中山大學 软件工程学院

SUN YAT-SEN UNIVERSITY SCHOOL OF SOFTWARE ENGINEERING

SSE316：云计算技术
Cloud Computing Technology

陈壮彬

软件工程学院

https://zbchern.github.io/sse316.html

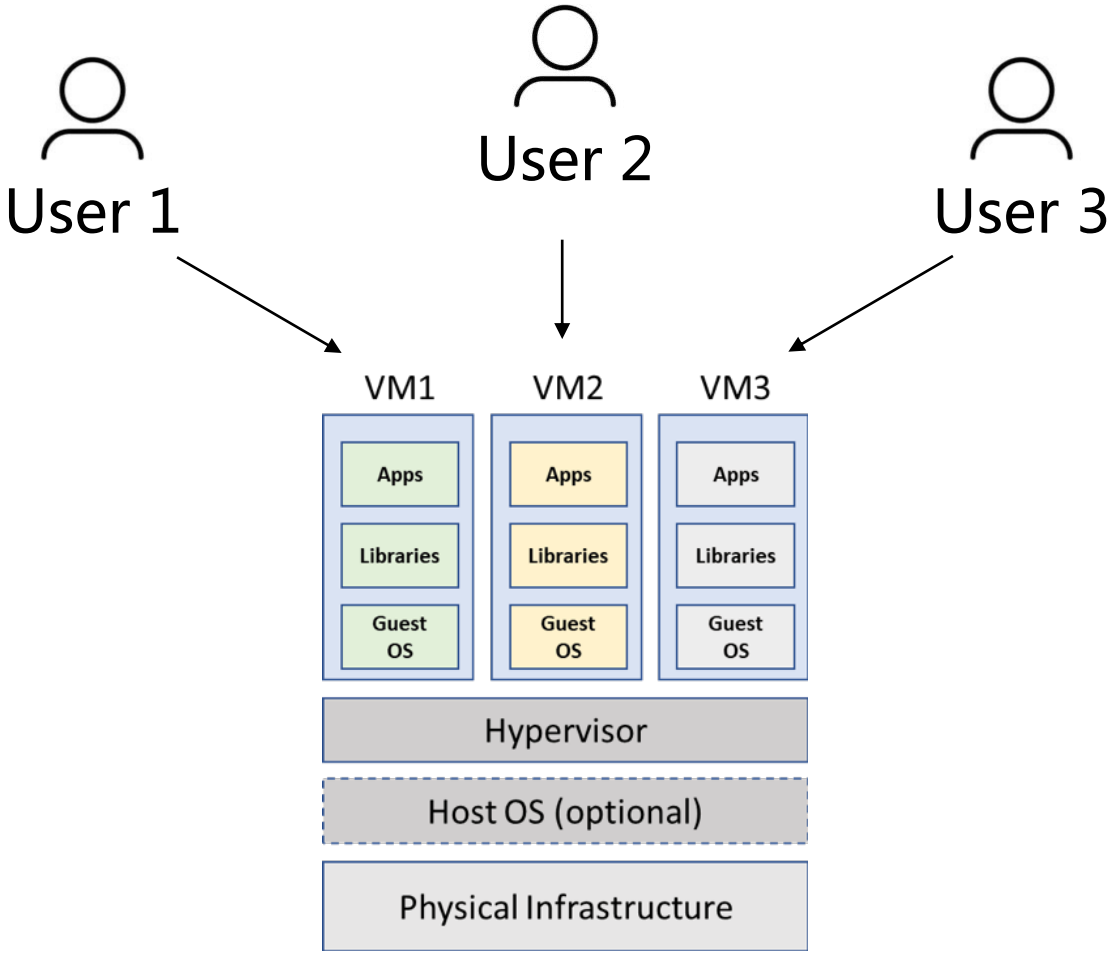
网络虚拟化技术

- 数据中心网络虚拟化
- 虚拟局域网（VLAN）
- 虚拟可扩展局域网（VxLAN）
- 网络功能虚拟化（NFV）

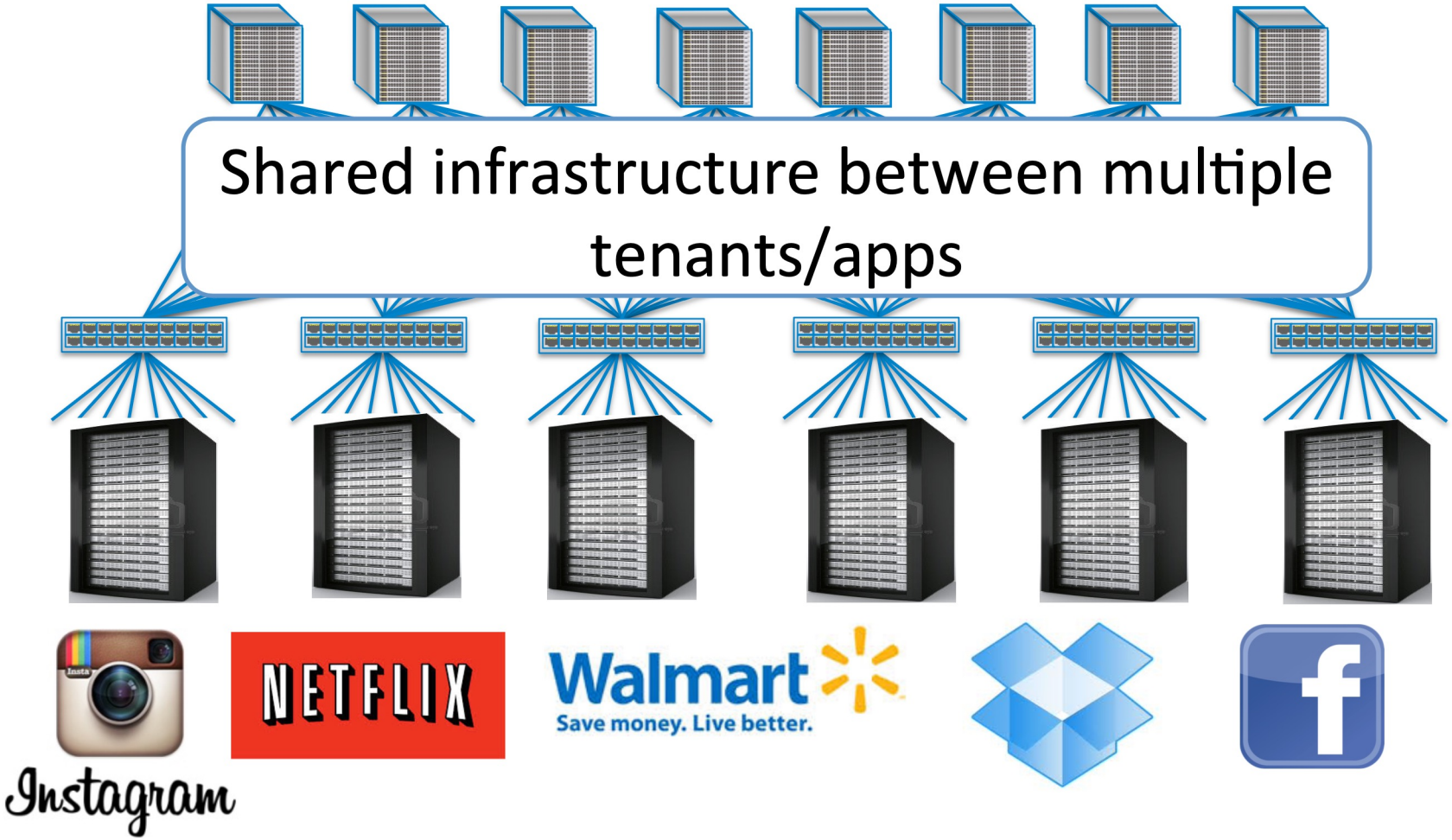
网络虚拟化技术

- 数据中心网络虚拟化
- 虚拟局域网（VLAN）
- 虚拟可扩展局域网（VxLAN）
- 网络功能虚拟化（NFV）

计算虚拟化


User 1
User 2
User 3
VM1
VM2
VM3
Apps
Libraries
Guest OS
Apps
Libraries
Guest OS
Apps
Libraries
Guest OS
Hypervisor
Host OS (optional)
Physical Infrastructure

云数据中心多租户

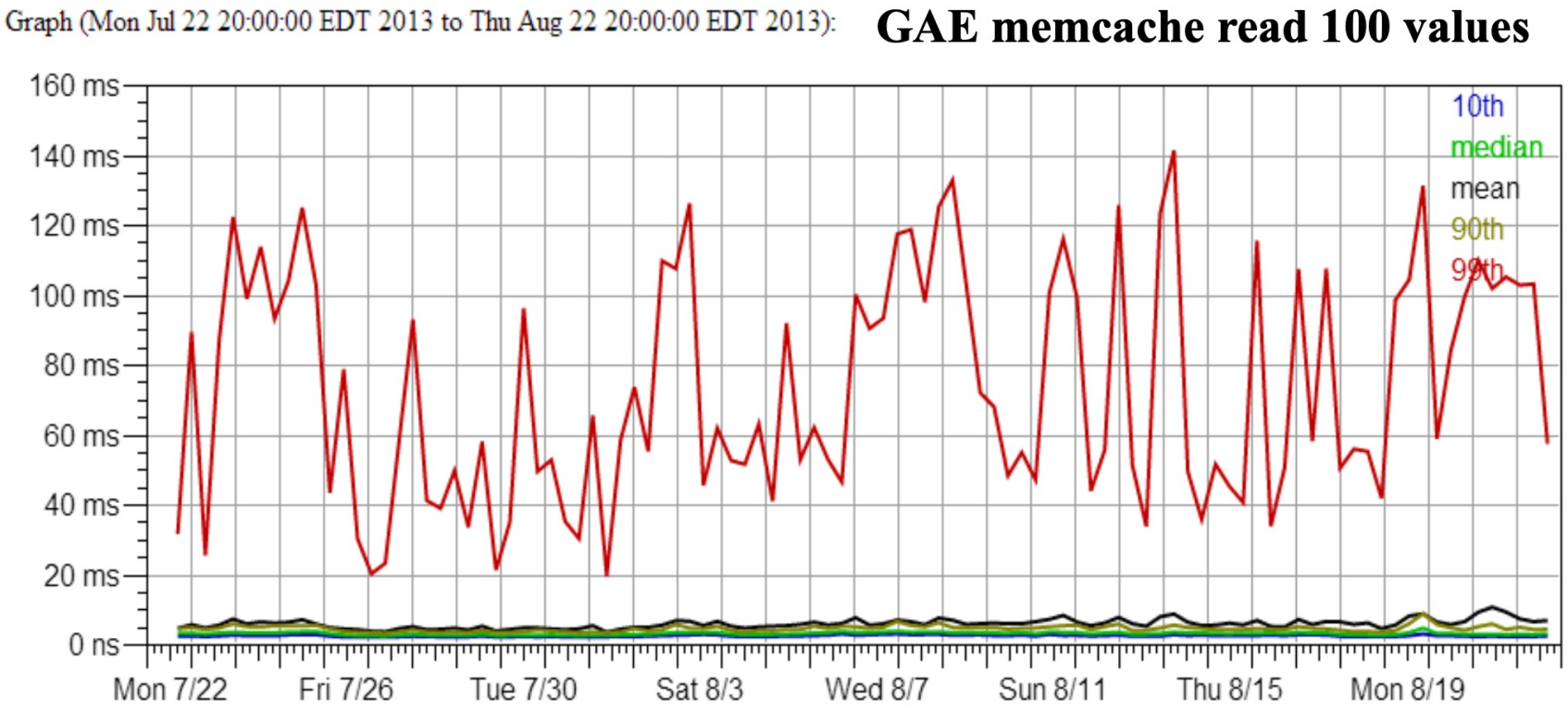

Shared infrastructure between multiple tenants/apps
Instagram
NETFLIX
Walmart
Save money. Live better.

网络性能难以预测

- 虽然云计算提供商保证了CPU和内存等资源的分配，但他们对网络资源不提供任何保证

- 由于缺乏网络保证，租户无法预测其应用程序性能的下限

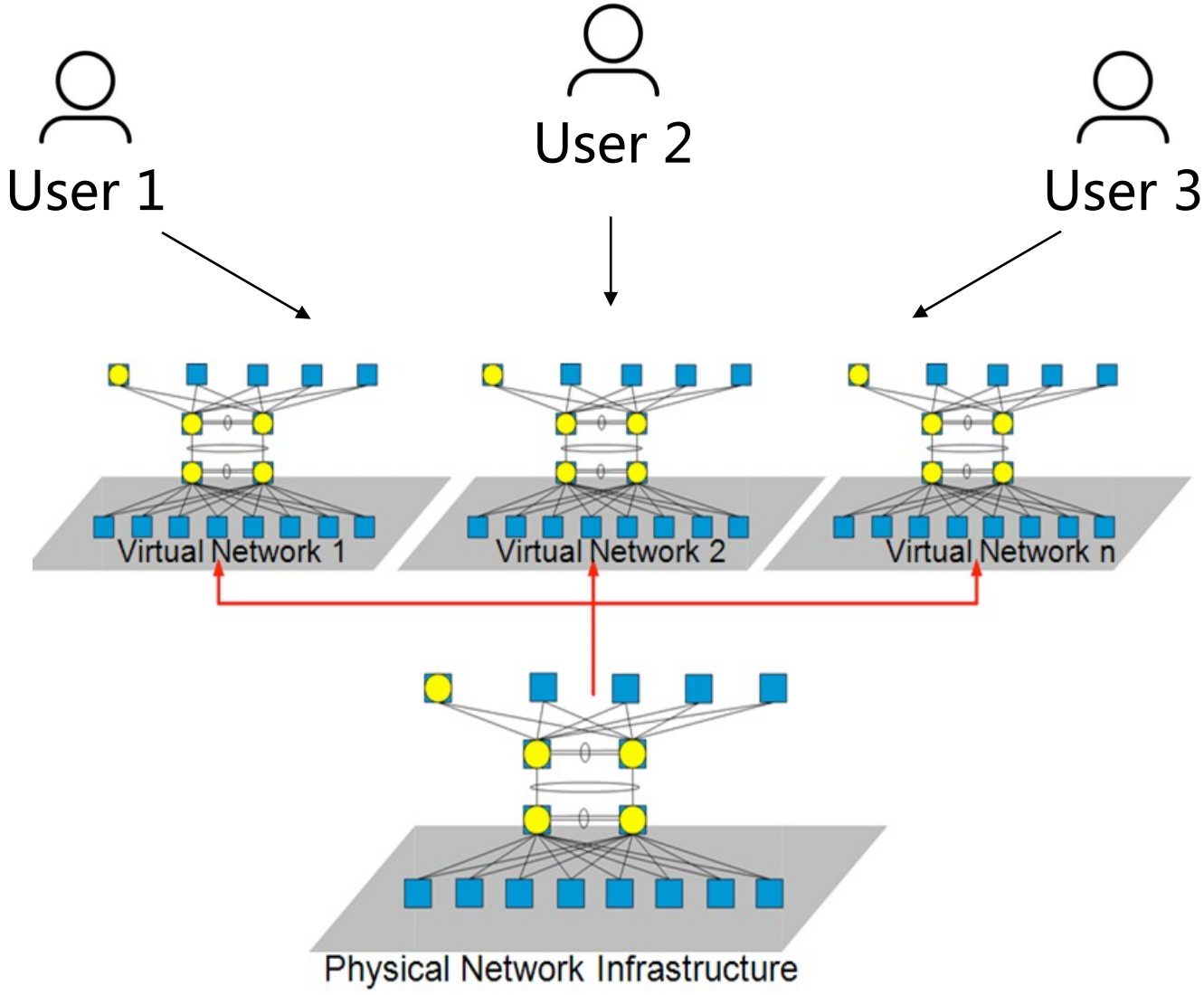
网络性能难以预测



Unpredictable performance, esp. at the tail

拥塞导致性能不可预测

网络虚拟化

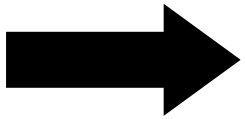

User 1
User 2
User 3
Virtual Network 1
Virtual Network 2
Virtual Network n
Physical Network Infrastructure

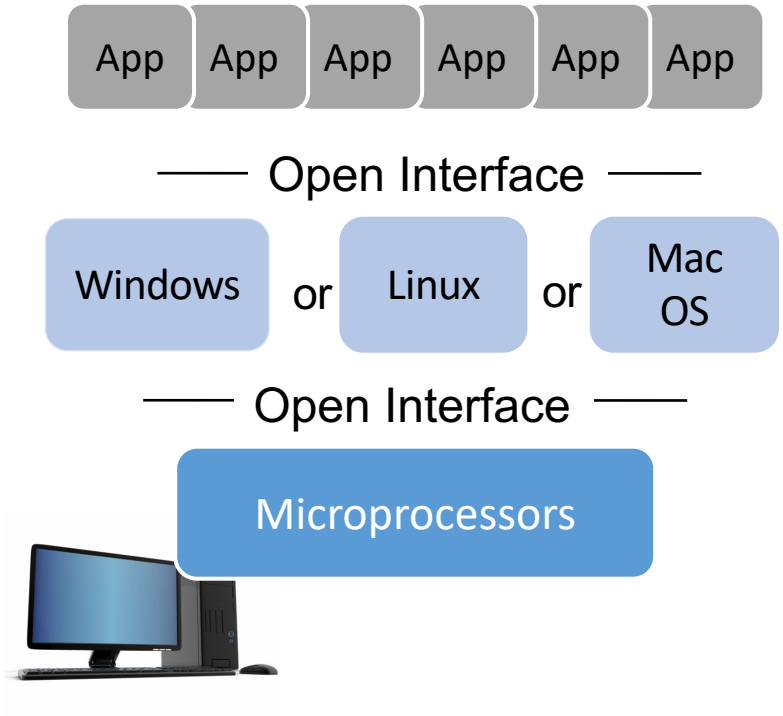
网络虚拟化定义

- 网络虚拟化是在单个共享物理基础设施之上创建多个逻辑隔离的网络的过程
- 它实现了网络资源的抽象，允许多个租户或应用程序共享相同的物理网络资源，且不会相互干扰
- 网络虚拟化将逻辑网络与底层物理基础设施解耦，从而简化了管理，提高了灵活性和资源利用率

计算资源虚拟化

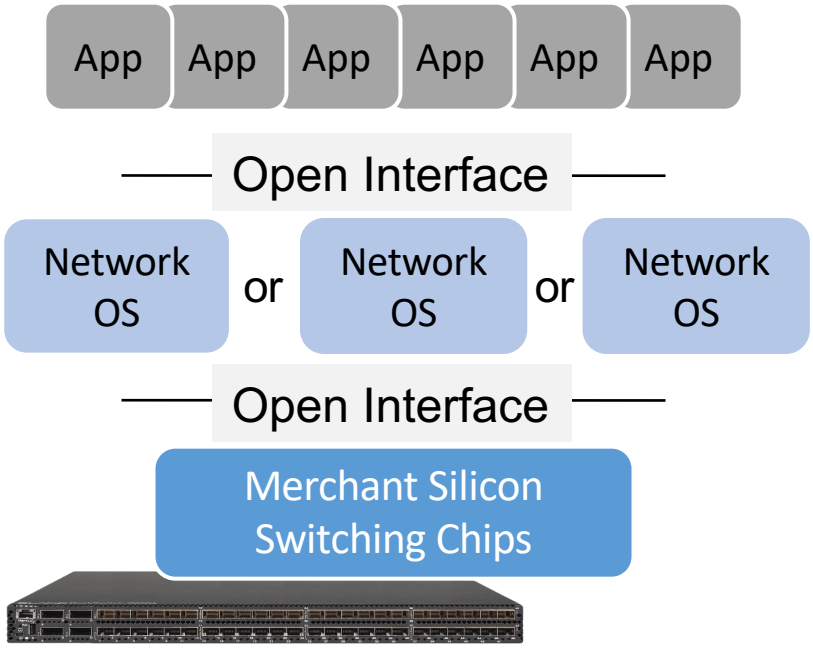

Specialized Applications
Specialized OS
Specialized Hardware




App
App
App
App
App
App
Open Interface
Windows
or
Linux
or
Mac OS
Open Interface
Microprocessors

网络资源虚拟化


Specialized Applications
Specialized OS
Specialized Hardware


App
App
App
App
App
App
Open Interface
Network OS
or
Network OS
or
Network OS
Open Interface
Merchant Silicon Switching Chips

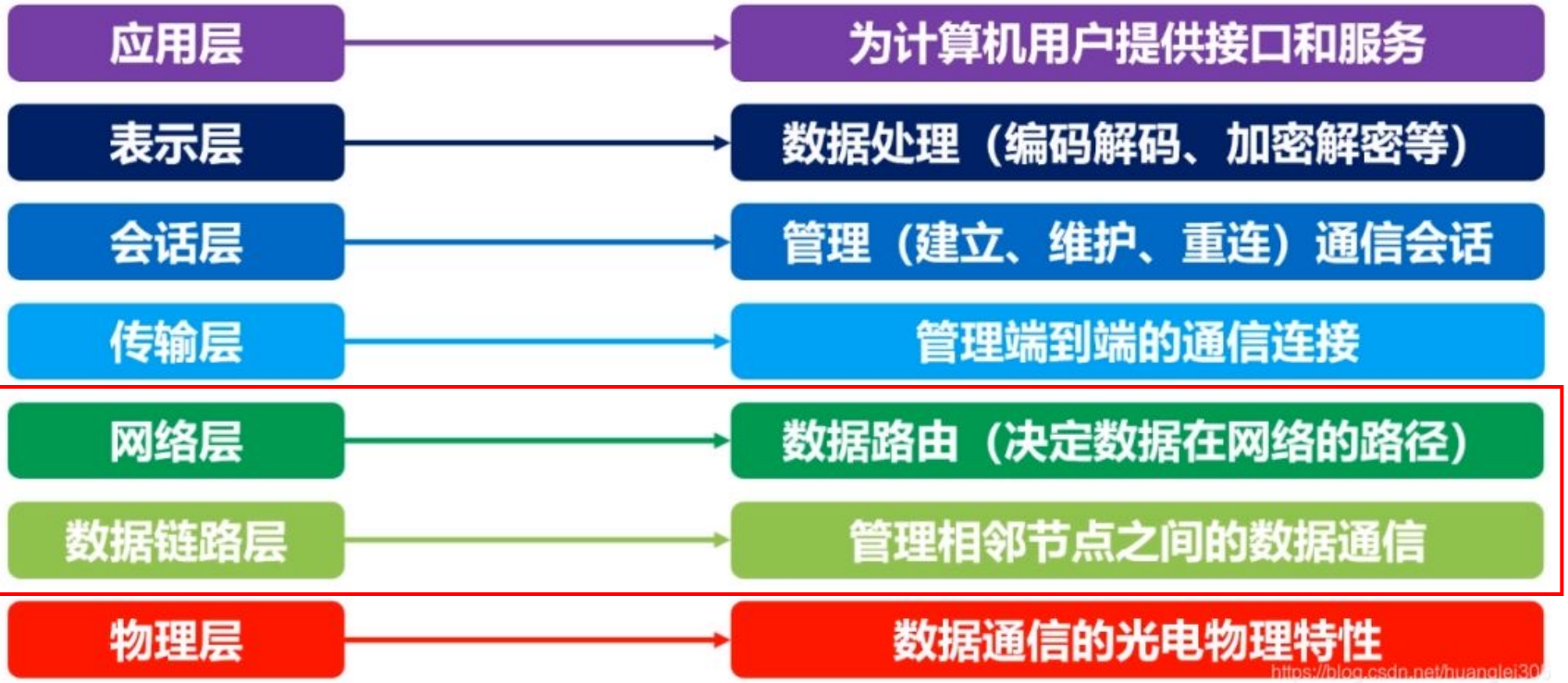
网络虚拟化的优点

- 多租户：支持多个客户或应用程序共享同一基础网络架构，而不会影响安全性或性能
- 网络隔离：保持虚拟网络之间的隔离，以确保隐私并防止未经授权访问敏感数据
- 可扩展性：轻松扩展网络资源，以适应不断增长的工作负载和不断变化的需求
- 灵活性：快速适应不断变化的业务需求，并根据需要动态分配资源
- 简化的管理：集中化网络管理，抽象底层物理基础设施的复杂性

网络虚拟化使能技术

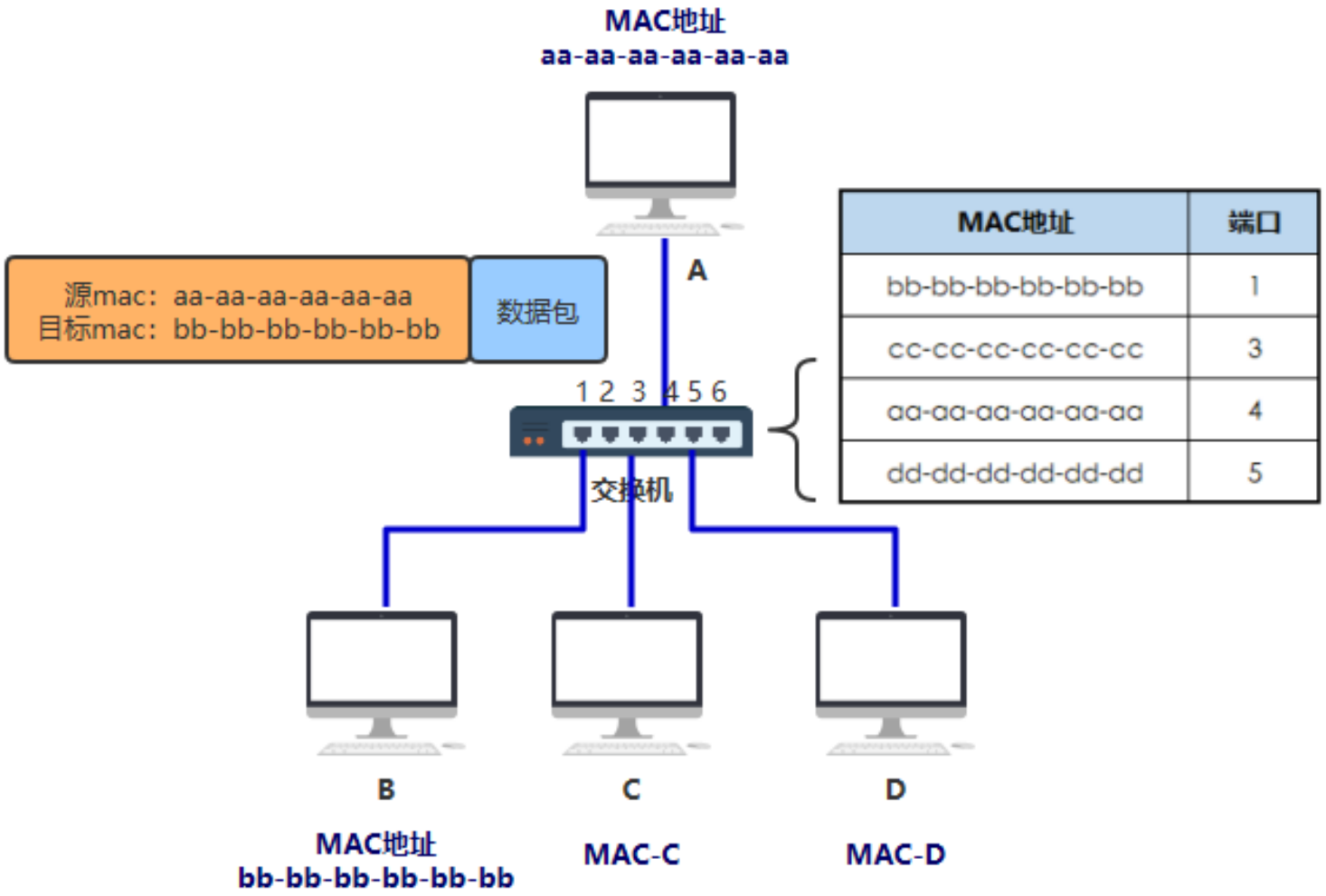
- VLAN
- VxLAN
- NFV

网络通信基础知识回顾


应用层
为计算机用户提供接口和服务
表示层
数据处理（编码解码、加密解密等）
会话层
管理（建立、维护、重连）通信会话
传输层
管理端到端的通信连接
网络层
数据路由（决定数据在网络的路径）
数据链路层
管理相邻节点之间的数据通信
物理层
数据通信的光电物理特性

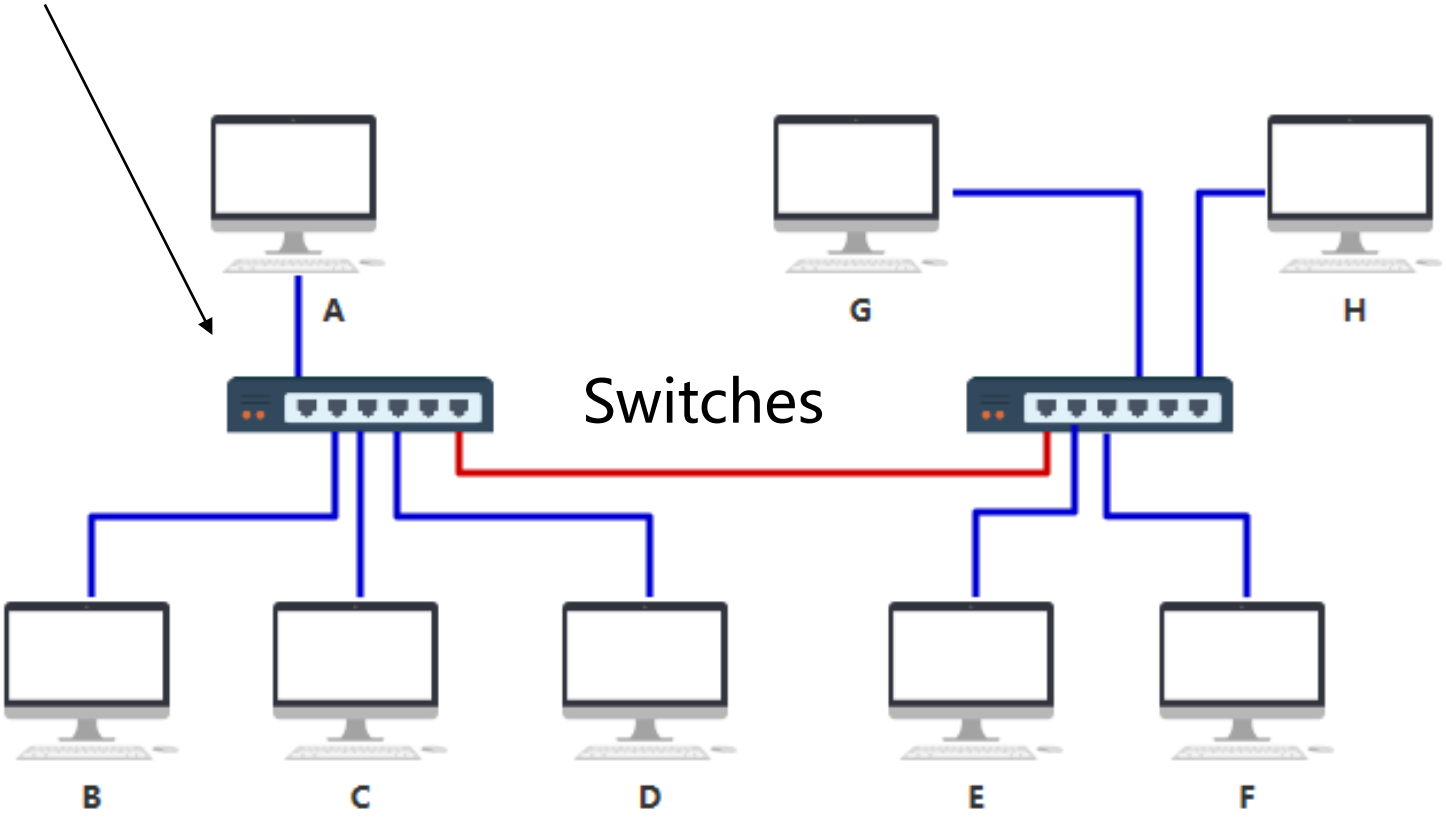
二层数据帧交换

- Media Access Control（MAC）地址



多个交换机

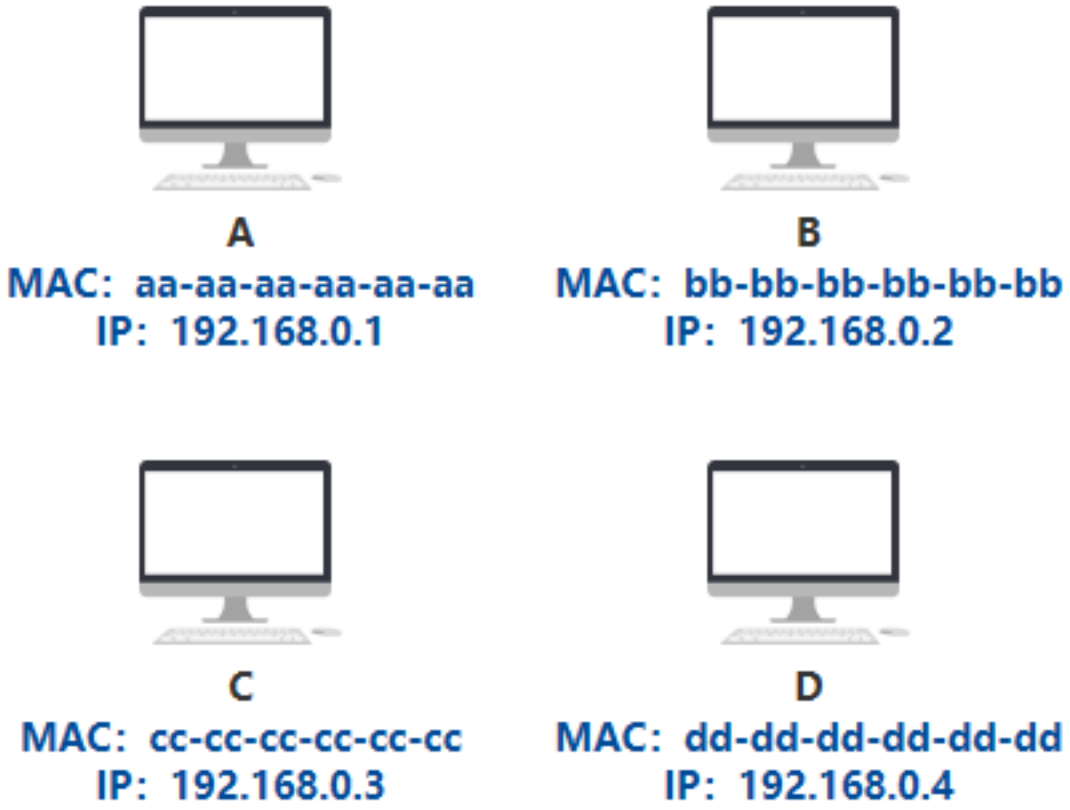
MAC 地址表怎么写？



冗余 MAC 地址项

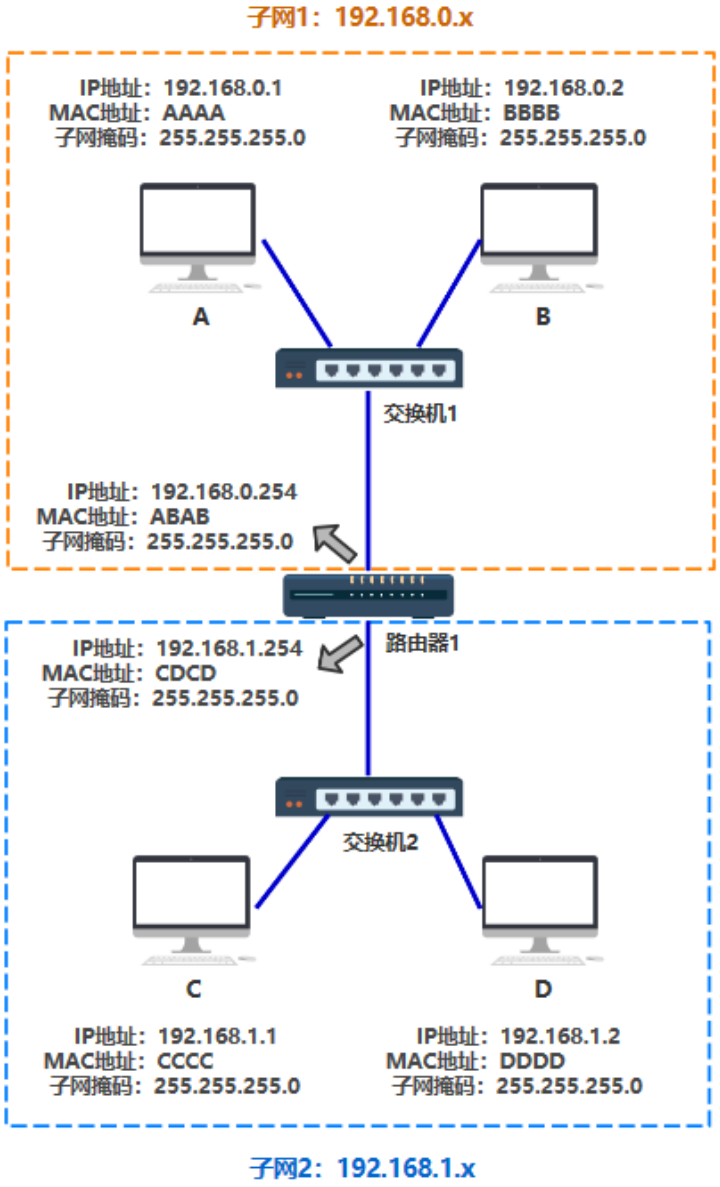
MAC 地址	端口
bb-bb-bb-bb-bb-bb	1
cc-cc-cc-cc-cc-cc	3
aa-aa-aa-aa-aa-aa	4
dd-dd-dd-dd-dd-dd	5
ee-ee-ee-ee-ee-ee	6
ff-ff-ff-ff-ff-ff	6
gg-gg-gg-gg-gg-gg	6
hh-hh-hh-hh-hh-hh	6

IP 地址



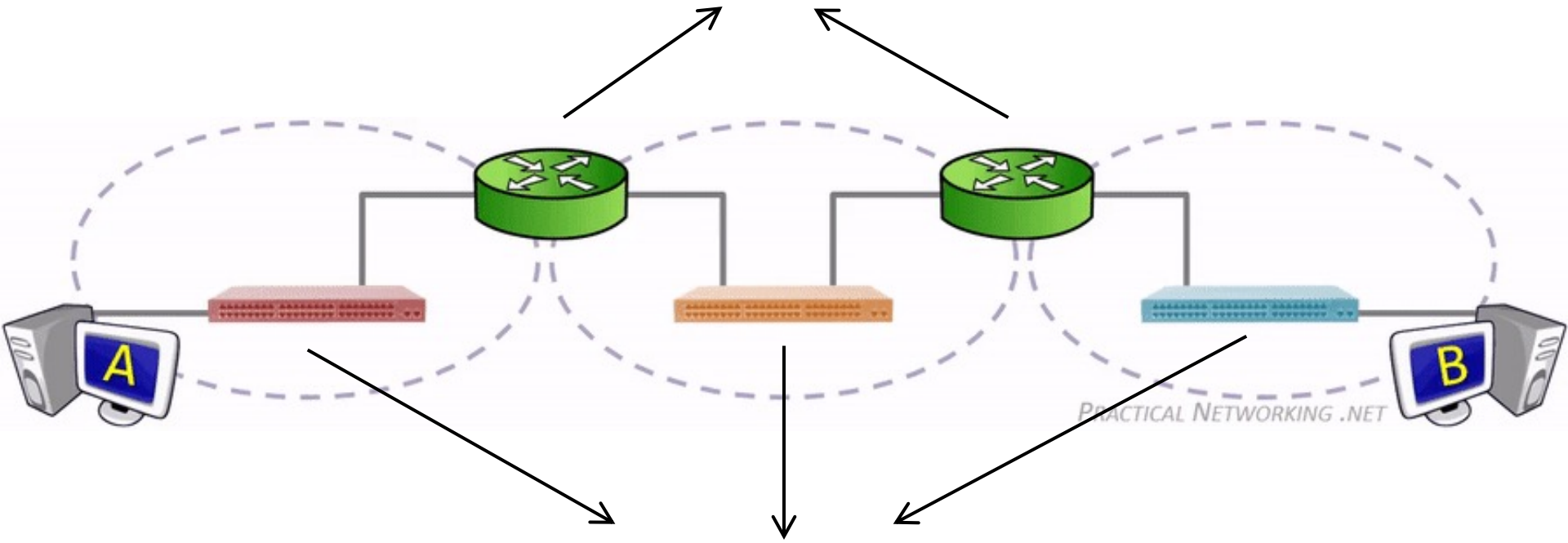
将 IP 地址为 192.168.0 开头的全部发送给到路由器！

三层数据包路由


子网1：192.168.0.x
IP地址：192.168.0.1
MAC地址：AAAA
子网掩码：255.255.255.0
IP地址：192.168.0.2
MAC地址：BBBB
子网掩码：255.255.255.0
A
B
交换机1
IP地址：192.168.0.254
MAC地址：ABAB
子网掩码：255.255.255.0
路由器1
IP地址：192.168.1.254
MAC地址：CDCD
子网掩码：255.255.255.0
交换机2
C
D
IP地址：192.168.1.1
MAC地址：CCCC
子网掩码：255.255.255.0
IP地址：192.168.1.2
MAC地址：DDDD
子网掩码：255.255.255.0
子网2：192.168.1.x

网络通信

网络第三层，路由器（router），基于 IP 地址

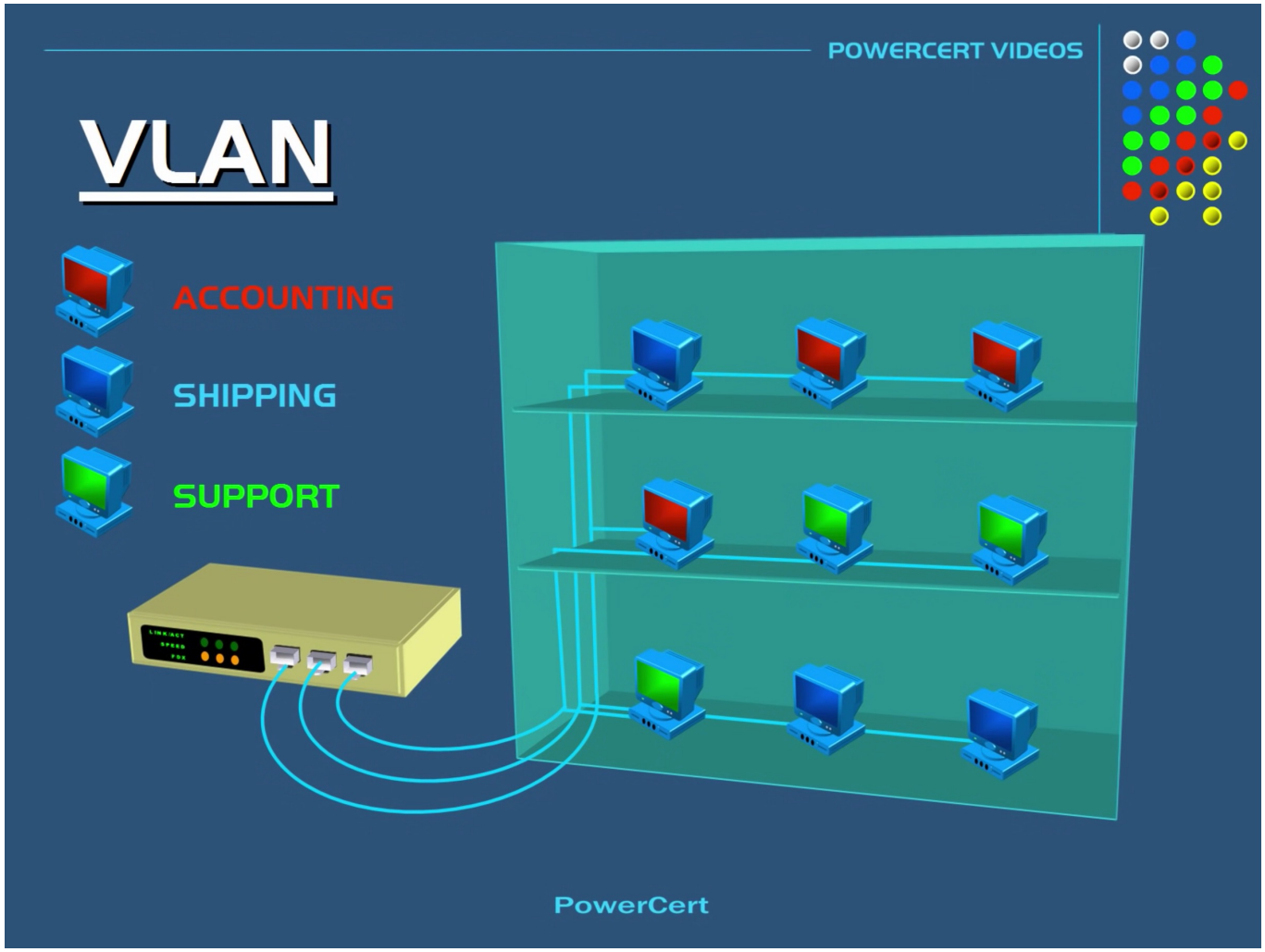


网络第二层，交换机（switch），基于MAC地址

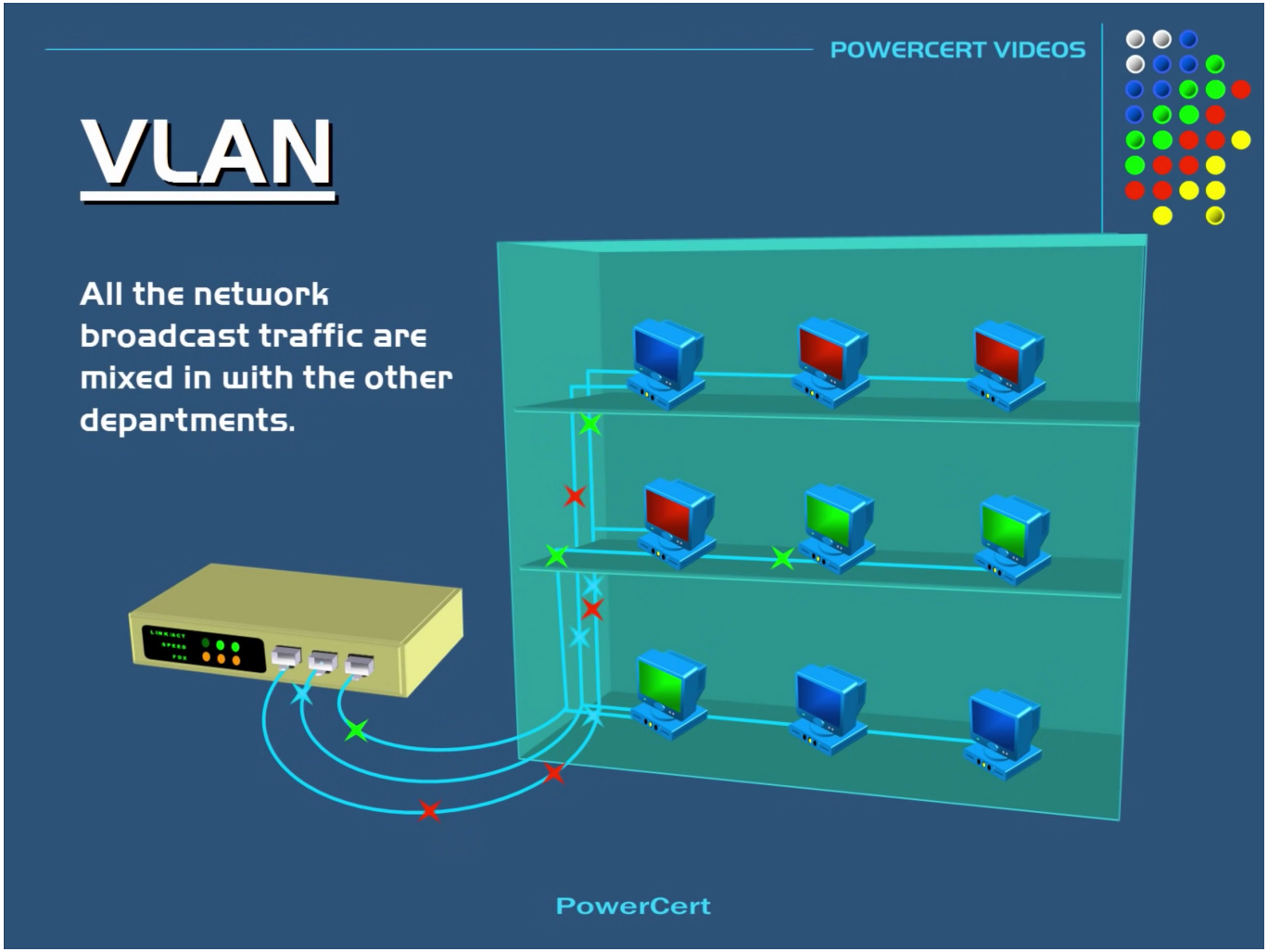
网络虚拟化技术

- 数据中心网络虚拟化
- 虚拟局域网（VLAN）
- 虚拟可扩展局域网（VxLAN）
- 网络功能虚拟化（NFV）

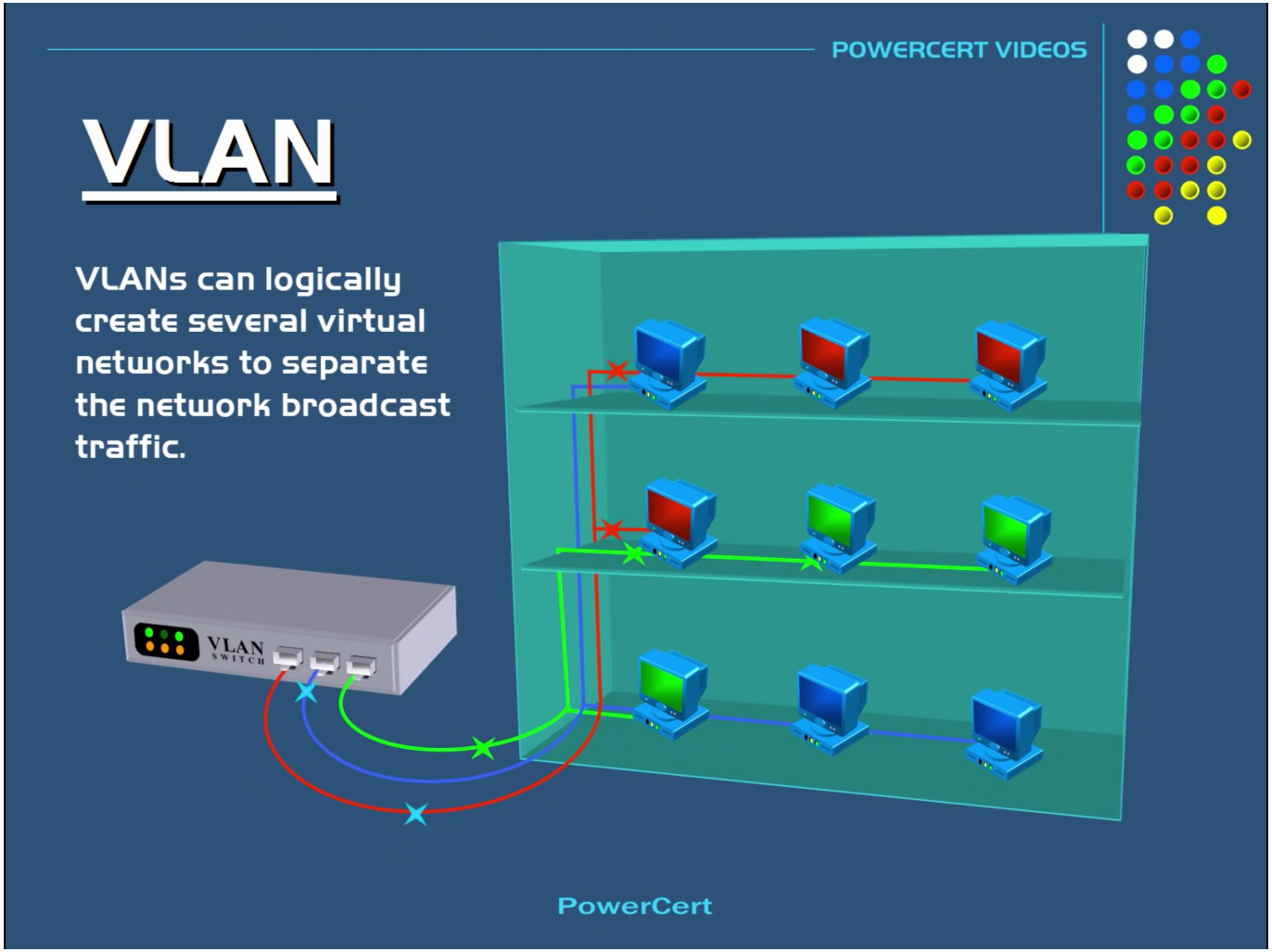
广播域（broadcast domain）



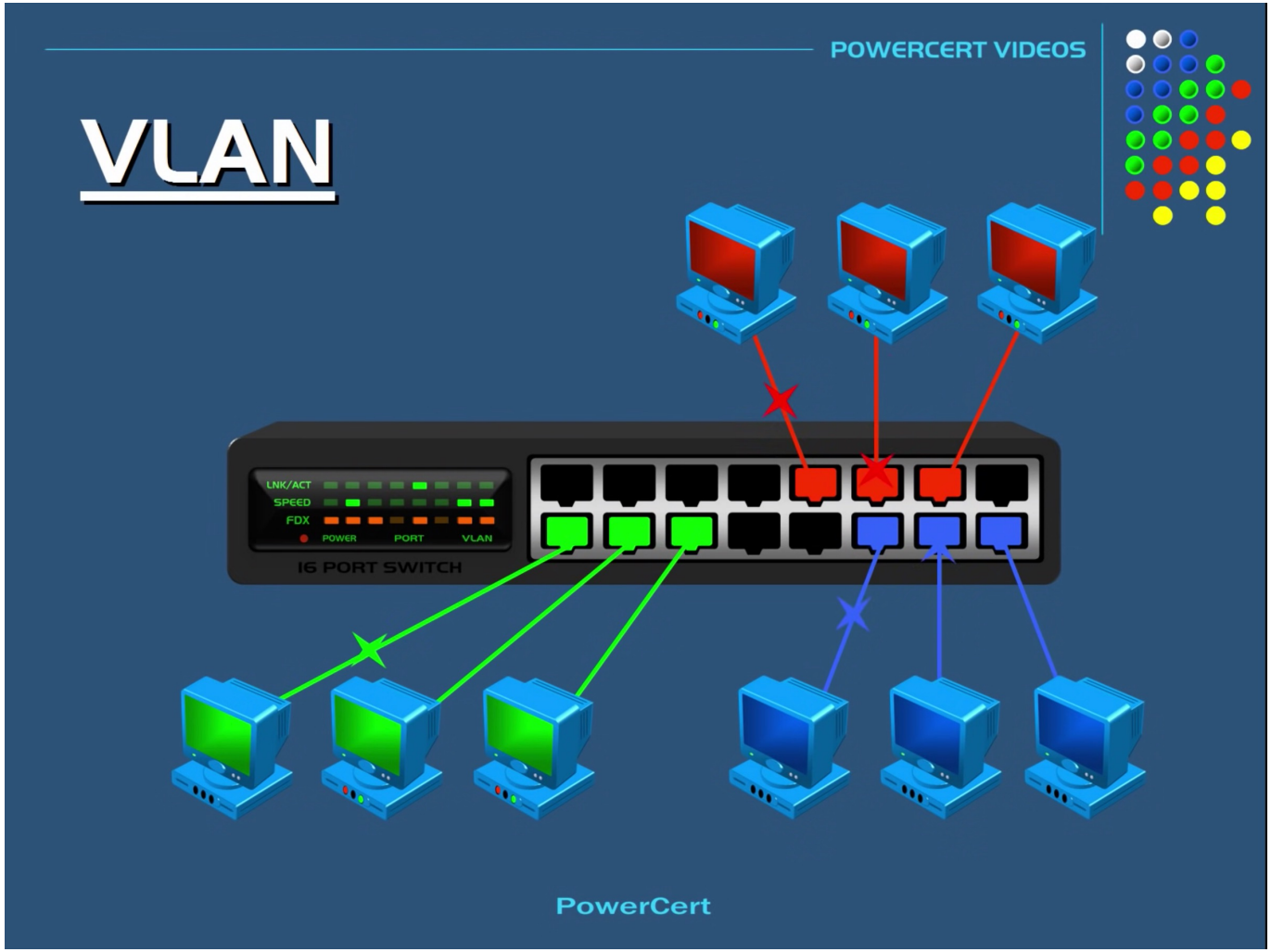
广播域（broadcast domain）



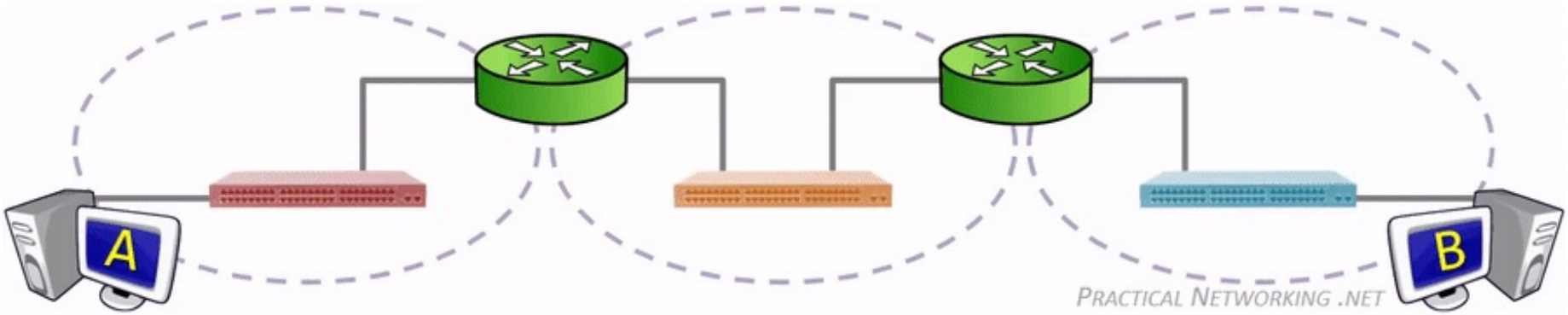
划分VLAN

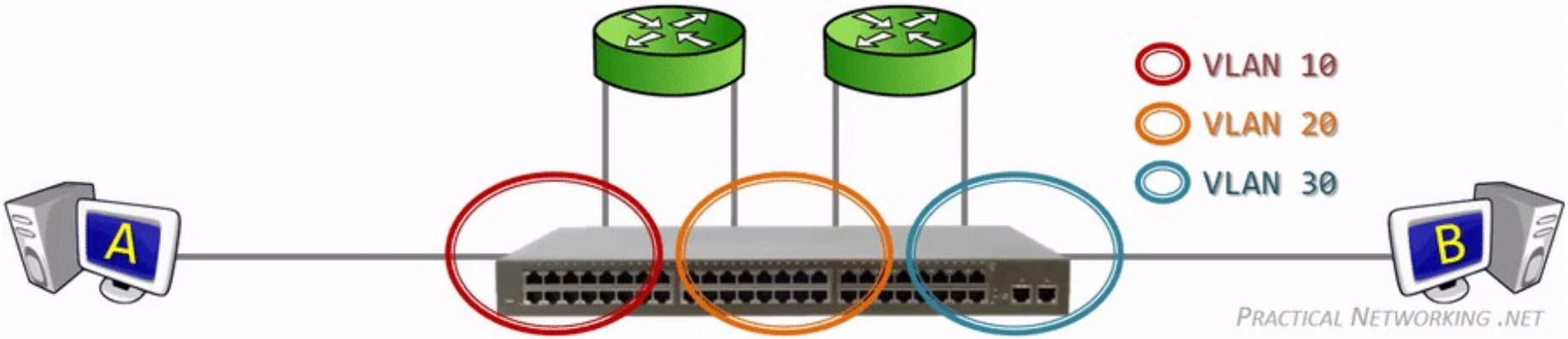

POWERCERT VIDEOS
VLAN
VLANs can logically create several virtual networks to separate the network broadcast traffic.
VLAN SWITCH
PowerCert

VLAN交换机

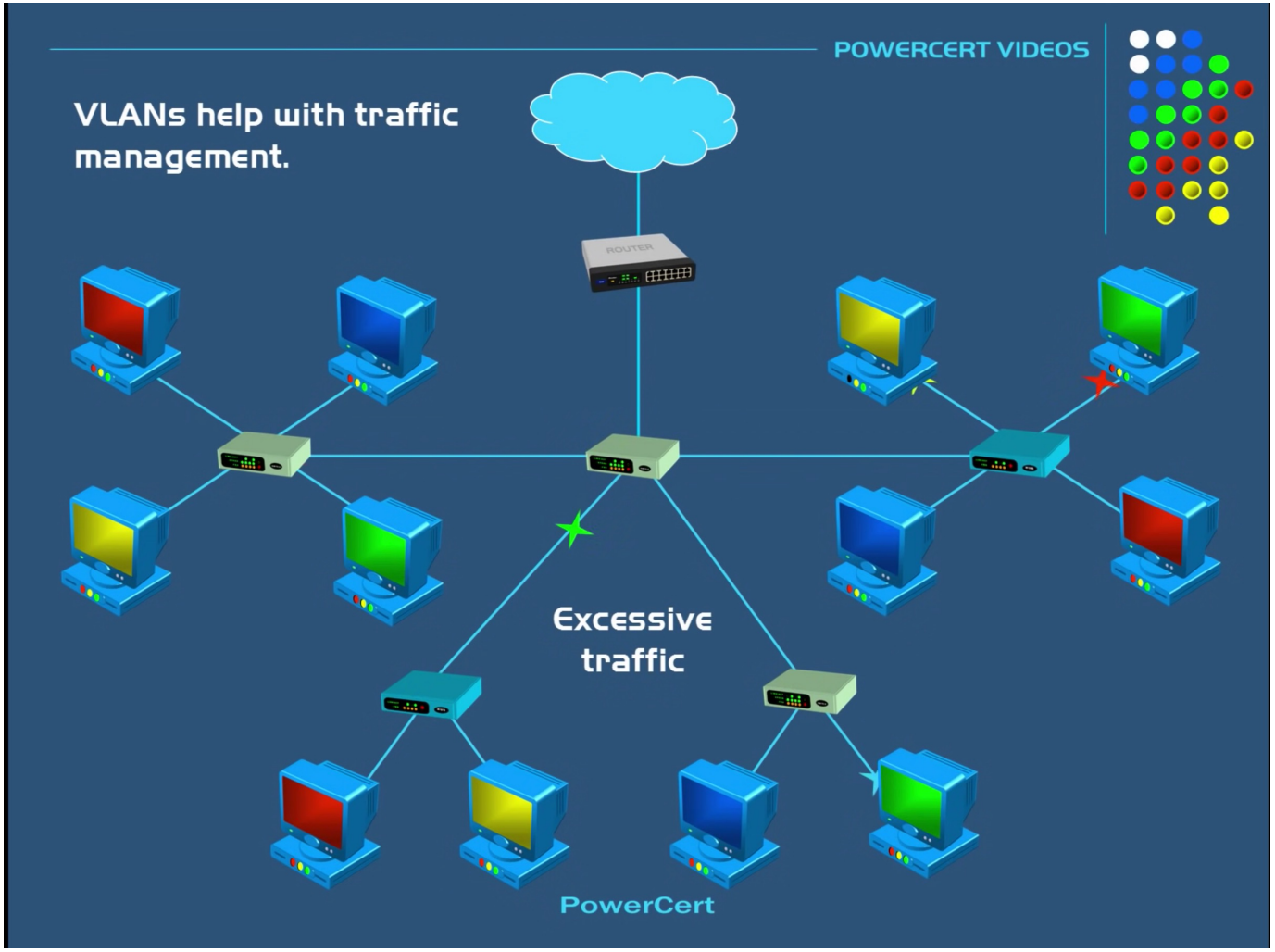

POWERCERT VIDEOS
VLAN
LNK/ACT
SPEED
FDX
POWER
PORT
VLAN
16 PORT SWITCH
PowerCert

VLAN交换机

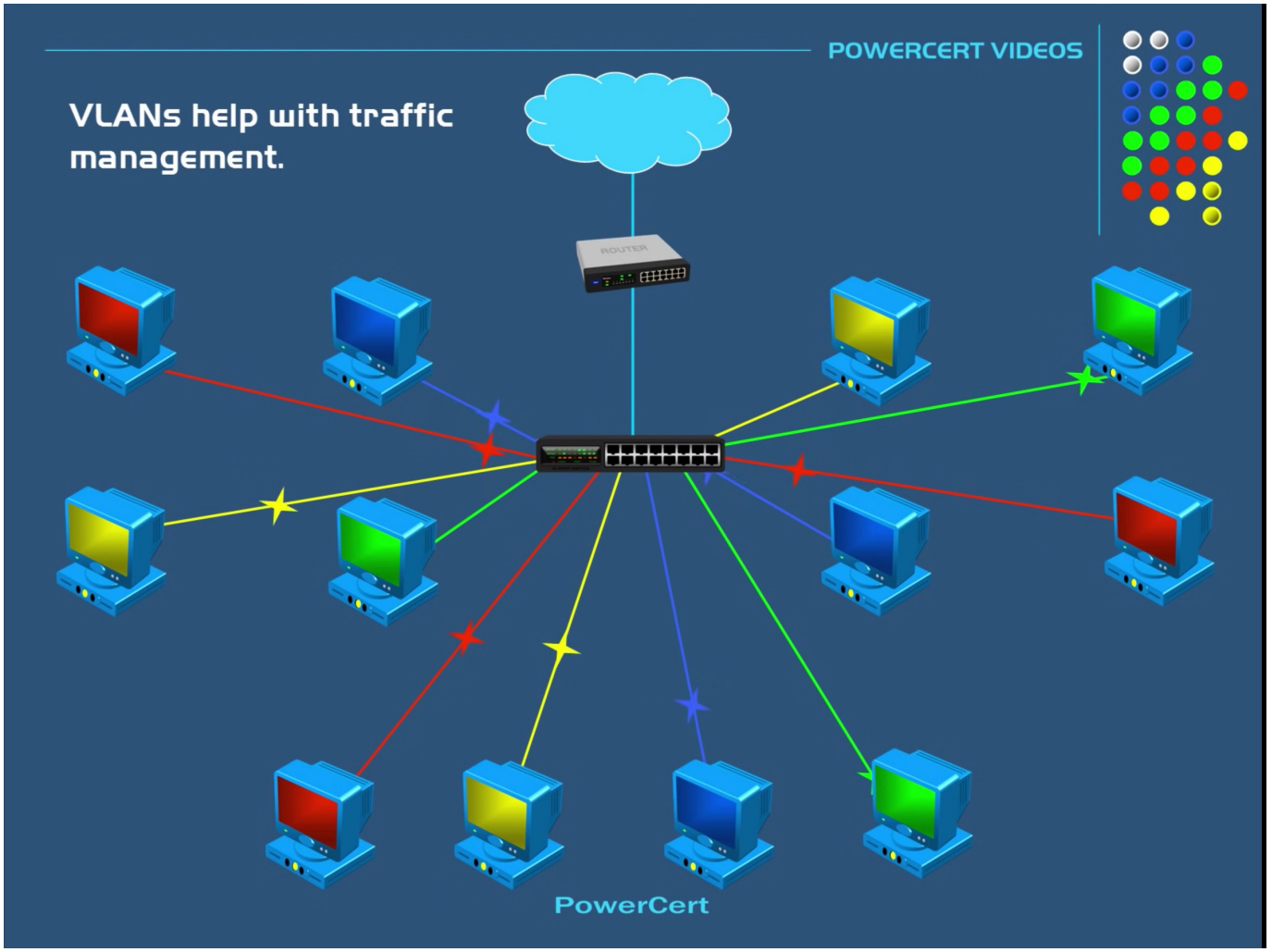

A
B
PRACTICAL NETWORKING .NET


VLAN 10
VLAN 20
VLAN 30
A
B
PRACTICAL NETWORKING .NET

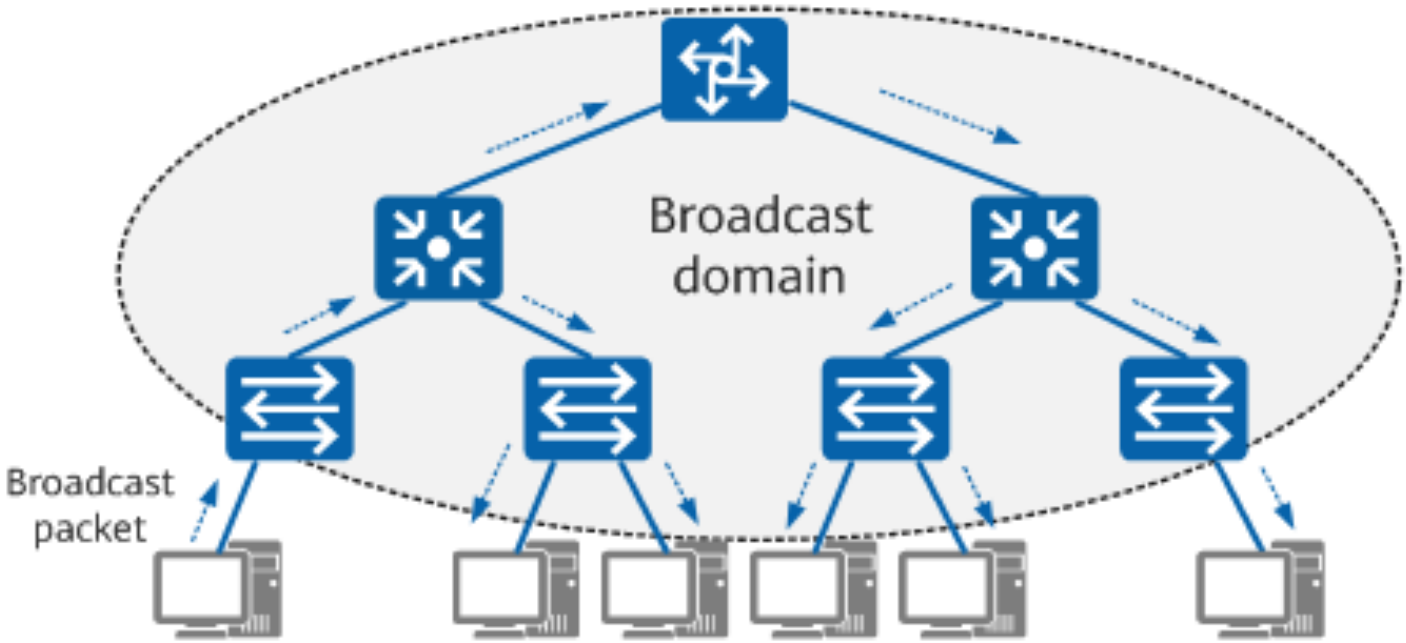
VLAN简化流量管理（1）



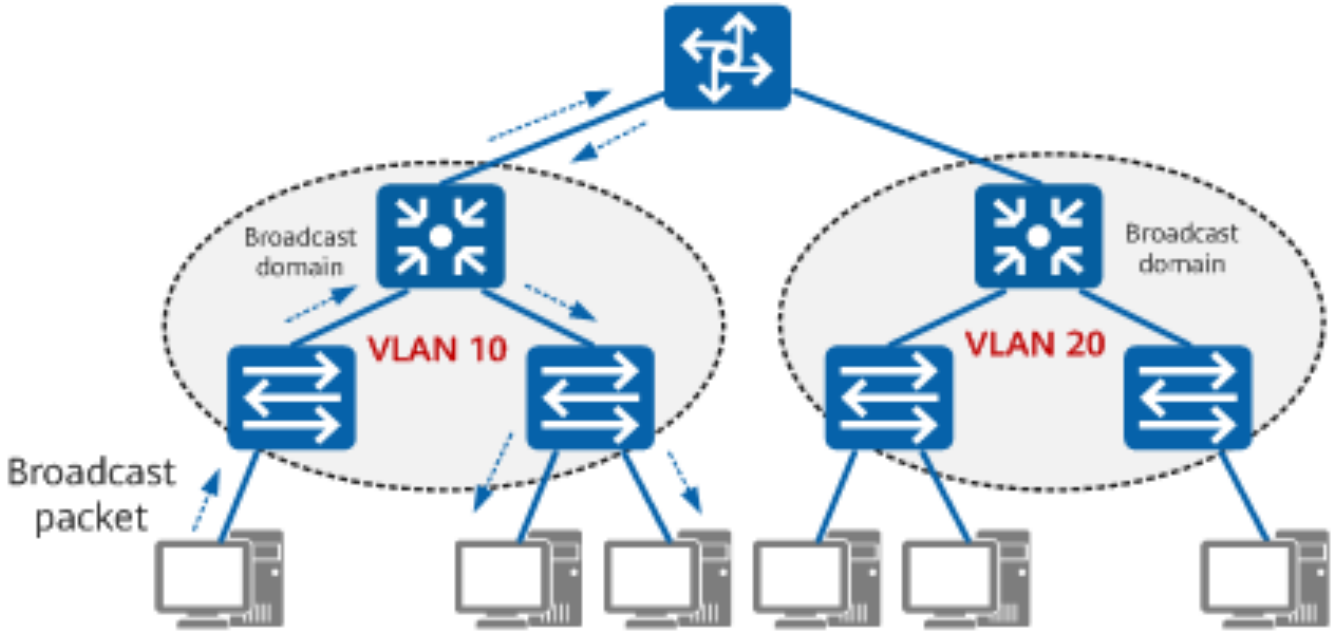
VLAN简化流量管理（2）



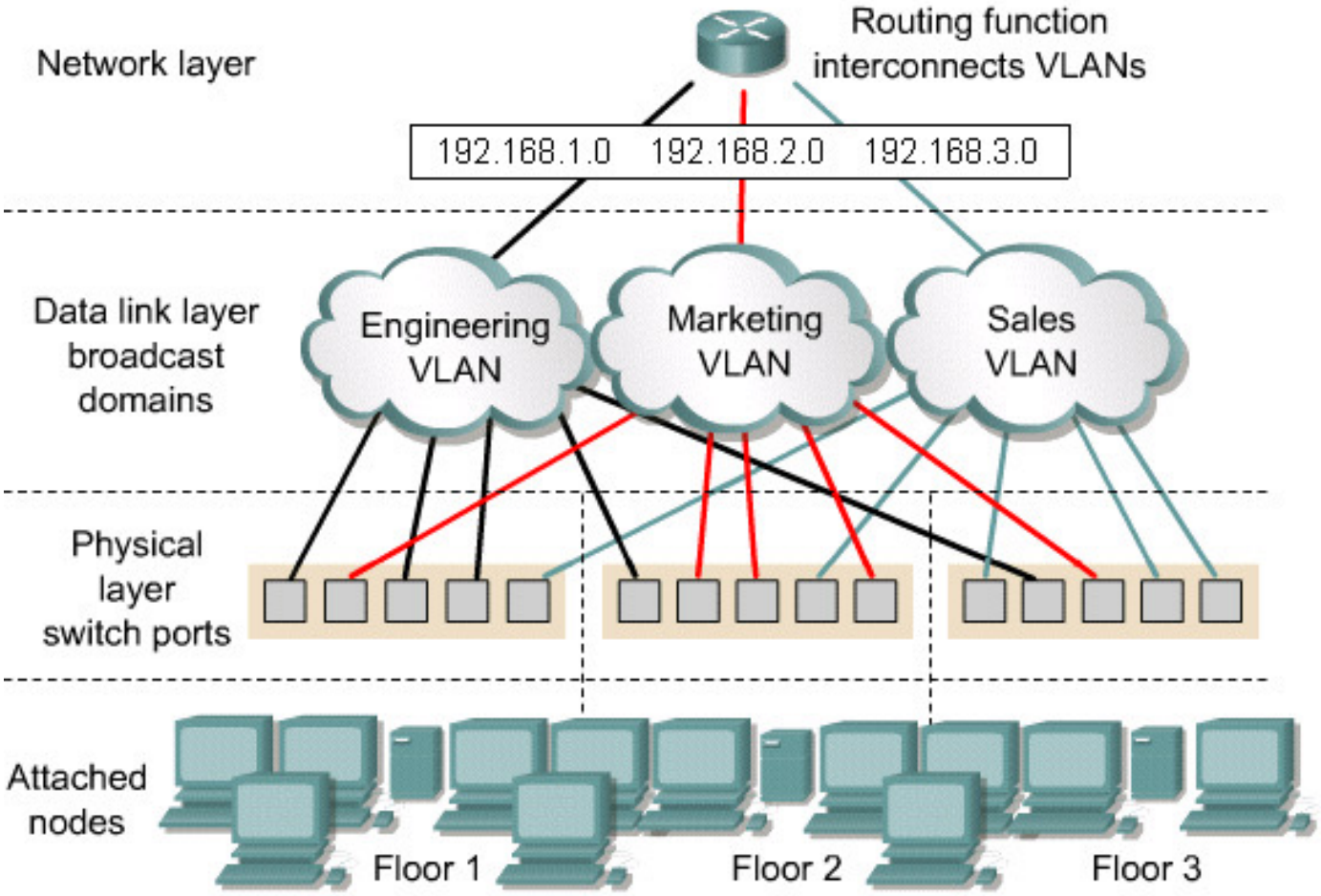
数据中心广播风暴



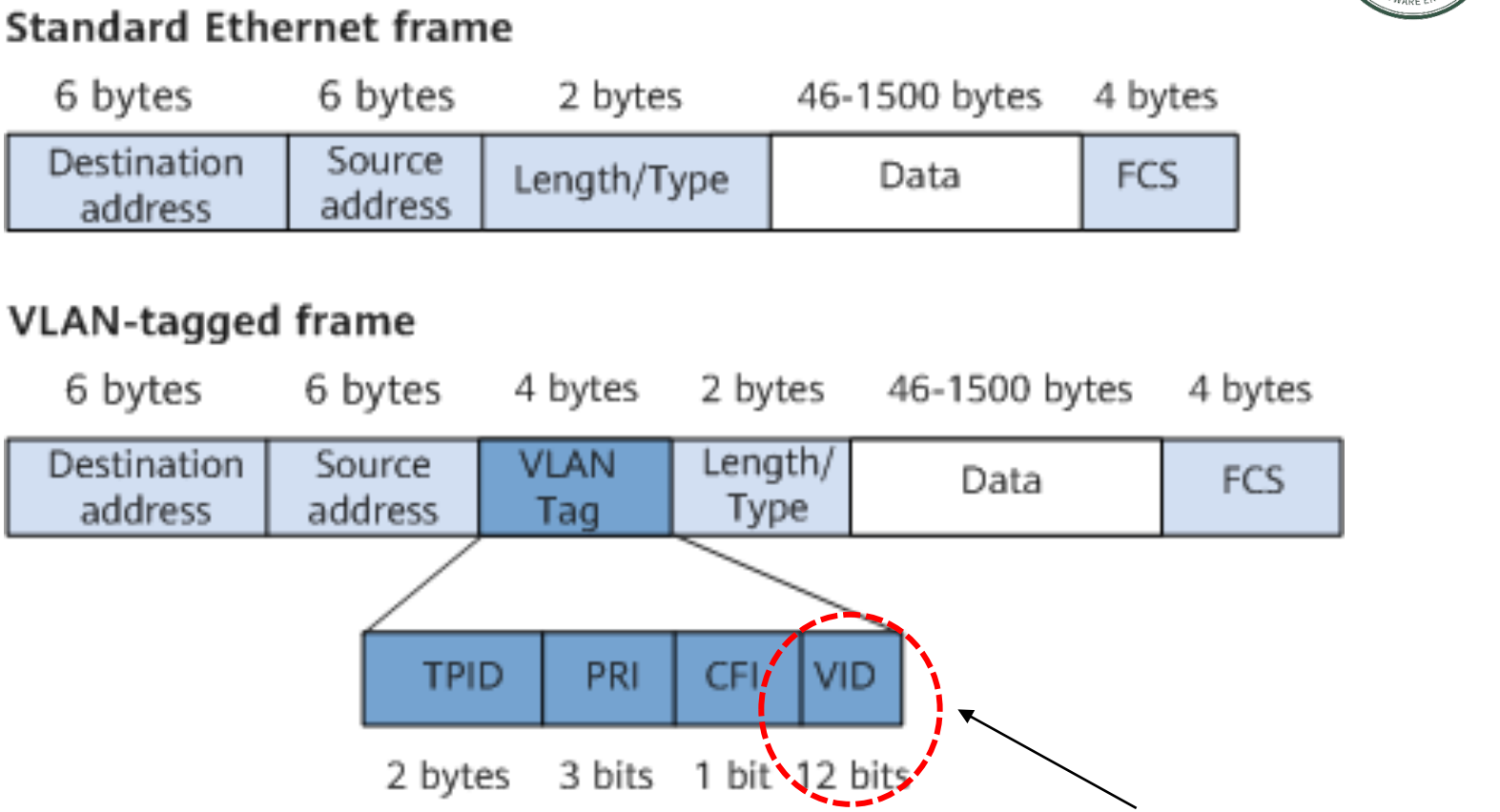
数据中心划分VLAN


Broadcast
domain
VLAN 10
Broadcast
domain
VLAN 20
Broadcast
packet

VLAN功能划分


Network layer
Routing function
interconnects VLANs
192.168.1.0
192.168.2.0
192.168.3.0
Data link layer
broadcast
domains
Engineering
VLAN
Marketing
VLAN
Sales
VLAN
Physical
layer
switch ports
Attached
nodes
Floor 1
Floor 2
Floor 3

VLAN Tag

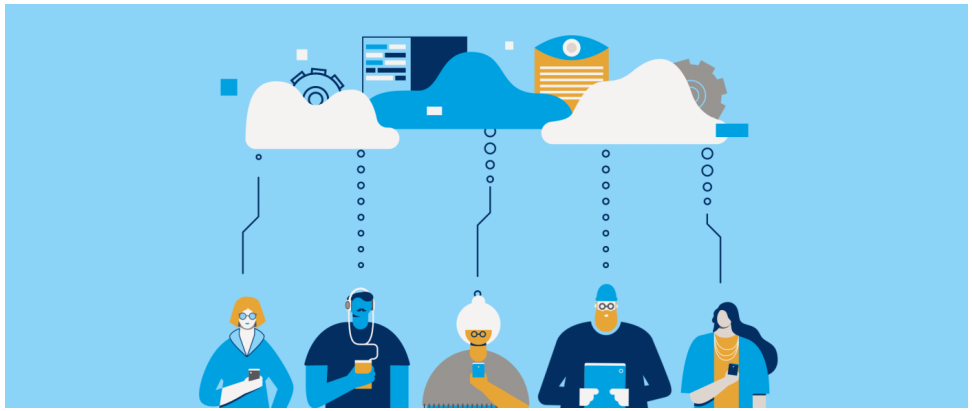


利用 VID 划分不同的 VLAN

VLAN存在的问题（1）

- VLAN 数量不够
  - ✓VLAN ID 为 12 个 bit
  - ✓4094 个不同的 VID（2 个保留 ID）

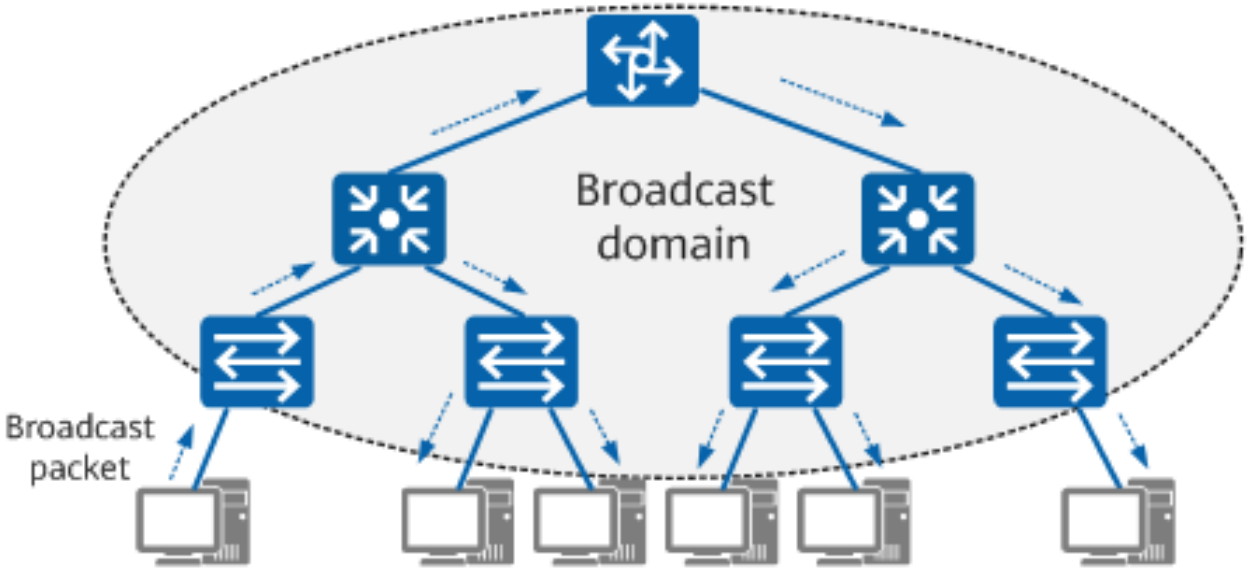
数据中心规模增大



租户数量增多（>4094）

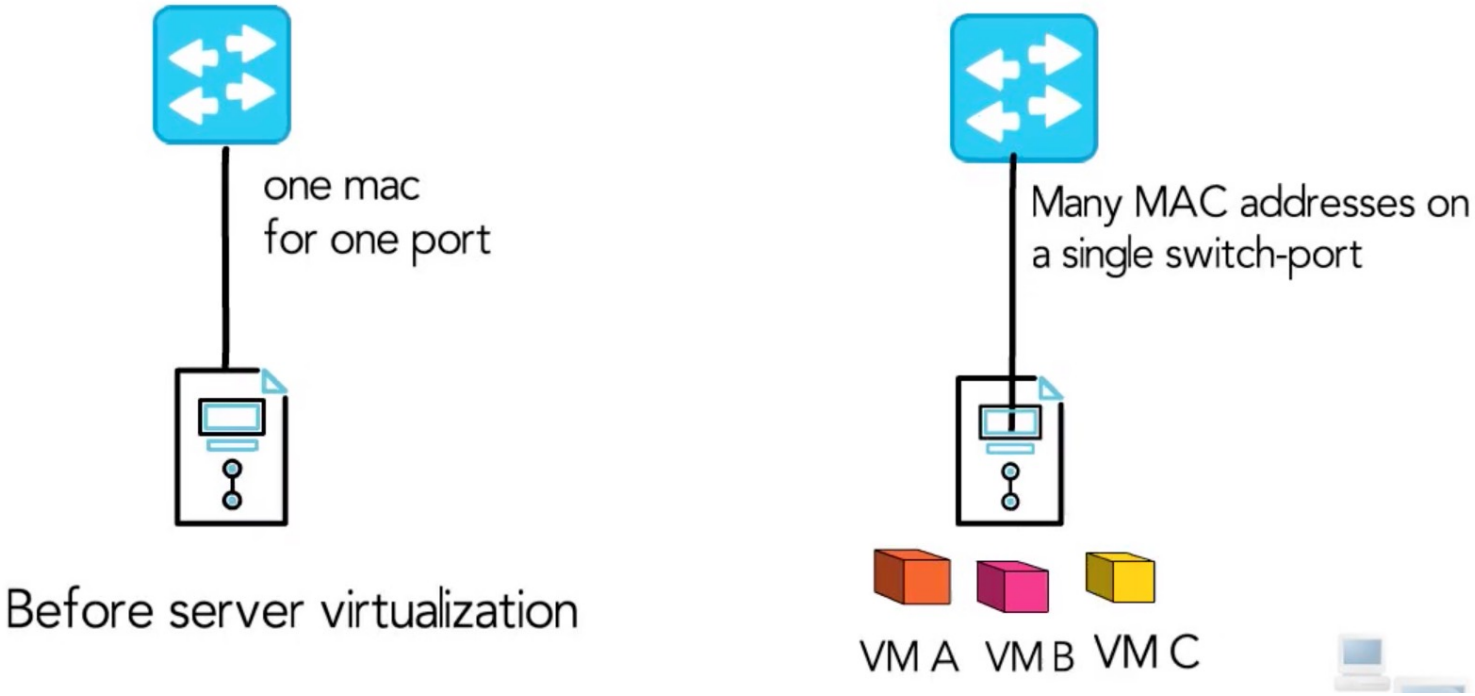
VLAN存在的问题（2）

- 地理局限性
  - ✓VLAN 中的设备必须在同一个广播域，受第二层网络边界的约束
  - ✓需要扩大物理边界



VLAN存在的问题（3）

- MAC 地址表过大

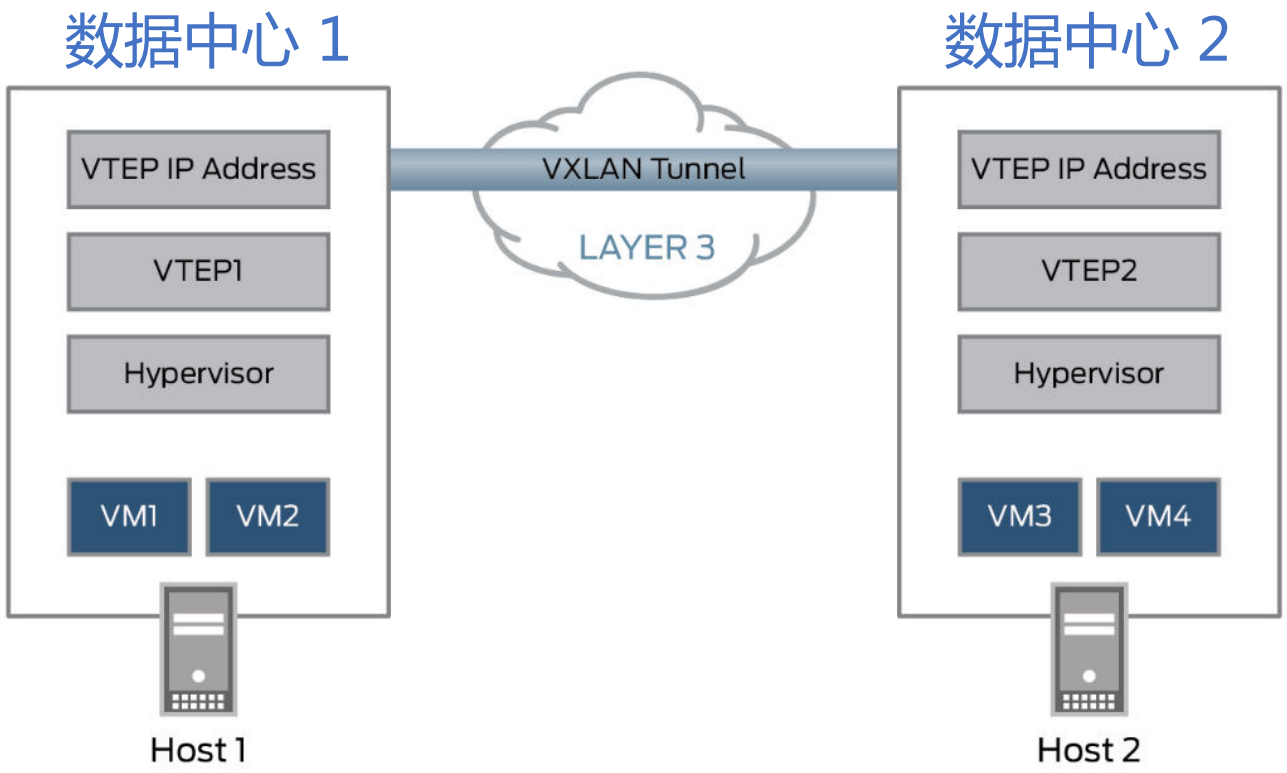


网络虚拟化技术

- 数据中心网络虚拟化
- 虚拟局域网（VLAN）
- 虚拟可扩展局域网（VxLAN）
- 网络功能虚拟化（NFV）

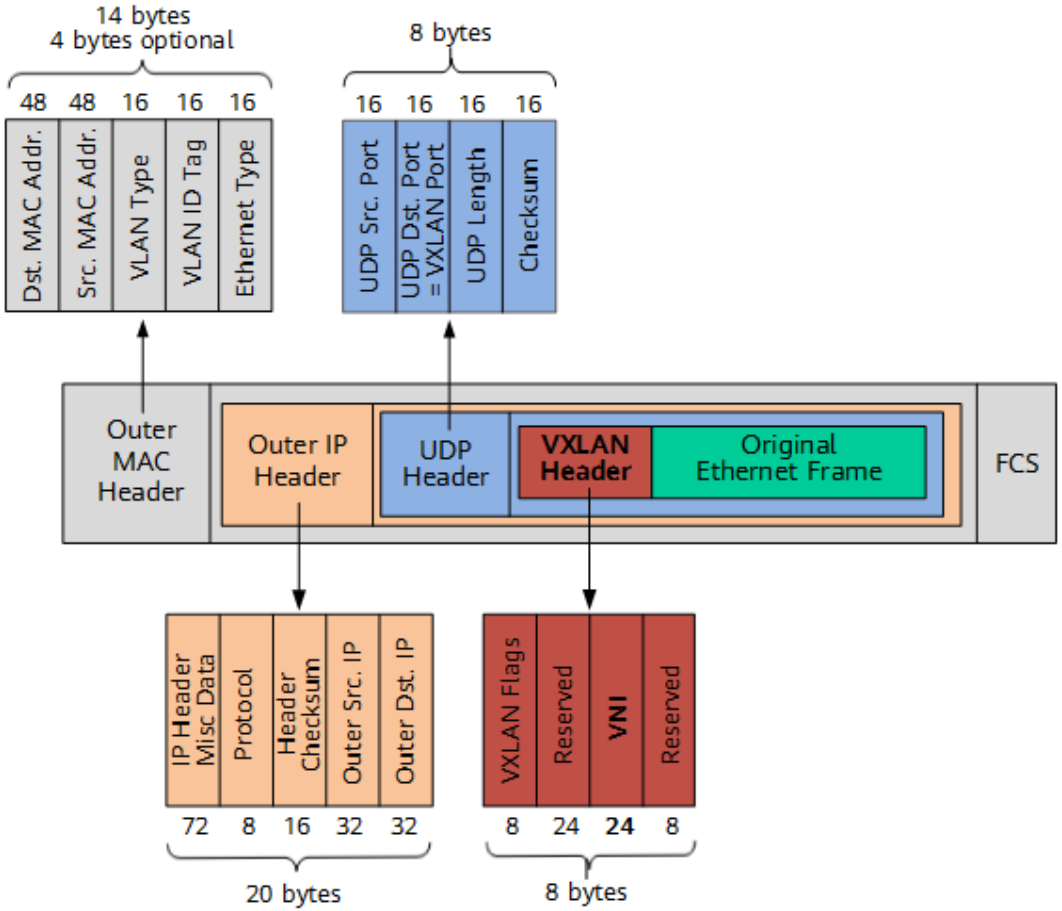
VxLAN

- VxLAN (Virtual eXtensible LAN)
  - ✓VxLAN是一种封装（encapsulation）协议，它使用隧道（tunneling）技术借助第 3 层网络上扩展第 2 层的连接



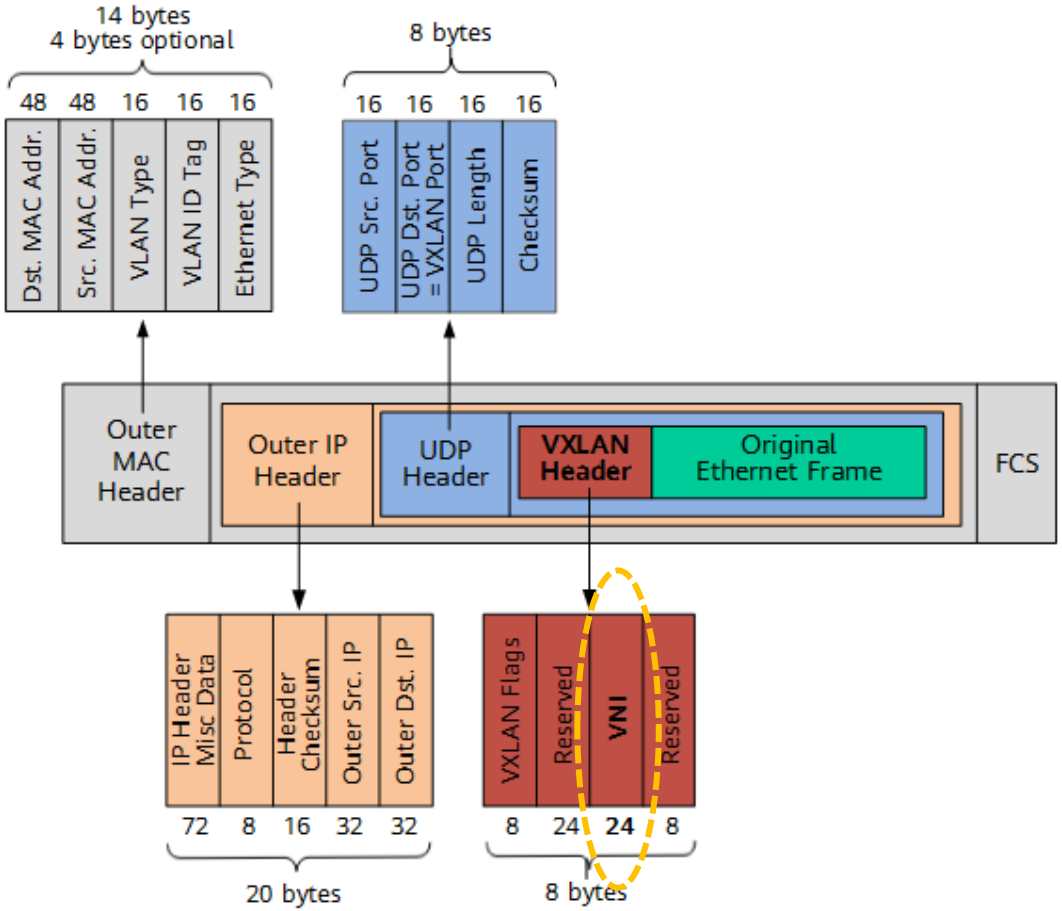
VxLAN 数据包

- 将虚拟机发送的原始以太网帧封装到UDP数据包中
- 然后，将 IP 报头和以太网报头封装为外部报头，使这些数据包能够像普通 IP 数据包一样在网络上路由



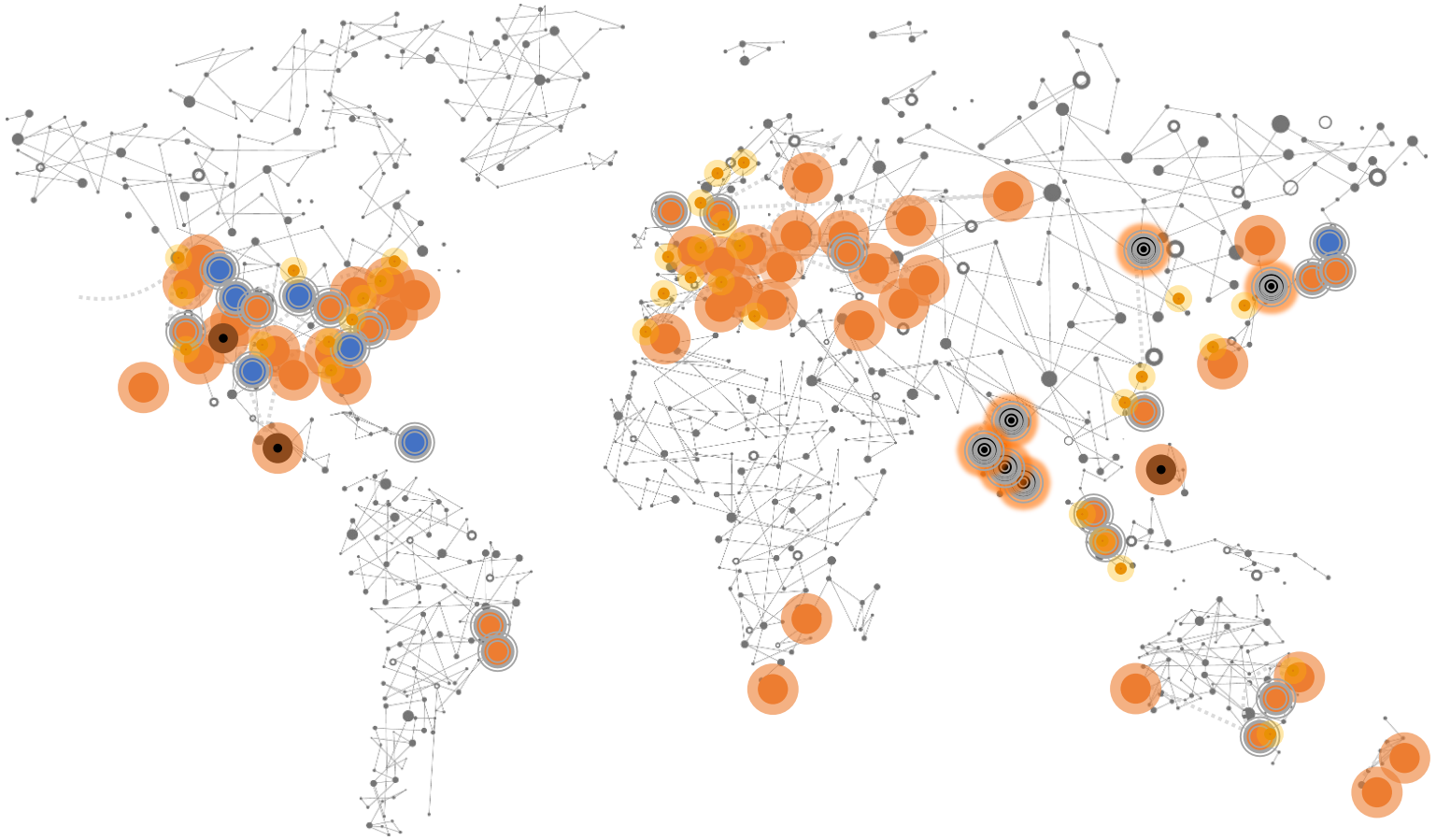
VxLAN 的优势（1）

- 更大的 ID 空间
  - ✓VxLAN ID (VNI) 有 24 bits，超过 16M 个不同的 ID



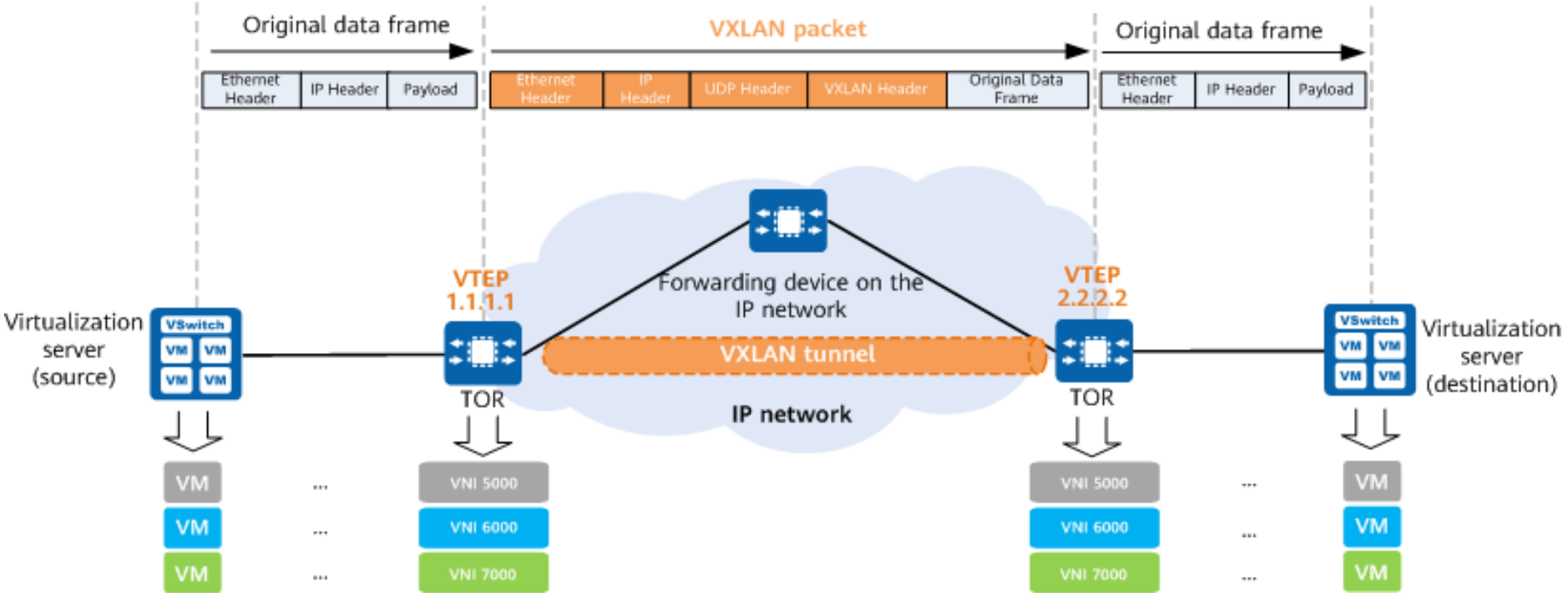
VxLAN 的优势（2）

- VxLAN 突破了网络第 2 层的边界限制



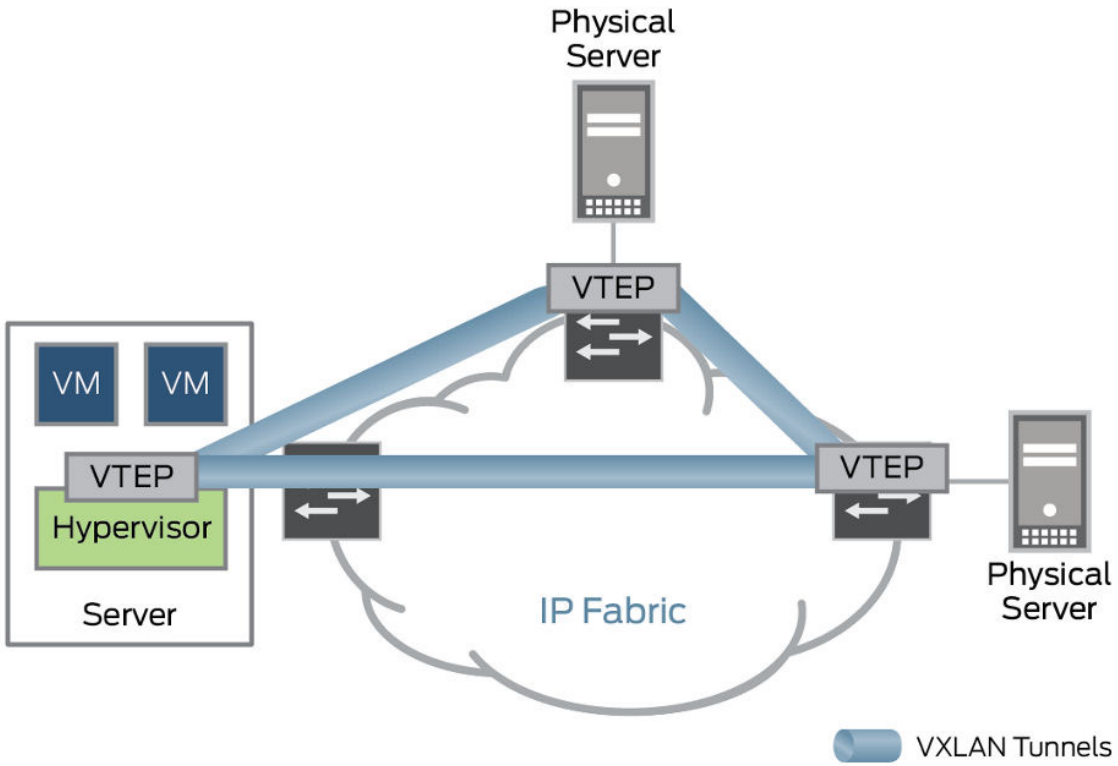
VxLAN 实现

- VxLAN 隧道在两个 TOR 交换机之间建立
- 将源服务器发送的原始数据帧封装成 VxLAN 数据包，从而使原始数据帧能够在 IP 网络上传输
- 当 VXLAN 数据包到达连接到目的服务器的 TOR 交换机时，TOR 交换机将这些数据包解封装为原始数据帧



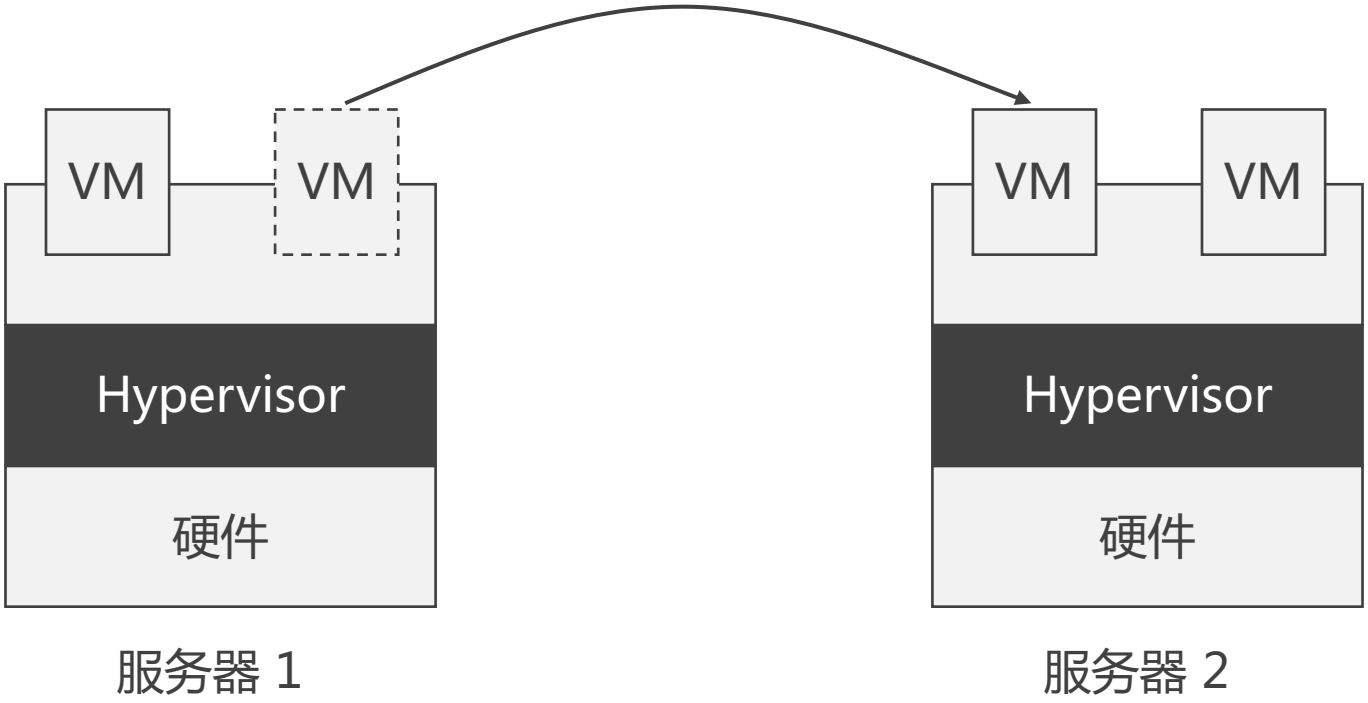
VTEP (VXLAN tunnel endpoint)

- VTEP 是 VxLAN 网络上的边缘设备，也是 VxLAN 隧道的起点或终点
- 源 VTEP 将源服务器发送的原始数据帧封装成 VxLAN 数据包，并将其发送到 IP 网络上的目的 VTEP
- 目的地 VTEP 将 VxLAN 分组解封装为原始数据帧，并将这些帧转发到目的地服务器



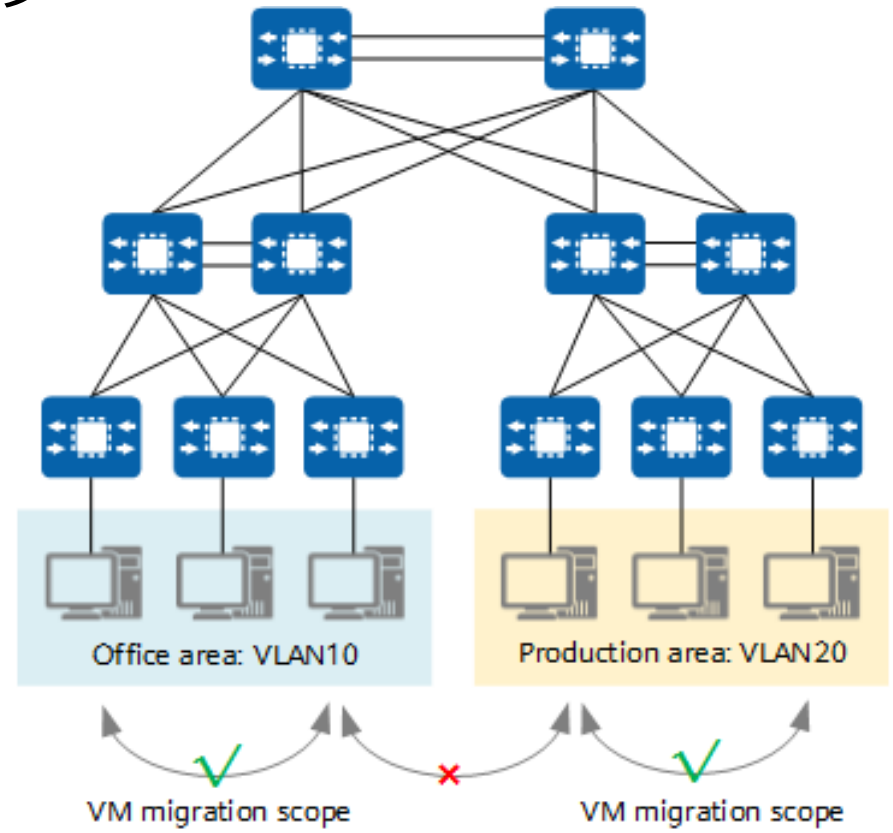
VM 迁移

- 将 VM 从一个物理机迁移到另一个物理机
- VM 的 MAC 地址和 IP 地址要保持不变（why?）



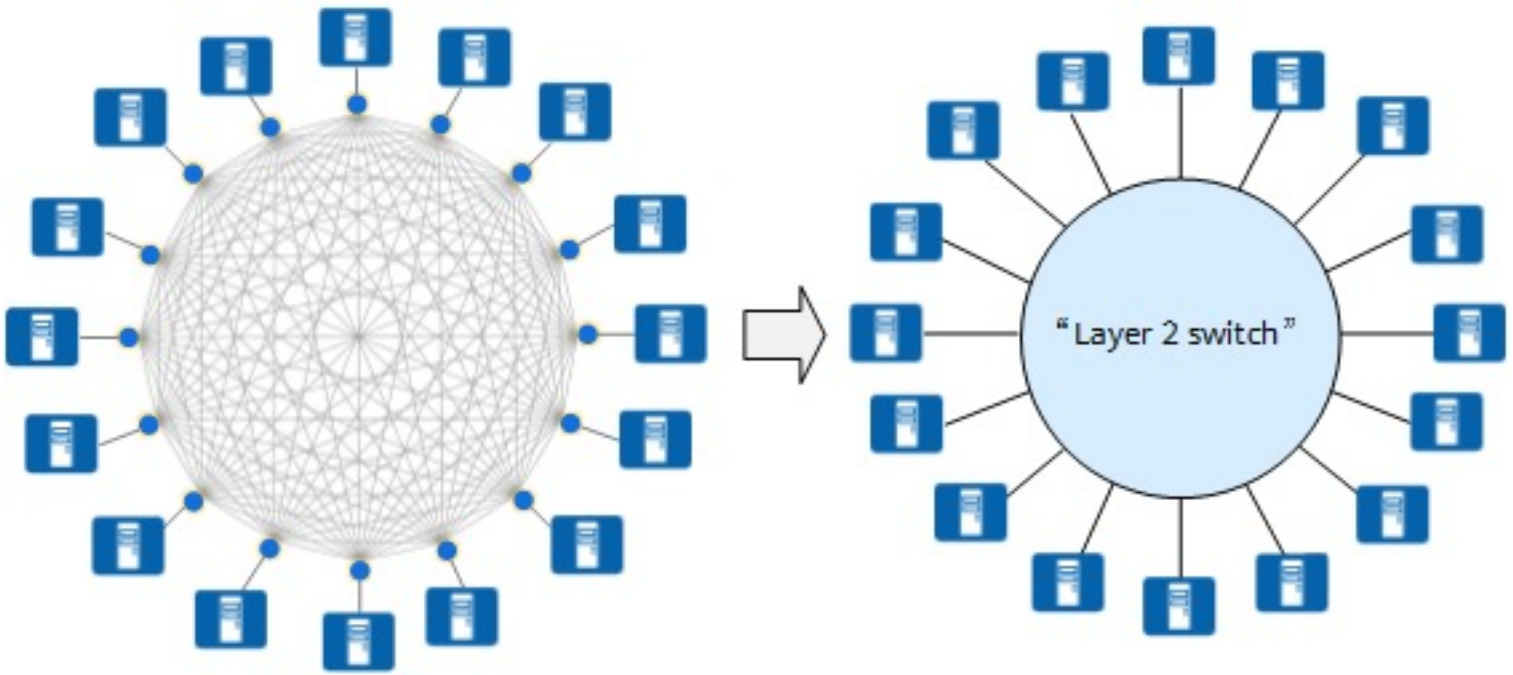
VLAN 中的 VM 迁移

- VLAN 中，VM 迁移的范围有限
- VM 的 MAC 地址不变，MAC 地址与虚拟网卡绑定，一同迁移
- 若 VM 要迁往另一个子网，则 IP 地址需要改变，因而不支持跨子网的 VM 迁移



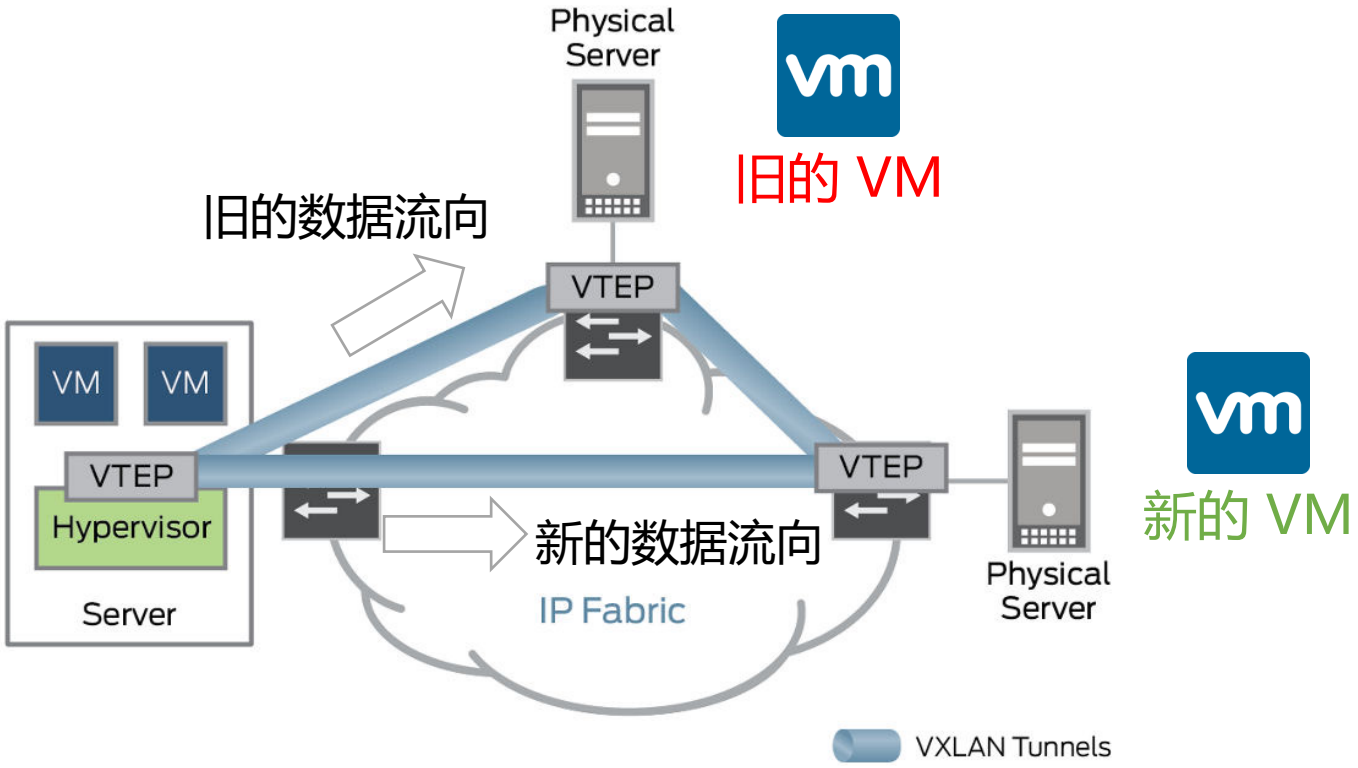
VxLAN 中的 VM 迁移（1）

- VxLAN 将整个基础设施网络虚拟化为一个大型的“2层虚拟交换机”
- VM 迁移的范围被极大扩展，VxLAN 使 VM 的 MAC 和 IP 地址保持不变的情况下迁移到别处



VxLAN 中的 VM 迁移（2）

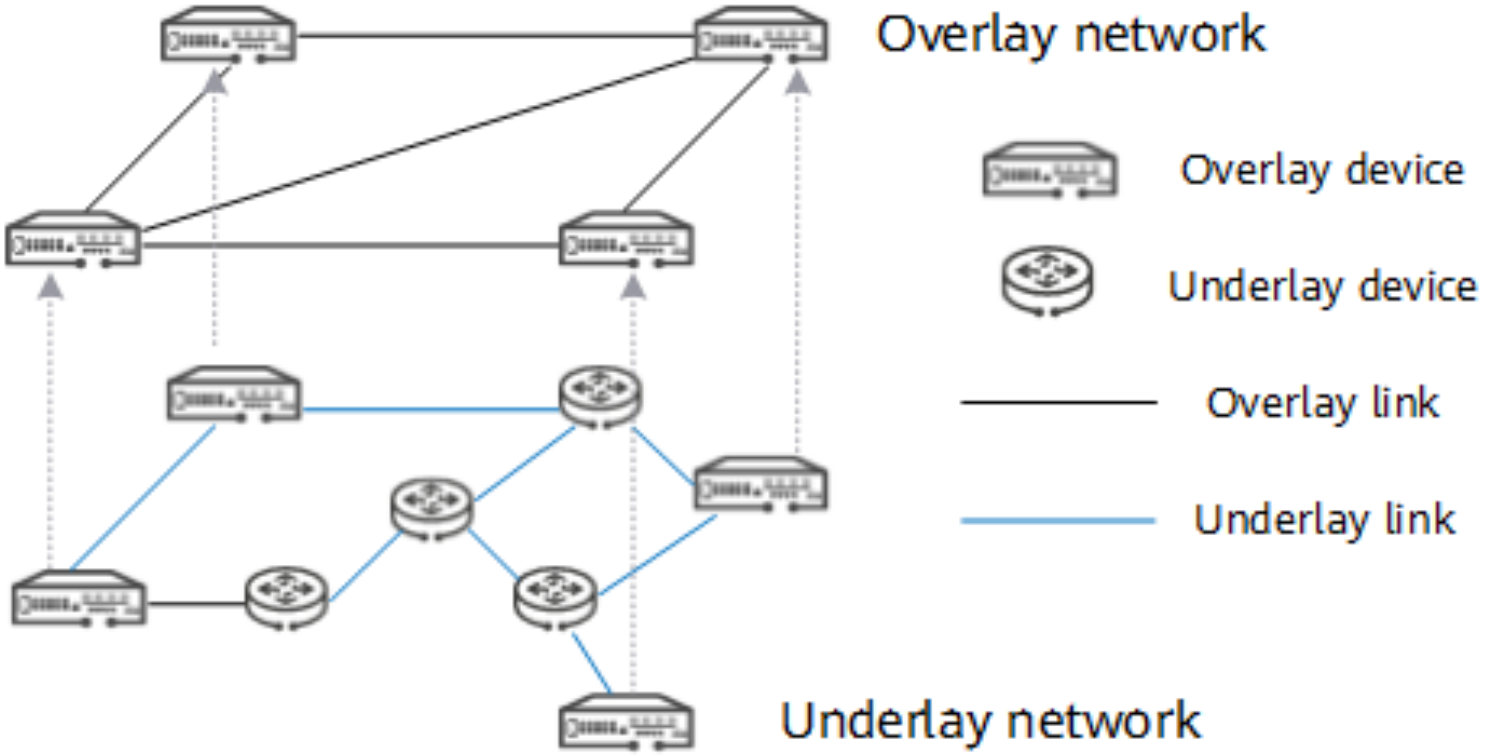
- 为什么 VM 在 MAC 地址和 IP 地址不变的情况下，数据可以正确被路由到新的 VM？
- 因为 VTEP 的 IP 地址 和 VM 的 MAC 之间的映射发生改变



覆盖网络（Overlay network）

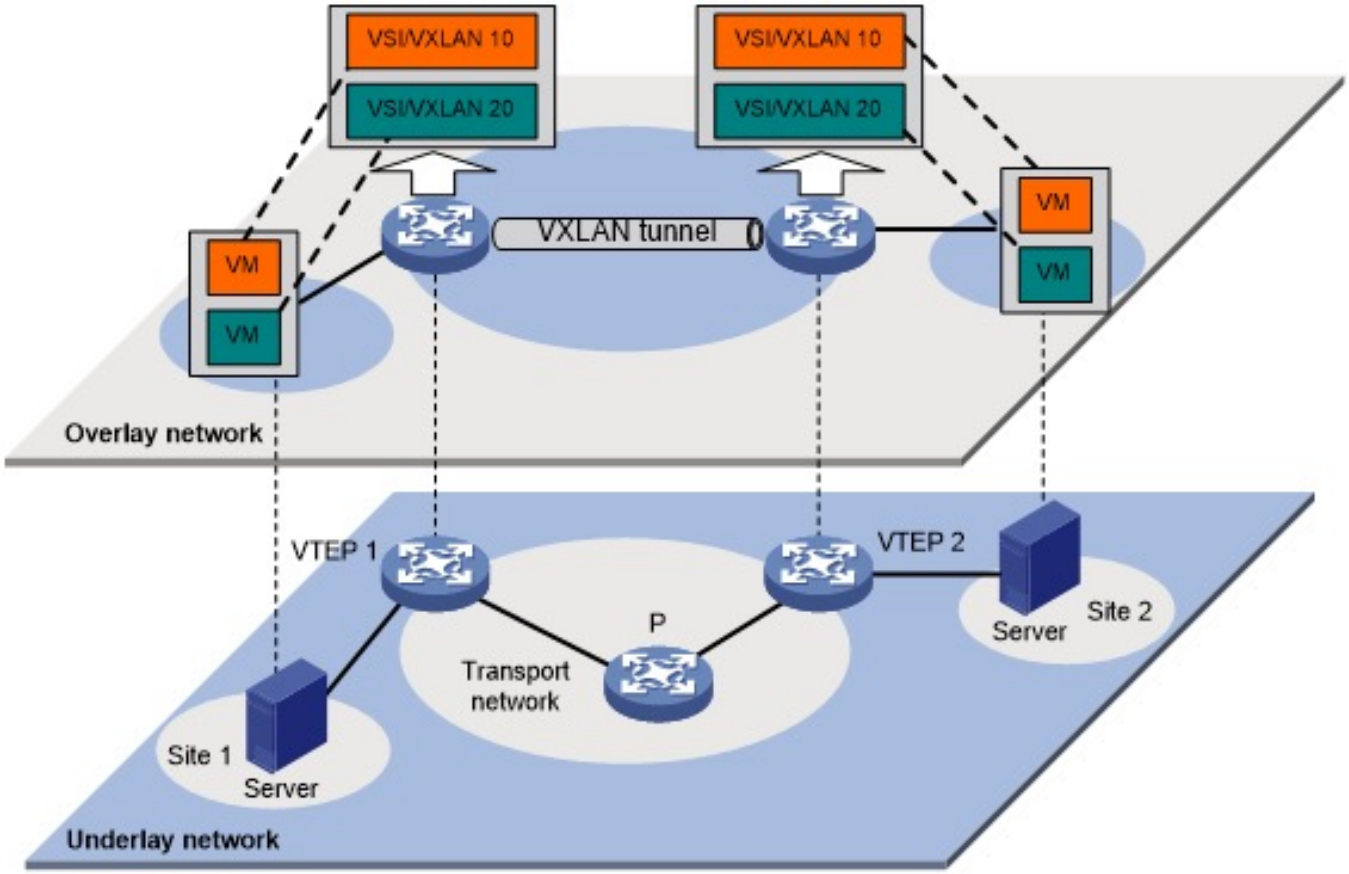
- 覆盖网络是建立在现有物理网络基础设施之上的虚拟网络
- 覆盖网络允许网络管理员独立于底层物理网络拓扑和配置创建、管理和维护虚拟网络
- 覆盖网络是使用各种隧道和封装协议创建的，这些协议将来自虚拟网络的流量封装在可以通过物理网络传输的数据包中
- 一些常见的覆盖网络协议包括VxLAN、NVGRE和GENEVE

覆盖网络图示


Overlay network
Overlay device
Underlay device
Overlay link
Underlay link
Underlay network

覆盖网络

- 通过 VxLAN 构建的 overlay network



虚拟私有云（Virtual Private Cloud）

HUAWEI CLOUD

Virtual Private Cloud (VPC)

Virtual Private Cloud (VPC) allows you to isolate online resources with virtual private networks. VPC enables your cloud resources to securely communicate with each other, the internet, and on-premises networks.

Alibaba Cloud

Virtual Private Cloud

A virtual private cloud service that provides an isolated cloud network to operate resources in a secure environment.

aws

Amazon Virtual Private Cloud (Amazon VPC)

Define and launch AWS resources in a logically isolated virtual network

Azure

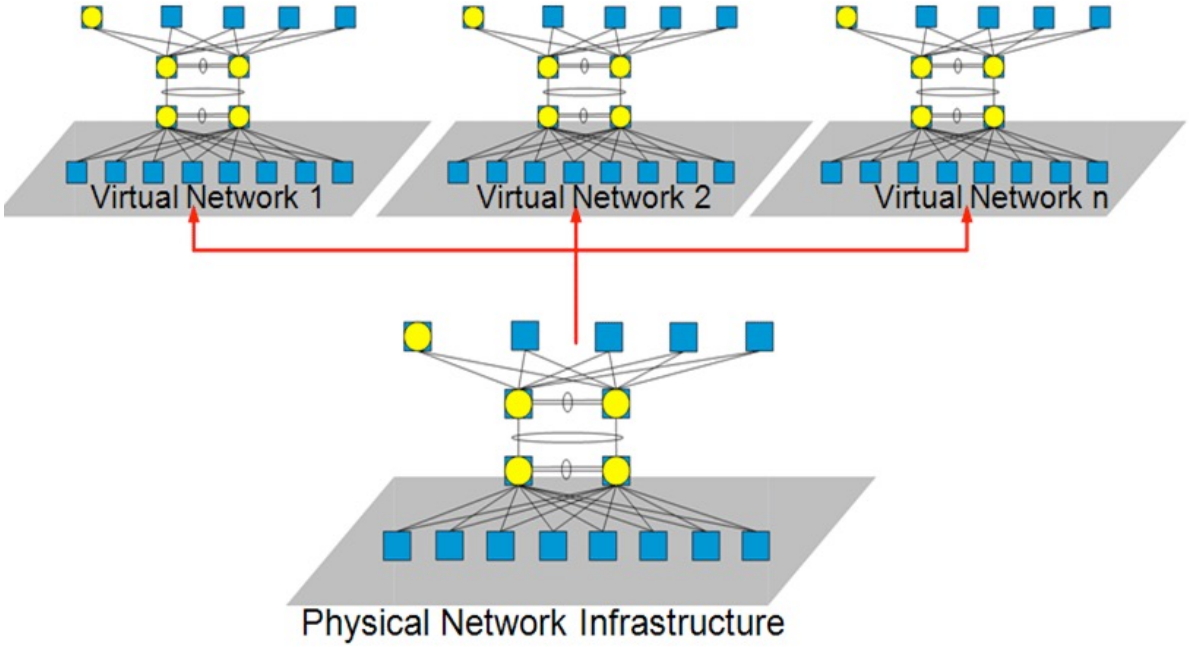
Azure Virtual Network

Create your own private network infrastructure in the cloud.

虚拟私有云

- 虚拟私有云（VPC）是一个租户专用的虚拟网络。它在逻辑上与云基础设施中的其他虚拟网络隔离。租户可以将实例、数据库和容器等云资源部署到虚拟私有云中。
- 租户可以定义自己的IP地址范围、创建子网以及配置路由表和网络网关的灵活性。租户能够构建一个类似于在自己的数据中心运行的传统网络的网络环境，并利用云的可扩展基础设施。

覆盖网络使能虚拟私有云



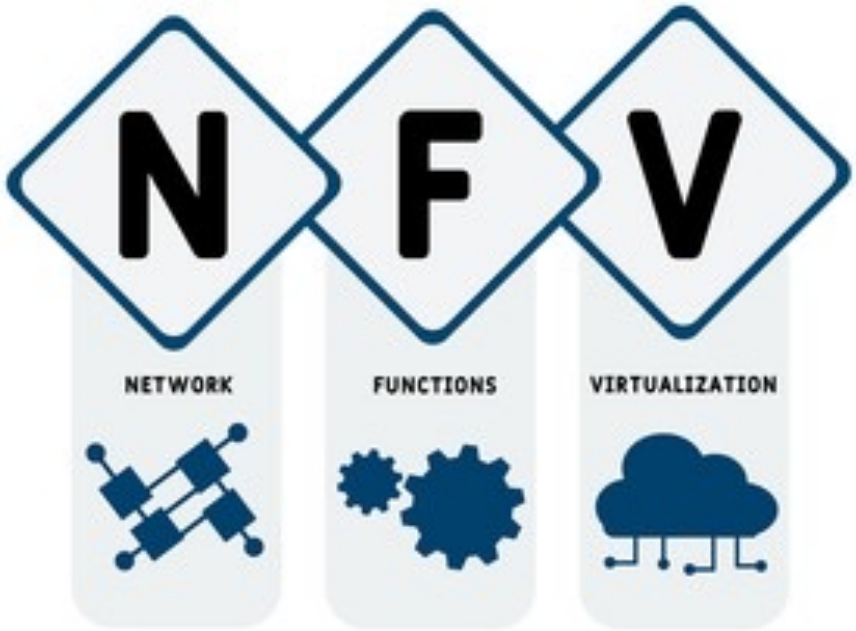
覆盖网络用于为每个租户的虚拟私有云创建隔离的虚拟网络。这些覆盖网络对流量进行封装和隔离，确保每个专有网络在共享基础设施上拥有自己的独立网络。

网络虚拟化技术

- 数据中心网络虚拟化
- 虚拟局域网（VLAN）
- 虚拟可扩展局域网（VxLAN）
- 网络功能虚拟化（NFV）

网络功能虚拟化

在传统网络中，路由、负载平衡、防火墙和入侵检测等功能通常由专用硬件设备执行。通过NFV（Network Function Virtualization），这些功能被虚拟化，可以在标准服务器上运行，从而减少了对专用硬件的需求。

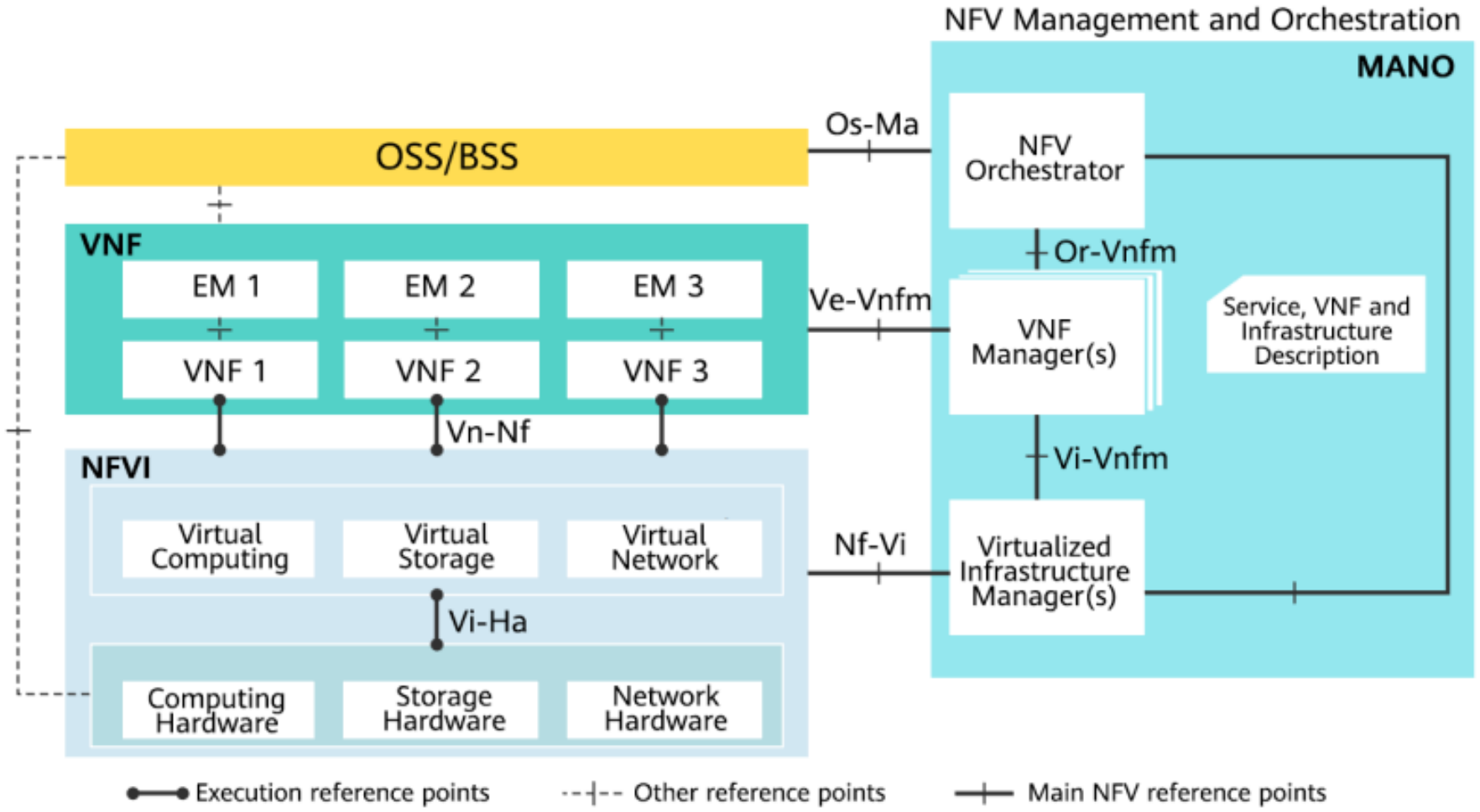


NFV 的优势

- 灵活的网络服务
  - ✓NFV提供了跨不同服务器运行VNF的灵活性，或者在需求变化时根据需要移动VNF，加快了网络功能和应用程序的交付。

- 更低的开销
  - ✓使用NFV，多个虚拟化网络功能可以在单个硬件服务器上运行。这意味着需要更少的物理硬件，降低硬件所需的空间和成本。

NFV 的架构


NFV Management and Orchestration
MANO
OSS/BSS
Os-Ma
NFV Orchestrator
Or-Vnfm
VNF
EM 1
EM 2
EM 3
VNF 1
VNF 2
VNF 3
Ve-Vnfm
VNF Manager(s)
Service, VNF and Infrastructure Description
Vn-Nf
Vi-Vnfm
NFVI
Virtual Computing
Virtual Storage
Virtual Network
Nf-Vi
Virtualized Infrastructure Manager(s)
Vi-Ha
Computing Hardware
Storage Hardware
Network Hardware
Execution reference points
Other reference points
Main NFV reference points

NFV 的架构

- 网络功能虚拟化基础设施（NFVI）
  - ✓与手机操作系统类似，NFVI 提供基础设施组件，以支持在硬件上运行网络应用程序所需的软件和容器管理平台。

- 虚拟网络功能（VNF）
  - ✓与手机上的应用程序类似，VNF 是提供网络功能的软件应用程序，如转发服务和 IP 配置。基于 NFVI 的标准化架构，VNF 可以独立于硬件运行。

- 管理和编排（MANO）
  - ✓MANO 提供了管理 NFVI 和 VNF 的框架，促进了运维人员的服务编排和设备管理。

中山大學 软件工程学院

SUN YAT-SEN UNIVERSITY SCHOOL OF SOFTWARE ENGINEERING

谢谢

陈壮彬

软件工程学院

https://zbchern.github.io/sse316.html