

手机淘宝无线网关的演进优化之路与高可用性实践

阿里巴巴—无线事业部

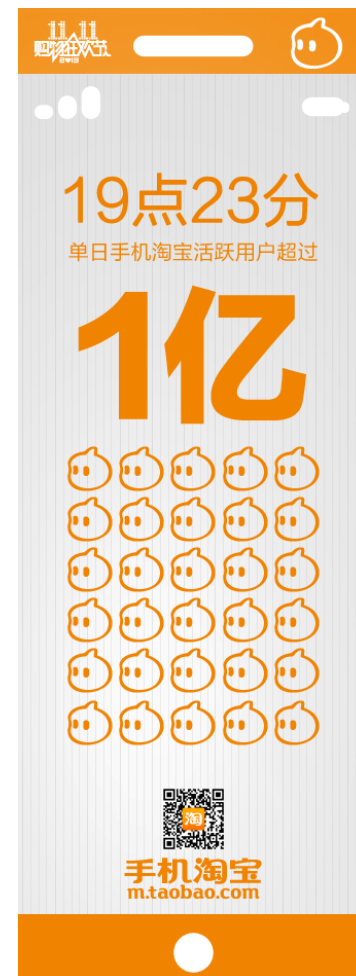
洪海(孤星)

weibo:孤星_T



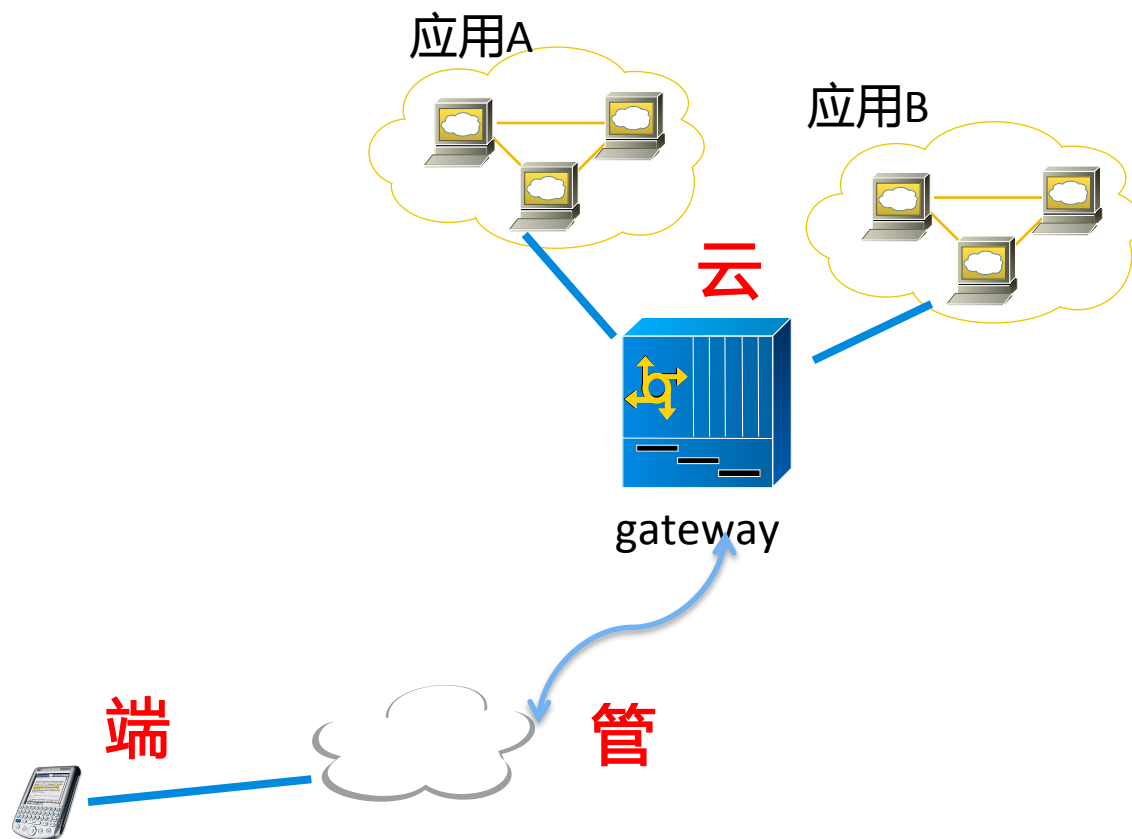


手机淘宝2013双11





手机淘宝无线网关





网关 —— 云与端的桥梁

- 统一的接入形式，为客户端开发提速
- 服务端快速向客户端输出业务能力
- 集中的API治理与基础设施能力

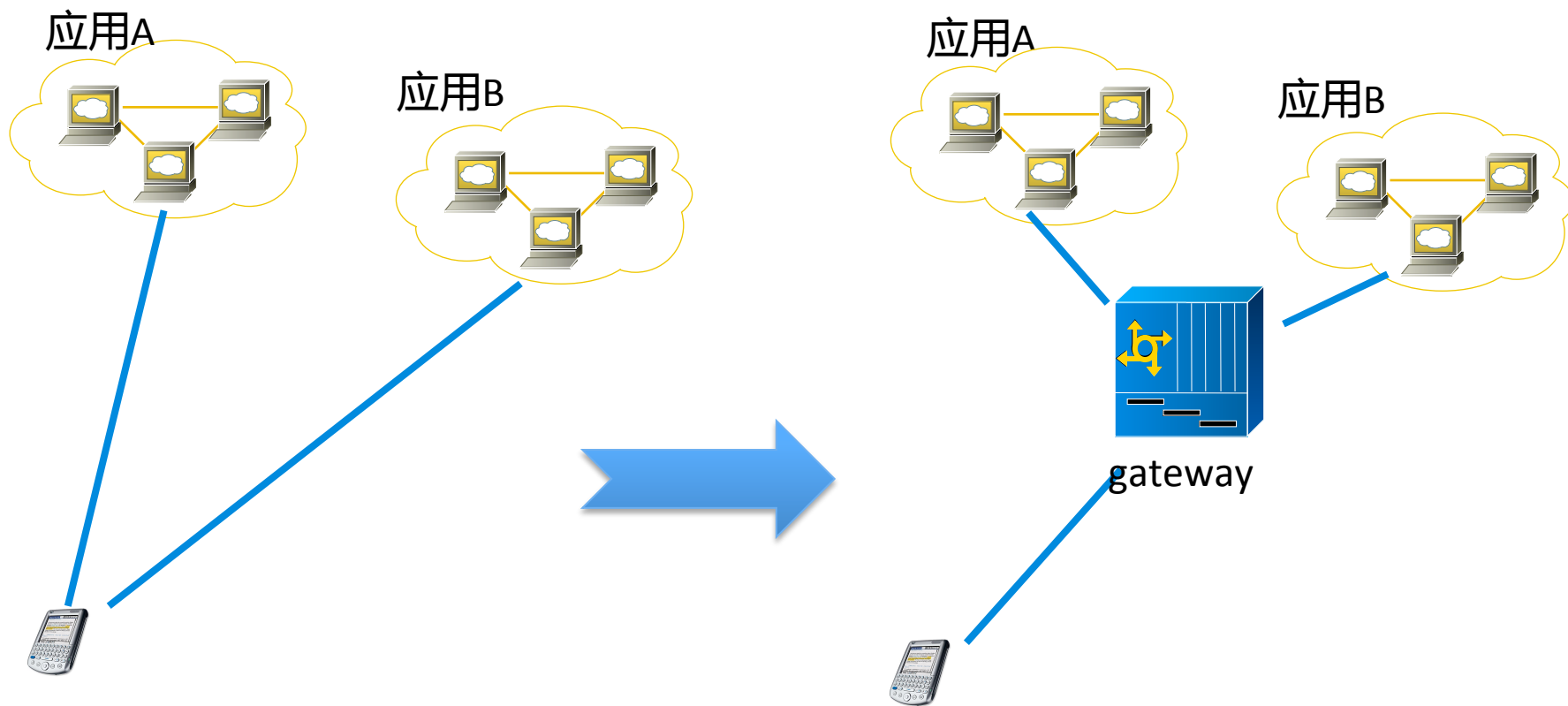


网关的演化—统一接入





统一接入





统一接入

- 统一的协议与标准化的SDK
- 集中的会话管理
- 集中的安全策略
- 统一的监控告警与报表
- 通用缓存服务
- 流控与过载保护
- 网络优化

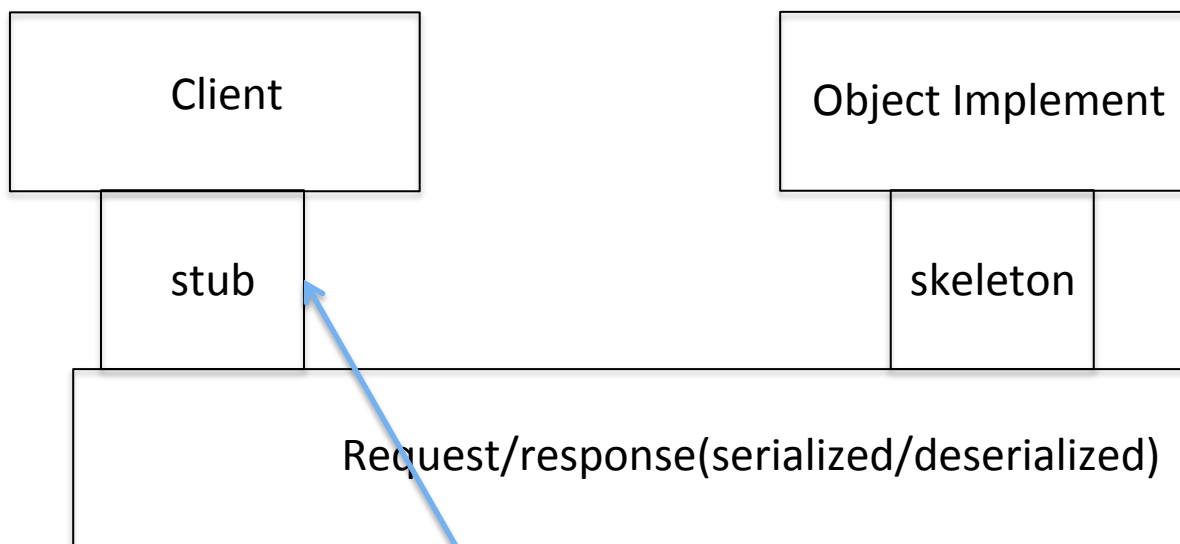


网关的演化—api动态发布





API动态发布 ---传统rpc的调用

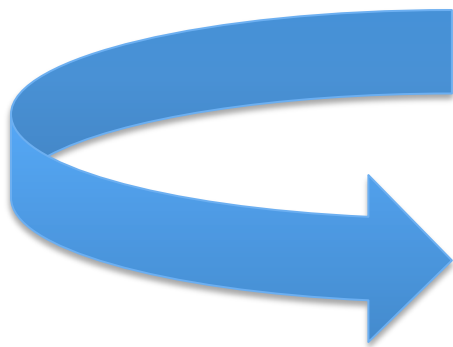
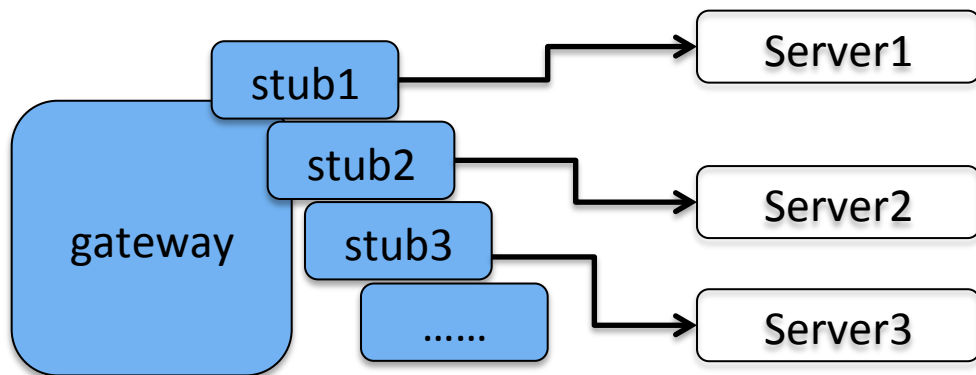


- 对服务端的类型的直接依赖或间接依赖



API动态发布

----传统RPC接入API方式的困境



1. 依赖复杂

大量的接口包，打包越来越复杂

2. 多分支与发布频率

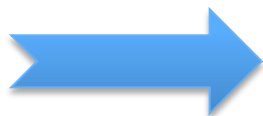
同时存在的大量的发布分支，网关频繁发布，稳定性受到挑战



Map结构的启发

Map:

```
{  
  "uid": 123,  
  "name": "张三",  
  "gender": "male"  
}
```



com.taobao.User:

```
{  
  "uid": 123,  
  "name": "张三",  
  "gender": "male"  
}
```

Map的Key--→ 对象属性

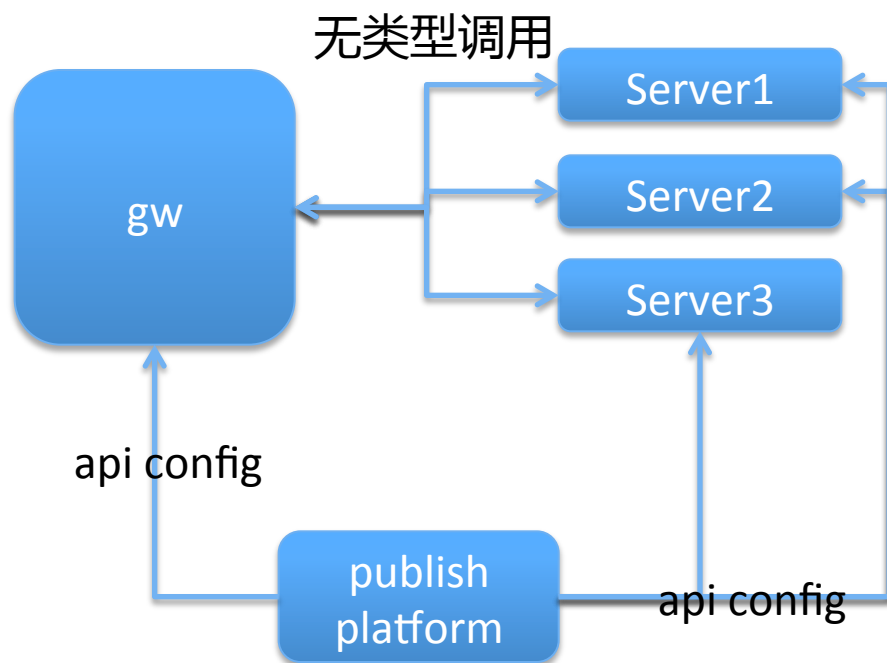


API接入动态化

- 无类型调用代替stub的依赖
- 映射配置文件的动态发布

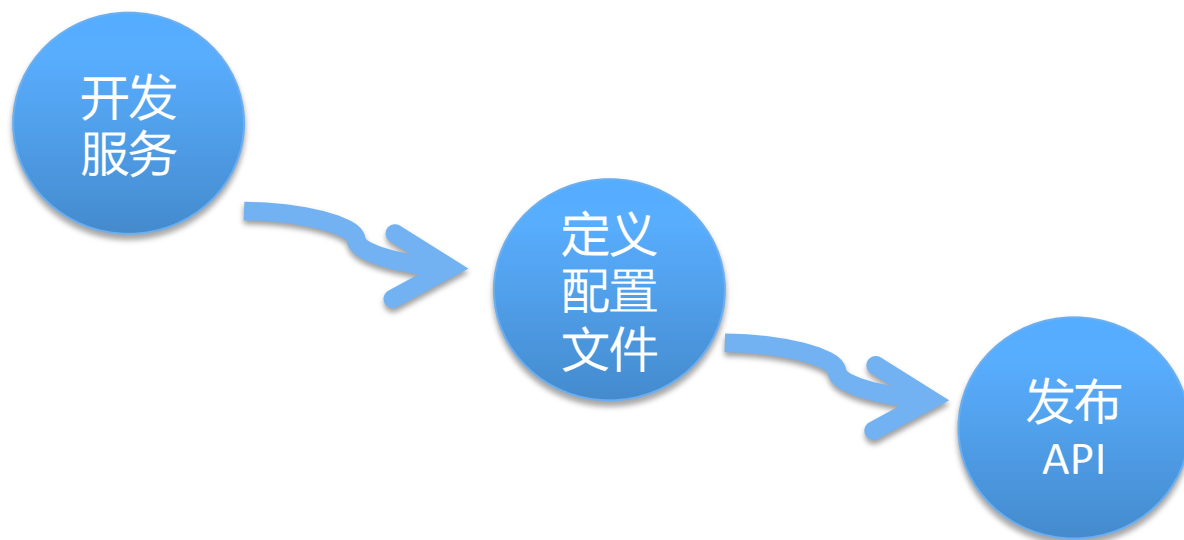


解决方式 - 类型解耦





API的发布流程





网关的演化—网关优化

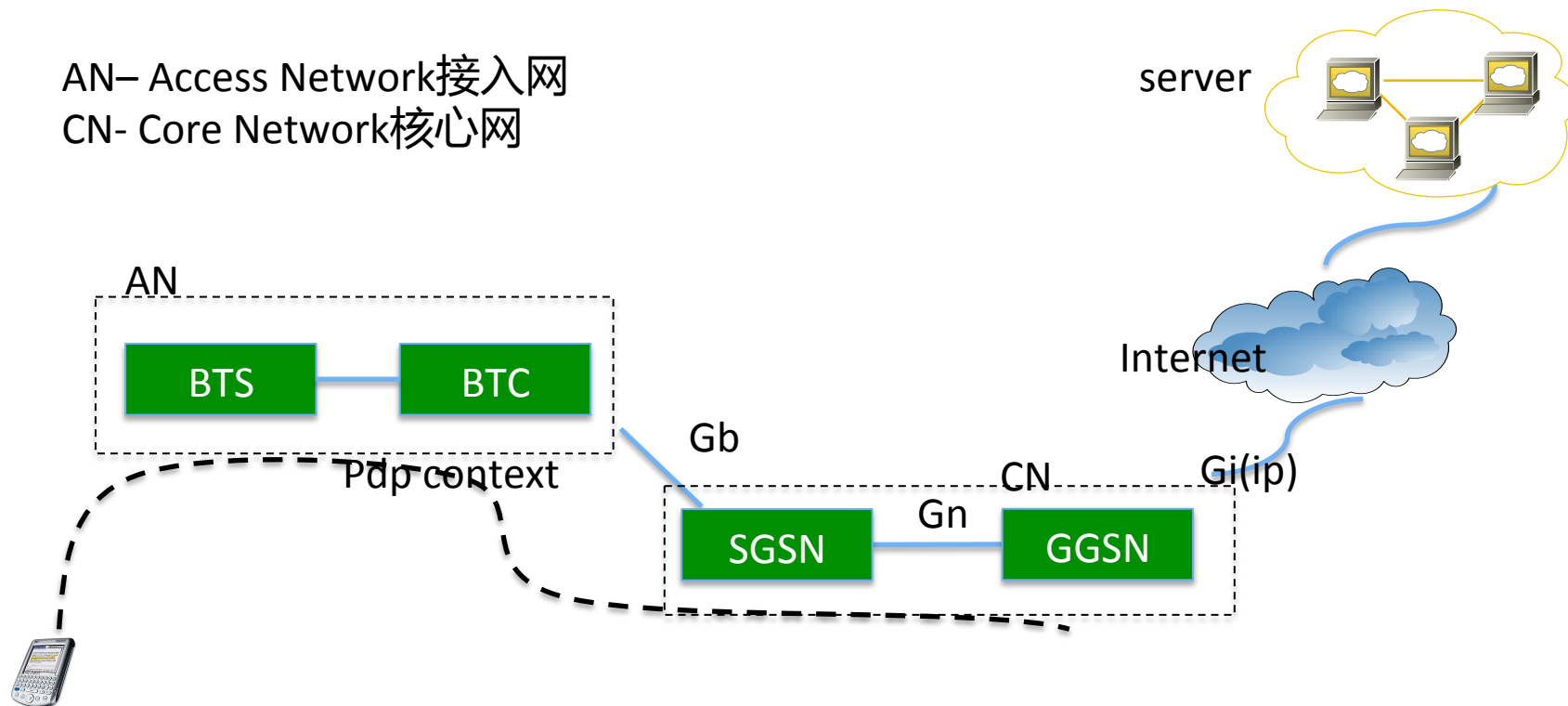




无线网关优化

AN- Access Network接入网

CN- Core Network核心网





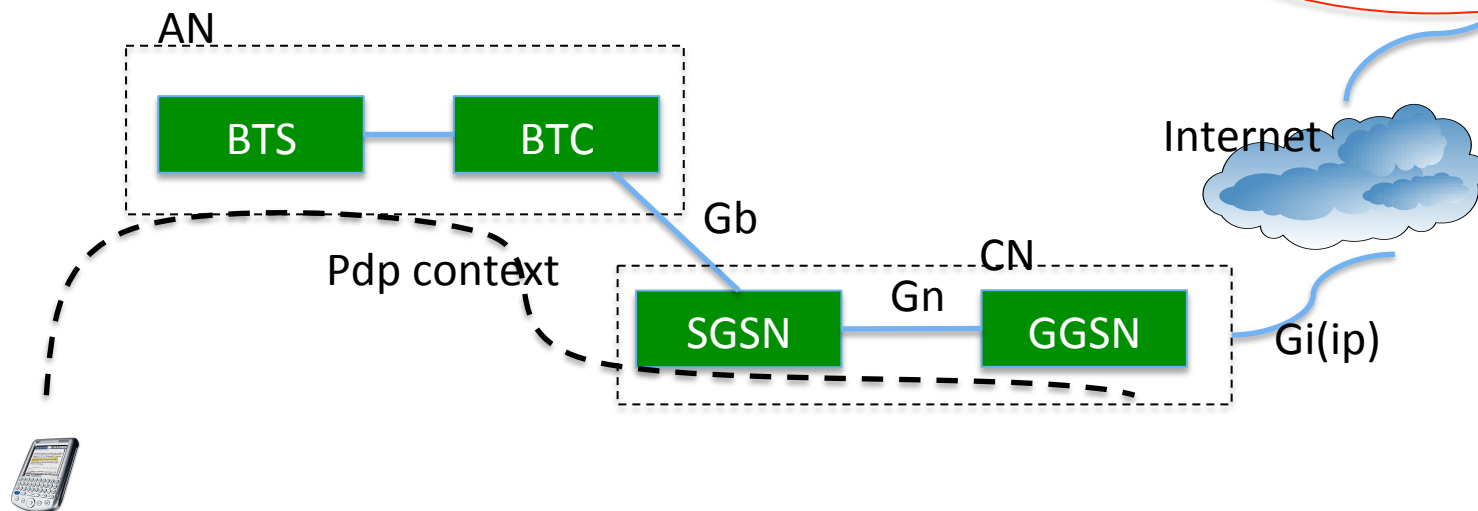
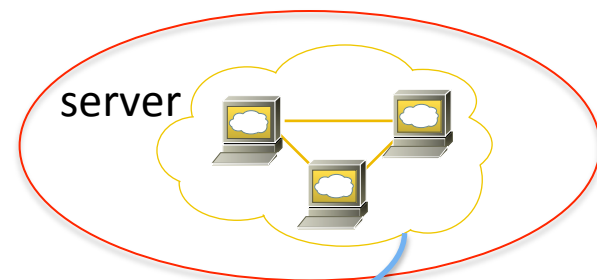
无线网关优化

- 创建PDPContext
- 数据通过无线信号传到接入网(AN)
- 数据从接入网通过核心网(CN)到达数据网关 GGSN(GTP)
- 在GGSN数据网关，GTP被转换为TCP/IP协议通过internet传递给应用网关
- 应用网关将请求转发给业务服务器



无线网关优化

服务端优化





服务端优化 - - 挑战

- 流量增长的挑战

随着手淘流量的加速增长，集群不断扩大。

- 高延时服务对网关的吞吐量的影响，无法合理的评估网关容量

业务延时差异从10ms—10s



继续加机器？还是应用优化？



问题：

- HSF同步调用，等待结果时线程资源浪费



解决方案：空间换时间(同时支持更多的连接)

- RPC的异步调用(callback)
- SERVLET 3.0异步

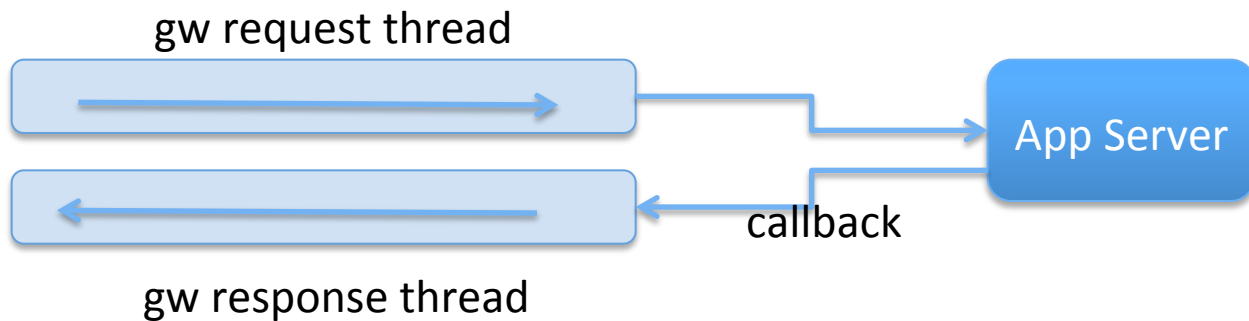


线程模型对比

- 原架构

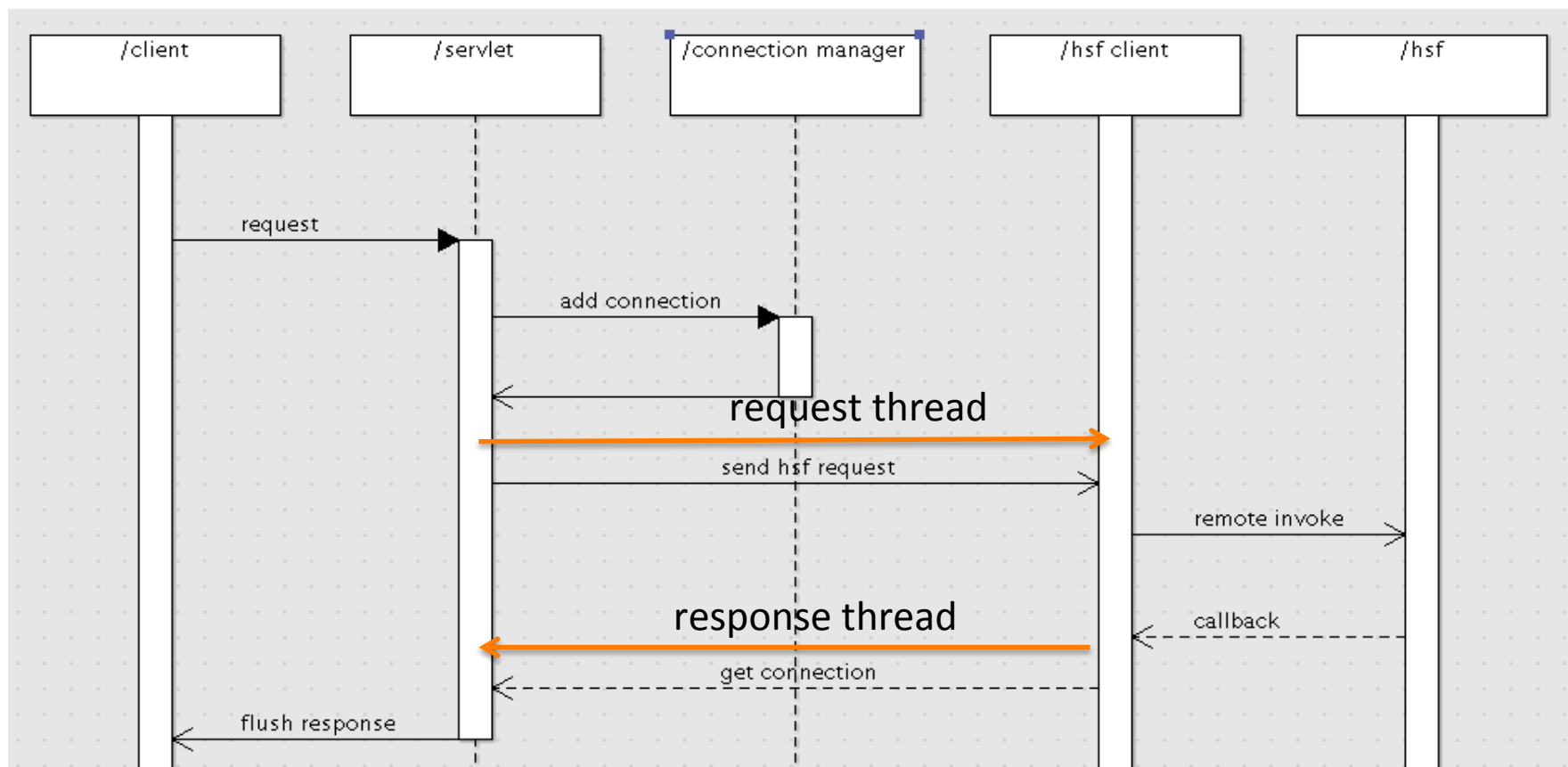


- 新架构





性能：服务端异步—时序图



性能：服务端异步--优化结果

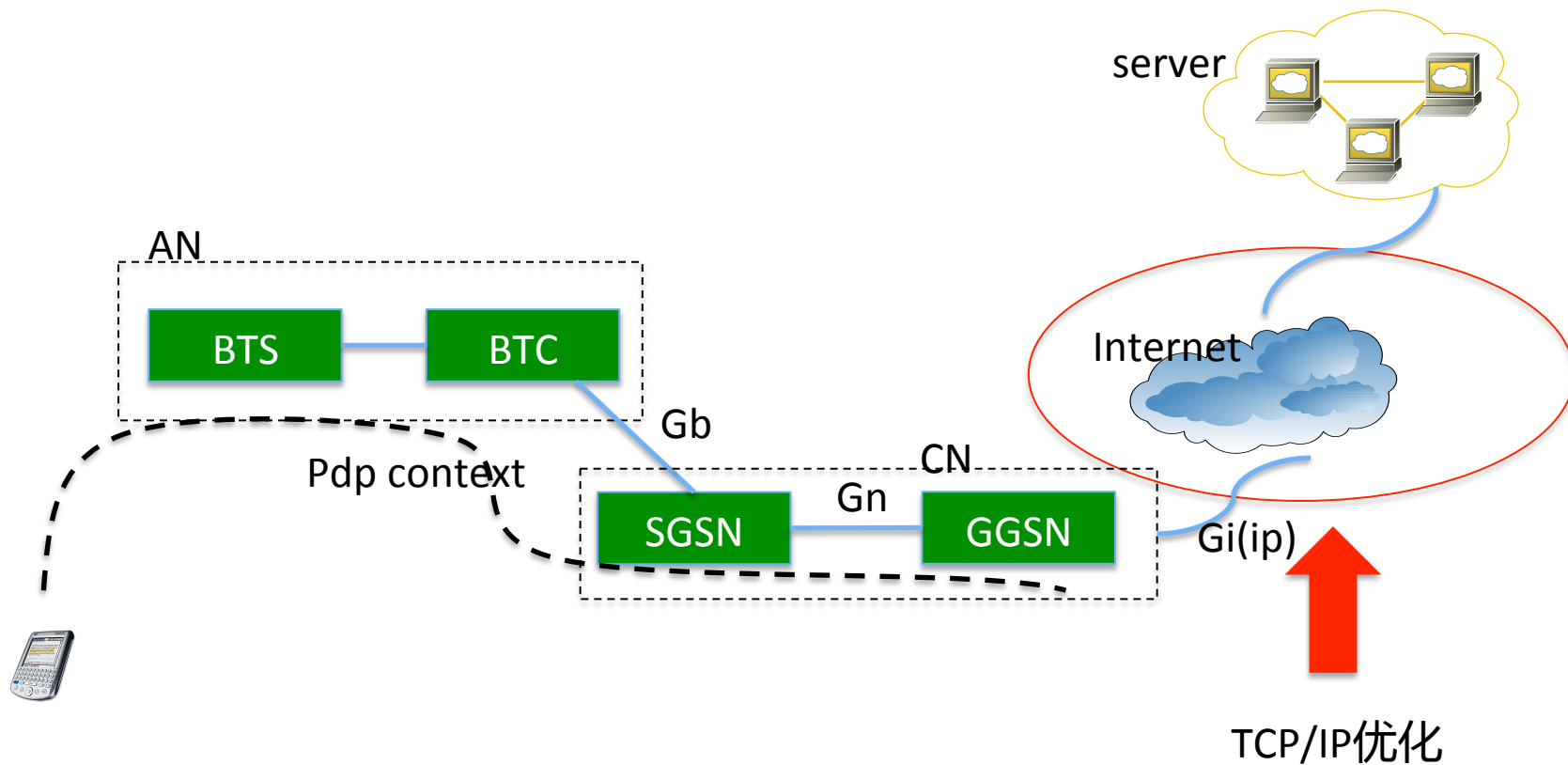
淘宝网
Taobao.com



- 在HSF 400ms延时时qps提升2.3倍，rt降低70%（请求堆积）
- 在HSF 15ms下提升qps提升7%，load降低40%、cpu使用率降低16百分点



无线网关优化





TCP/IP优化—参数调整

- TCP拥塞控制
 - 调大初始窗口到10
 - `tcp_slow_start_after_idle`
- 积极主动的重传策略
- MSS调整

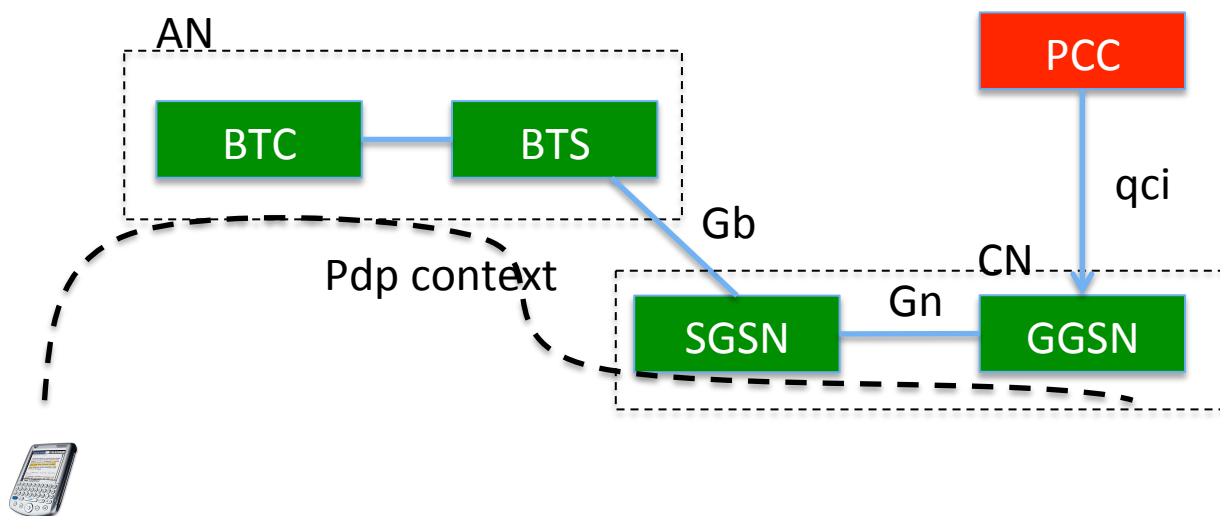


TCP/IP优化—减少握手损耗

- spdy协议与长连接
 - 减少TCP连接创建，提高tcp连接的利用率
- slight-SSL
 - 优化的ssl握手过程，ssl握手0耗时



运营商网络优化—PCC优化



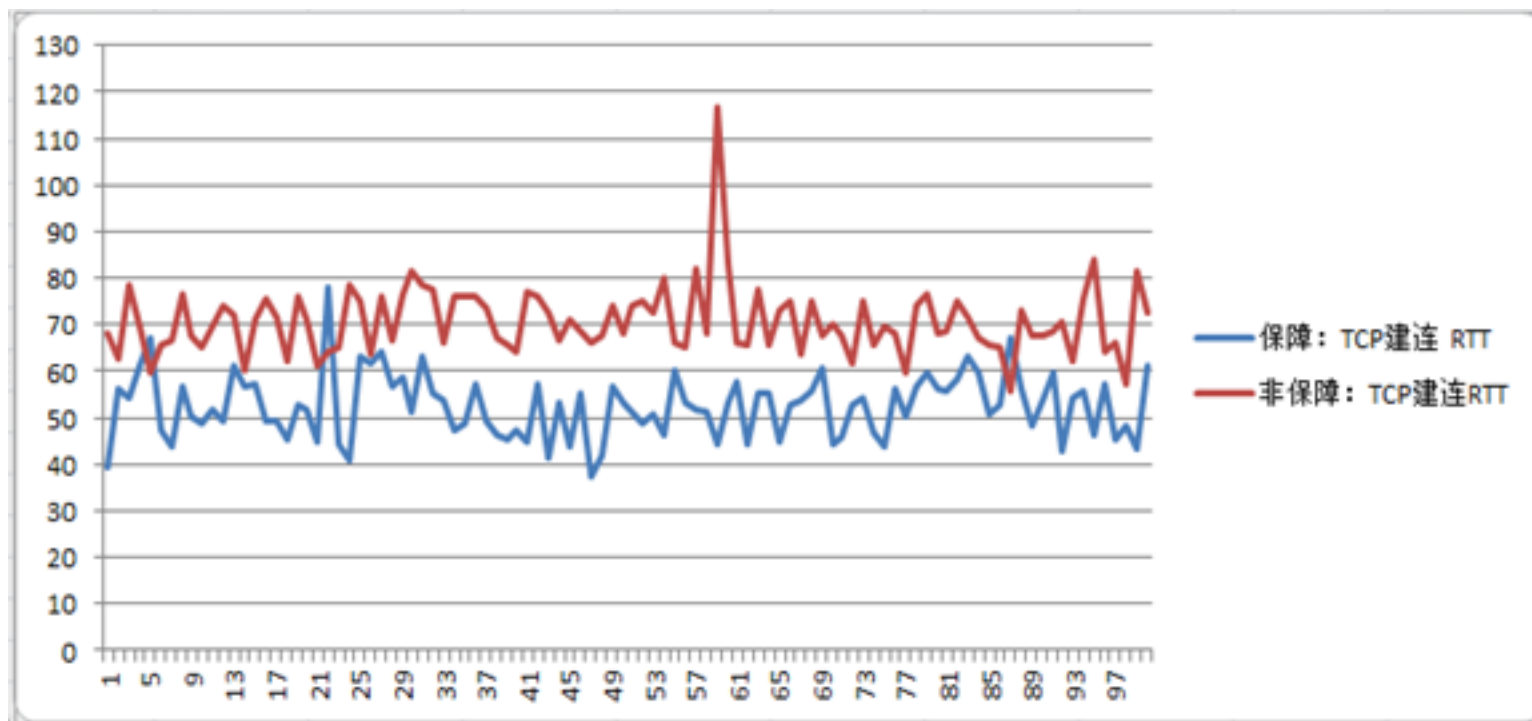


运营商网络优化—PCC优化

QCI	Resource Type	Priority	Packet Delay Budget (NOTE 1)	Packet Error Loss Rate (NOTE 2)	Example Services
1 (NOTE 3)	GBR	2	100 ms	10^{-2}	Conversational Voice
2 (NOTE 3)		4	150 ms	10^{-3}	Conversational Video (Live Streaming)
3 (NOTE 3)		3	50 ms	10^{-3}	Real Time Gaming
4 (NOTE 3)		5	300 ms	10^{-6}	Non-Conversational Video (Buffered Streaming)
5 (NOTE 3)	Non-GBR	1	100 ms	10^{-6}	IMS Signalling
6 (NOTE 4)		6	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
7 (NOTE 3)		7	100 ms	10^{-3}	Voice, Video (Live Streaming) Interactive Gaming
8 (NOTE 5)		8	300 ms	10^{-6}	Video (Buffered Streaming) TCP-based (e.g., www, e-mail, chat, ftp, p2p file sharing, progressive video, etc.)
9 (NOTE 6)		9			



运营商网络优化—PCC优化





客户端的优化

- 预加载
 - wifi下的数据预加载
- 本地缓存
 - 业务数据的本地缓存
 - 缓存优先展示和二次重绘



无线高可用性的延伸

App的可用性不只是云端

- 云上的可用性
- 管道的可用性
- 端的可用性



云上的可用性 - 隔离

- 隔离
 - 按业务分级形成物理隔离的集群
- 堵截
 - 软件防火墙(TMD)系统的防恶意攻击保护
 - 应用上的并发流控，保护后端应用的稳定性



管道的可用性

- HTTP DNS与IP直连
 - 防止DNS劫持
 - 加速DNS解析
- 自适应协与端口适配
 - sdpy优先， failover到http(代理无法穿透或端口问题)
 - 端口由云端控制动态调整，端口不可用时的动态切换
- 自适应MSS



端的可用性

客户端的动态更新和patch技术

- Hotpatch
 - wax
 - Xposed
- atlas

Q & A

