

SDCC 2016

中国软件开发者大会

SOFTWARE DEVELOPER CONFERENCE CHINA

QQ空间后台架构优化之路

嘉宾：王帅

讲师简介

- 讲师：王帅
- 部门：腾讯SNG社交平台部
- 经历：武汉大学计算机专业研究生毕业；曾在通讯行业从事了几年后台开发工作；2012年加入腾讯社交平台部，后台高级工程师，主要负责QQ空间后台的相关架构与优化工作。

目录

- 一. 空间简介
- 二. 系统架构
- 三. 系统优化
 - 动态拉取优化
 - 提速与省流量
- 四. 容灾调度
 - SET部署
 - 数据同步
- 五. 柔性可用
 - 柔性含义
 - 实践

空间主场景：好友动态、我的空间和与我相关





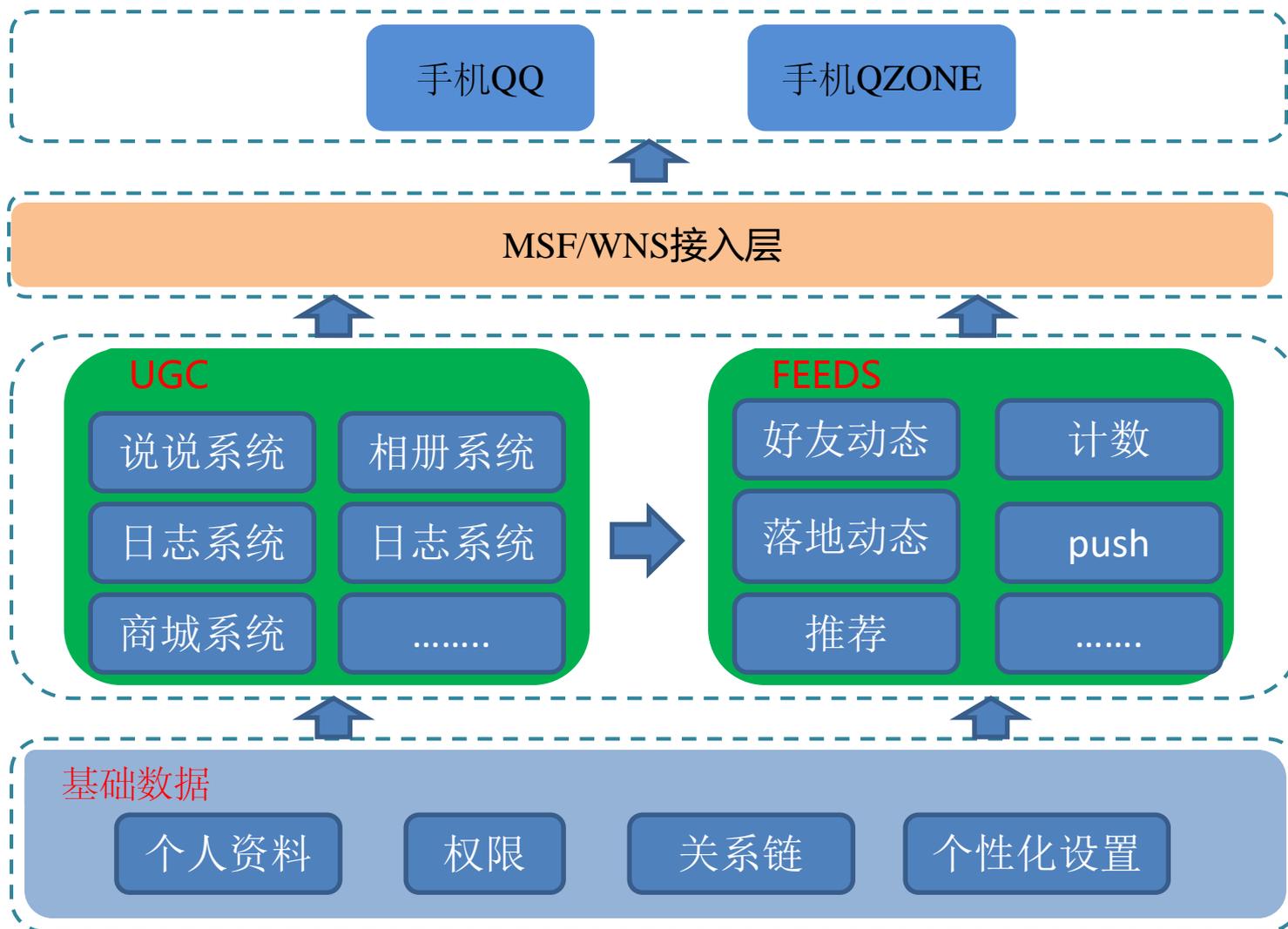
空间的量有多少？

-  月活跃用户超6亿，日活跃用户超3亿
-  UGC日发表量超2亿，日互动量超7亿。
-  照片上传单日峰值超7亿，照片总存储超 200p
-  动态拉取峰值超10W/s，空间首屏日打开量超30亿

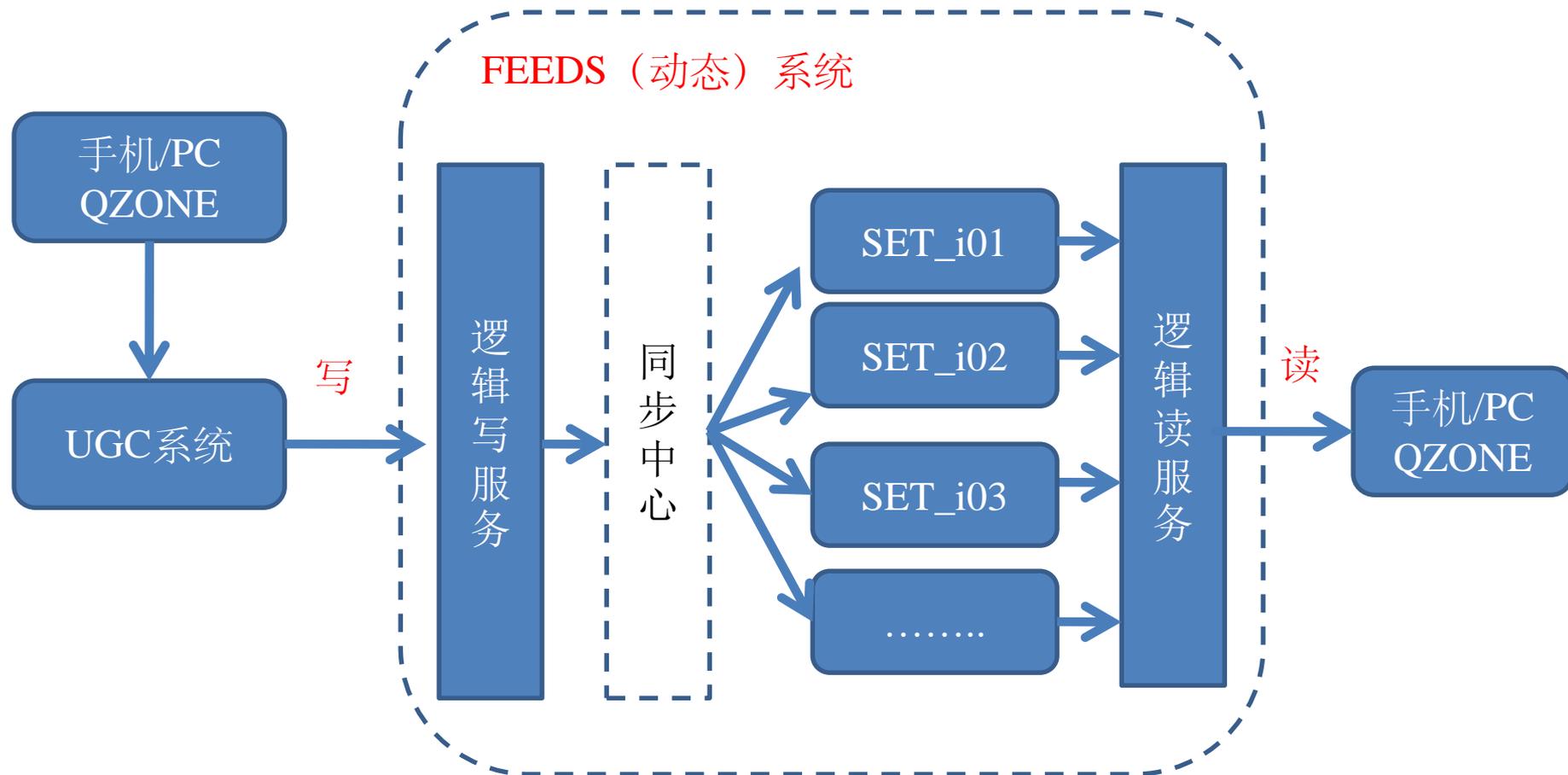
目录

- 一. 空间简介
- 二. 系统架构
- 三. 系统优化
 - 动态拉取优化
 - 提速与省流量
- 四. 容灾调度
 - SET部署
 - 数据同步
- 五. 柔性可用
 - 柔性含义
 - 实践

主要功能模块



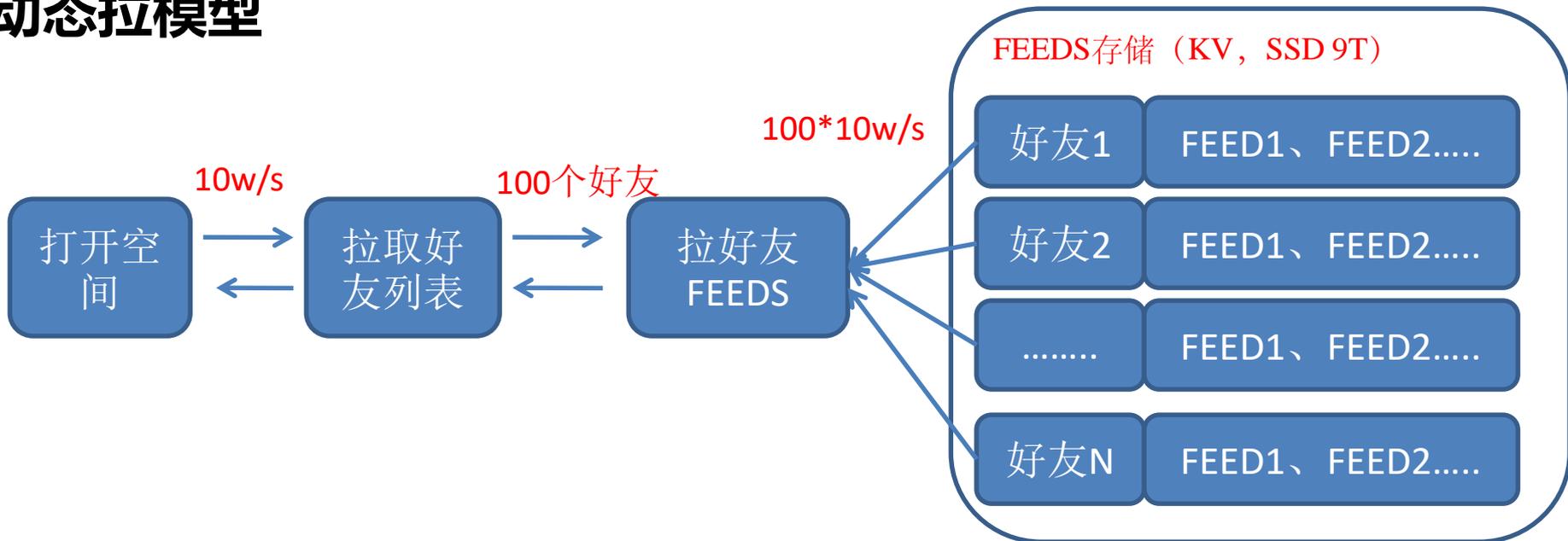
FEEDS系统



目录

- 一. 空间简介
- 二. 系统架构
- 三. 系统优化
 - 动态拉取优化
 - 提速与省流量
- 四. 容灾调度
 - SET部署
 - 数据同步
- 五. 柔性可用
 - 柔性含义
 - 实践

动态拉模型



假设每个用户平均有100个好友，每人存20条，则：

访问量： $100 * 10w / s = 1000w/s$

流量： $100 * 10w/s * 20条/人 * 1kB/条 = 1.6Tbps$

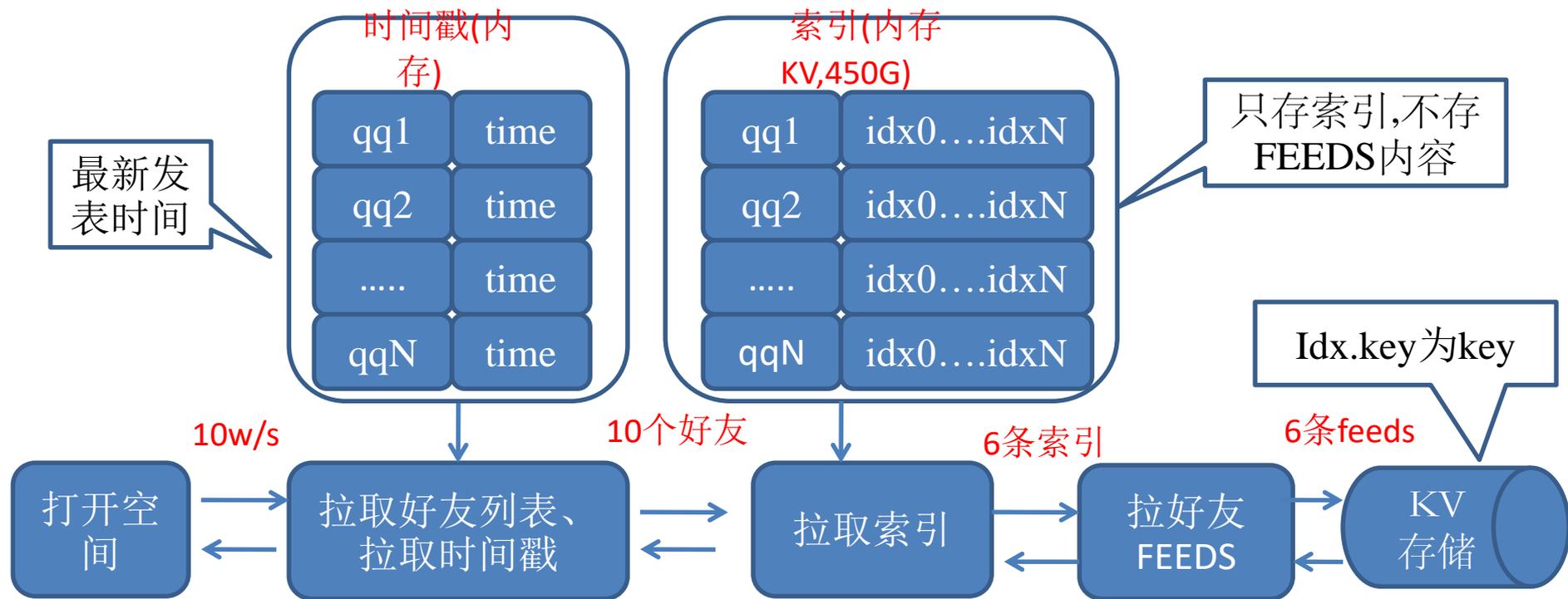


业务特点：好友动态流按时间排序

如何减轻对后端的压力、省流量？



持续优化后



现按照时间戳排序，再按照索引排序，分屏拉取（6条一屏），则

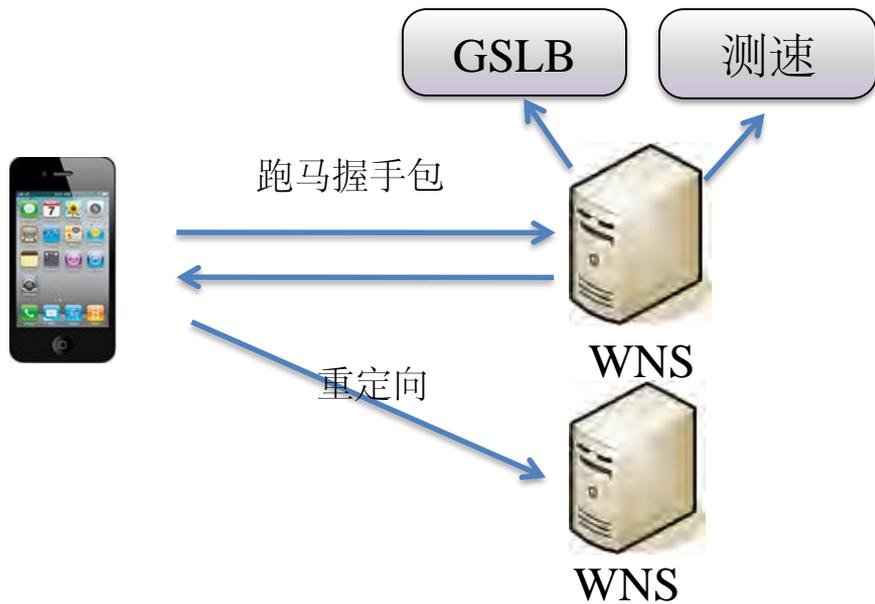
访问量： $10w/s * 6 = 60w/s$

流量： $10w/s * 6条 * 1KB/条 = 4.9G$

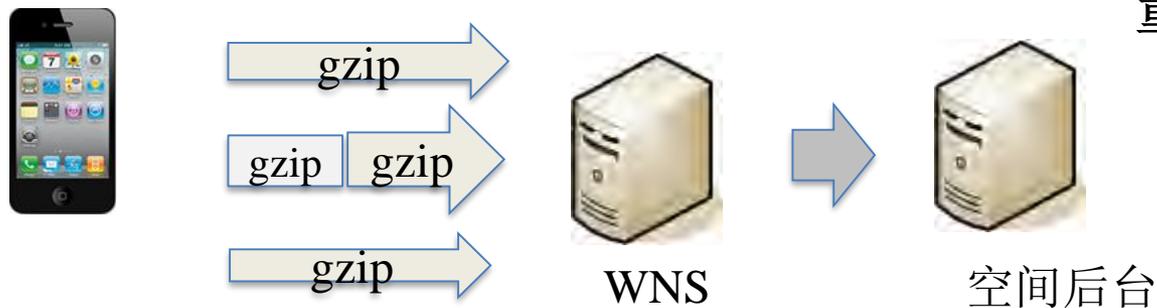
提速与省流量

WNS接入优化

最优接入:

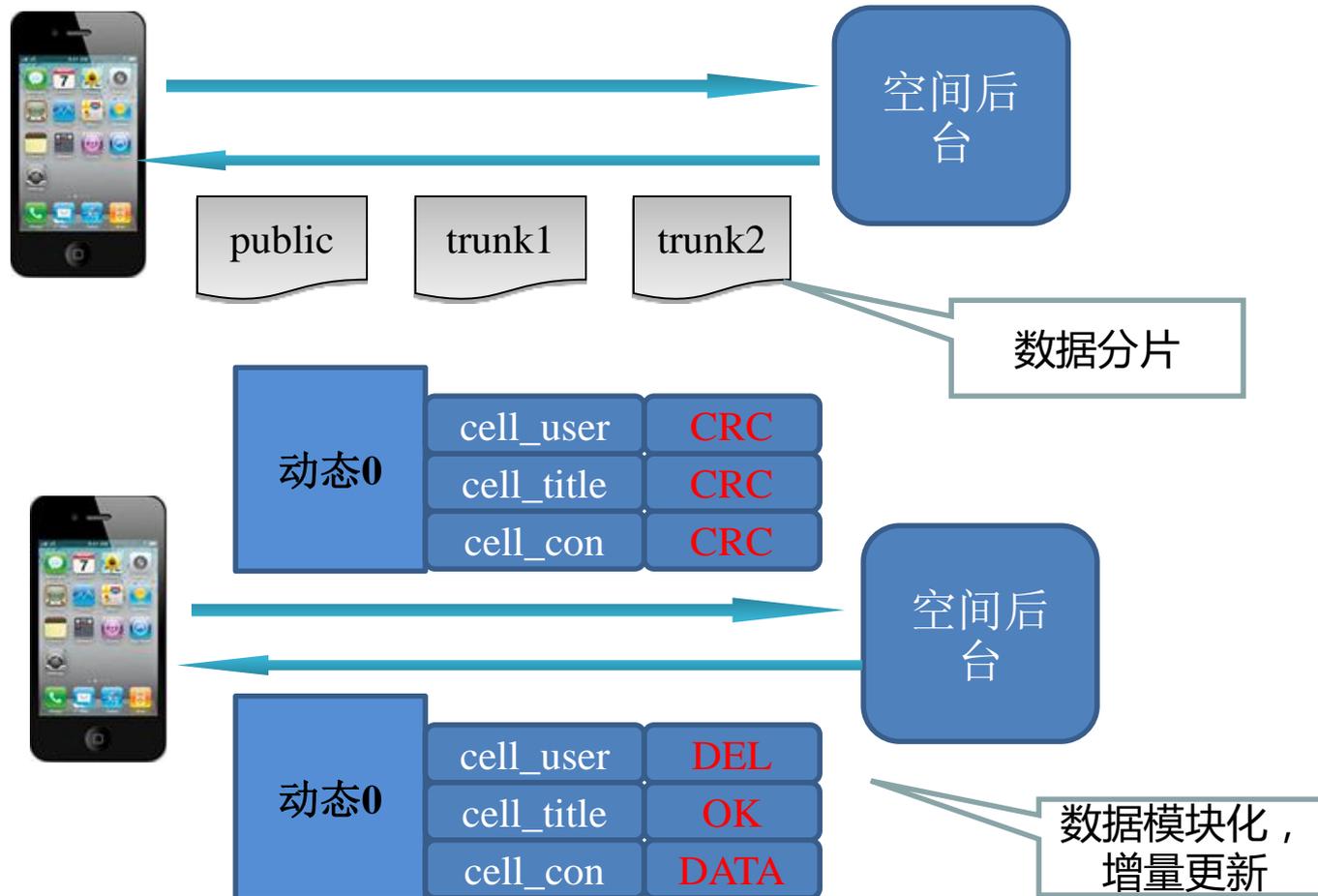


最小流量:



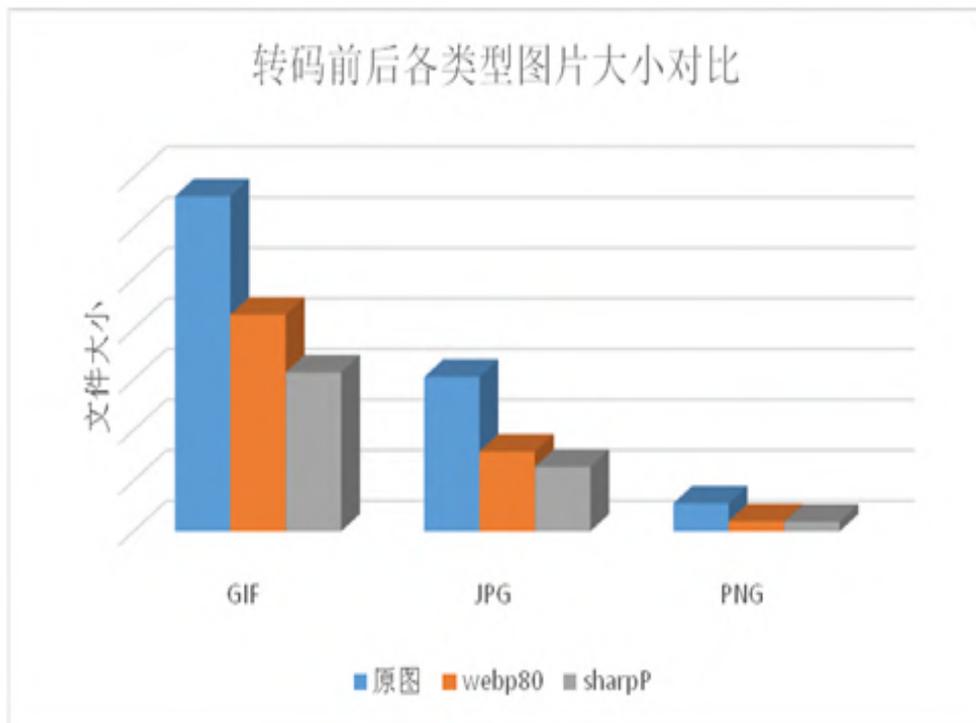
提速与省流量

后台业务优化：



提速与省流量

图片优化：适配服务、sharpP



场景	像素	图档	
浮层	$X \geq 720$	大图	中图
	$480 = < X < 720$	大图	中图
	$X < 480$		小图
Feeds单图	$X \geq 720$	中图	中图
	$480 = < X < 720$		小图
	$X < 480$		
详情	$X \geq 720$	中图	中图
	$480 = < X < 720$		小图
	$X < 480$		
Feeds多图	不区分	小图	

目录

- 一. 空间简介
- 二. 系统架构
- 三. 系统优化
 - 动态拉取优化
 - 提速与省流量
- 四. 容灾调度
 - SET部署
 - 数据同步
- 五. 柔性可用
 - 柔性含义
 - 实践

容灾调度

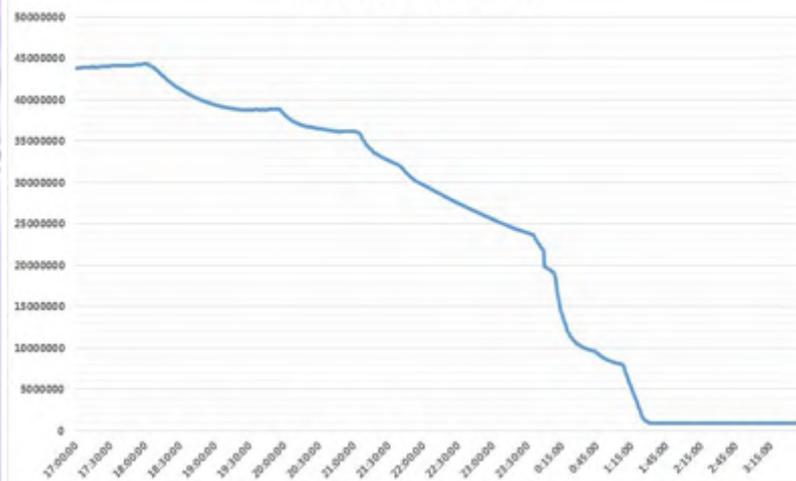
3:34:33

天津塘沽爆炸事件，造成腾讯天津数据备份中心，共62台服务器 5台交换机受影响（H12列机架掉电包含一台外网、内网、管理网交换，G13列一台外网交换机，C10一台内网交换机（无业务））。

目前因天津机房附近爆炸产生的故障已经初步恢复，目前正在密切关注事态发展。由于很可能是危险品爆炸，政府已对相关区域进行紧急管控，极端情况下可能会出现人员需要撤离等情况，机房可维护性会降低。所以业务需提前做好准备。对具备容灾能力的重点业务提前迁移。



天津手Q在线(13-14号)



各位同事：

2016年10月25日17时47分，深圳滨海到深圳同乐一条波分主干光线路中断（主干光名称：深圳滨海0701-R08B至深圳同乐0401-0302B-FIBER-01，专线编号：滨海腾讯-同乐02），网络架构冗余，线路切换期间导致深圳内网短暂波动，网络平台已经联系运营商紧急处理中，目前故障原因不明，恢复时间待定。

容灾调度成为必须！

容灾调度

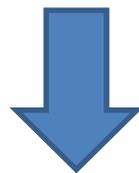
全网调度的痛点:

链路有短板，撑不住

容量不足，撑不住

没底气，评估慢

Cache命中下降



解决方案:

SET标准化

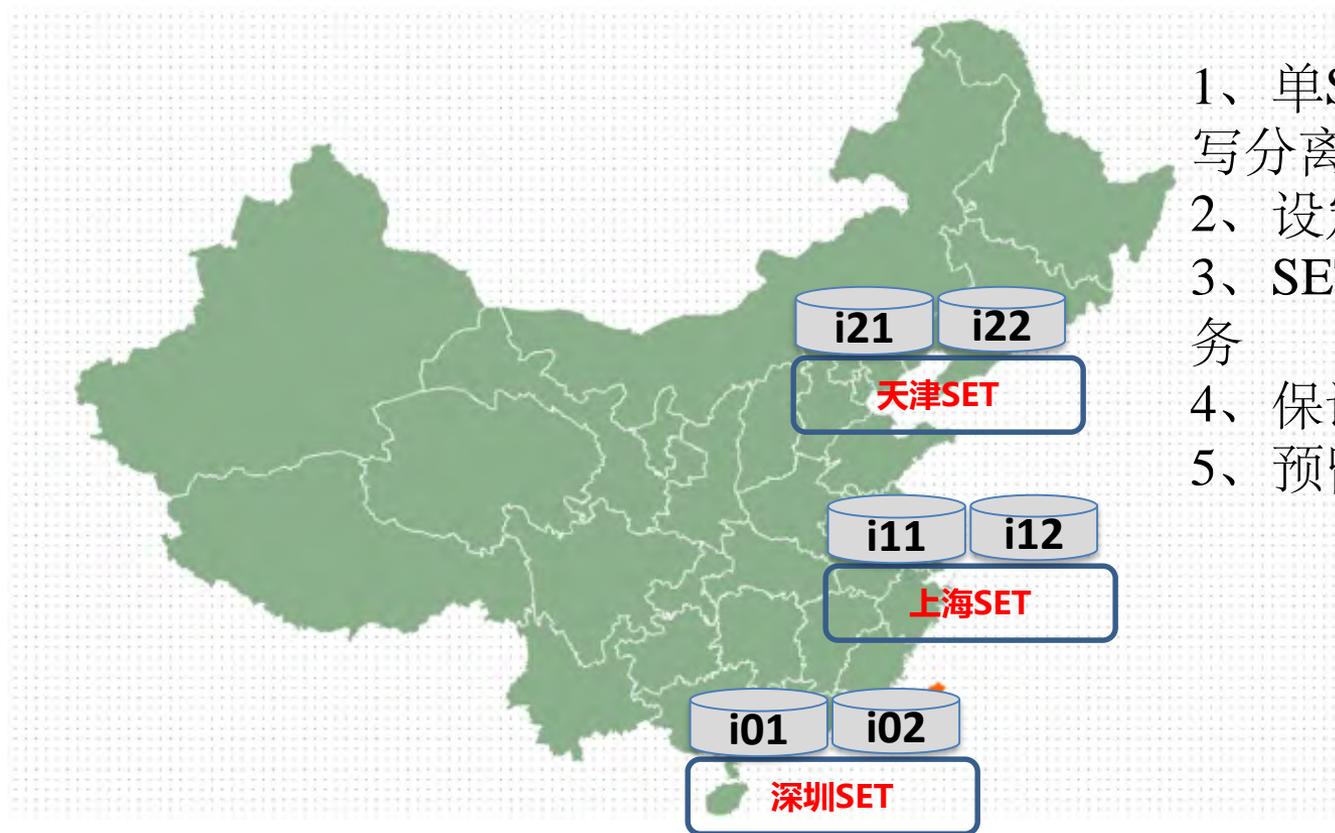
数据同步

降级服务

定期演习

容灾调度

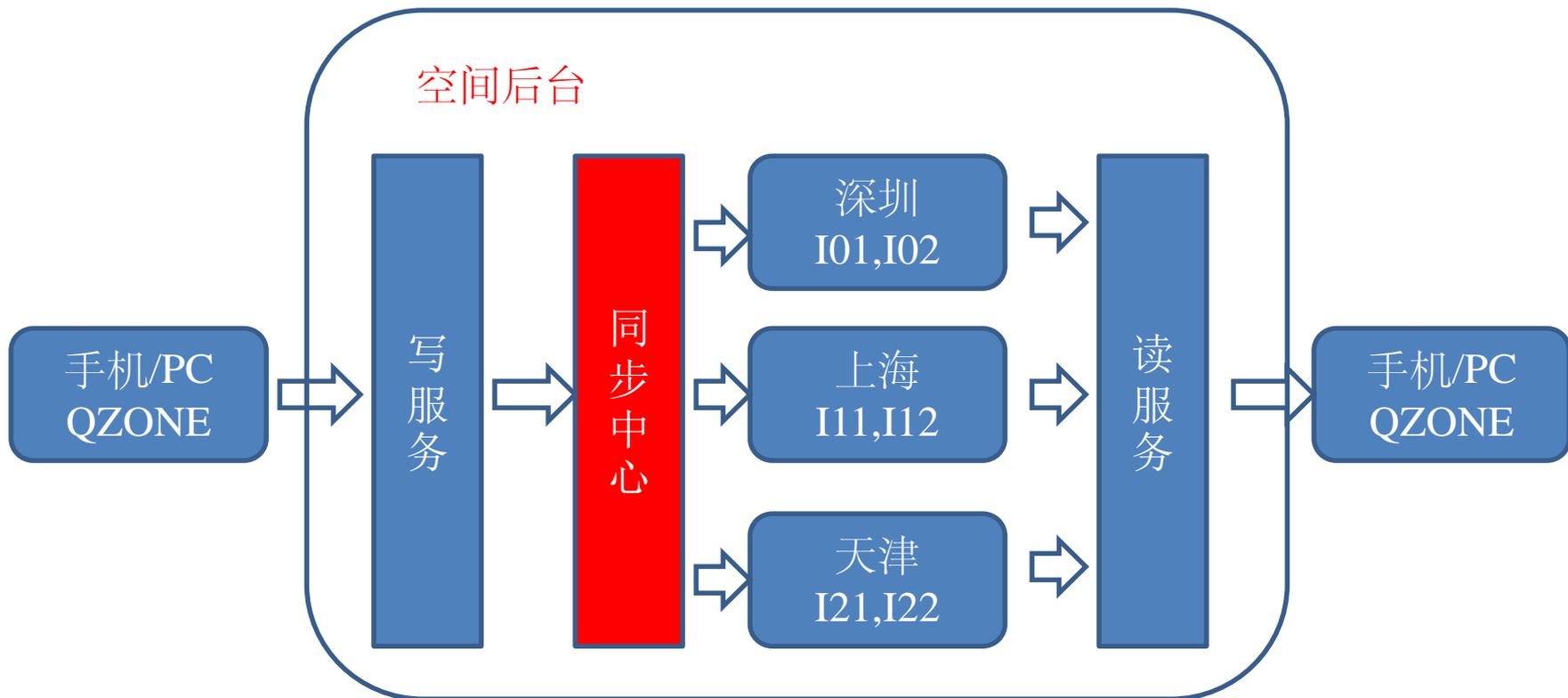
SET部署



- 1、单SET: 负载均衡、读写分离、cache、分库分表
- 2、设定SET大小标准
- 3、SET物理隔离，独立服务
- 4、保证各SET是同构的
- 5、预留buffer

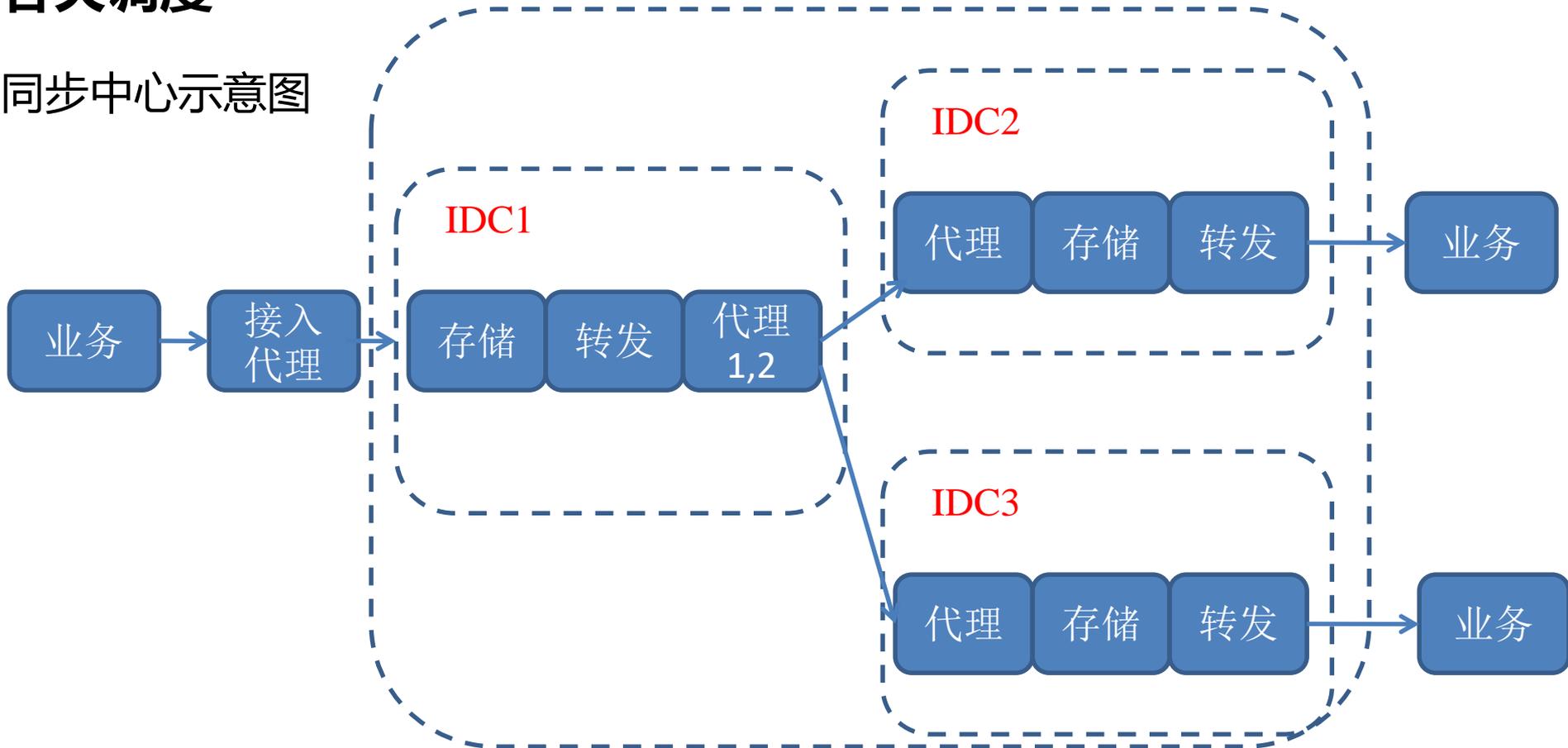
容灾调度

SET同步，存储/cache一致



容灾调度

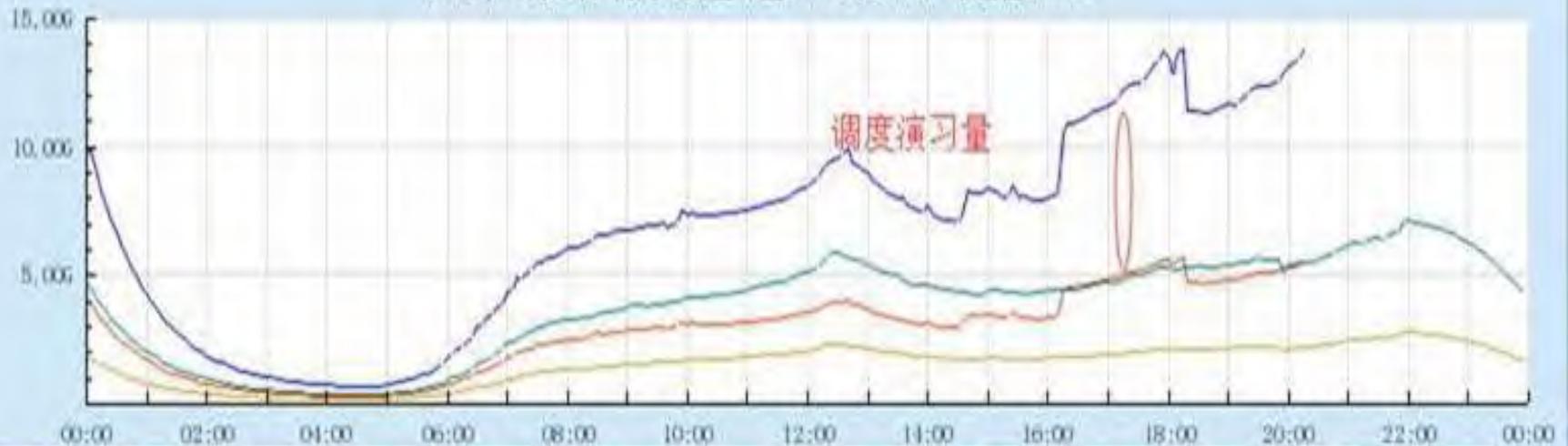
同步中心示意图



容灾调度

定期演习压测，做到心里有底

2015年08月20日 内网出流量峰值5.73G(18:15) 最新值5.64G



— 本周期 内网入流量 最新值:13.71G 平均值:6.68G 峰值:13.88G(18:15)	— 本周期 内网出流量 最新值:5.64G 平均值:2.77G 峰值:5.73G(18:15)
— 上周期 内网入流量 最新值:4.33G 平均值:3.82G 峰值:7.15G(22:00)	— 上周期 内网出流量 最新值:1.69G 平均值:1.54G 峰值:2.83G(22:00)

目录

- 一. 空间简介
- 二. 系统架构
- 三. 系统优化
 - 动态拉取优化
 - 提速与省流量
- 四. 容灾调度
 - SET部署
 - 数据同步
- 五. 柔性可用
 - 柔性含义
 - 实践

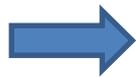
柔性可用—什么是柔性可用？

●柔性可用

在资源有限的情况下，不要刚性的依赖某个环节或者特性，提供最合适的服务。

●关键点

核心诉求、局部放弃、降级服务



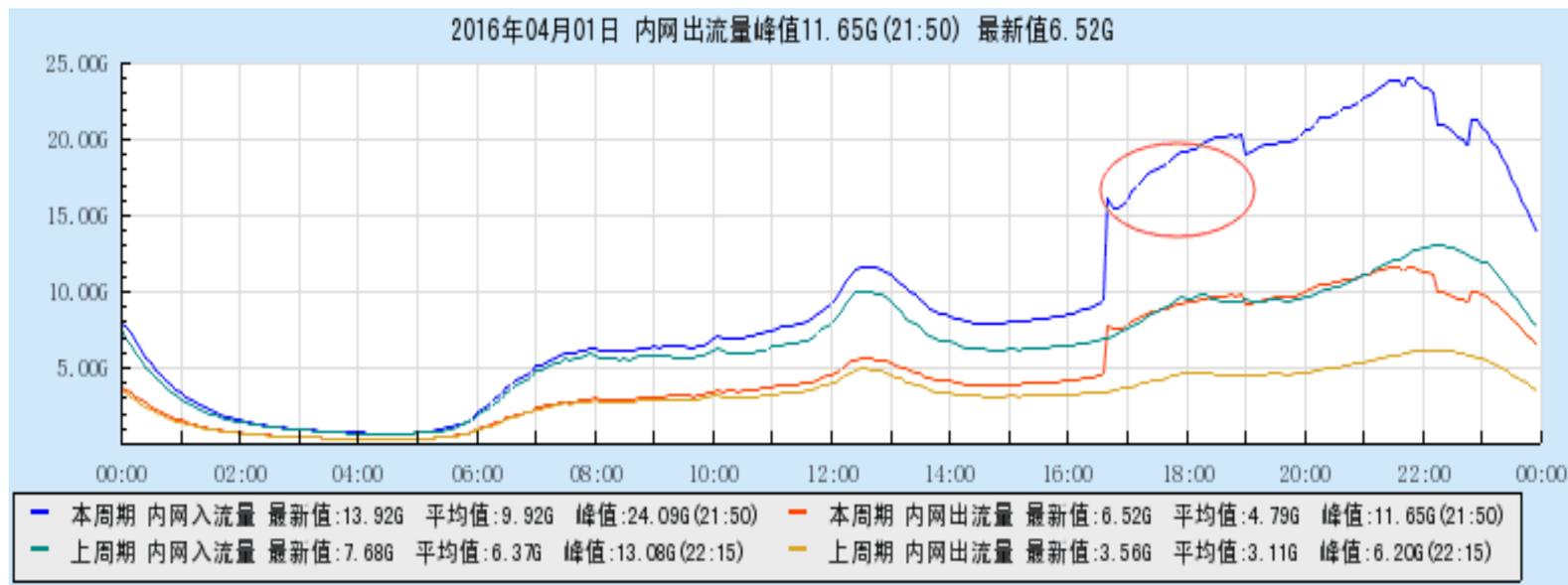
柔性可用---使用场景

元旦0点，动态的读写量，比平时陡增三倍多，容量不够，撑不住？



柔性可用---使用场景

故障恢复期间，临时调度，资源有限，撑不住？



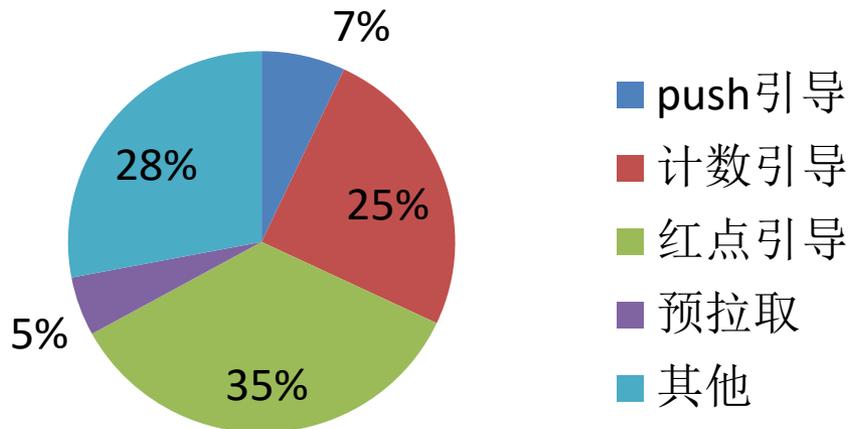
柔性可用—实践

●请求量：

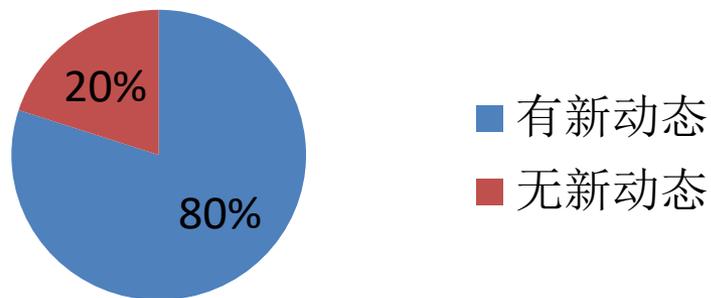
平时首屏打开量6.5W/s，UGC写1.5W/s，UGC互动5.5W/s；节假日峰值涨到3倍。

●请求特征

读请求来源占比



拉到新动态占比



反复刷频率



柔性可用—实践

●动态读柔性

体验降级:

预拉取
关闭

10s重复刷返
回304

关闭特定号
段push

关掉特点号
段红点计数

技术侧调优:

关闭非关
键路径

适当调小
超时

合理的重
试

效果: 首屏打开量18w/s→9w/s, 达到削峰效果

体验有损:

- 1, 10s的好友动态变更无感知
- 2, Feeds变更不实时
- 3, PUSH、计数推送不及时

柔性可用—实践

●动态写柔性

体验降级:

削量 (被动引入的量)

延迟写 (互动性弱操作)

合并写 (时间聚集操作)

技术侧调优:

关闭非关键路径

数据层 cache

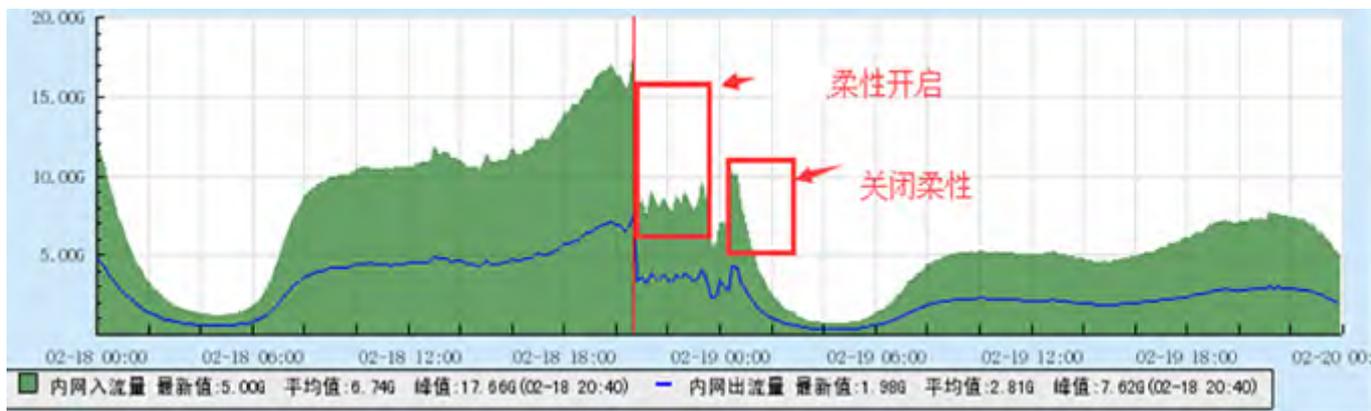
效果: UGC互动从14.5w/s→8w/s, 达到削峰效果

体验有损:

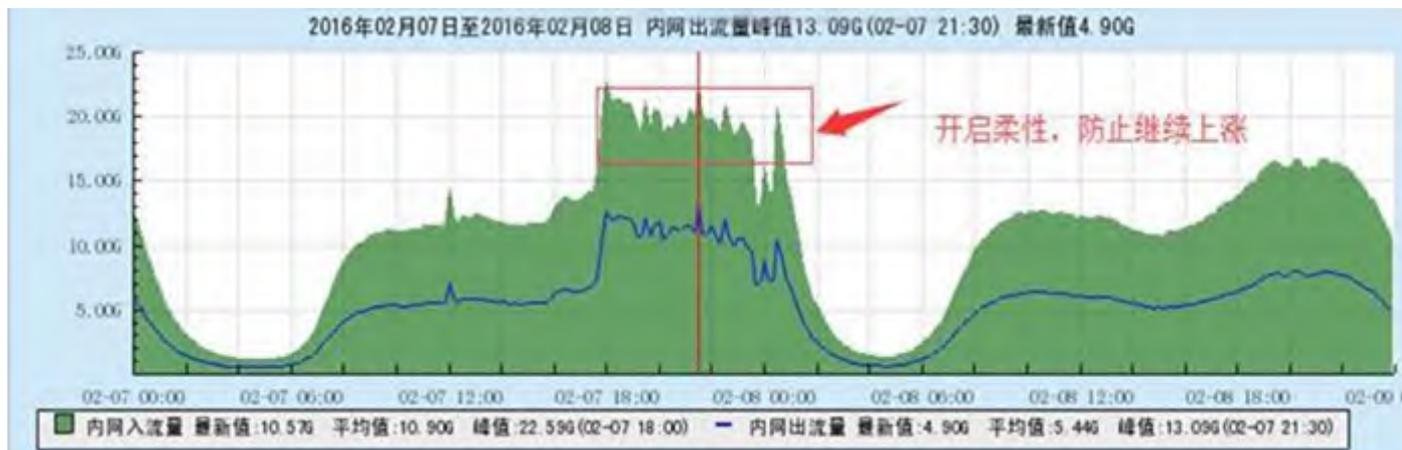
- 1, 赞等互动通知会合并通知
- 2, 个人档、UGC详情内容不实时
- 3, PUSH、计数推送不及时

柔性可用—实践

效果：



柔性演习



春节开启柔性



SDCC 2016

中国软件开发者大会

SOFTWARE DEVELOPER CONFERENCE CHINA

谢谢！