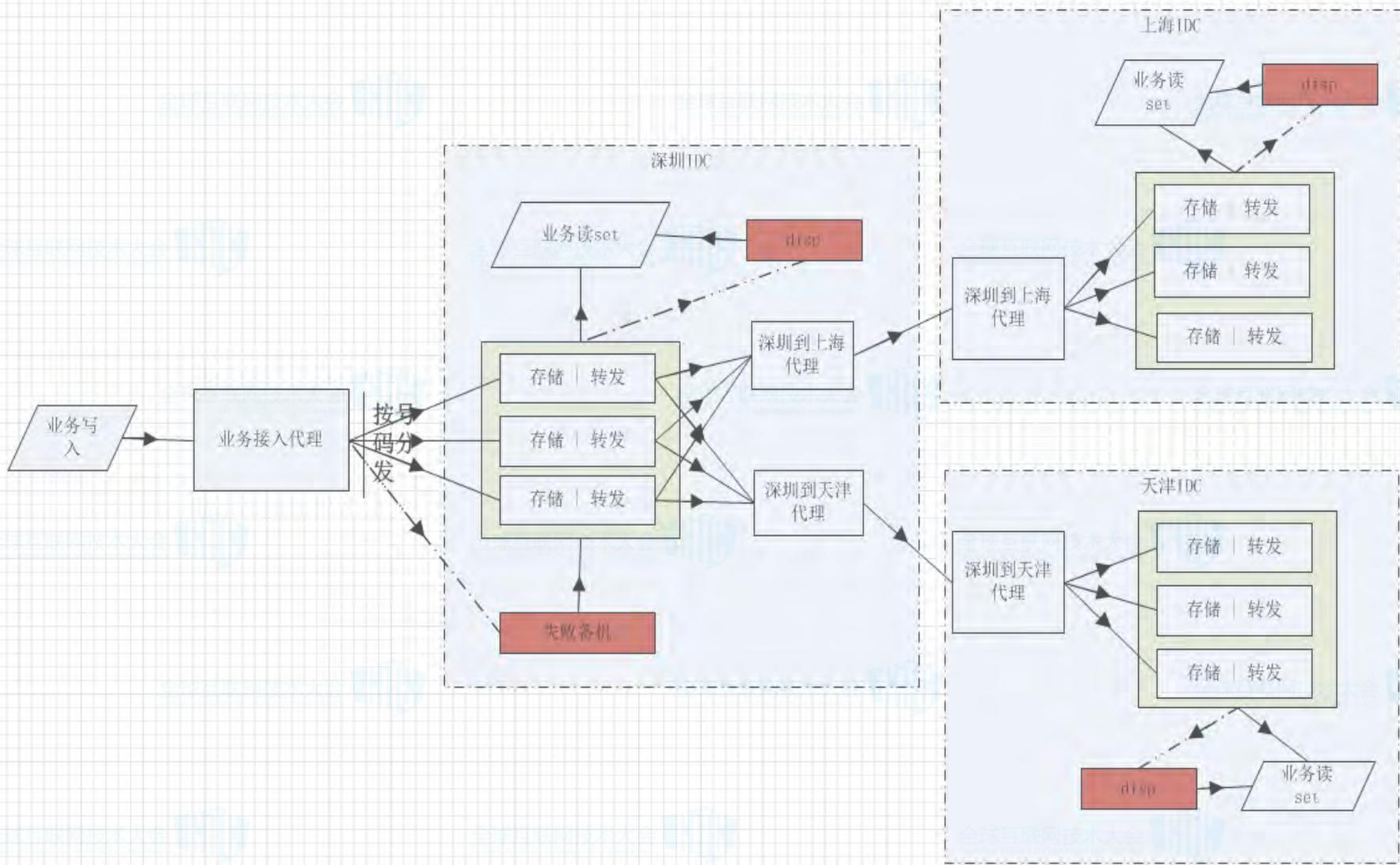
## 几年前的同步方式

- 各业务把流水落本地磁盘，通过agent方式把数据同步各SET内。



重复开发  
维护成本高  
故障容错能力弱  
监控缺乏、混乱

# 同步中心--框架



# 同步中心-接入

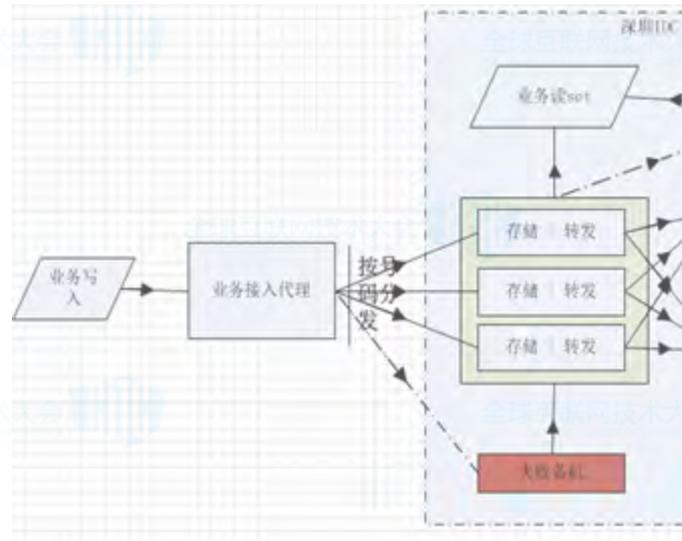
宗旨：对业务透明

存储选型：内存vs硬盘

硬盘：成本低；硬盘受宕机影响小，数据更持久；读取写入延时在us级，在整个数据同步中占比小。

接入方式：本地vs网络

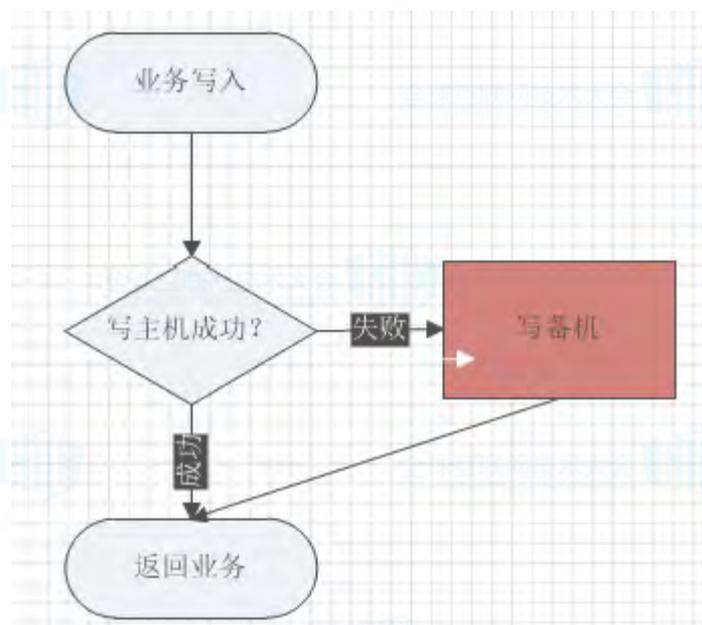
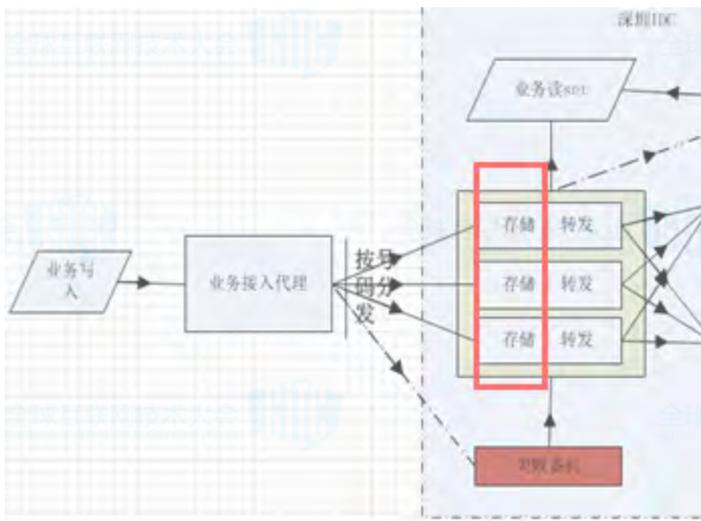
网络：不占用业务机硬盘，不需要装agent；按号段有状态路由到云端，不受磁盘大小影响，流水保存时间长，易于统一维护。



# 同步中心-接入优化

## 问题:

- 接入存储机定期磁盘清理I/O高，导致接入成功率波动。
- 接入存储机故障/宕机，如何容错。



# 同步中心-分发

## 序列号:

不需要: 对业务透明, 业务数据存储不需要额外支持序列号。

## 通知类型: 内容vs信号

内容: 同步每次操作请求, 业务主/备之间完全同构;

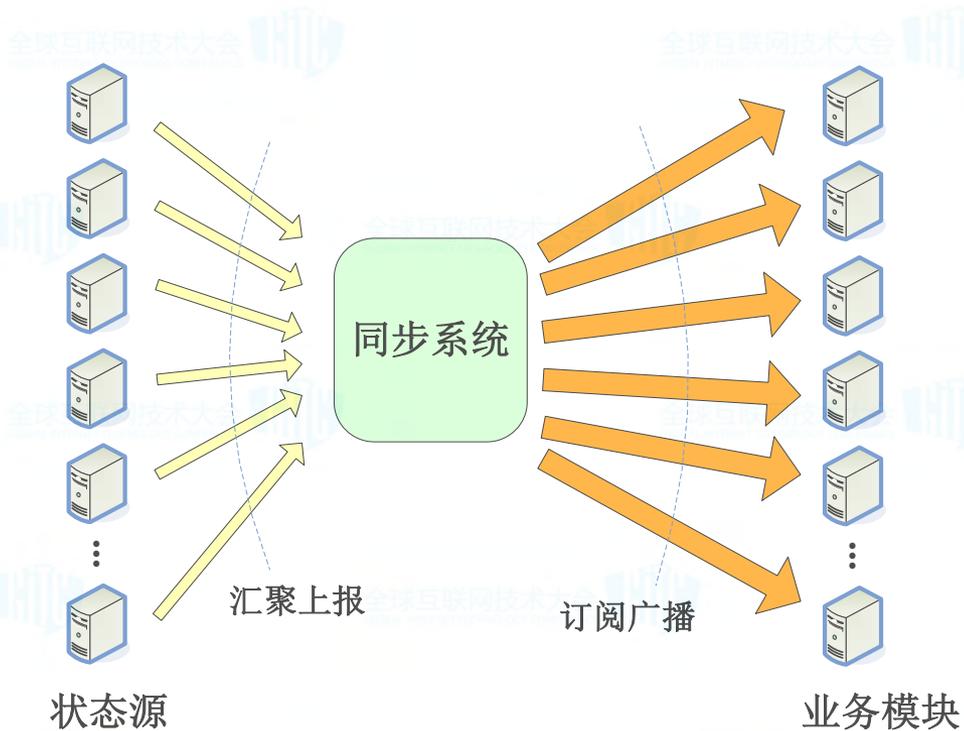
## 分发方式:

同步时序: 保证时序, 同一个用户只会在同一个进程中同步。

# 同步中心-分发优化

## 数据分发瓶颈

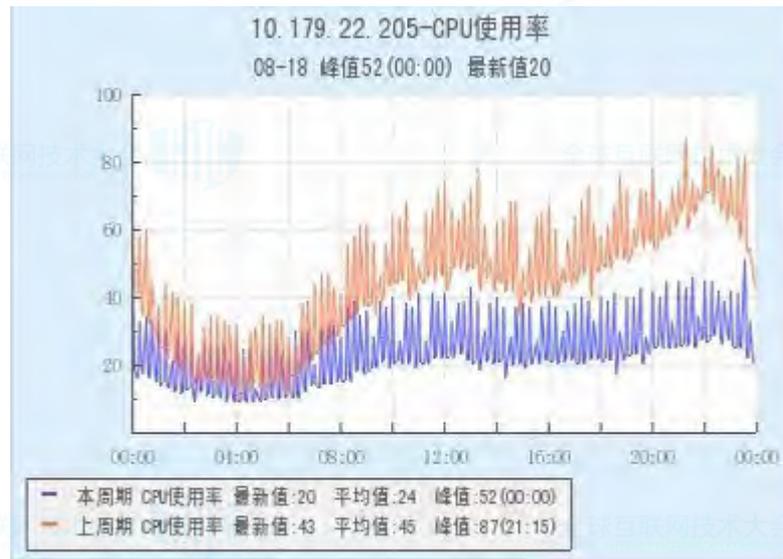
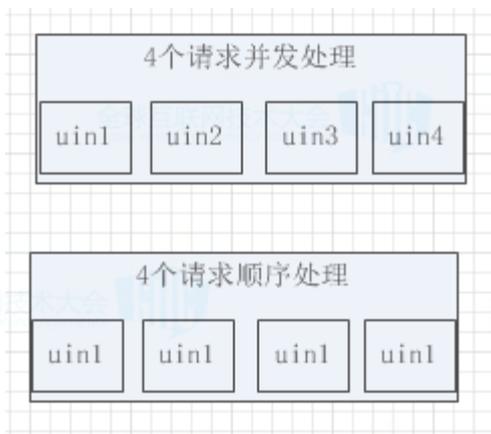
- 需大量进程，耗机器CPU
- 受被同步服务延时影响
- 大量磁盘I/O操作



# 同步中心-分发优化

## 方案

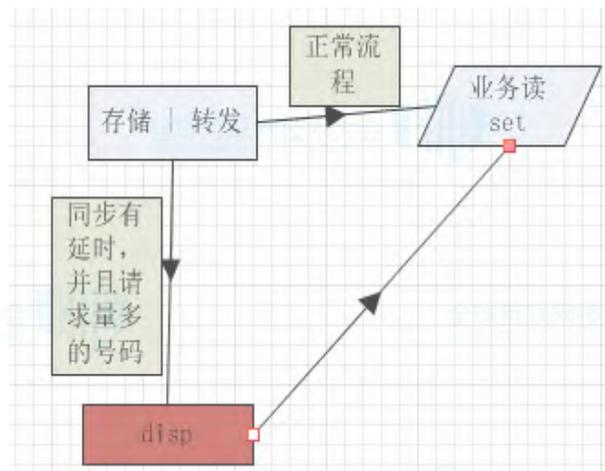
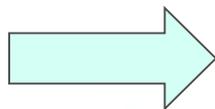
半异步：不同用户并发，同一用户顺序



# 同步中心-分发优化

## 号码聚集，延时增加

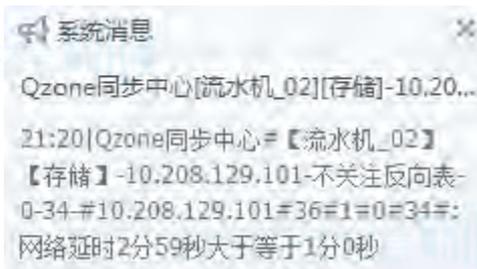
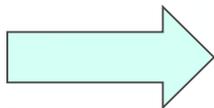
- 微云备份
- 相册删除



# 同步实时性和一致性监控

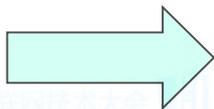
## 延时监控上报

- 数据源写入时间和同步出去时间差值上报。



## 数据一致性对账

- 全量对账
- 增量对账
- 抽样对账
- 数据修复



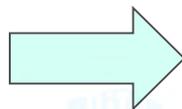
| 子应用名称        | 总成功率    | 检测总   | 错误数 | 超时数 | 成功数    |
|--------------|---------|-------|-----|-----|--------|
| (1-13):qzone | 100%    | 1150  | 0   | 0   | 1150   |
| (30-246):ugc | 100.00% | 5E+06 | 238 | 0   | 5E+06  |
| (30-240):手   | 100.00% | 4E+05 | 11  | 0   | 365451 |
| (30-186):权   | 100.00% | 68008 | 1   | 0   | 68007  |
| (30-217):用   | 100.00% | 2E+07 | 119 | 0   | 2E+07  |
| (30-235):ugc | 100%    | 4E+05 | 2   | 0   | 422446 |
| (30-238):ugc | 100%    | 2E+06 | 0   | 0   | 2E+06  |
| (30-236):ugc | 100%    | 5E+06 | 0   | 0   | 5E+06  |
| (31-14):权限   | 99.99%  | 5E+06 | 360 | 0   | 5E+06  |

# 目录

- Qzone简介
- 整体架构
- 网关接入
  - 接入层介绍
- 异地分布
  - SET分布
  - 同步中心
- 系统容灾
  - 全网调度

# 全网调度--这些年全网调度遇到的坑

- ❶ 链路上有短板，切过去，撑不住
- ❷ 容量不足，切过去影响更大
- ❸ 切换起来没底气，临时找各模块负责人评估，耗时长
- ❹ 异地切换过来的用户cache命中率，透传到后端请求量增加几十倍，后端撑不住



# 全网调度—艰难调度

I: 压测: 如纸上谈兵

覆盖单个系统, 或者一条线, 很难覆盖到一个面到

演习: 实践出真知

定期SET间通过接入层进行调度, 单SET用户请求真实翻倍



# 全网调度—柔性调度

II: 8.12天津爆炸，Qzone天津数据中心随时有断电危险，访问切到深圳和上海；cache失效，切换后需开启柔性支持。

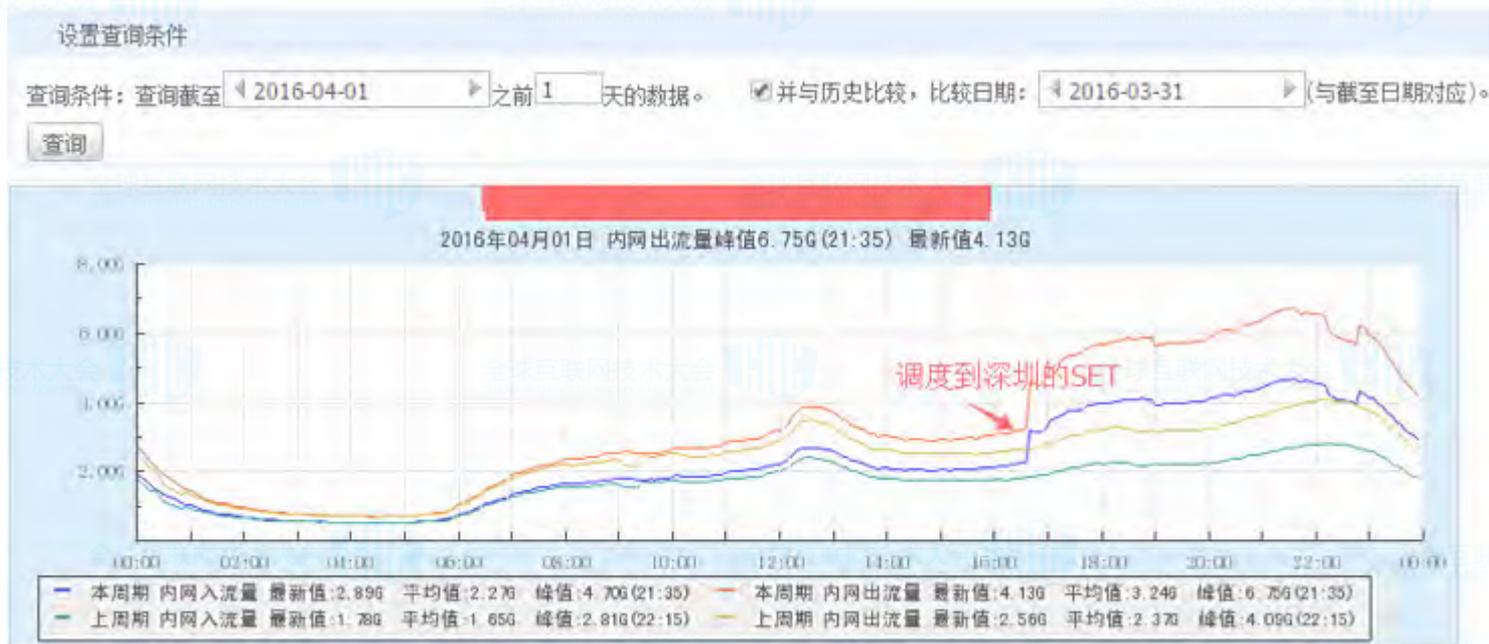
## Cache数据三地同步

定期异地SET间通过接入层进行调度，单SET内新用户请求真实翻倍



# 全网调度--完美调度

III: 愚人节天津数据中心掉电, 故障发生→发现→决策→实施  
总耗时不到1小时



# END

全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



全球互联网技术大会  
GLOBAL INTERNET TECHNOLOGY CONFERENCE



# Thanks