

网易云实践
即时通讯云服务之路

SPEAKER

阙杭宁



促进软件开发领域知识与创新的传播



关注InfoQ官方信息
及时获取QCon软件开发者
大会演讲视频信息



[北京站] 2016年12月2日-3日
咨询热线: 010-89880682



[北京站] 2017年4月16日-18日
咨询热线: 010-64738142

关于我

- 阙杭宁

- 2007 商务领航SAAS平台开发
- 2010 中国电信协同通信能力开放平台技术负责人
- 2012 中国电信翼聊服务端技术负责人
- 2013 易信公众平台、易信能力开放平台技术负责人
- 2014-至今 网易云信 技术负责人

互联网+下的即时通讯



即时通讯把人与人连接起来

即时通讯架构特点的演进

- PC时代

- 强连接
- 状态敏感
- 流量不敏感
- 计算量不敏感
- 运行环境简单

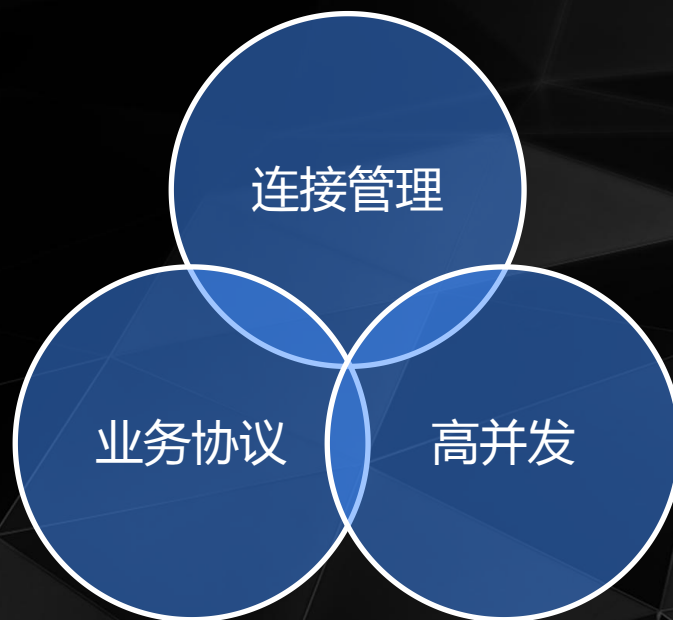
- 移动互联网时代

- 弱连接
- 状态不敏感
- 流量敏感
- 计算量敏感
- 运行环境复杂

即时通讯研发关键点

研发面临的实际情况：

技术难度大
开发周期长
业务细节多
维护复杂度高



传统即时通讯开发模式已无法适应互联网+下的团队节奏

即时通讯云服务

- 以基础服务形态开放即时通讯能力
 - 更大的数据量级
 - 更全的网络覆盖
 - 更高的兼容性应对
 - 更复杂的业务处理
 - 更快的扩容速度要求
 - 不同应用之间的隔离

网易云信

网易云服务系列推出的第一个平台

2015.10.13

344天



为**100000+**开发者
提供服务



覆盖**4亿+**用户



覆盖**196个**国家
567个地区

云信平台实践之路

稳定那些事儿

实时性的挑战

稳定——云服务的生命线

常规外的影响因素



不可掌控的业务
峰谷及增长速度



云上的资源竞争



客户端实现不可控状况下
的服务升级兼容性难题

业务容量

足够的容量

如何定义云服务不同阶段的容量模型特征

初期		发展期		成熟期	
特征	应对	特征	应对	特征	应对
用户体量小 投入阶段 与增速密切相关	按产品预期规划容量建设预算 关注增长速度	有一定体量 快速增长期 大客户对容量冲击感知明显	维持中低水位以保障突发业务量 对水位影响的关键因子保持关注	平台容量大 日常业务量大 增速斜率较平缓	维持中等水位以提高资源使用率 关注水位毛刺异常

突发峰值应对

不可预知的峰值，某APP忽然爆红，或者大客户做活动。



快速扩容

单元化服务

快速扩容

- 触发扩容

- 子服务集群容积和水位
- 任务积压

- 环境准备与服务部署

- 网易蜂巢容器云
- 服务镜像
- 配置注入

单元化服务域



不按套路出牌的接入者

- 不合理的调用方式



资源侵略



兼容性

不按套路出牌的接入者



消息撤回 功能升级的案例

实时性——即时通讯云的核心价值

- 复杂网络的连接稳定性

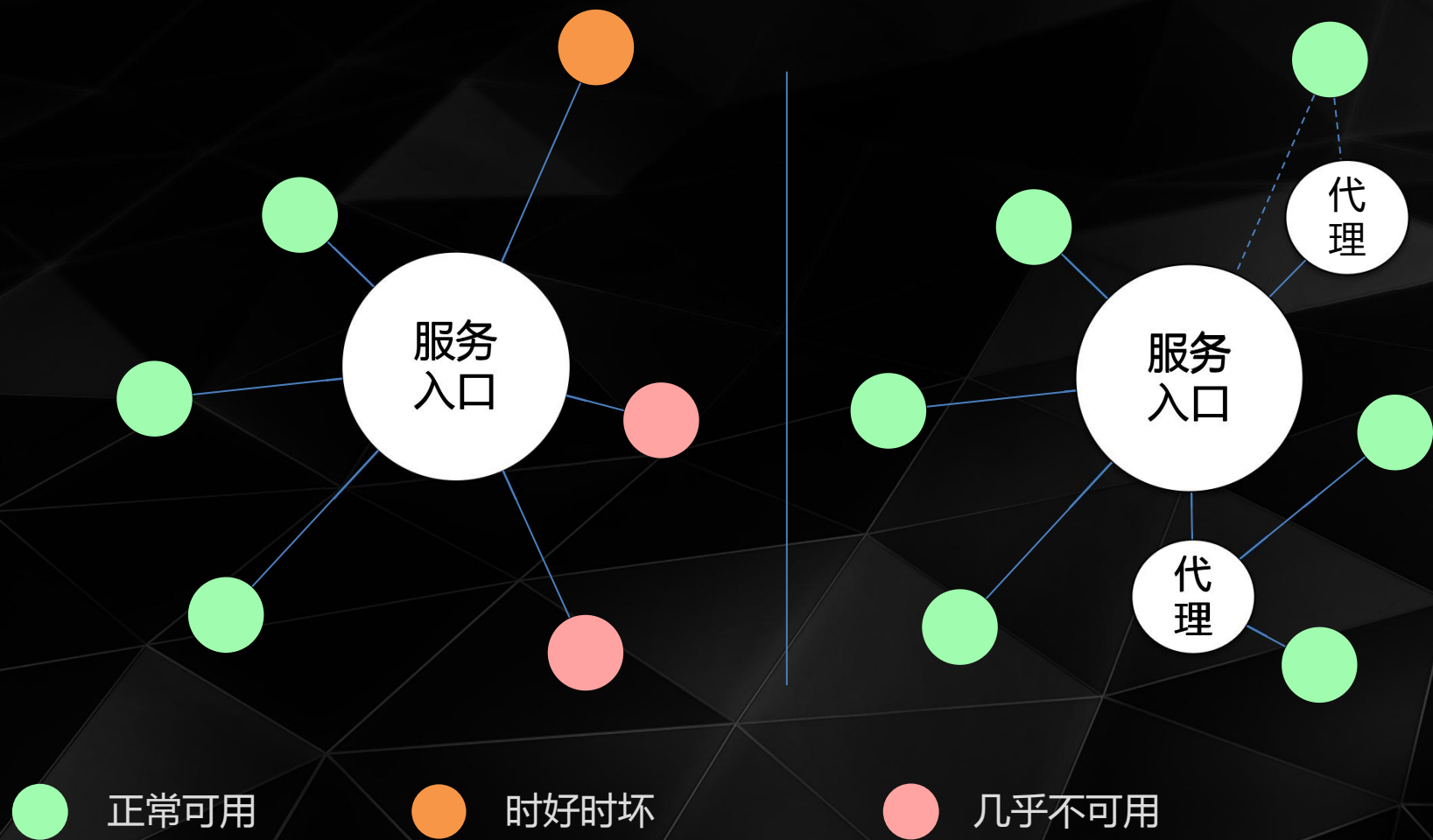
- 问题本质

不可控的客户端网络无法由单一的网络入口满足
不可避免的“弱网”条件下的连接保持和使用

- 应对策略

用更可靠的服务端网络代替部分特殊客户网络
优化连接探活和重连效率
精简数据包大小，增强对弱网的支持

服务网络入口拓扑



典型业务的实时性要求



IM



实时音视频



直播聊天室

实时互动业务的优化

业务特征



低延迟



数据有时效性



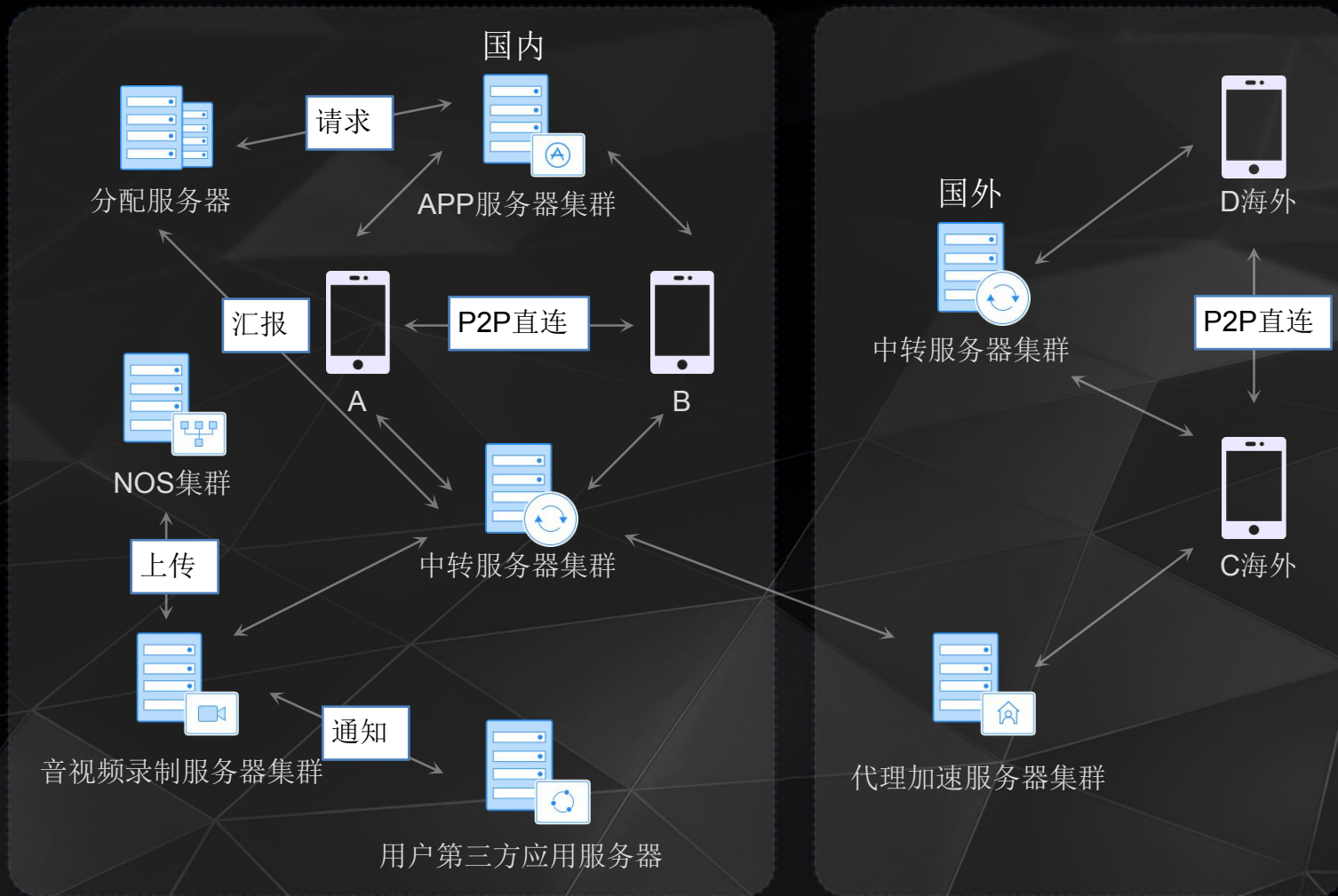
高数据流量

距离远，相对慢；距离近，却不一定快

实时互动业务的架构目标

- 本质:网络传输质量
 - 网络调度调优
 - 缩短传输距离
 - 多通路，网络探测择优选取
 - 适配不同网络情况
 - 抗抖动算法
 - 动态码率

实时通信拓扑



Dispatcher

- 调度因子
 - 用户IP
 - 用户ISP
 - 全国ISP和位置到各地机房的数据
- 分配算法
 - $\text{rate} = \text{distance} * \text{priority} * \text{rand} / \text{isp}$

海量实时消息——超大聊天室

业务特征



发言人数中低



观众人数高



消息量大



消息实时性要求高



离线消息关注度低

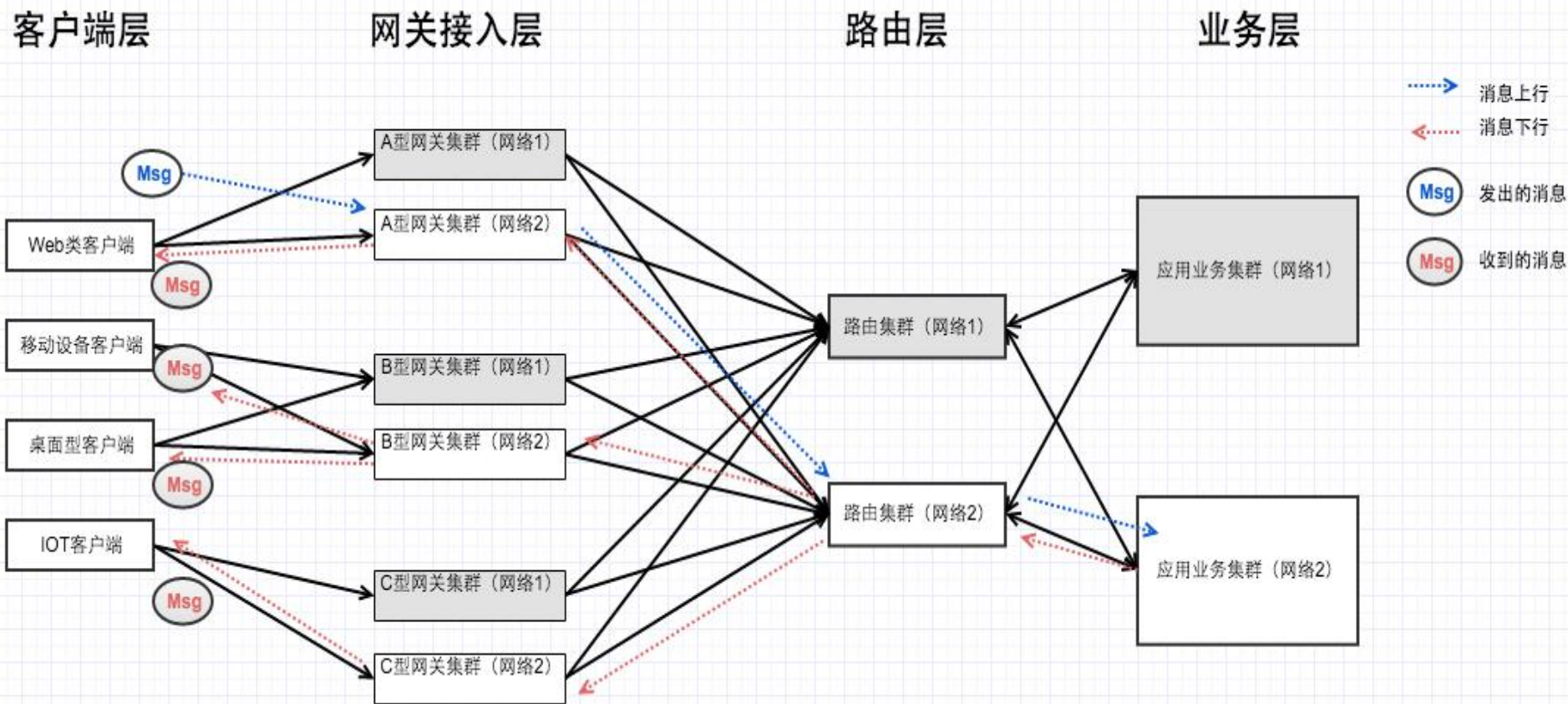


历史消息关注度较低

超大聊天室的架构目标

- 本质:组播性能
 - 组容量扩展性
 - 分布式并行架构，可横向无损扩展
 - 组播性能
 - 减少消息放大中间环节的数据量

聊天室架构



多层架构

客户端层

SDK多平台适配
移动弱网络优化
安全加密
压缩

网关接入层

长连接管理优化
支持平滑升级
支持跨网络切换
广播分包

路由层

支持服务化后端
提升横向
扩展能力

业务层

高可用
弹性扩展

即时通讯云的差异点

- 一对一 => 一对多
 - 需求复杂度
 - 稳定性复杂度
 - 运维复杂度

即时通讯云的差异点



“全网”覆盖



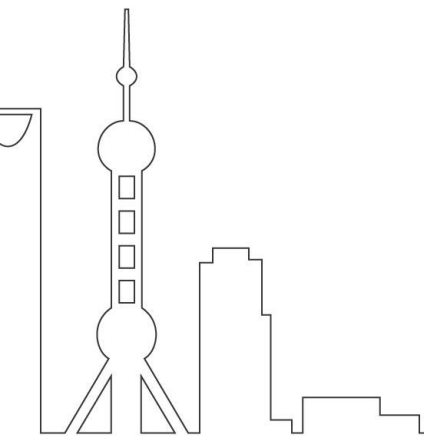
“全平台”覆盖

即时通讯云的差异点

- “不一样”的高数据量挑战
 - 客户端实现不可控
 - 多客户的云上“共存与隔离”

即时通讯云服务

- 从客户的终端用户的角度理解即时通讯场景
- 从接入开发者的角度立场思考解决问题



Thanks!

International Software Development Conference