



IT大咖说
知识分享平台

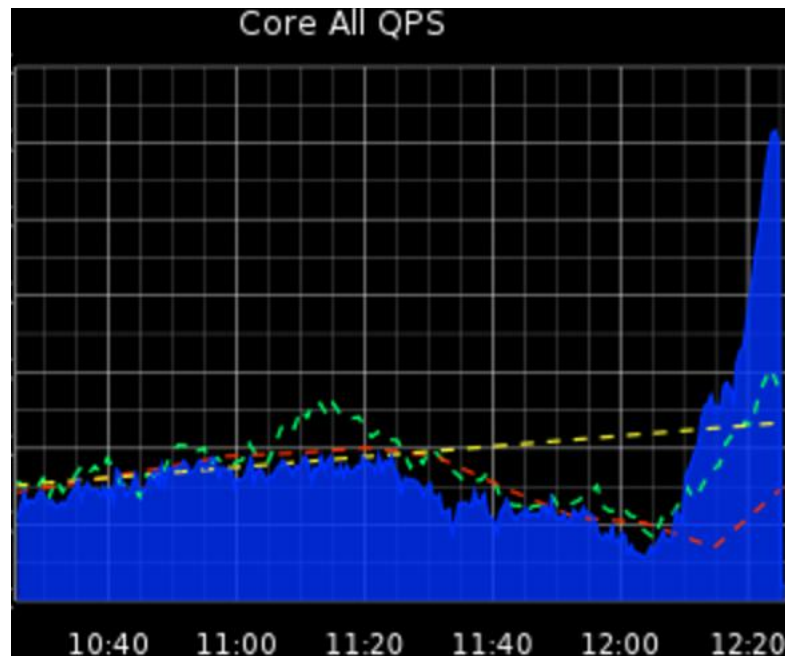
新浪&微博数据中心网络创新与实践

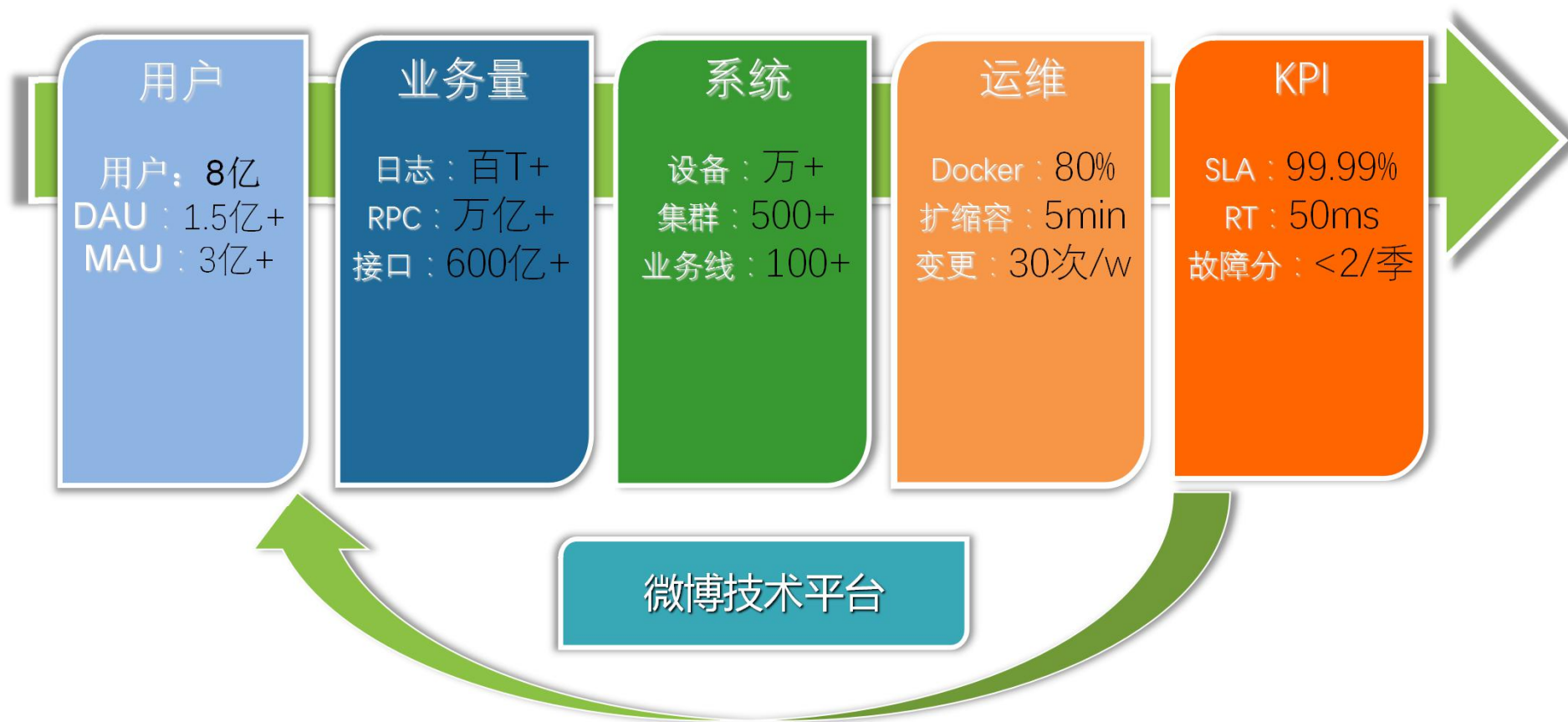


目录

- 一．高速混合云连接
- 二．重构DCI
- 三．全盒式设备的IP Fabric
- 四．跨数据中心大二层

- 白百合、李晨娱乐事件等热点突发峰值应对
 - 突发性强无预期、无准备
 - PUSH常规化，短时间大量设备扩容需求





十亿级PV、千亿级数据、2000台以上的服务器规模、20个以上的大小服务模块、百亿级数据HBase存储、千台以上的Docker混合云集群，持续不断的技术挑战；

■ 混合云进展：

容器数：5000+

晚高峰自动扩容500+

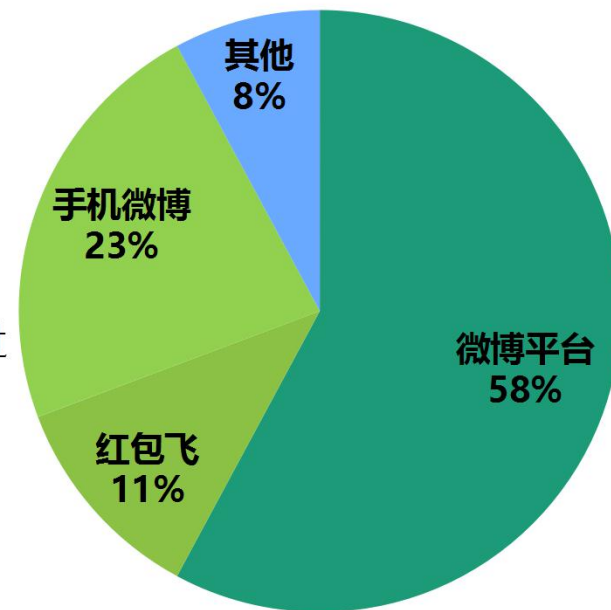
■ 春晚备战：

10分钟混合云扩容1000节点技术能力

2017春晚峰值历史新高，完成4700台阿里云ECS扩容，实现无降级平滑过渡，公有云高峰支持微博50%主体流量。完成feed流、红包飞、手机微博公司各主要业务线均完成上云支持

主要业务方

■ 微博平台 ■ 红包飞 ■ 手机微博 ■ 其他



目录

一 . 高速混合云连接

二 . 重构DCI

三 . 全盒式设备的IP Fabric

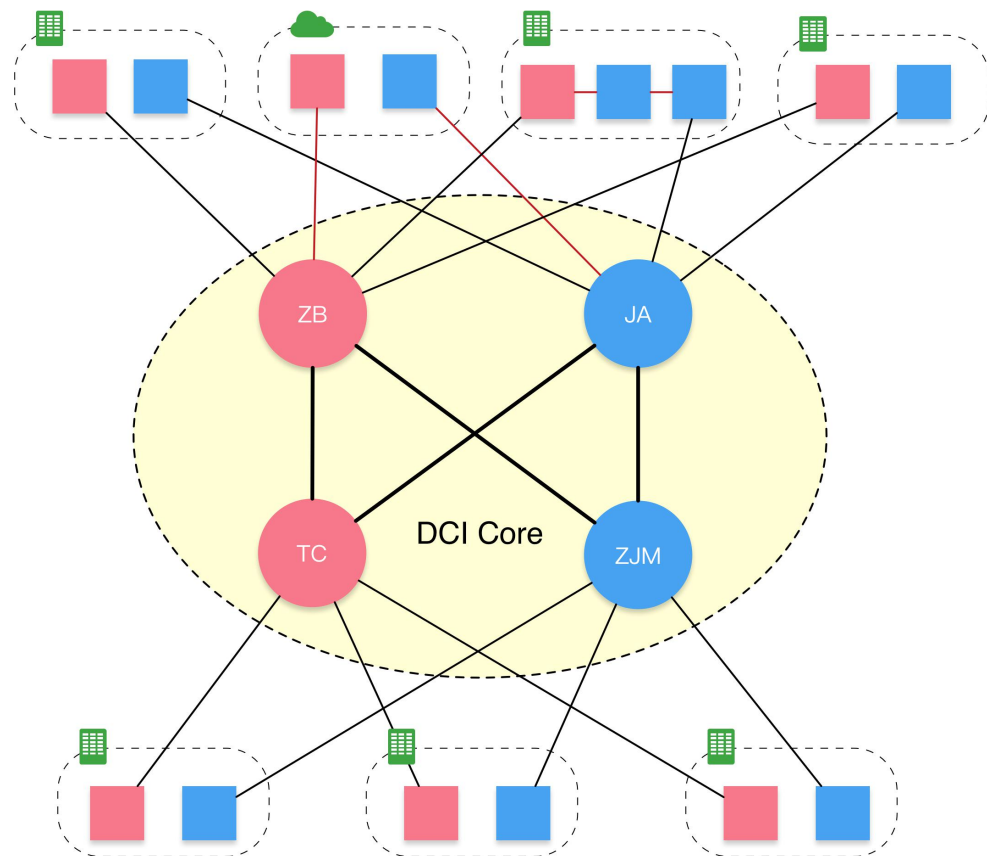
四 . 跨数据中心大二层

- 网络没有层次，DCN和DCI边界不清晰
- 专线资源虽多但杂乱无章
- 从拓扑图无法快速识别流量路径

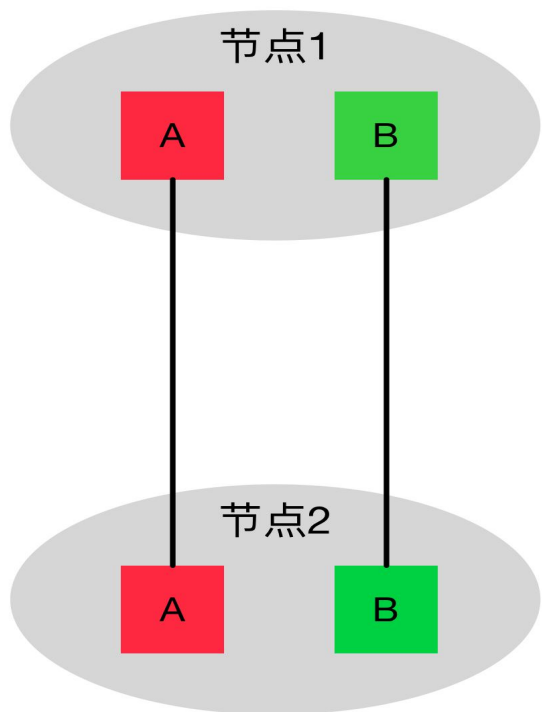
分布式路由计算环境下，基于链路权值(加性参数)计算最短路径，很难灵活调度流量。



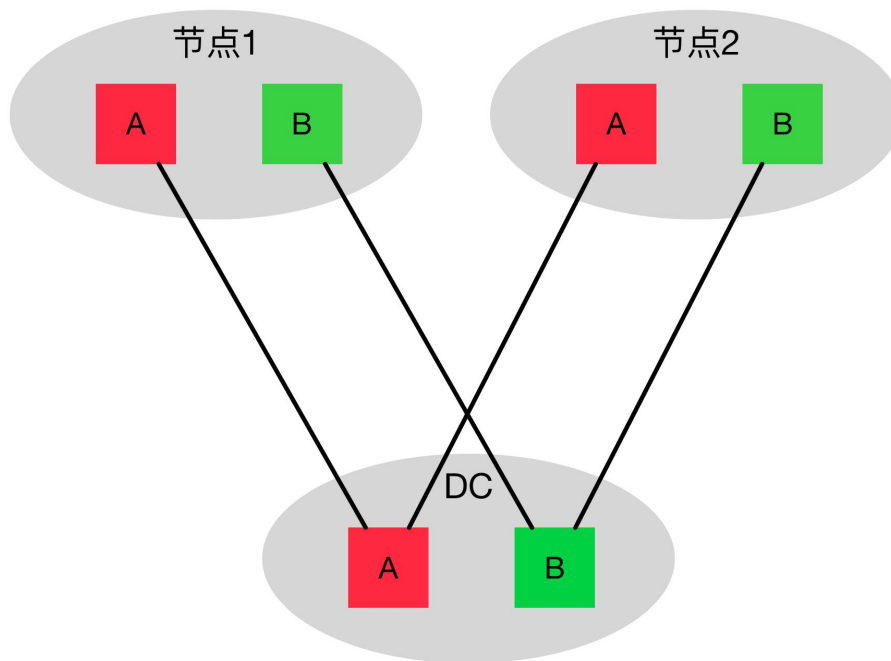
DCI 2.0



充分利用现有资源
尽量保证网络简单
流量流向清晰



核心层互联



DC接入上联

- 设备及链路分为AB两个平面
- 核心层节点互联形成2个链路组
- 接入层上联形成4个链路组
- 网内及网间均采用对称结构

- 网络结构：数据中心上联、核心层内部，均衡、对称
- 路由一致：路由条目一致、路由属性一致
- 虚拟下一跳

虚拟下一跳实现方式

- C路由器向RR广播路由时将BGP Next-hop修改为虚拟下一跳地址；
- 在C路由器将该VNH地址发布到IGP；
- 通过IGP实现负载均衡。

目录

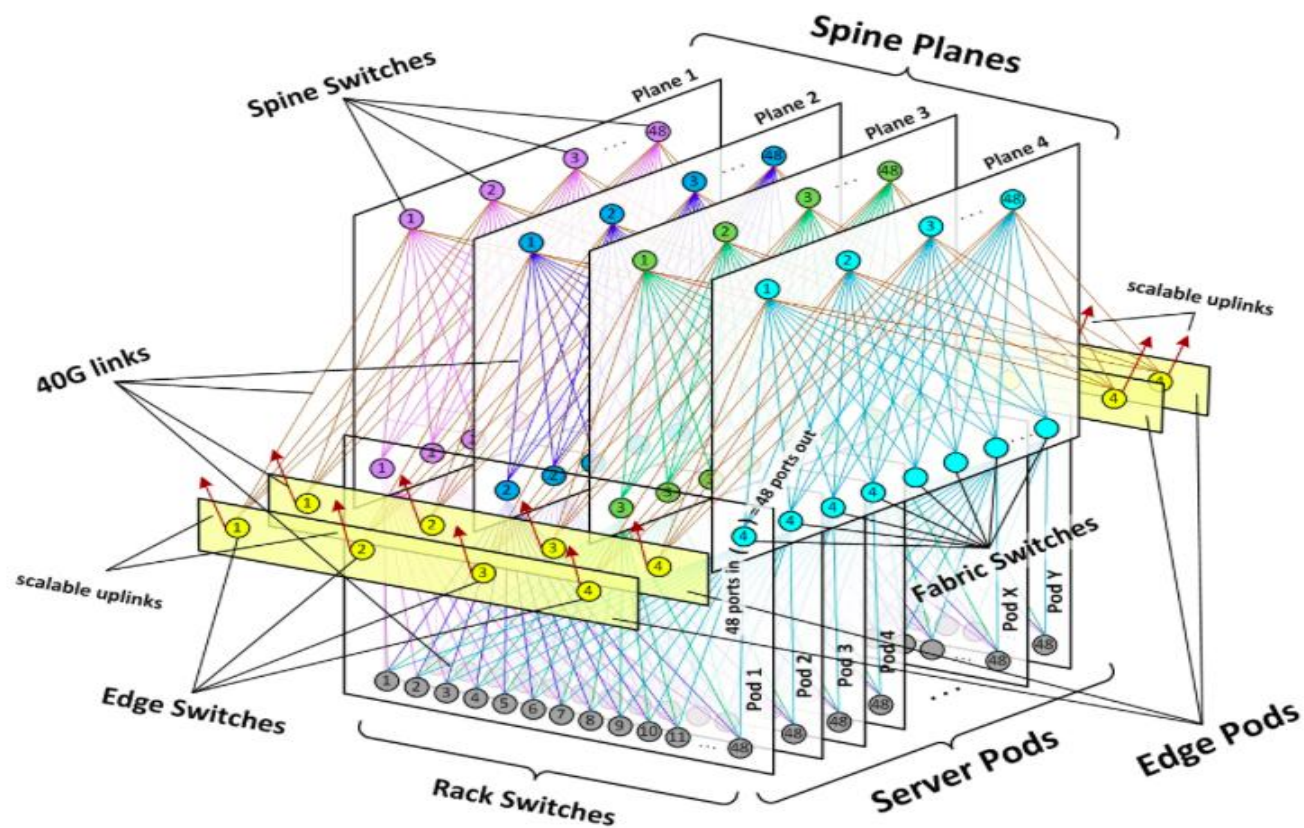
一 . 高速混合云连接

二 . 重构DCI

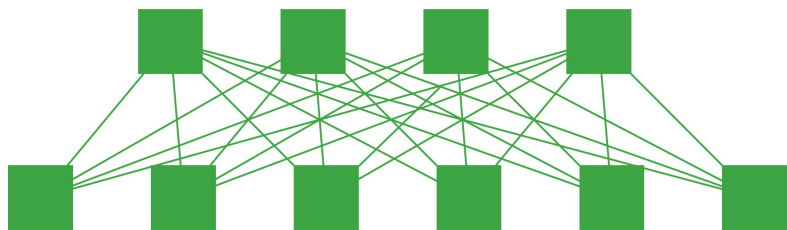
三 . 全盒式设备的IP Fabric

四 . 跨数据中心大二层

Facebook IP Fabric



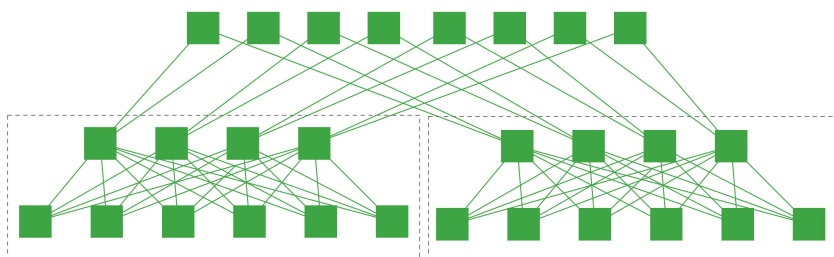
全盒式设备的IP Fabric



1. 传统网络架构的规模受限于交换机端口密度
2. 高性能机框式交换机价格昂贵
3. 传统机框式交换机故障时板卡更换难度大、风险高



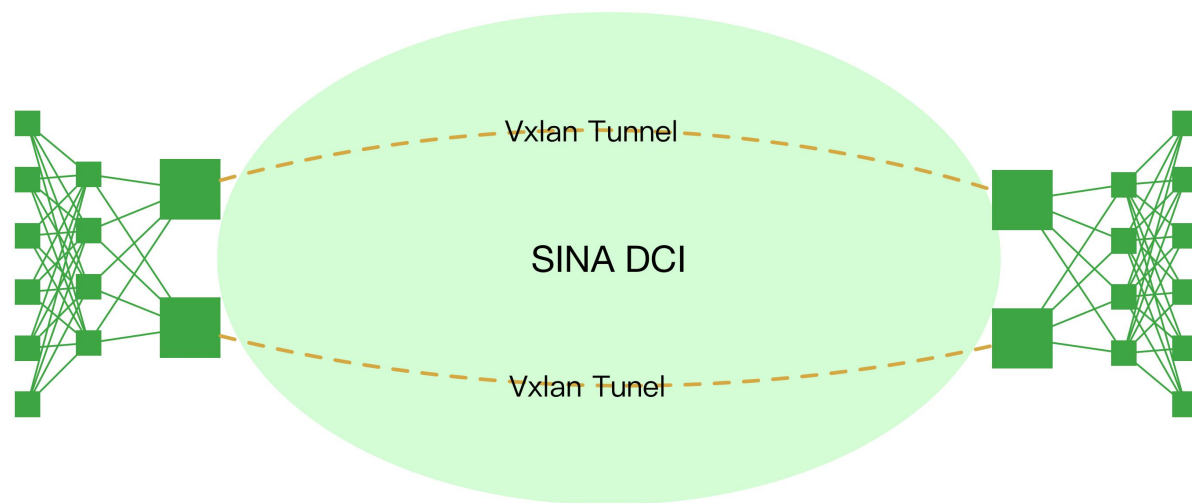
1. ECMP等价路径保证网络高可靠
2. Spine节点设备故障可快速替换，无需考虑兼容性
3. 横向扩展不受限于单台设备能力



目录

- 一 . 高速混合云连接
- 二 . 重构DCI
- 三 . 全盒式设备的IP Fabric
- 四 . 跨数据中心大二层

第一阶段：两个IDC BR之间互相通告EVPN路由，实现跨机房VXLAN 隧道建立。



第二阶段：各IDC BR设备具备全网EVPN信息，可在任何机房间建立VXLAN 隧道。

