



# 云计算数据中心网络



# Agenda

- 云时代的网络挑战
- 网络虚拟化技术
- 数据中心二层网络技术
- 跨数据中心二层网络技术
- 非以太网网络技术
- 总结

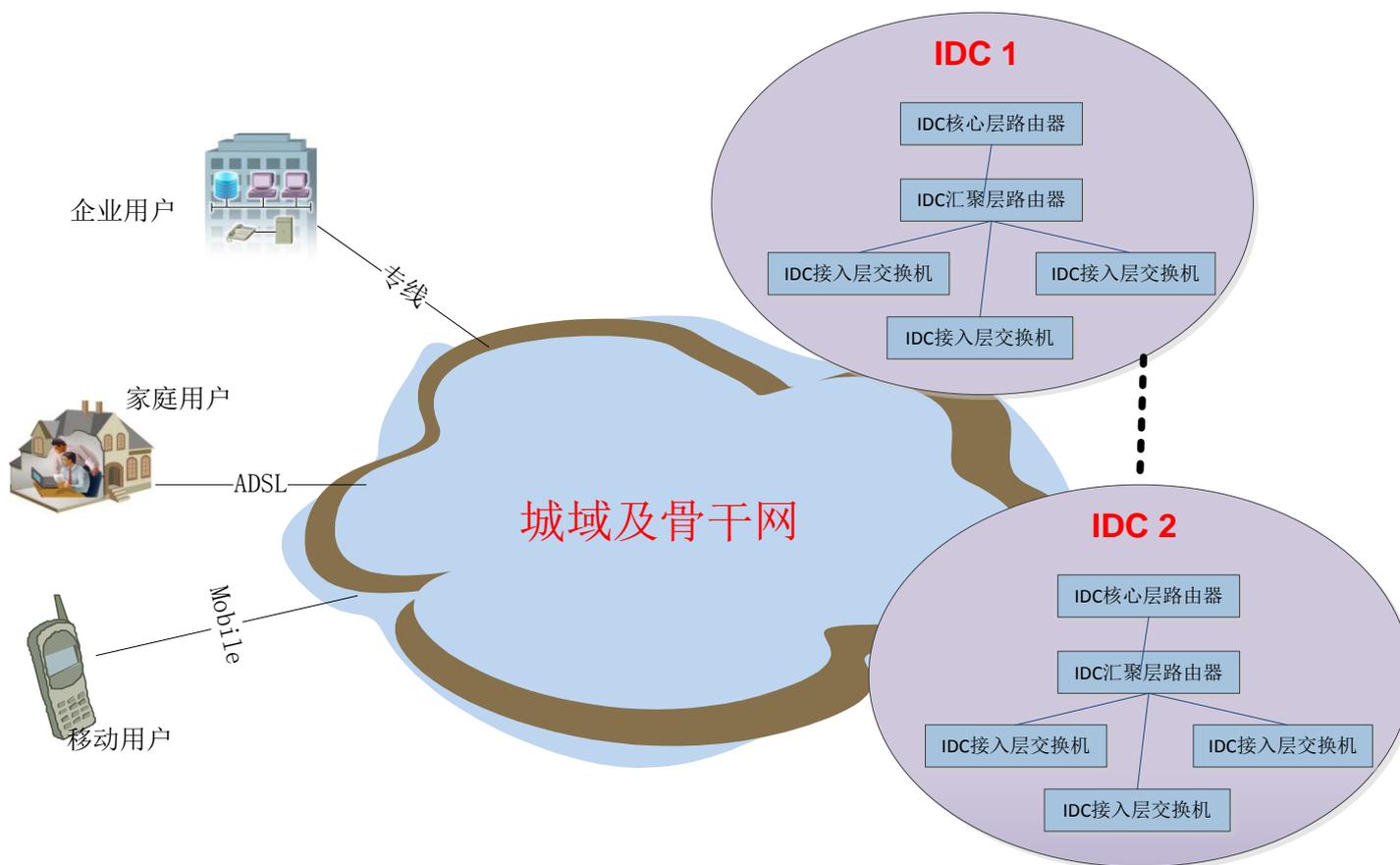


# Agenda

- 云时代的网络挑战
- 网络虚拟化技术
- 数据中心二层网络技术
- 跨数据中心二层网络技术
- 非以太网网络技术
- 总结



# 云时代的网络





# 网络需求的变化

- 大型化/异构化
  - 更多的节点数
  - 存储网络节点
  - 跨数据中心
- 虚拟化
  - 同一物理服务器的虚拟机之间的通信
  - 跨物理服务器的虚拟机通信时的虚拟机标识
  - 虚拟机迁移



# 数据中心网络三大挑战

- 虚拟机的通信与迁移问题
  - 同一物理服务器内部的虚拟机通信、虚拟机通信时的标识、以及虚拟机迁移时网络配置属性的同步迁移问题。
- 提高数据中心内部二层以太网的传输效率问题
  - 实现二层以太网无环路的多链路传输，以及减轻网络设备节点内部MAC地址存储的压力。
- 在跨数据中心之间需要建立大二层互连网络的问题
  - 以支持跨数据中心的虚拟机迁移操作和分布式集群的跨数据中心任务调度



# Agenda

- 云时代的网络挑战
- 网络虚拟化技术
- 数据中心二层网络技术
- 跨数据中心二层网络技术
- 非以太网网络技术
- 总结

# 虚拟机通信及迁移的相关技术

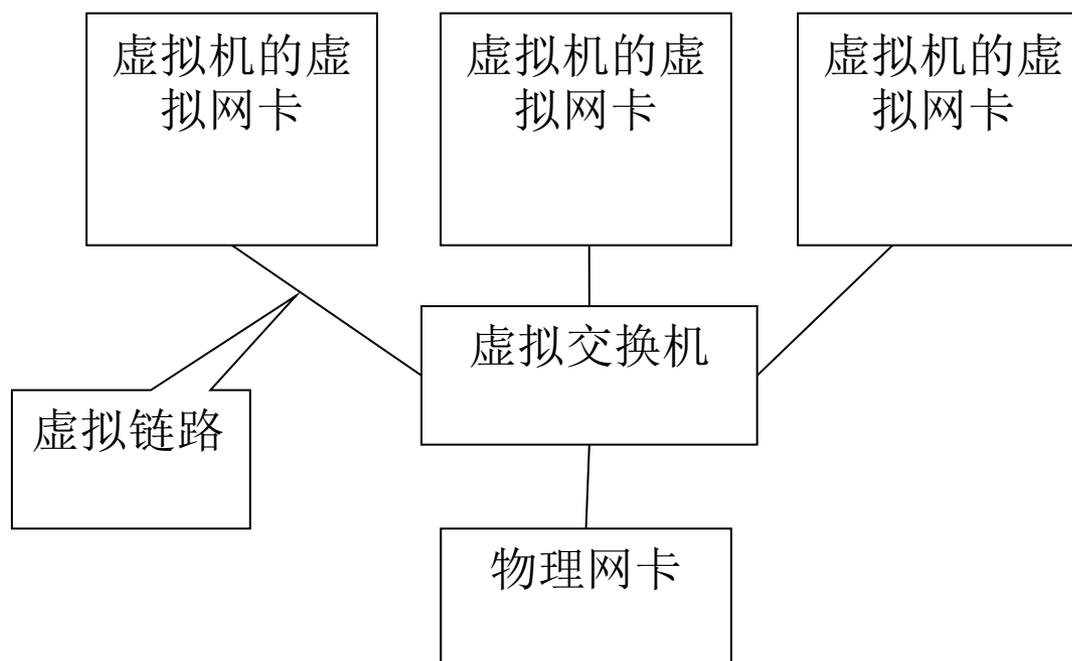
- 三个问题
  - 同一物理服务器内部虚拟机之间的通信问题
  - 虚拟机通信时的标识问题，即如何以虚拟机作为网络管理（比如流量控制、访问控制列表等）的基本单元
  - 虚拟机迁移时网络配置的跟随迁移
- 两类方案
  - 虚拟软件交换机
  - 更新硬件交换机

# 虚拟交换机

- 虚拟交换机（vSwitch）
  - In Hypervisor
  - 与虚拟网卡一起在服务器内部建立虚拟网络
  - 虚拟网卡上有虚拟的MAC地址作为虚拟机的网络标识
  - 将多个虚拟网卡绑定到物理网卡、并且可以对虚拟网卡的流量进行管控



# 虚拟交换机结构图





# 虚拟交换机功能

- 通过共享一个物理网卡，多个虚拟网卡能够对外表现为多块独立的网卡，有自己独立的MAC地址和IP地址
- 同一物理服务器内部虚拟机之间数据交换可以在物理服务器内部完成，不必经过外部的物理交换机
- 可以划分VLAN，实现不同VLAN间的数据隔离
- 可以进行流量管控



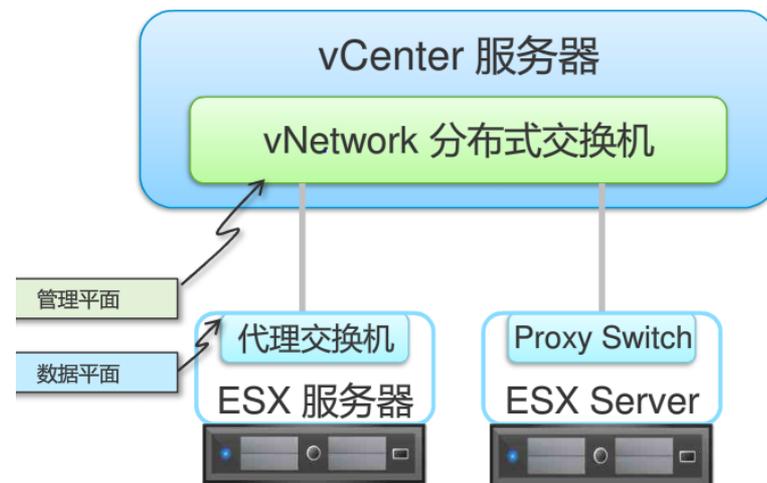
# 虚拟交换机的不足

- 扩展受限
  - 物理交换机的MAC地址表的容量压力
  - 服务器的资源开销
- 虚拟交换机的管理
  - 管理结构分离
    - 虚拟化管理软件只能管理虚拟交换机而不能管理网络上更多的物理交换机
  - 传统诊断工具故障定位困难
- Vmotion需要在不同的物理服务器上配置相同的vSwitch



# 分布式虚拟交换机

- vNetwork Distributed vSwitches (vDS或vNDS)
- 数据平面
  - 分布在各物理服务器之上的代理交换机，主要实现与各虚拟机的虚拟网卡vNIC以及物理服务器物理网卡的连接
- 控制平面
  - vCenter服务器上建立的分布式虚拟交换机控制平面
  - 第三方厂家的控制平台，如Cisco Nexus1000V的VSM(Virtual Supervisor Module)。

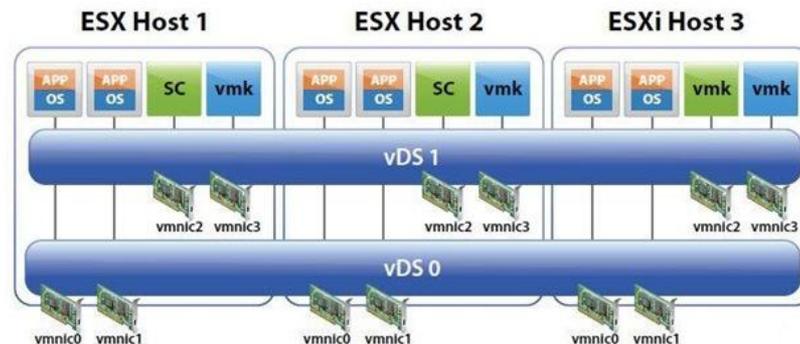
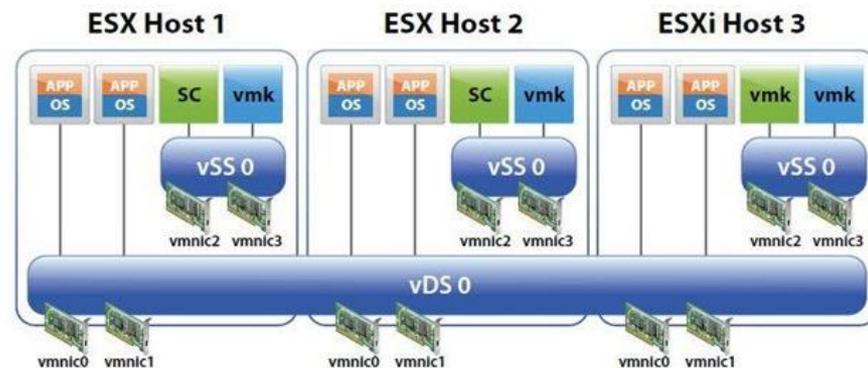


# 分布式虚拟交换机工作原理

- 状态管理
  - vCenter服务器集中管理
  - vDS运行基于本地缓存
- 虚拟机迁移
  - 分布式端口组
    - 集中式管理端口的集合，即端口属性的配置模板
  - 虚拟端口（DVPort）纳入至分布式端口组中，则完成虚拟端口的属性配置。
  - 虚拟端口（DVPort）绑定虚拟网卡（vNIC）
  - DVPort信息复制后就可以完成迁移

# 分布式虚拟交换机使用模型

- vSwitch和vDS一起使用
  - vSwitch连接VMkernel、ESX Service Console和ESXi Management
  - vDS实现虚拟机的网络连接
- 在每一台物理服务器上使用多个vDS





# 802.1Qbh BPE

- 动机：服务器和网络管理边界重叠
  - 802.1Qbh BPE（Bridge Port Extension）
  - 802.1Qbg EVB(Edge Virtual Bridge)
- BPE
  - Cisco: Fabric Extenders技术
  - IEEE: Standard for Local and Metropolitan Area Networks-Virtual Bridged Local Area Networks - Bridge Port Extension

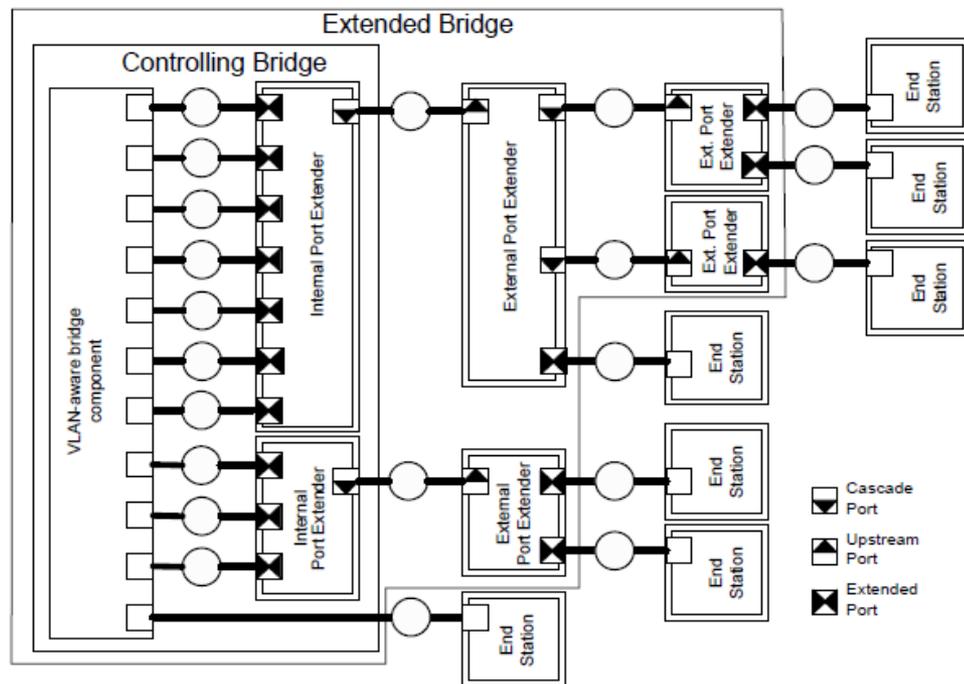


# BPE原理

- 主要思想
  - 把虚拟机的流量转发管理全部交给接入交换机，由交换机统一控制物理和虚拟网络接口
- 接入交换机新增功能
  - 控制平面，能够动态建立逻辑端口以对应新加入的物理或者虚拟接口
  - 数据平面，需要在数据包中加入扩展标记E-Tag，用来表征虚拟接口的信息
- 核心内容
  - 控制平面上的扩展桥（Extended Bridge）定义，包括一个控制桥（Controlling Bridge）加一或多个桥端口扩展（Bridge Port Extender）。
  - E-Tag定义，以识别扩展桥中各端口的标签格式。
  - 扩展桥的工作原理，即在桥端口扩展连接至控制桥时，如何在控制桥中形成逻辑端口，以及如何给数据流打标签进行数据的隔离。

# 扩展桥结构

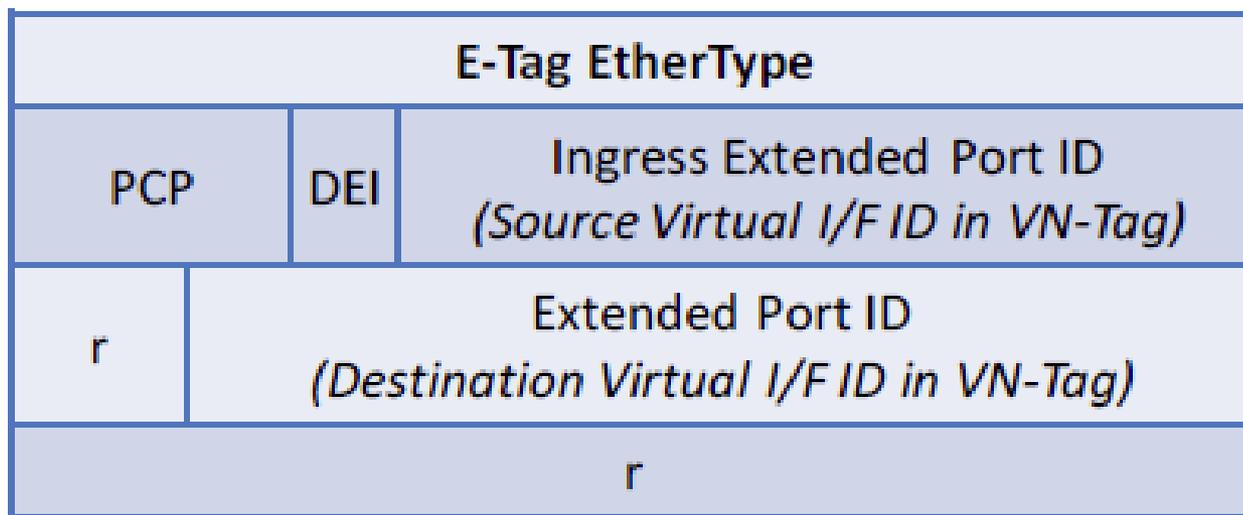
- 控制桥（Controlling Bridge）
  - 类似于交换机中带转发芯片的主控板，作为扩展桥中的控制设备，负责对所有数据流的转发进行控制
  - 可识别VLAN的桥组件
  - 内部端口扩展
- 外部桥端口扩展（External Port Extender）
- 上行端口
- 级联端口
- 扩展端口：





# E-Tag标签格式

- 将桥端口扩展中的虚拟端口映射为控制桥中的逻辑端口，
- 实现虚拟扩展端口之间点对点、点对多点的信息传输





# 扩展桥工作原理

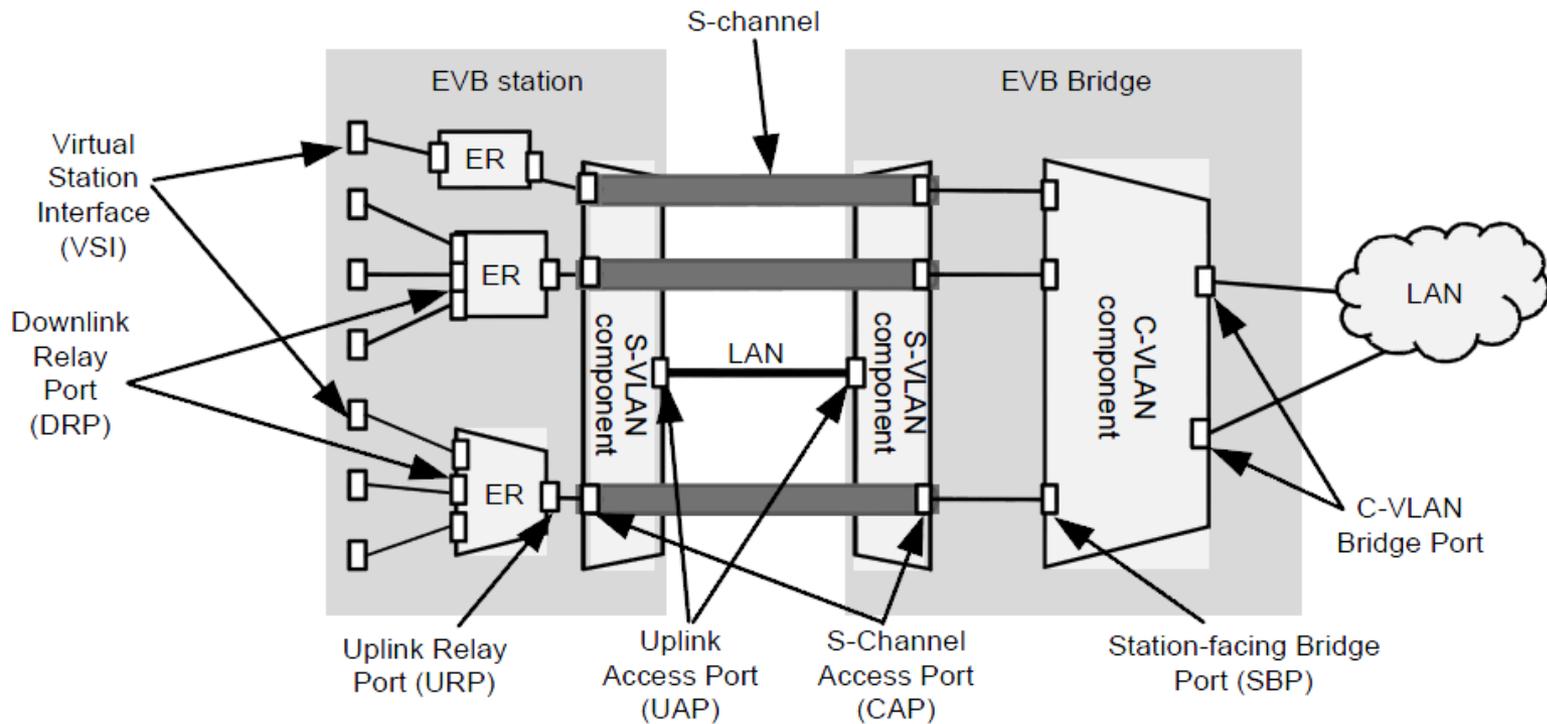
- 扩展端口的标识
  - 生成逻辑端口号对应扩展端口
- 流量的隔离
  - 逻辑端口会分配一个唯一的E-Channel标识
- 多播流的传送
  - 远端复制
  - 点对多点的E-Path



# 802.1Qbg EVB

- EVB
  - IEEE Standard for Local and Metropolitan Area Networks---Virtual Bridged Local Area Networks - Amendment: Edge Virtual Bridging
  - 802.1Qbh引入E-tag使得交换机必须更换硬件转发芯片，网络设备芯片厂商占据优势
  - HP等服务器厂家不愿受制于人
- 特点：在于充分利用现有的协议，尽量减少交换机设备功能的改动
  - 没有引入新的包格式，E-TAG的功能使用802.1Q中已提供的S-Vlan / C-Vlan Tag来实现
  - 它同时支持VEB和VEPA两种转发方式，分别对应于虚拟机流量的vSwitch内部转发和接入交换机聚合转发
  - 交换机设备记录了虚拟接口的信息用于虚拟机迁移，但管理和维护的权力仍在vSwitch上。
- 依赖于三个子协议，包括VDP、CDCP和ECP
  - 其中VDP和CDCP分别为VSI和S-Channel的发现和配置协议
  - 而ECP则是这两个协议的通信协议

# EVB架构





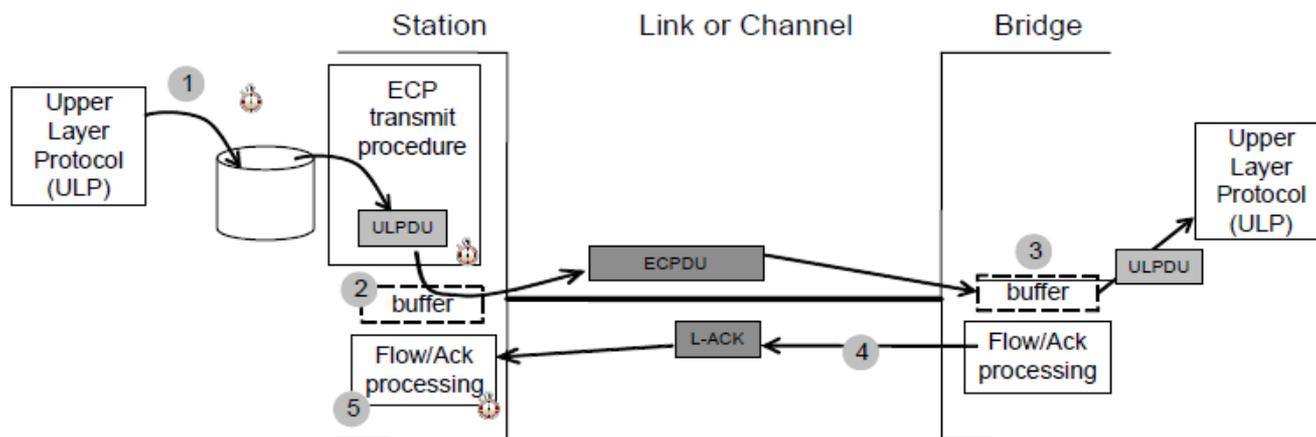
# EVB组件

- EVB Station虚拟化主机
  - 一个或多个ER（Edge Relay）： vSwitch
    - VSI(Virtual Station Interface)
    - 虚拟机连接至vSwitch的端口即是DRP（Downlink Relay Port）
    - vSwitch的上行端口URP（Uplink Relay Port）也就是主机中的物理网卡
  - 可选S-VLAN组件
    - 组件有两类端口，一类是连接ER的CAP（S-Channel Access Port），也就是S-Channel的端点，在此端口上进行S-Channel标签的处理工作，即打标签和去除标签；另一类是连接EVB Bridge的UAP（Uplink Access Port）。
- EVB Bridge
  - 需要有一个C-VLAN组件
  - 可选S-VLAN组件
  - C-VLAN组件面向EVB Station的端口称为SBP（Station-facing Bridge Port），连接其它交换机的端口称为C-VLAN Bridge Port。



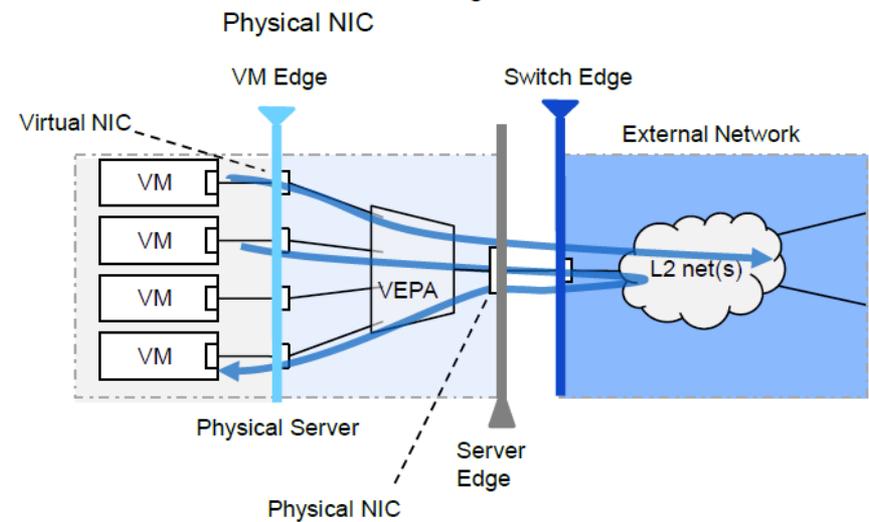
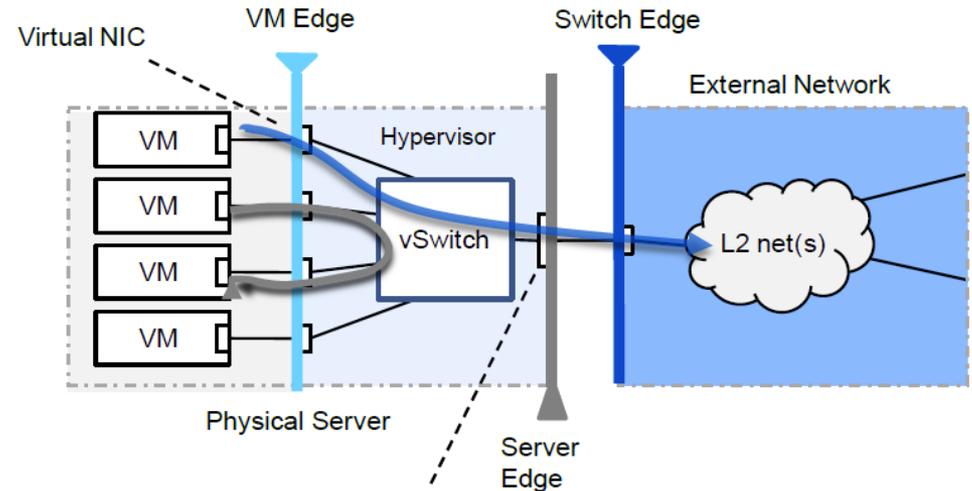
# EVB协议

- VDP (VSI Discovery and Configuration Protocol)
  - VSI实例的发现与配置
  - 将VSI的实例与EVB Bridge中SBP端口进行关联, 使得虚拟机的网络配置保存在EVB Bridge中
  - 虚拟机需要迁移时, 能快速将虚拟机的网络配置同步迁移
- CDCP (S-Channel Discovery and Configuration Protocol)
  - 用于在主机与交换机之间建立S-Channel
- ECP (Edge Control Protocol)



# EVB-ER工作模式

- VEB (Virtual Ethernet Bridge)
- VEPA (Virtual Ethernet Port Aggregator)





# Agenda

- 云时代的网络挑战
- 网络虚拟化技术
- 数据中心二层网络技术
- 跨数据中心二层网络技术
- 非以太网网络技术
- 总结

# 数据中心二层网络效率问题

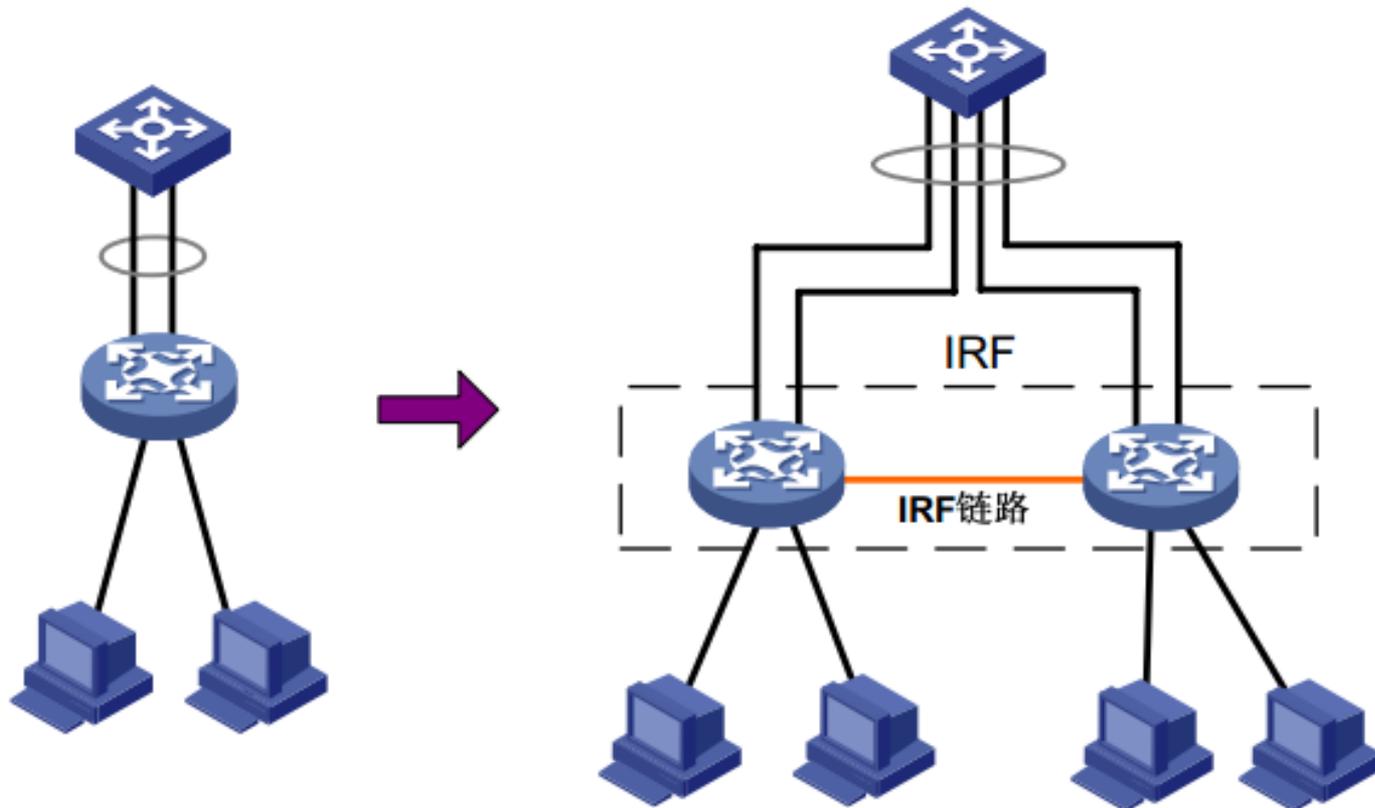
- 为什么是二层
  - 二层网络是虚拟机迁移的基础
  - 三层网络上的迁移很难做到应用透明
- STP的单路径之痛
  - 端到端环路避免环路，
    - 环路会造成广播风暴和整网瘫痪
  - 带宽的浪费
    - 链路阻塞：两点间只有一条路径可达
  - 延迟的增加： STP整网只有一棵生成树
  - MAC地址存储的压力
    - 传统的二层MAC地址没有层次化的概念

# 局部多路径—控制平面多虚一

- 多路径打响对STP第一枪
- 在控制平面上对多网络设备进行虚拟化
  - 状态复制，在逻辑上让它们看来就像是一台设备，因此这种方式又叫做控制平面多虚一
  - 既然表现为一台设备，那么不管其内部怎么互连，整个大网的STP需求都能得到满足，而构成这台大虚拟设备的物理设备之间的多路径也可以同时发挥作用了，带宽的利用率自然也就提高了
  - 还简化了管理、提高了服务的可靠性
- 核心：控制平面的主备模式和数据平面的跨设备链路聚合
  - 让一台设备的控制平面负责控制和管理网络协议的运行，如VLAN协议、ARP协议、各种路由协议和组播协议等，形成一个逻辑设备
  - 链路聚合使此逻辑设备与每个接入层设备之间只有一条逻辑链路连接，使得整个网络逻辑拓扑为无环的树状连接结构。
  - 只是将少量核心节点作为虚拟化对象，所以它的多路径只能在局部产生效果。
  - IRF和VSS/vPC

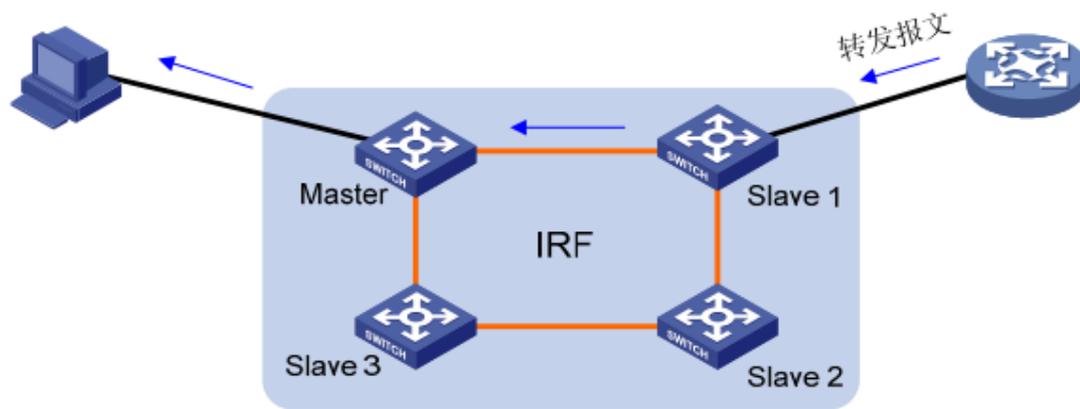
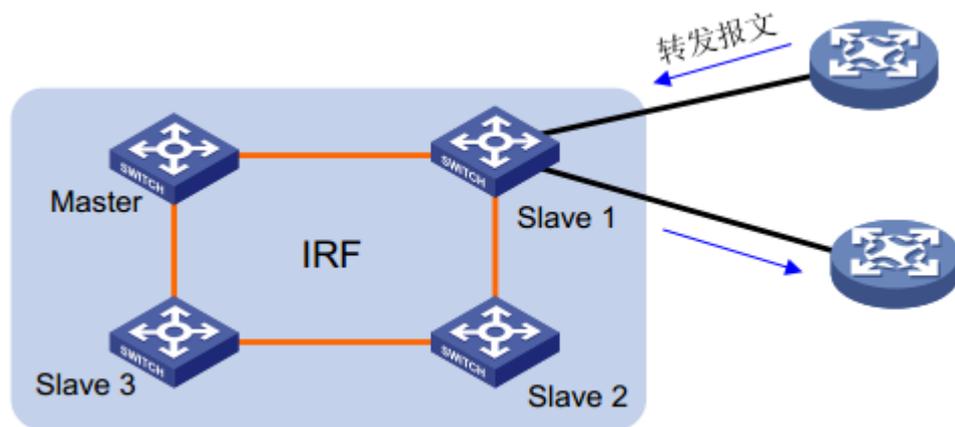
# IRF

- H3C 的IRF（Intelligent Resilient Framework），智能弹性架构



# IRF架构

- Master/Slave架构
  - 拓扑收集
  - 选举Master
- IRF专用链路和端口
  - IRF内部报文的转发



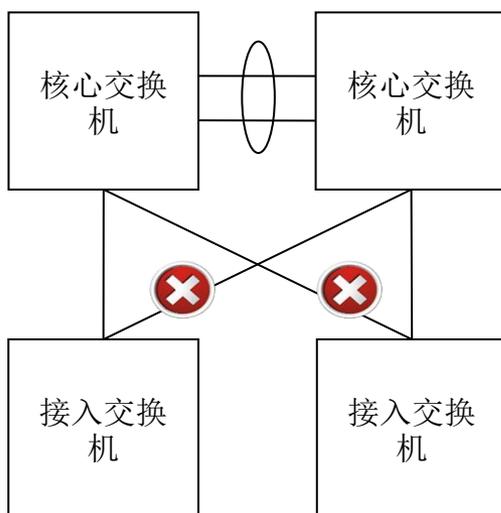
# VSS

- VSS（Virtual Switching System）
  - 将多个交换机虚拟成一台交换机的问题
  - 思科公司的Catalyst 6500系列交换机、7600系列路由器中已经被支持
- VSS将使两台物理 Cisco Catalyst 6500系列交换机作为单一逻辑虚拟交换机运行，称为虚拟交换系统 1440（VSS1440）
  - 从管理上看，只有一个管理平面，从而简化了对设备的管理。同时，两台设备相互冗余，增强了不间断通信能力
  - 将系统带宽容量扩展到1.4Tbps

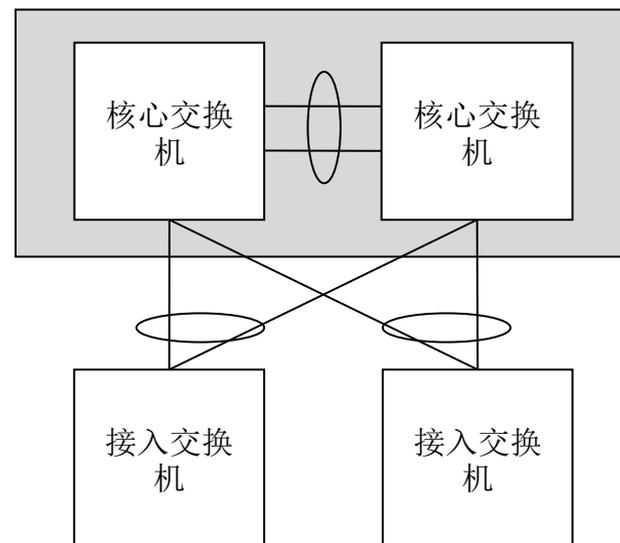
# VSS架构

- MEC（多机箱以太通道）互连
  - 跨机箱链路捆绑
  - 无环路的拓扑结构

传统二层/三层交换机



VSS 1440 物理视图





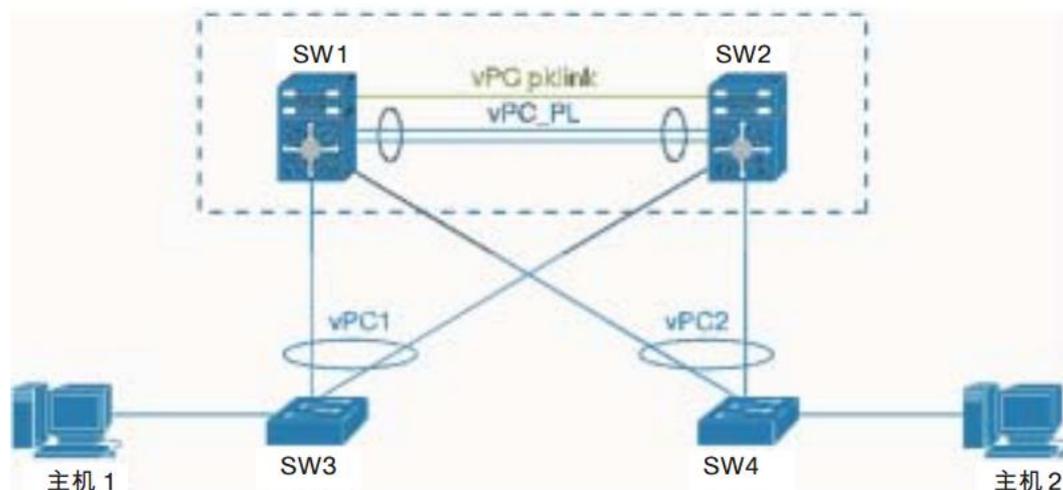
# Vpc(virtual Port-Channel)

- 专门针对二层多虚一的方案
  - VSS的一个精简技术
  - 保留了通过MEC实现的跨机箱链路捆绑
  - 减弱了虚拟控制平面两台交换机之间精确同步，也就是说两台交换机控制平面会同时活跃，不再有主备之分



# vPC网络组成

- vPC 对等交换机
- vPC 对等链路
- vPC 对等设备保持激活链路
- vPC 成员端口
  - 虚拟端口绑定组



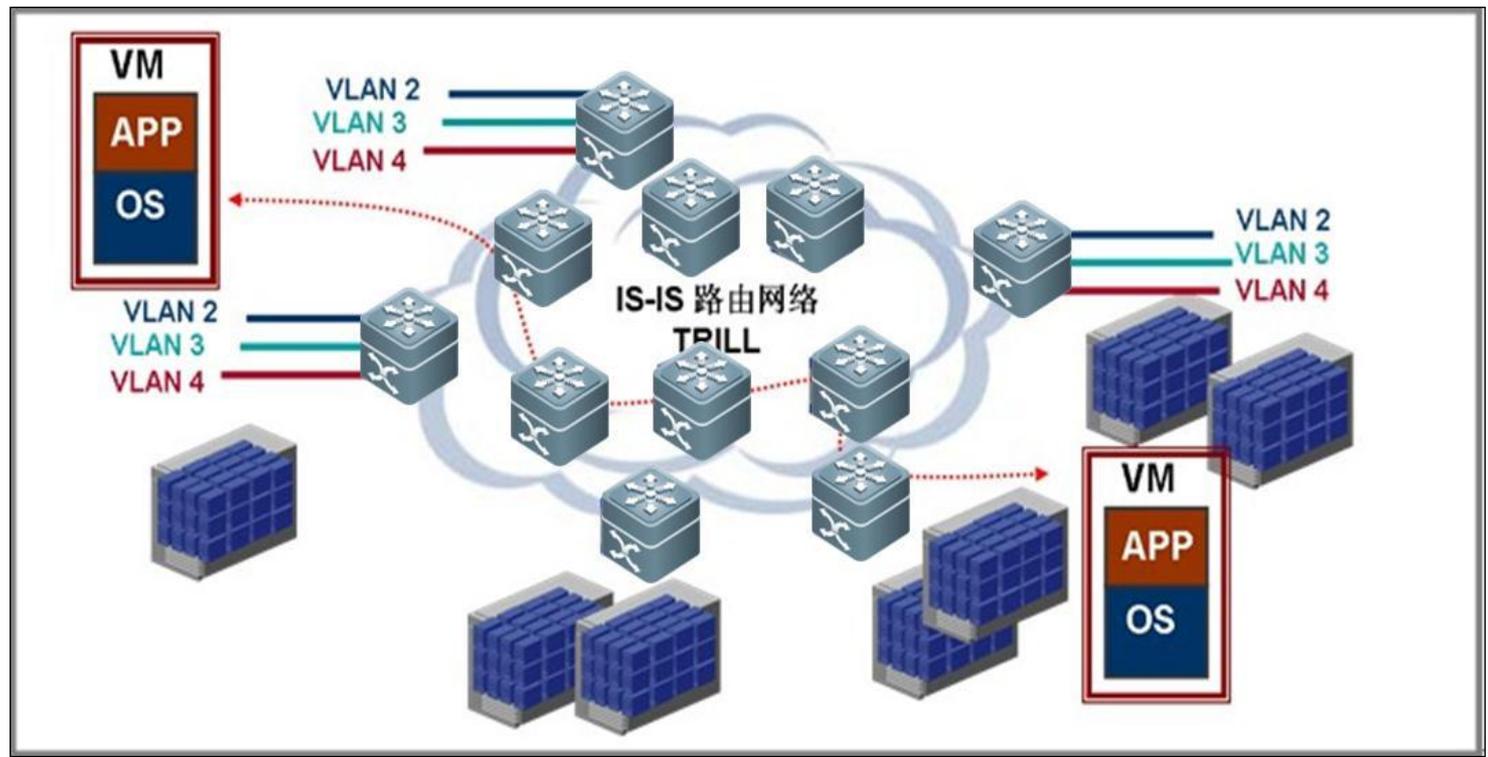
# 控制平面虚拟化技术的评价

- 主-从架构
  - 由“主”交换机负责控制平面的功能，并且把策略等复制到“从”交换机上，并对外显示出一台逻辑交换机的特点。
- 优点
  - 增加带宽容量
  - 简化管理
  - 提高可靠性
- 缺点
  - 有限的多路径技术，只能针对局部节点来虚拟化，所以它的扩展性潜力有限。
  - 内部节点间大量的协议状态同步工作会消耗很多系统资源
  - “双活”问题，造成地址混乱。

# 全面的多路径—数据平面多虚一

- 全面取代STP是趋势
- 二层转发技术+三层路由的链路状态发现协议
  - 实现整网无环路转发，
  - 既保持了二层配置的灵活性
  - 又像三层网络一样更好地支持网络融合和规模扩展
- 整个交换机群都可以被视作一个单交换机，因此这种技术也叫做数据平面多虚一
- IETF: TRILL预标准规范
  - STP协议作者
- 思科: FabricPath。
- IEEE: SPB

# TRILL (Transparent Intercon... Lots of Links, 多链接透明互联)





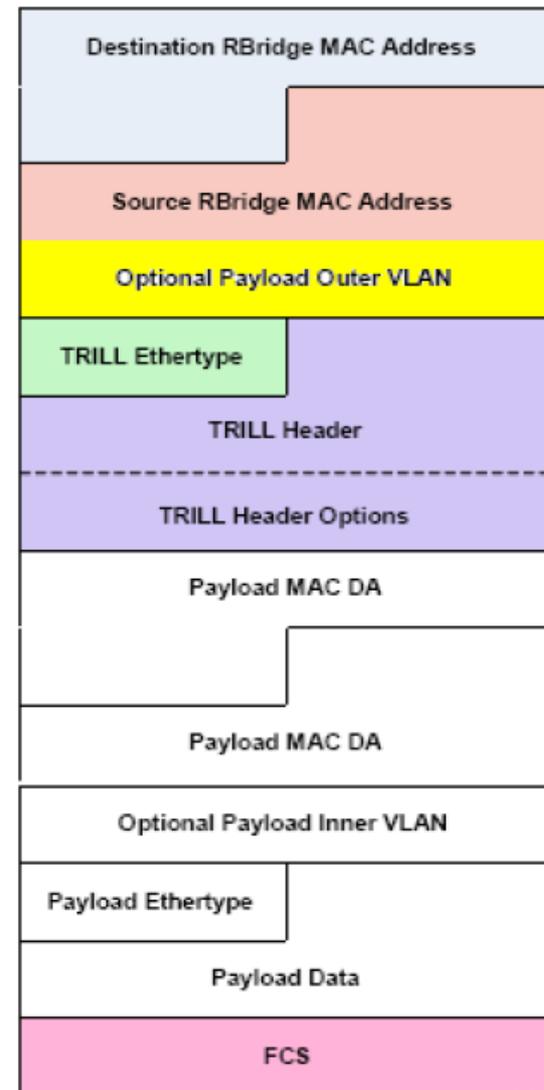
# TRILL网络核心设备

- Rbridge
  - 具有路由转发特性的网桥设备
  - 扩展ISIS协议去识别TRILL域（TRILL Campus）的拓扑
  - 使用最短路径树算法生成从该RBridge到达TRILL Campus里的各个RBridge的路由转发表
    - TRILL路由表
  - 通过ESADI(End Station Address Distribution Information, 终端地址交互协议)交互学习各自的Mac地址，
    - 获得了终端Mac究竟属于哪个RBridge的信息



# TRILL数据帧格式

- 源和目的RBridge Mac
- TRILL报头
  - 版本(ver)
  - 多播标记 (M)
  - 选项长度 (Option Length)
  - TTL (Hop Count)
  - 出口RBridge (Egress RBridge Nickname), 入口RBridge (Ingress RBridge Nickname)
- Outer-VLAN, Inner-VLAN,
  - 分别对应于为TRILL转发使用的VLAN和进入TRILL Campus时生成的VLAN





# TRILL原理

- 数据帧转发方式
  - 单播数据帧
  - 广播帧和多播帧
    - 是链路协议避免二层环路的重点
    - TRILL报头中引入了TTL字段，这样即使出现了环路问题TRILL报文也不会被无限转发引起风暴。
    - 多个分发树（Distribution Tree）的规则，类似于STP的生成树，有一个Root Bridge
- 路由信息生成
  - ISIS收集拓扑
  - 生成路由表和分发树



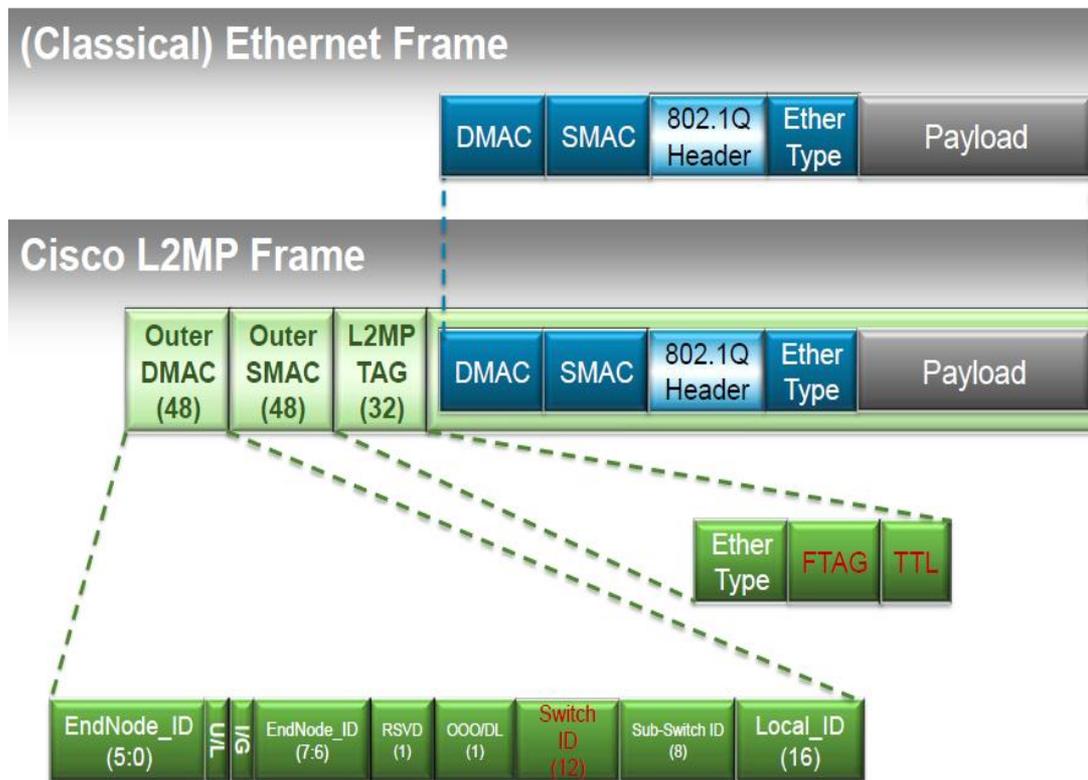
# FabricPath

- TRILL正式标准化之前，思科推向市场的“Pre-Standard”技术
- 基本内容与TRILL相同
- 增加了“基于会话的MAC地址学习”、“Vpc+”和“多重拓扑”等高级功能
- FabricPath架构与TRILL完全兼容



# FabricPath帧格式

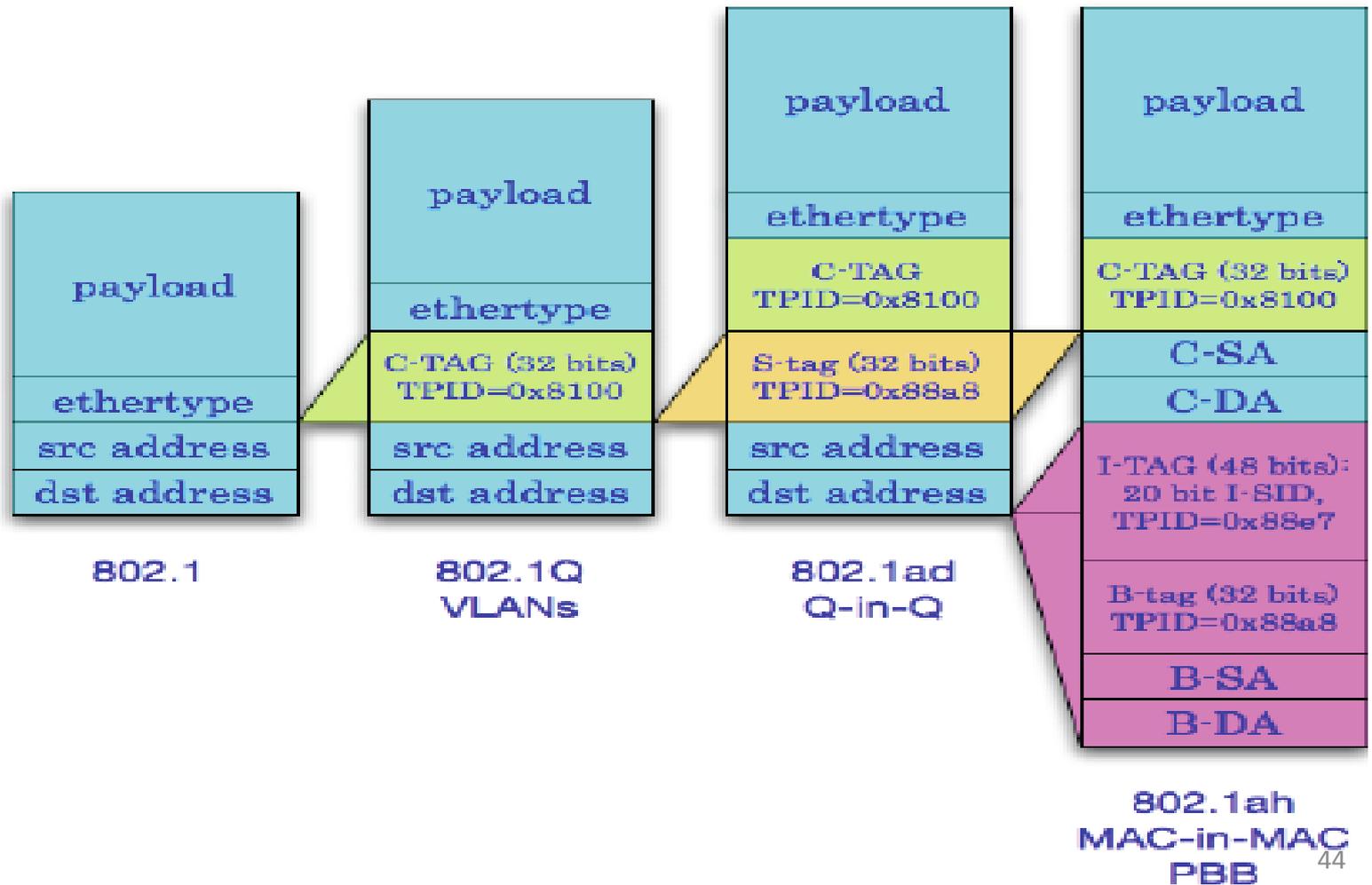
- 源地址(Outer SMAC)和目的地址(Outer dMAC)
  - switch ID标识一台交换机的身份，也是节点之间进行路由寻址的依据
- FTAG（Forwarding TAG）标示不同的多播树Graph



# SPB

- SPB( Shortest Path Bridging)最短路径桥接，是IEEE为解决二层网大规模扩展而提出的预标准。
  - 起源于 PBB (Provider Backbone Bridging)
    - 为运营商城域以太网定义了一整套MAC-in-MAC的转发机制，仍然需要依靠传统的STP进行环路避免和转发控制。
  - 2009， 802.1Qay PBB-TE(Provider Backbone Bridge Traffic Engineering)
    - 通过手工方式配置一堆指定路径取代STP的自动收敛。
- 分为SPBV(VLAN Q-in-Q)和SPBM(MAC-in-MAC)两个部分，目前主要用到的是SPBM。
- 同TRILL相比， SPB的特点是纯软件实现
  - 对标准二层以太网设备的兼容性，根据标准的以太网头进行标准的以太网转发，因此不需要更新转发芯片去支持

# SPB协议帧变迁



# SPB特性

- 依靠外层MAC做传统的二层转发
  - 而外层以太网报头中的源B-SA和目的B-DA则指明了SPB区域边缘的设备
  - I-SID被用来区分多个拓扑，I-SID信息在数据报文中以I-TAG（外层Ethernet报头中的VLAN Tag）形式携带
- 环路抑制机制
  - 节点间路径的双向对称性



# 数据平面虚拟化技术评价

- 优点
  - 提升网络带宽，降低传输延迟
  - 最短路径转发和多路径的流量负载分担
  - 减少了广播和MAC地址表的大小
  - 配置、维护和管理简单
  - 收敛速度快
  - 支持存储融合的网络结构
- 缺点
  - 技术尚未成熟
  - 增加了交换机成本
  - VLAN边界问题：在TRILL网络中跨越多个VLAN操作的开销很大
  - 路由计算的压力：主要是SPB的问题
  - 端到端阻塞管理复杂



# Agenda

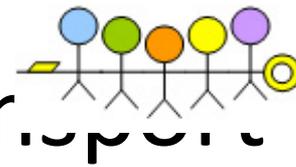
- 云时代的网络挑战
- 网络虚拟化技术
- 数据中心二层网络技术
- 跨数据中心二层网络技术
- 非以太网网络技术
- 总结



# 跨数据中心二层网络

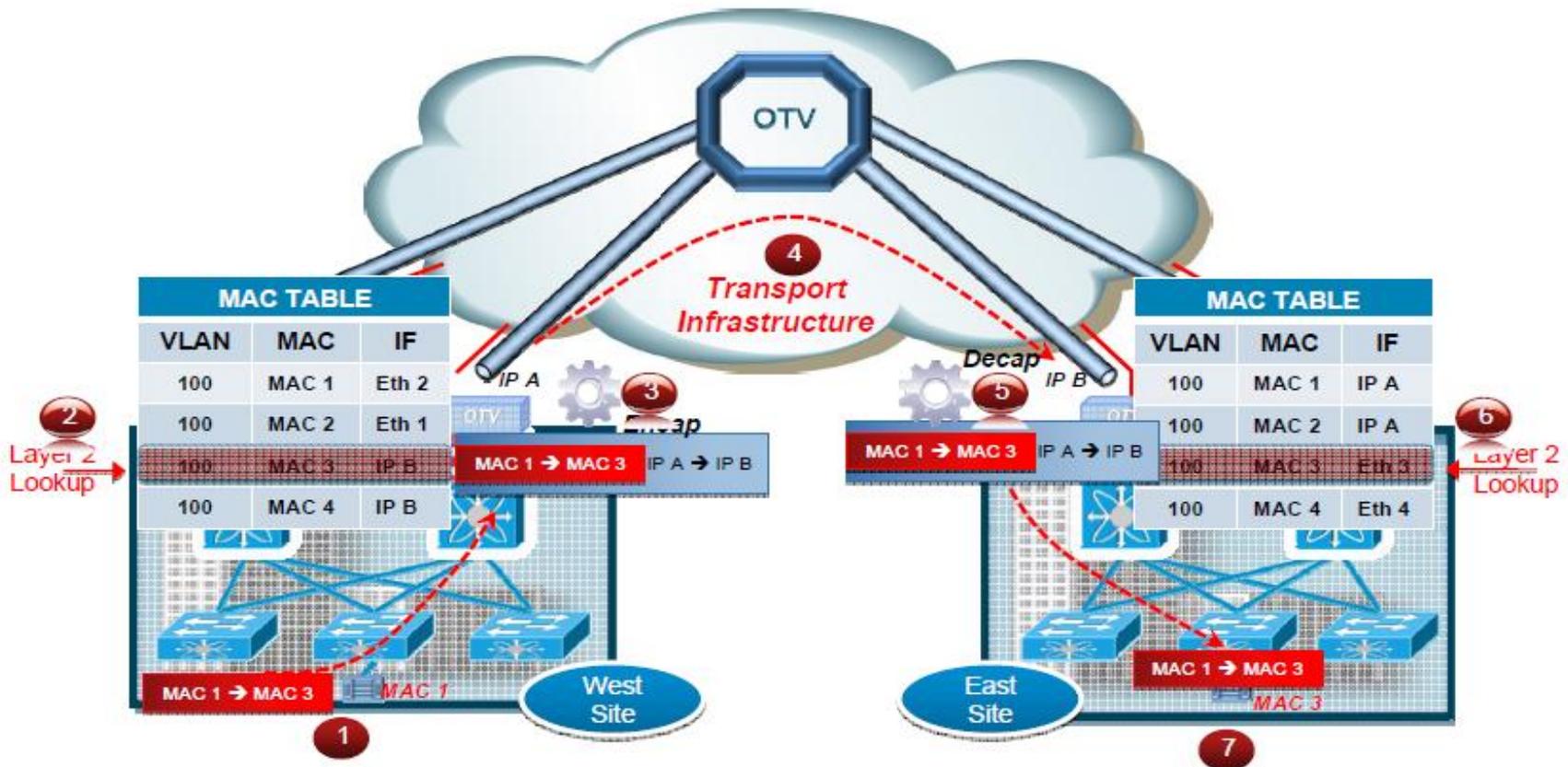
- 目的：在不同跨数据中心链路上构建虚拟大二层VLAN，以覆盖不同地理位置上分布的虚拟化和集群服务器
- 光纤直联的互联方案
  - 星型和环型拓扑
- 承载于MPLS网络的VLL或VPLS方案
  - MPLS VLL（MPLS Virtual Leased Line）：实现两数据中心的点到点二层互联。
  - VPLS（Virtual Private LAN Service）：实现多个数据中心的点到多点二层互联。
- 承载于IP网络的VLL Over GRE或VPLS Over GRE方案
  - GRE（Generic Routing Encapsulation，通用路由封装）协议能对某些网络层协议（如IP和IPX）的数据报文进行封装，使这些被封装的数据报文能够在另一个网络协议（如IP）中传输。

# 思科OTV（Overlay Transport Virtualization）

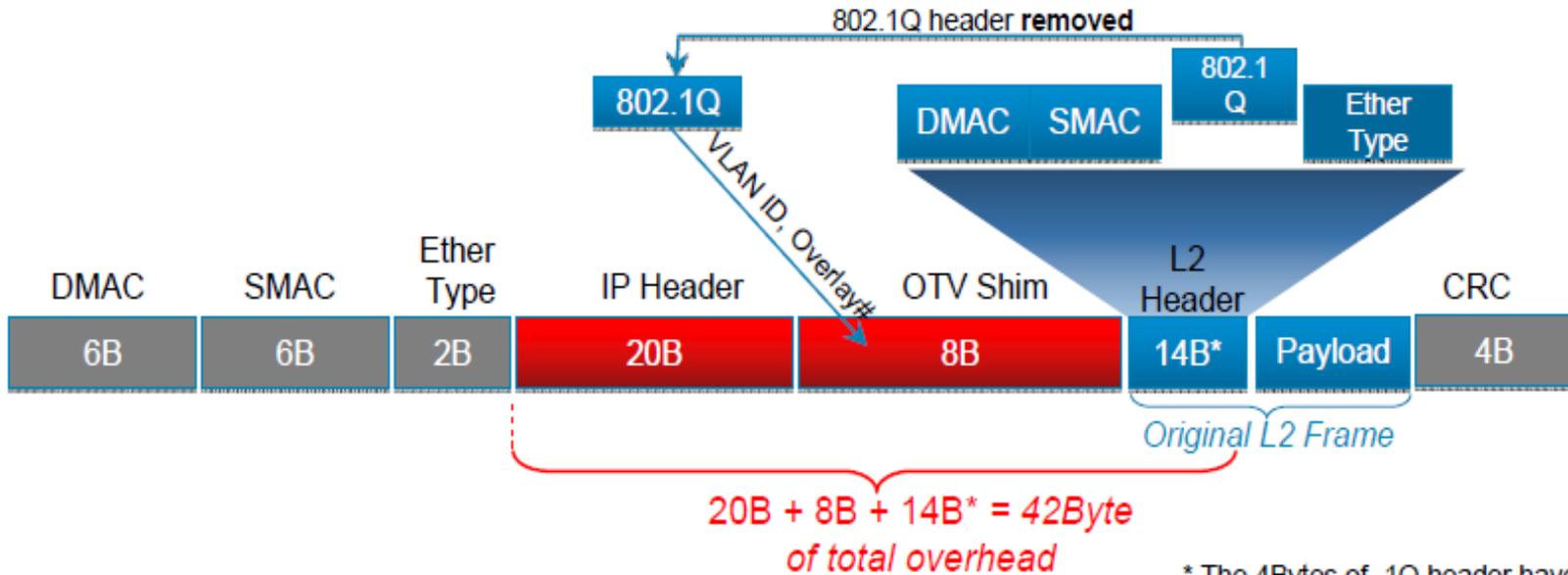


- 采用了隧道技术穿越三层网络实现二层网络的互连互通
- 实现不依赖于任何特殊的底层网络设施
- 核心思想是采用“MAC in IP”的传输方式

# OTV工作原理



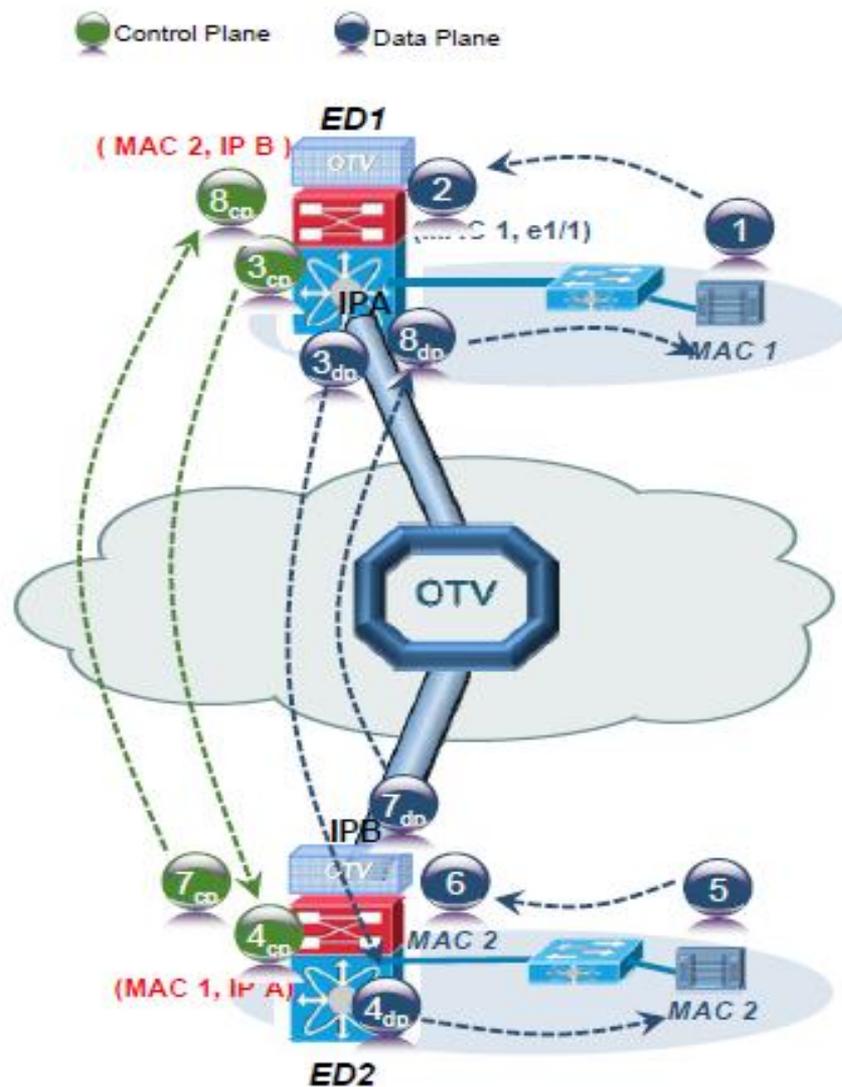
# OTV帧格式



\* The 4Bytes of .1Q header have already been removed



# OTV帧转发方式





# Agenda

- 云时代的网络挑战
- 网络虚拟化技术
- 数据中心二层网络技术
- 跨数据中心二层网络技术
- 非以太网网络技术
- 总结



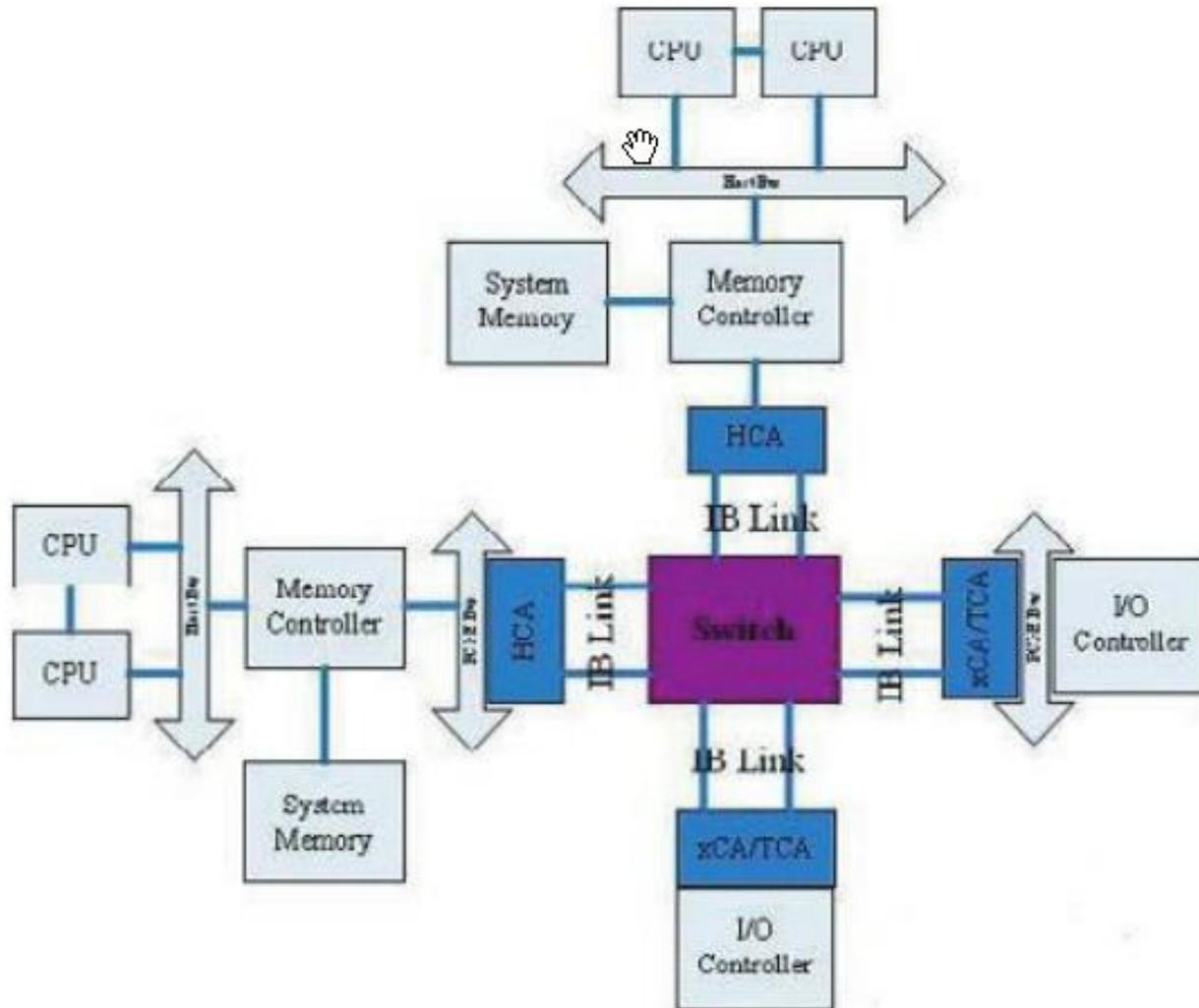
# 非二层网络方案

- 其他传输领域的先进技术
  - 高性能计算领域的InfinBand
- 创新的网络传输模式，让控制平面和数据平面完全分离
  - OpenFlow
- 二三层网络的架构完全打碎重组，简化为单层的高效网络
  - Juniper的QFabric

# 他山之石，可以攻玉—InfiniBand

- 服务器端的高性能互联
  - 服务器互连、服务器与存储互连、存储网络在内的统一的互联结构体系支持，
  - 高带宽、低延迟
- 历史
  - 网络的交换思路被引入IO传输技术当中，诞生了IB
  - IO的作用被PCIE取代
  - 内存访问的优势使它高性能计算机节点互连中大放异彩
  - 数据中心对带宽和延迟的渴求

# IB硬件架构





# IB硬件组成部分

- IB信道适配器，用于IB结构同其他设备之间的连接
  - 主机信道适配器（HCA）
  - 目标信道适配器（TCA）
- IB交换机
- IB线缆
  - 1x，4x，8x和12x
  - 1x可提供2.5Gps的全双工连接
- 子网管理器



# IB协议栈



起始定 界符	LRH	GRH	BRH	ERH	Payload	I-Data	ICRC	VCRC	结束定 界符
-----------	-----	-----	-----	-----	---------	--------	------	------	-----------

LRH: 本地路由报头  
GRH: 全局路由报头  
BRH: 基本传输报头  
ERH: 扩展传输报头  
I-DATA: 即时数据  
ICRC/VCRC: 不变/可变CRC



# 传输层工作队列

- 工作队列(Work Queue): IB执行数据通信任务的处理机制。
- 发送和接收工作队列成对出现
- 传输事务对应工作队列单元(WQE, Work Queue Element)
  - 信道适配器 (Channel Adapter) 负责解释每个WQE, 并执行指定的操作
  - 在处理完成后, 信道适配器从队列中移除WQE, 并在完成队列中生成该WQE的完成队列单元(CQE, Completion Queue Element), 表示完成用户的该项请求工作



# IB优势分析

- 高带宽
  - 单对差分线上基本的串行信号速率是2.5 Gbps（单倍速, SDR, Single Data Rate），
  - 双倍速DDR（Double Data Rate）、四倍速QDR（Quad Data Rate）
  - 并行通道（Lane），三种并行通道（1x, 4x, 12x），
- 低延迟
  - 直通转发（Cut-Through）的机制
  - kernel bypass
- 高效率
  - IB采用基于信用（credit）的端到端流控机制来确保连接的完整性，数据包极少丢失，并消除了丢包而带来的重发延迟，因而极大的提升了效率和整体性能。
- 多虚拟通道支持
  - IB提供多个独立的基于信用的端到端流控的通道，目前IB提供了16级可映射到16个服务层的虚拟通道。
- 可扩展性好
- RDMA支持
  - RDMA（远程直接存储访问）技术和transport offload技术,使对远程节点内存和存储的访问不需要经过远程主机节点的处理，从而减少了处理器的负担

# IB缺点分析

- 硬件价格昂贵
  - 仅HCA卡就需要4000元到5000元人民币，而一个交换机端口也达到3000元以上，这远远超过普通以太网网络设备的价格
- 接口尺寸大
  - IB线缆接口尺寸明显大于以太网接口
- 缺乏虚拟化支持
  - 相对于以太网设备纷纷加入了虚拟化支持，IB这方面的工作才刚刚开始。
- 大规模互连并不成熟
  - 用于子网连接的IB路由器尚未实用，因此纯IB跨子网的实现还需要等待。

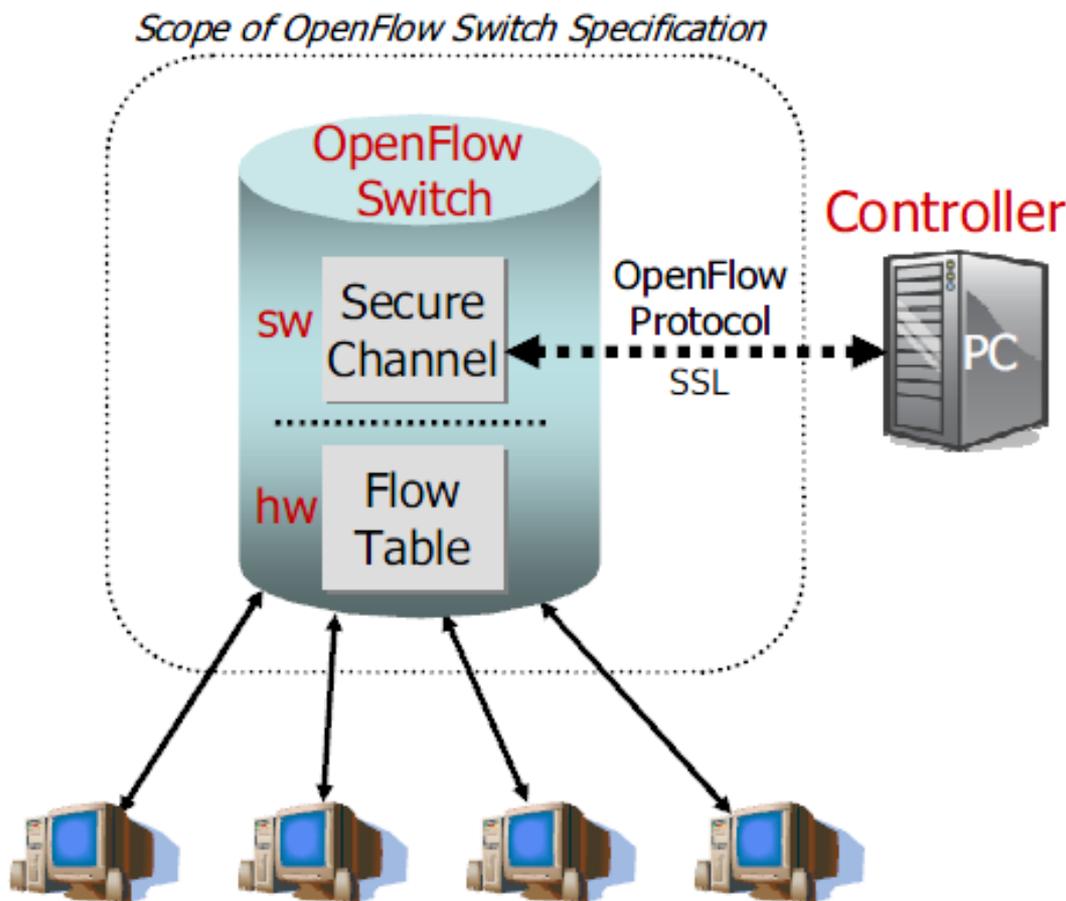


# 创新网络架构OpenFlow

- 斯坦福大学的Nick McKeown教授在2008年4月ACM Communications Review上发表的一篇文章《OpenFlow: enabling innovation in campus networks》
- 控制转发分离架构
  - 将控制逻辑从网络设备盒子中引出来
  - 可以对其进行任意的编程从而实现新型的网络协议、拓扑架构而无需改动网络设备本身



# OpenFlow概念架构



- 流表
  - 流表项
  - 包头域
  - 活动计数器
  - 执行动作 (actions)
- 控制器
- 安全通道

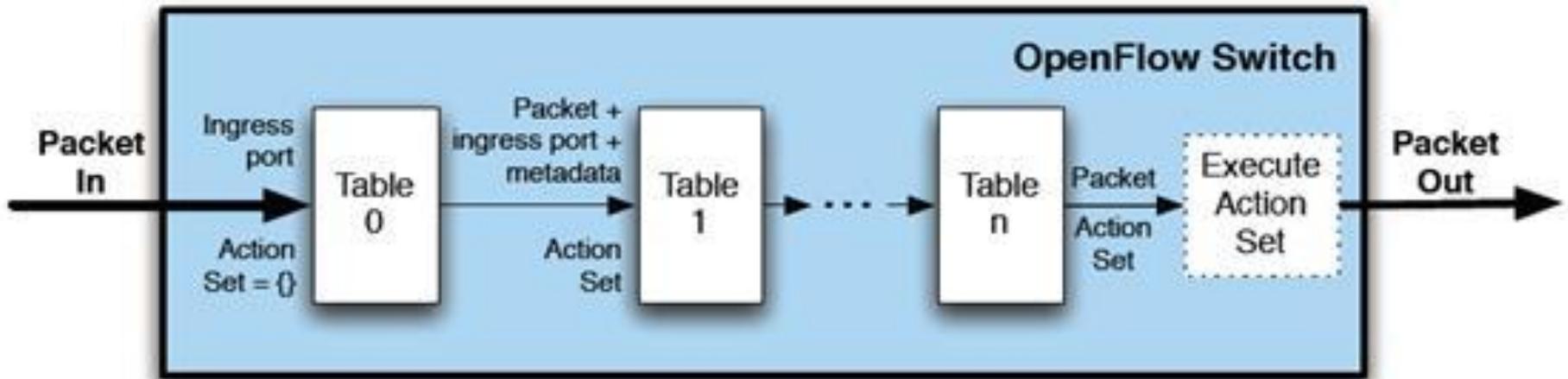


# 基于流表的处理动作

- 流表是交换机进行转发策略控制的核心数据结构。
  - 交换芯片通过查找流表表项来决策
  - 对每一个包进行查找，如果匹配则执行相关策略，否则通过安全通道将包转发到控制器
- 必备动作
  - ALL 转发到所有出口（不包括入口）
  - CONTROLLER 封装并转发给控制器
  - LOCAL 转发给本地网络栈
  - TABLE 对要发出的包执行流表中的行动
  - IN\_PORT 从入口发出
- 可选动作
  - NORMAL 按照传统交换机的2层或3层进行转发处理。
  - FLOOD 通过最小生成树从出口泛洪发出，注意不包括入口。
  - 入队（Enqueue）将包转发到绑定到某个端口的队列中。

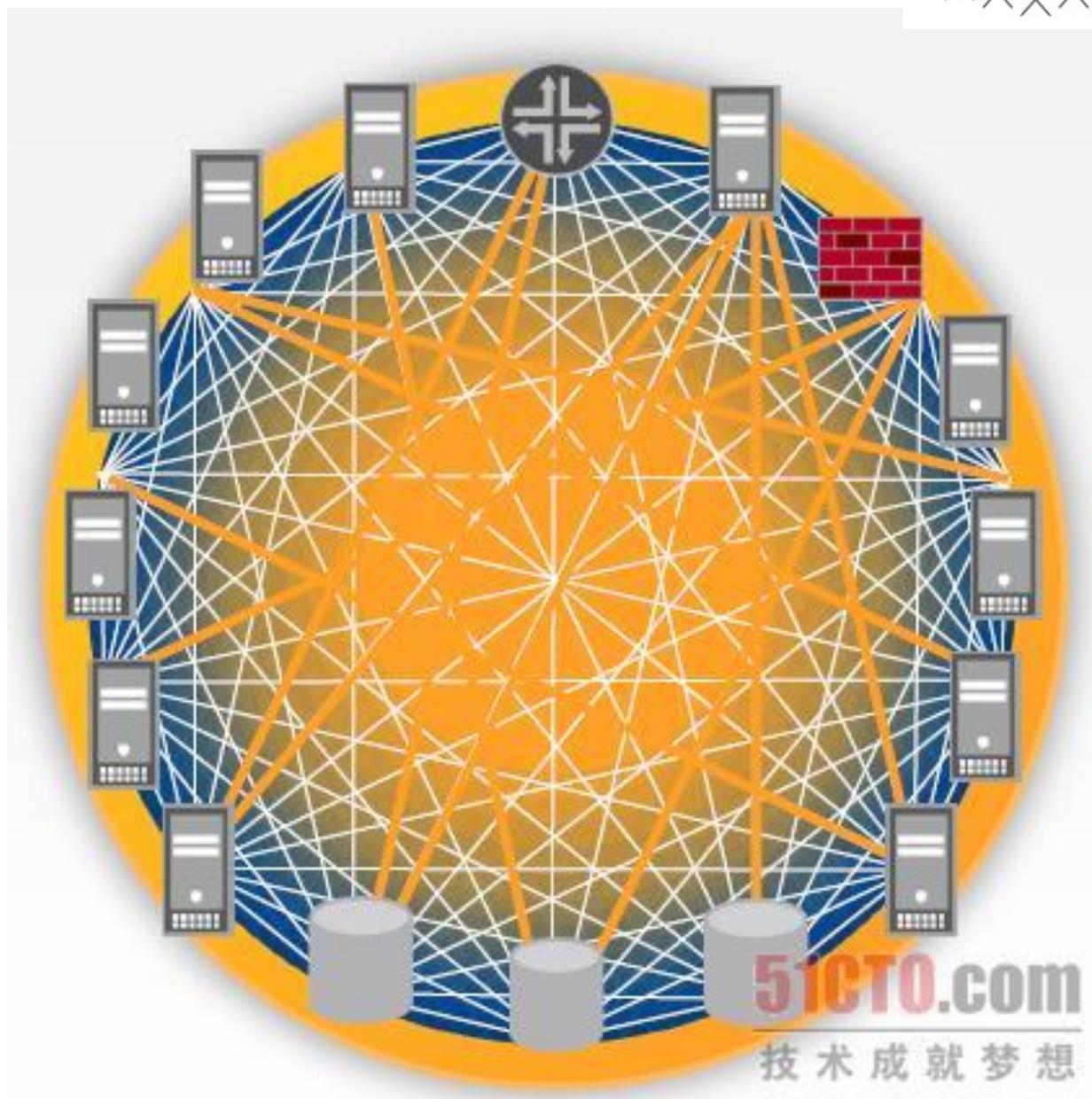
# OpenFlow1.1

- 多级流水化流表
  - 流表处理开销过大



# QFabric

- 瞻博网络——“层云计划”（Project Stratus）
  - 私有化技术
- 扁平化，全互连
  - 任意两点间都是一跳
- 数据、控制、管理3个平面分离
  - 数据平面，任意2点之间都是以40G传输速率相连，传输时延相同。
  - 控制平面是通过带外连接的
    - QF/Director通过1GE端口以星形连接方式连接到所有的QF/Node和QF/Interconnect
  - 管理平面也是通过带外连接的，管理平面N=1





# Agenda

- 云时代的网络挑战
- 网络虚拟化技术
- 数据中心二层网络技术
- 跨数据中心二层网络技术
- 非以太网网络技术
- 总结



# 三大核心挑战→技术

- 虚拟机的通信与迁移问题
  - vSwitch & Vds
  - BPE&EPB
- 提高数据中心内部二层以太网的传输效率问题
  - IRF & VSS
  - TRILL/FabricPath
  - SPB
- 在跨数据中心之间需要建立大二层互联网络的问题
  - OTV
- 替代以太网的新技术
  - InfiniBand
  - OpenFlow
  - Qfabric



# 未来趋势

- 融合
  - 二层三层转发融合
  - 物理层一致性
  - 存储网和传输网的同构化
- 分化
  - 标准分化：服务器厂商v.s. 网络设备商
  - 私有协议大行其道
  - 吃掉整个蛋糕—整体化方案
- 虚拟化厂家的话语权

# 谢谢

## Q & A