

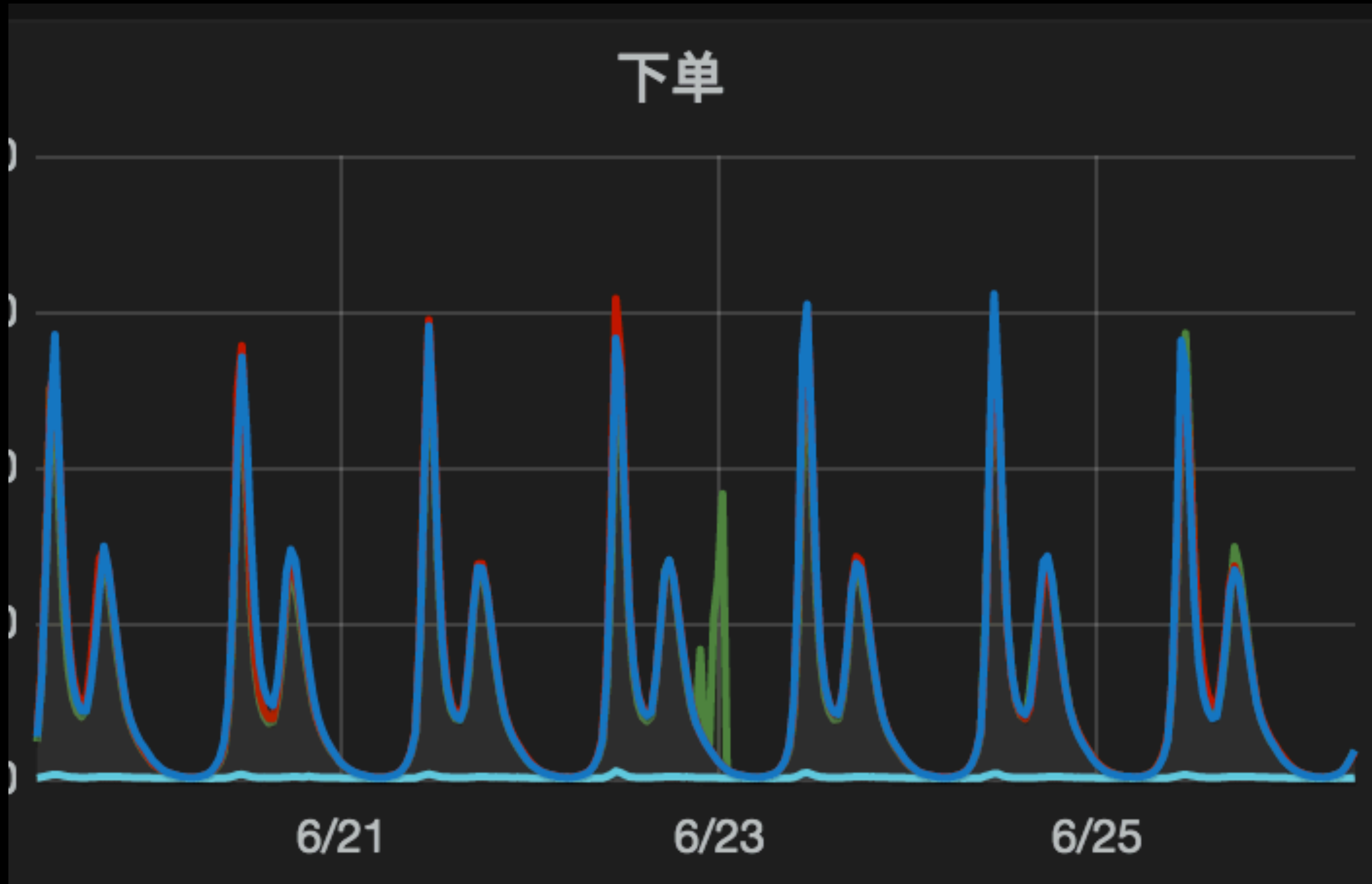
—— “2007-2017” ECUG 十年高峰论坛 ——

饿了么混合云架构探索

# 日程

- Challenge
- Architecture
- Topology & Data
- Doing & Planning

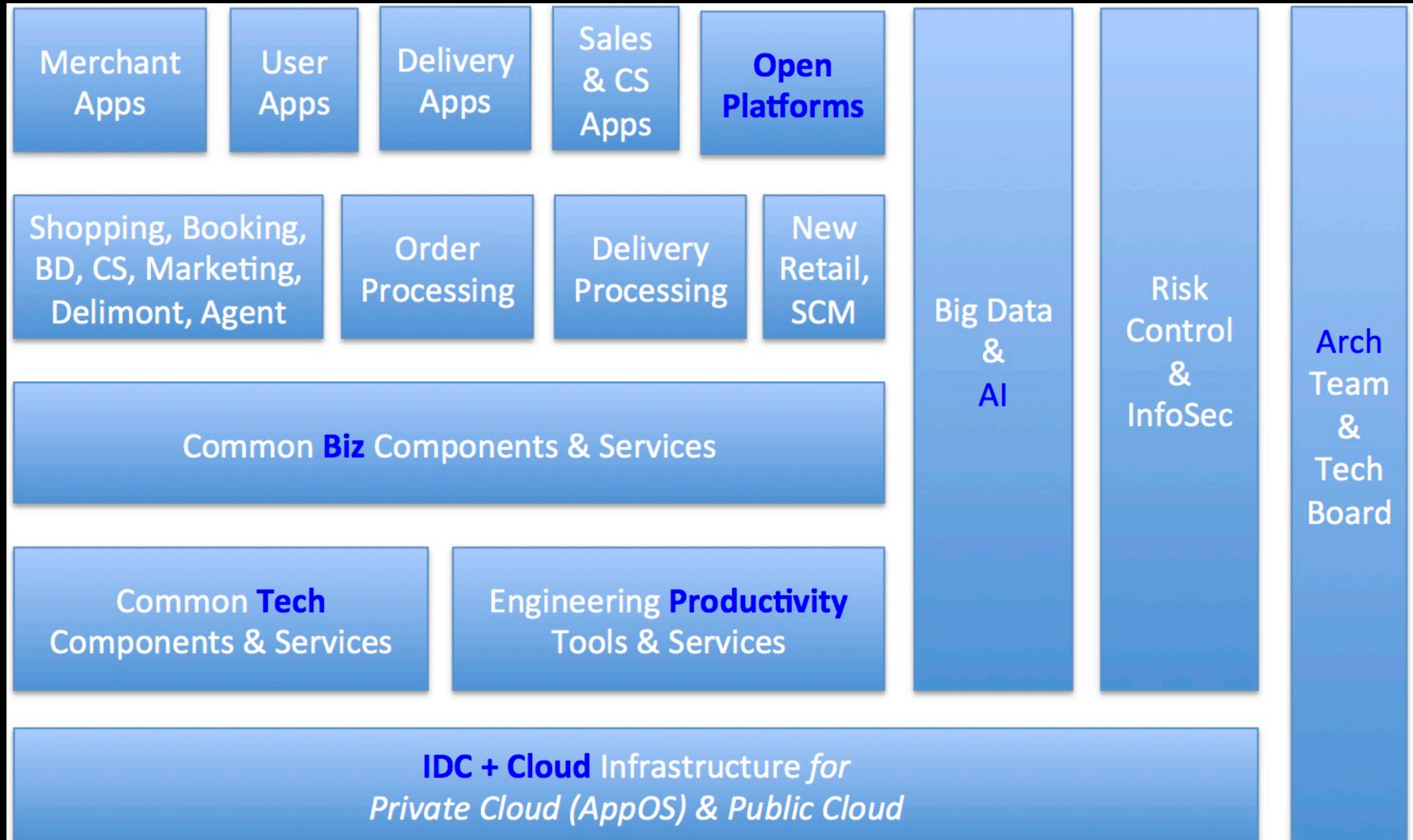
# Challenge (1)



# Challenge (2)

- 订单：高并发、**瞬时冲击**、517、秒杀
- 运单：快、**准**、**全**
- 全链路 (订单 + 运单)：**准时达**、**0元起送**
  - 下单预测、接单预测、**配送费实时计算**
  - **出餐预测**
  - 配送调度、路径规划
  - **降级处理**、异常处理、区域负载均衡、骑手负载均衡
- 难点：**实时**订餐领域，一环扣一环，即使解决了订单高并发或瞬时冲击，运单**闭环**没搞定，客人/商户体验还是不好。

# Architecture (1)



# Architecture (2)

## 多活架构 Overview



**ezone:** 一个ezone对应一个数据中心；一个订单在一个ezone中完成；ezone中的数据相互复制

**global zone:** 对强一致性有要求的数据放在global zone，global zone在单个数据中心写入，复制到其他数据中心，global zone的服务在每个机房都需要部署

# Architecture (3)

## Sharding 划分 (大数据提供)



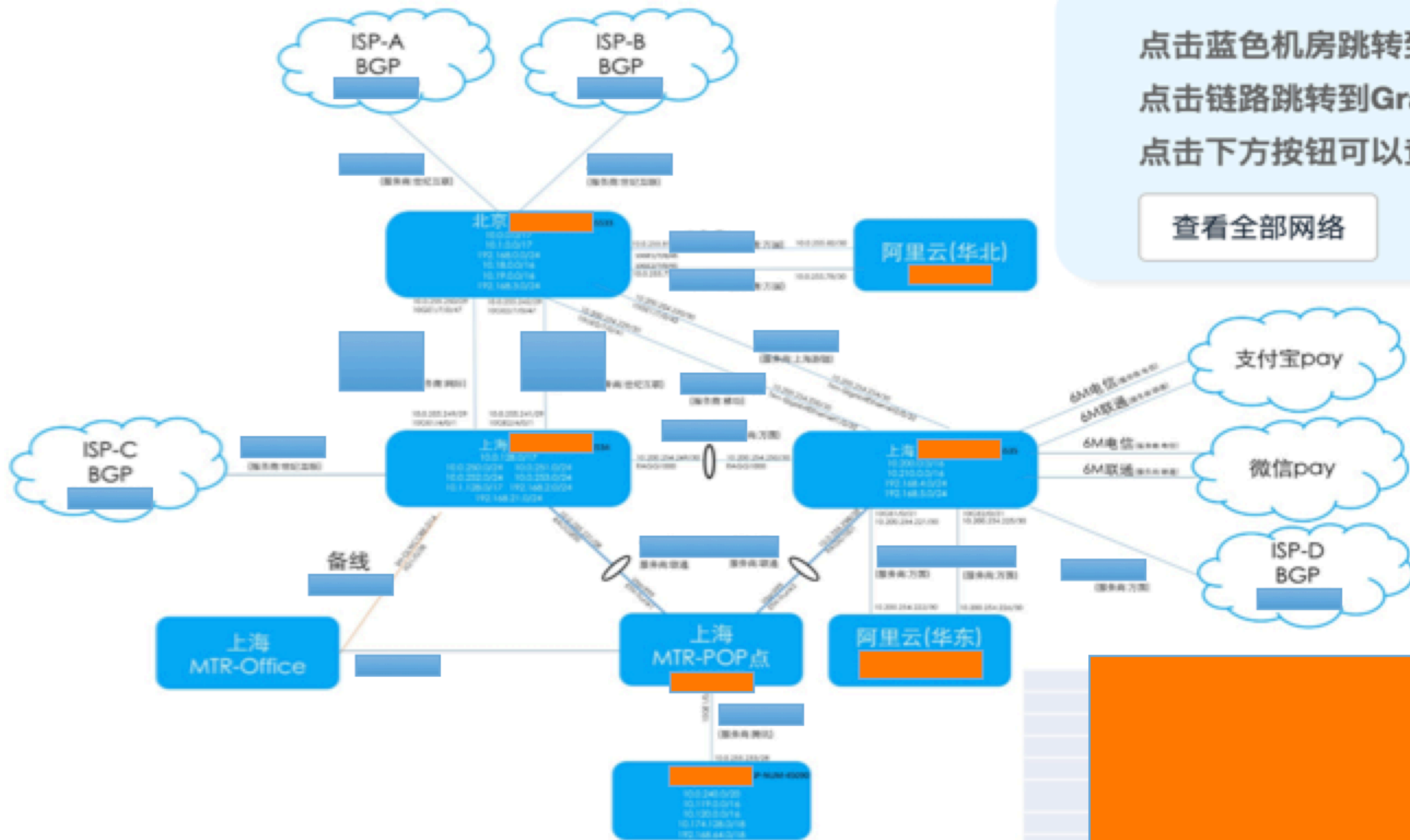
1. 以地理位置相近为优先原则，将全国31个省级行政区(除港澳台)划分为流量相当的24个虚拟区域，根据区域流量占比最高原则确定了14个区域中心
2. 以南北分界线做分割线，分割线以北归为北京站点，分割线以南归为上海站点

# Architecture (4)

- IDC
  - VM over ZStack
  - Docker over Mesos (二次开发)
  - OLAP over YARN (二次开发)
  - AI (*elearn*, 自研) over k8s
- Cloud
  - ECS + RDS + Docker
  - Mesos (二次开发)



# Topology & Data (1)



点击蓝色机房跳转到机房视图  
点击链路跳转到Grafana监控  
点击下方按钮可以查看所有链路

[查看全部网络](#)



# Topology & Data (2)

- 逻辑数据中心: 6 (4 传统 IDC + 阿里云 + 百度云)
- 物理数据中心: 8 (含阿里云 SH、阿里云 BJ、百度云 GZ-1 & 2)
- 服务器节点: 25K+
  
- 每日业务数据增量: 120T+ with replica, 40T+ no replica
- 每日所有存储增量: 400T+ with non-biz data
- 每日业务数据处理量: 3P+
- 总业务数据量: 12P+
  
- SLOC: 130M+ (master/default-only)
- App ID: 2000+
  
- Commit/Push: 3800+/天
- Rolling/Release: 350+/天

- **核心技术战略：IDC + Cloud**
  - 4 IDC + 1 Cloud
  - **多活** (世界级难题): 2 IDC e-zones + 2 Cloud e-zones
- **拥抱 Cloud Native Architecture**
  - 技术运营: 从 SSH 到 API 的巨大**理念转变**
  - 其他团队: 架构设计时就要考虑 Cloud-Affinity
- **展望 Serverless & Auto-Scaling Architecture**
  - 2017 尝试一个相对独立应用上线
  - 2018 尝试在 Job 类应用中推广及部分场景**按需使用**

# Doing & Future (混部尝试)



- 业务波峰期 (或各种事件 & 活动)
  - OLAP 集群运行 OLTP 服务
  - OLAP 任务延迟执行，但要确保可追上大盘
- 业务波谷期
  - OLTP 集群运行 OLAP 任务
  - OLAP 任务需要适配 OLTP 机型
- Mesos & YARN & ZStack 相互适配？

Over, 谢谢!