Virtualization

EE450: Introduction to Computer Networks

Professor A. Zahid

Virtualization

- Virtualization
  - Abstraction between the physical resources and their logical representation
  - Can be implemented in various layers of a computer system or network
    - Server Virtualization: Server virtualization refers to the partitioning of the resources of a single physical machine into multiple execution environments each of which can host a differentserver
    - Network Virtualization: Allows heterogeneous virtual networks that are isolated, independently managed to coexist over a shared physical network infrastructure. We have already studied VLANs, VPNs, etc.

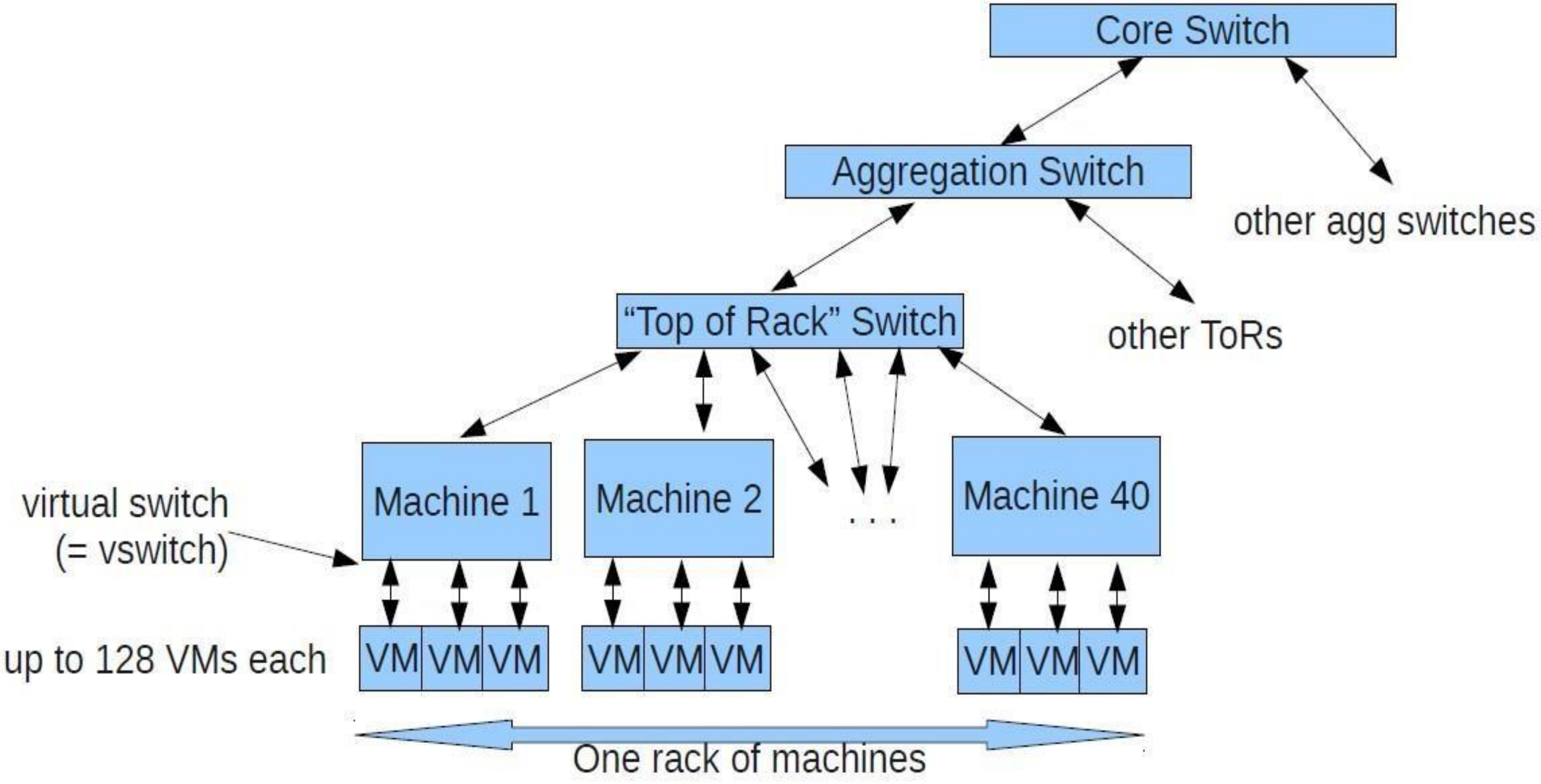
Motivation for Virtualization

- Data Centers consists of multiple physical servers. Measurement studies on these server farms noted that individual server usage was often as low as 15 percent for various reasons
- The consequence of this server sprawl with low usage was large financial outlays, both Capital and Operating Expenditures (Power, Cooling, etc…)
- Hence Virtualization came into the Picture.

Motivation for Virtualization (Cont.)

- Diverse operating systems
  - Running software for obsoleteplatforms
  - Research, experimentation, and testing
- Sharing a single host
  - Server consolidation (lower cost, energy)
  - Isolation of applications or customers/tenants
- Fast provisioning of new servers
- Snapshotting system state
  - Backup and redeployment
  - Migrating a VM to a different host machine

DCN with Server Virtualization



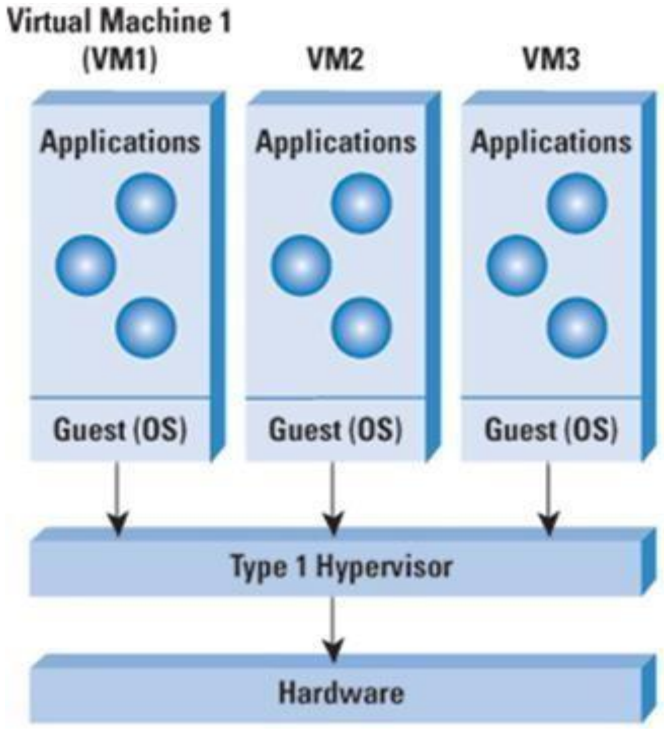
Server Virtualization

- Server Virtualization is a piece of software used to run multiple Virtual Machines (VMs) on a single physical server to provide the same functions as multiple physical machines.
- Known as a Hypervisor, the virtualization software performs the abstraction of the hardware to the individual VMs.

Hypervisor and Virtual machines

- A hypervisor is implemented on a server either directly running over the hardware (Type 1 hypervisor) or running over an operating system (Type 2 hypervisor).
  - Supports the running (and scheduling) of multiple VMs
  - Providing them a unified access to the CPU, memory, and I/O resources on the physical machine.
- A VM typically runs an operating system and applications. The applications are not aware that they are running in a virtualized environment .
  - Note that the OS inside the VM may (or may not be) aware of Virtualization. If it is, it would require modifications to run over a hypervisor

Hypervisors in Virtualization



Virtual Machine 1 (VM1)

VM2

VM3

Applications

Applications

Applications

Guest (OS)

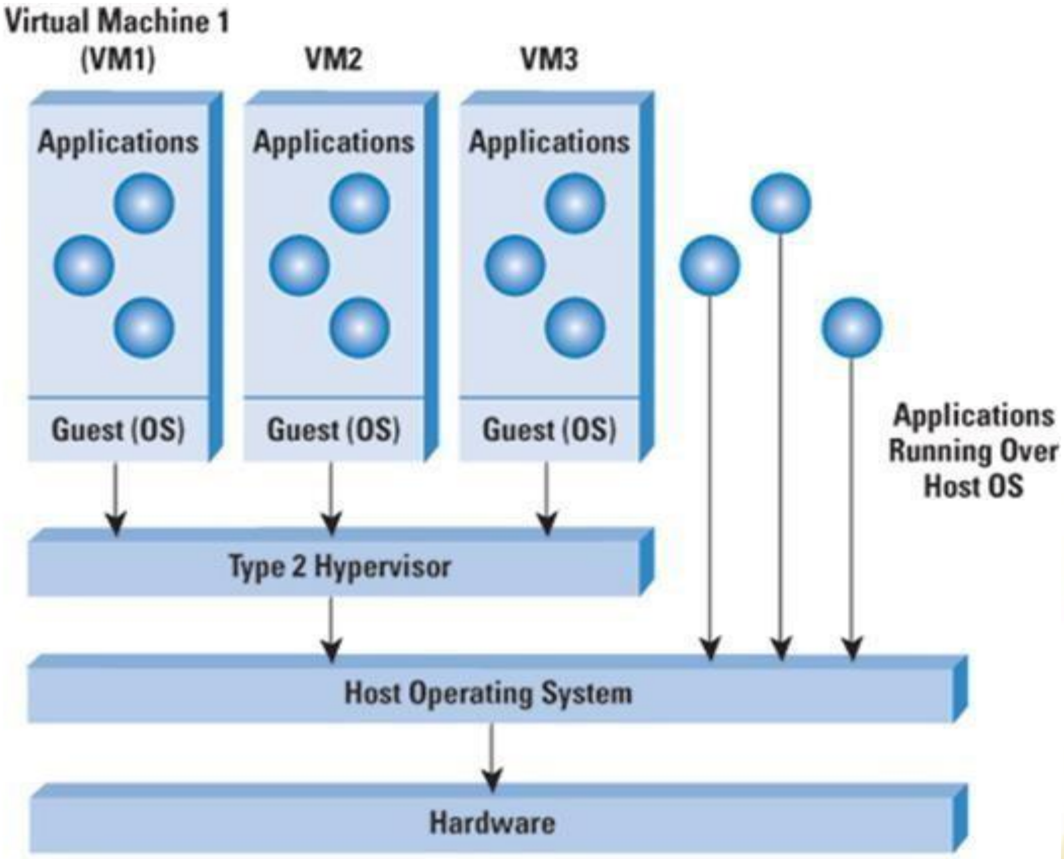
Guest (OS)

Guest (OS)

Type 1 Hypervisor

Hardware

Type 1 Hypervisor-Based Virtualization



Virtual Machine 1 (VM1)

VM2

VM3

Applications

Applications

Applications

Guest (OS)

Guest (OS)

Guest (OS)

Applications Running Over Host OS

Type 2 Hypervisor

Host Operating System

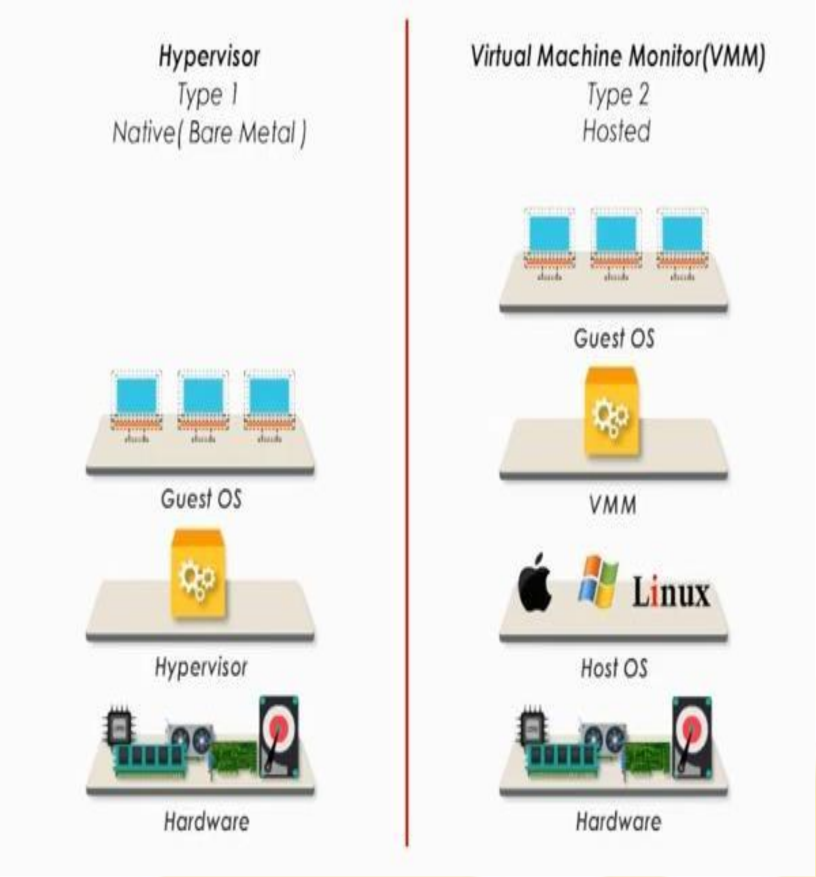
Hardware

Type 2 Hypervisor-Based Virtualization

Types of Hypervisors

Type I hypervisor runs directly over the host hardware. It control the hardware and manage the VMs. They act like OSs Vmware ESXI, Microsoft Hyper V, Citrix's Zen etc…

Type II hypervisor (VMM) runs over the host OS (Like an application program). VM are created and managed by VMM through the host OS. Virtual Box, Vmware workstation



Migration of VMs

- VM Migration allows you to move an entire VM (including Application and OS) from one host machine to another another and continue operation of the VM on the second machine.
- This advantage is unique to virtualized environments because you can take down physical servers for maintenance with minimal effect on running applications.
- Migration can be performed after suspending the VM on the source machine and restart after moving its attendant information to the target machine
- To lower the downtime, Migration can be performed while the VM is running and resuming its operation on the target machine after all the state is migrated.

Problem with VM Migration

- VM migration introduces its own set of problems. The most common scenario is when a VM is migrated to a different host on the same Layer 2 topology (or VLAN)
- Consider the case where a VM with open TCP connections is migrated. TCP connections will not see any downtime except for a short "hiccup." However, after the migration, IP packets destined for the VM will need to be resolved to a different MAC address or the same MAC address but now connected to a different physical switch in the network so that the connections can be continued without disruption.
- Proposed solutions include an unsolicited Address Resolution Protocol (ARP) request from the migrated VM so that the switch tables can be updated. Another proposed solution is to use a pseudo-MAC address for the VM that is externally managed

Advantages/Disadvantages of VMs

Cost effective
Simplified Management
Threat Isolation
Backup & Recovery
Comprised Performance
- Guest and host share limited resources

Increased Complexity
Increased Risk
- If host machine fail, all VMs fail as well

License Cost

Virtual Network Components

- Virtual Switch: a virtual switch lets the VM on the same host communicate with each other using the same protocols that would be used over physical switches.
- Virtual switches are created as needed by the Hypervisor. You can connect one or more virtual machines to a switch.
- You connect a virtual machine to a switch by selecting the Virtual Network Adapter (vNIC)you want to connect in the Virtual Machine Control Panel, then configuring it to use the desired virtual network. A VM can be configured with one or more VNICs each of which has its own IP address and MAC address ⇒ From a Networking standpoint, VMs have the same properties as Physical machines

Virtual Network Components (Cont.)

- Host Virtual Adapter : It allows communication between the host computer and the VMs on that host computer. The host virtual adapter is used in "Host-only" and "NAT-based" configurations".
- The host virtual adapter is not connected to any external network unless you set up special software on the host computer - a proxy server, for example - to connect the host-only adapter to the physical network adapter.
- The software that creates the host virtual adapter is installed when you install the Hypervisor. A host virtual adapter is then created automatically when you boot the host computer.
- You can set up additional host virtual adapters as needed.

Virtual Network Components (Cont.)

- Virtual Bridge - The bridge connects the VMs to the external LAN used by your host computer. It connects the VNIC in the VM to the PNIC in host computer.
- The bridge is installed during the installation of the Hypervisor. It is set up automatically when you create a new VM using “Bridged Networking” configuration.
- Additional Virtual Bridges can be set up for use in configurations that require connections to more than one PNIC on the host computer

Virtual Network Components (Cont.)

- NAT device - The NAT device allows you to connect your virtual machines to an external network when you have only one IP network address on the physical network, and that address is used by the host computer.
- You can, for example, use NAT to connect your virtual machines to the Internet through the host PNIC or wireless Ethernet adapter.
- The NAT device is set up automatically when you install the Hypervisor. On a Linux host, you must choose to make NAT available to your VMs
- DHCP server - The DHCP server provides IP network addresses to VMs in “Host-only” and “NAT-based” configurations”. External DHCP is used for Bridged configurations

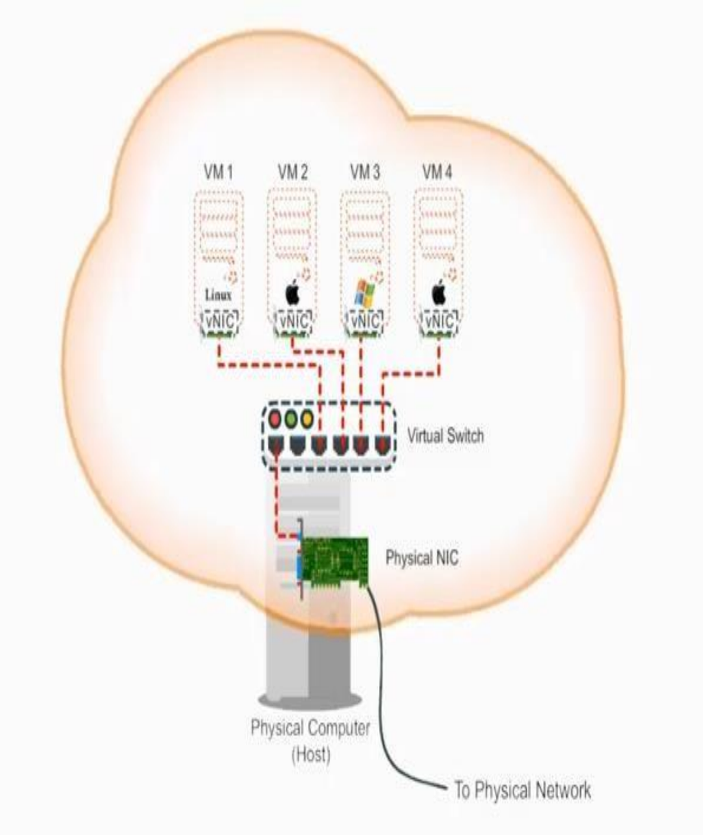
How do VMs Communicate with the Network?

VMs can communicate on the same host or with other physical machines on the network. They need two virtual items namely

- Virtual NIC
- Virtual Switch/Bridge

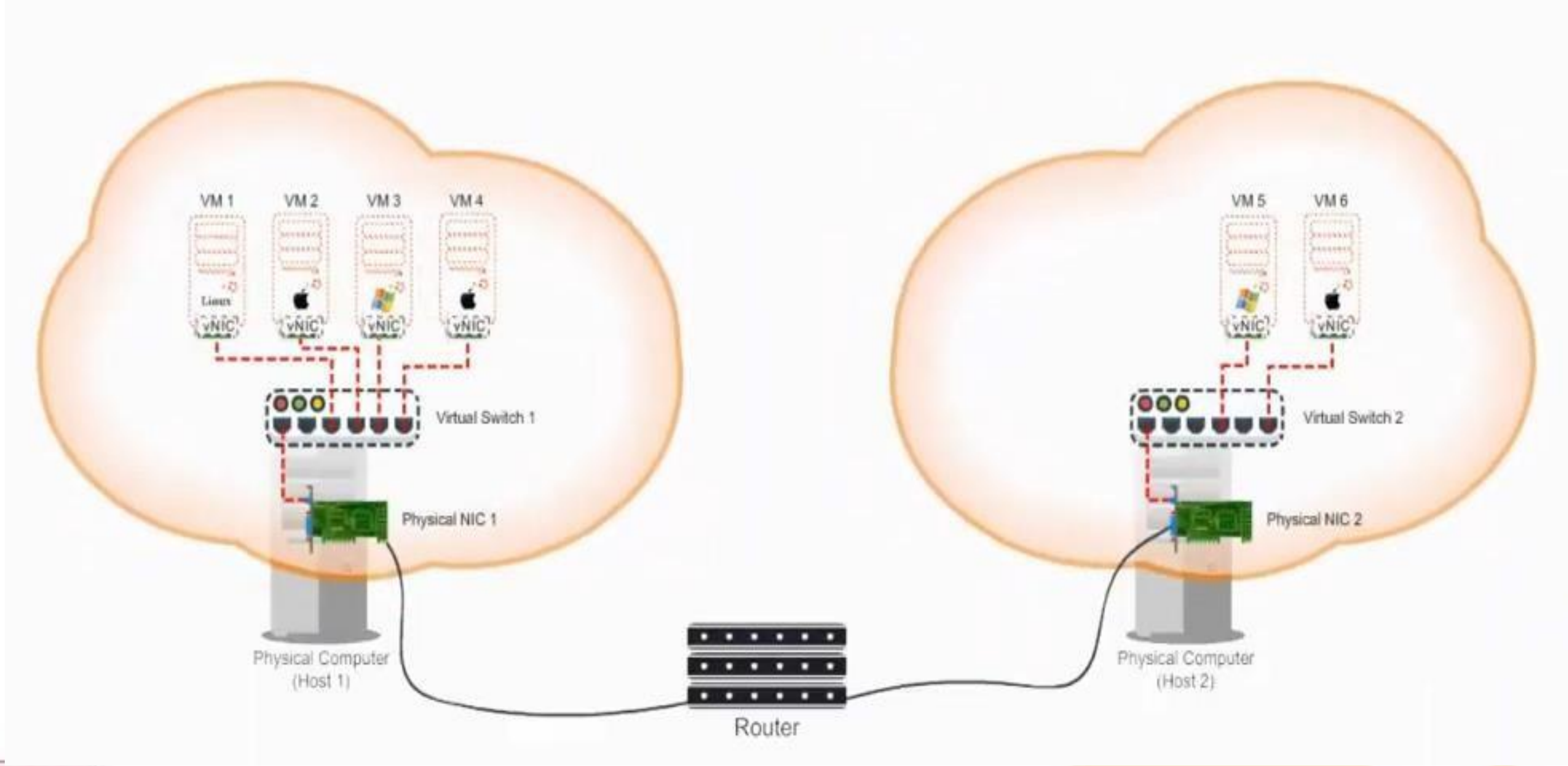
A Virtual switch is a logical Layer-2 device that passes frames between nodes. Virtual NICs are connected to Virtual Ports on Virtual Switch

Virtual switch is connected to the physical network through physical NIC



Physical Switches vs. Virtual Switches

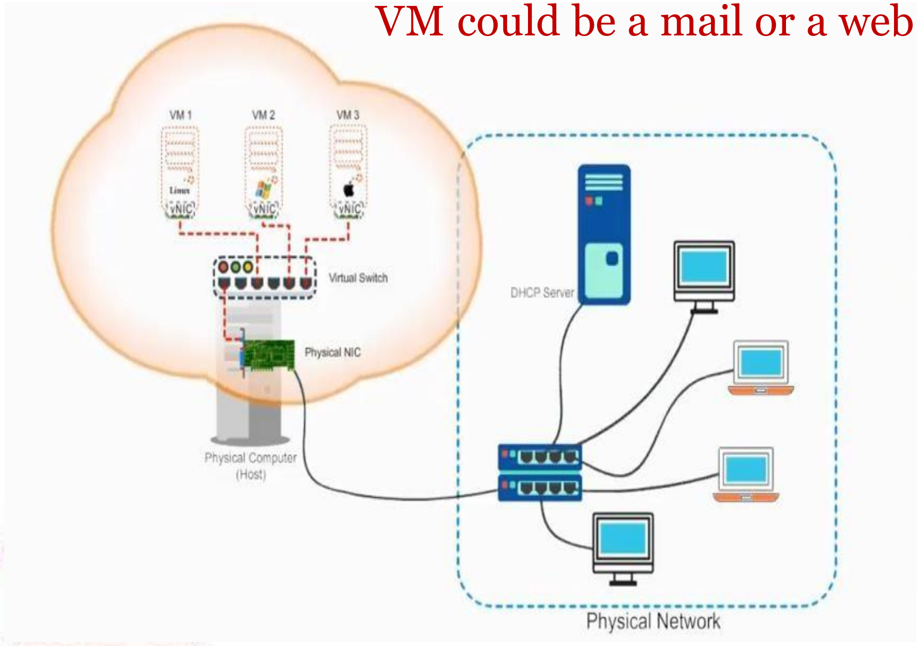
Each Virtual Switch creates a separate broadcast domain
To connect two broadcast domains we need a Router

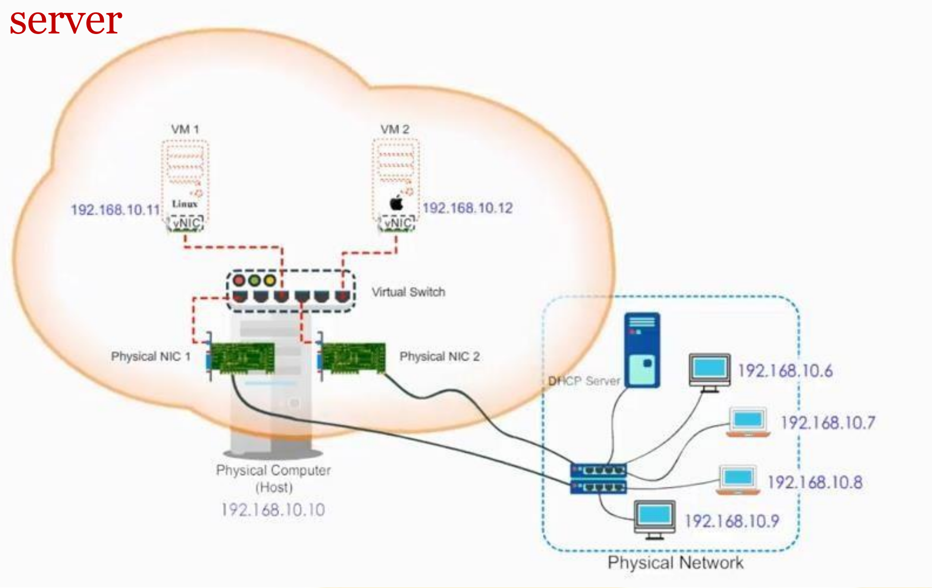


Bridged Connection

In a bridged connection mode, the VM connects to the physical network directly via the PNIC. The host NIC is acting as a bridge to all VMs. VMs obtain their IP addresses from a DHCP server on the physical network. The VM appears just like any other node on the network. Each VM can also be connected via a dedicated PNIC. IP addresses of VM are visible to any device

VM could be a mail or a web server



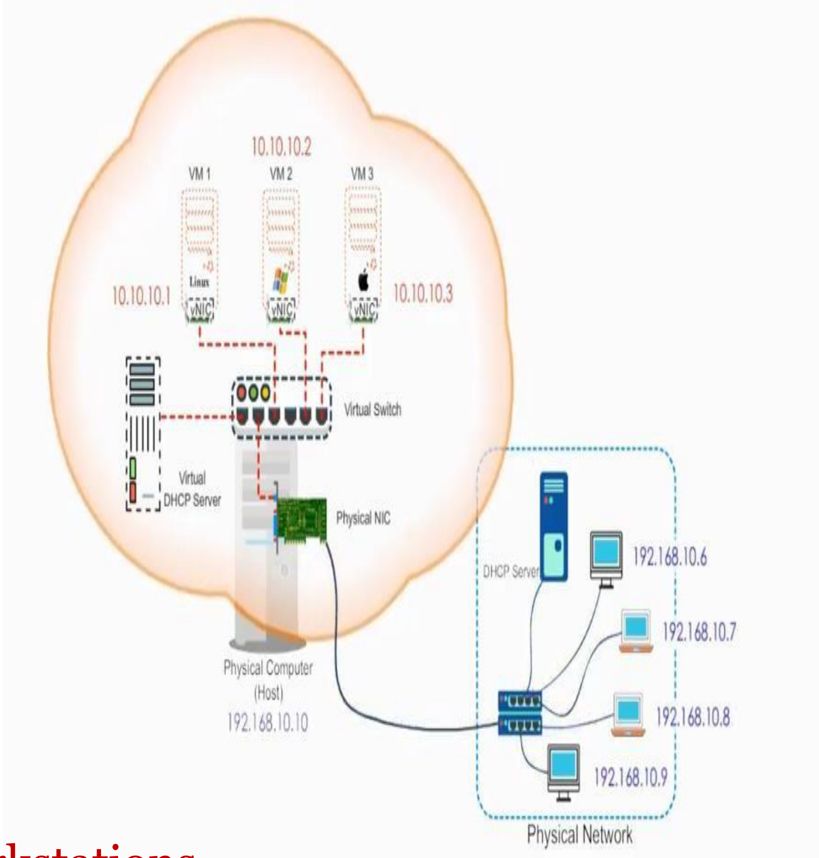


NAT-based Connection

VM rely on the host to act as a NAT device

A Virtual DHCP server will assign private IP addresses to VMs. They form a separate network.

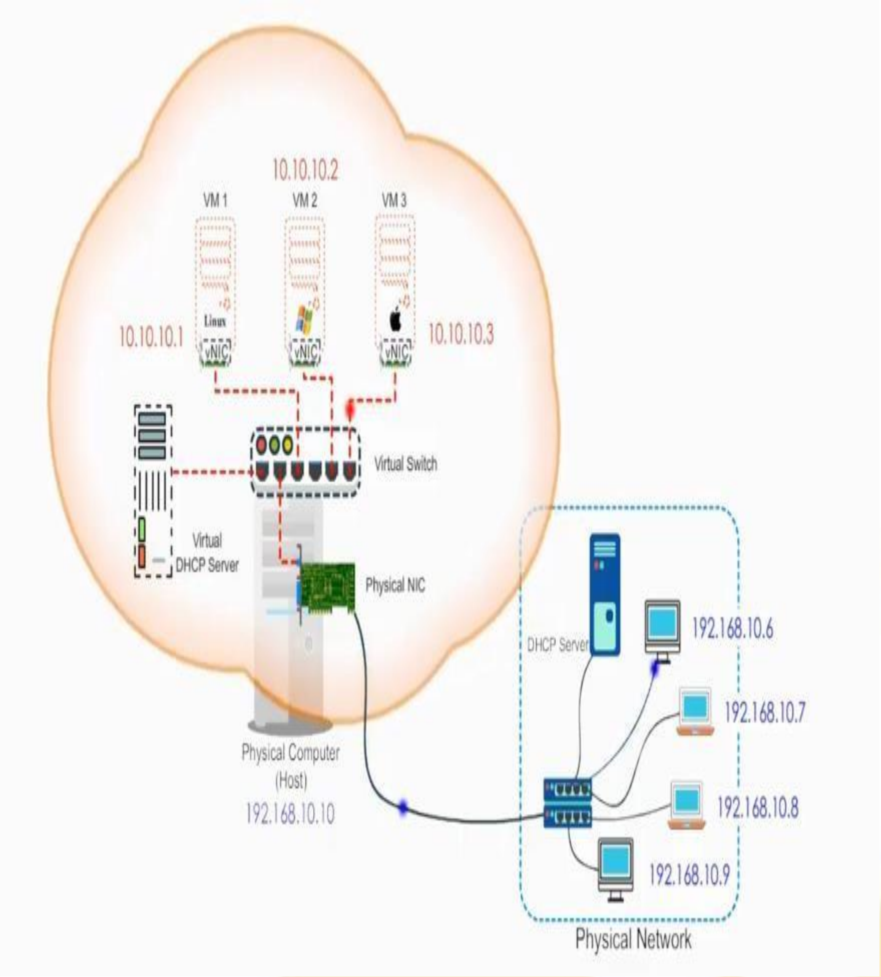
The host is sitting between the two networks and translate IP addresses. Physical devices see VM traffic as if it is coming from host.



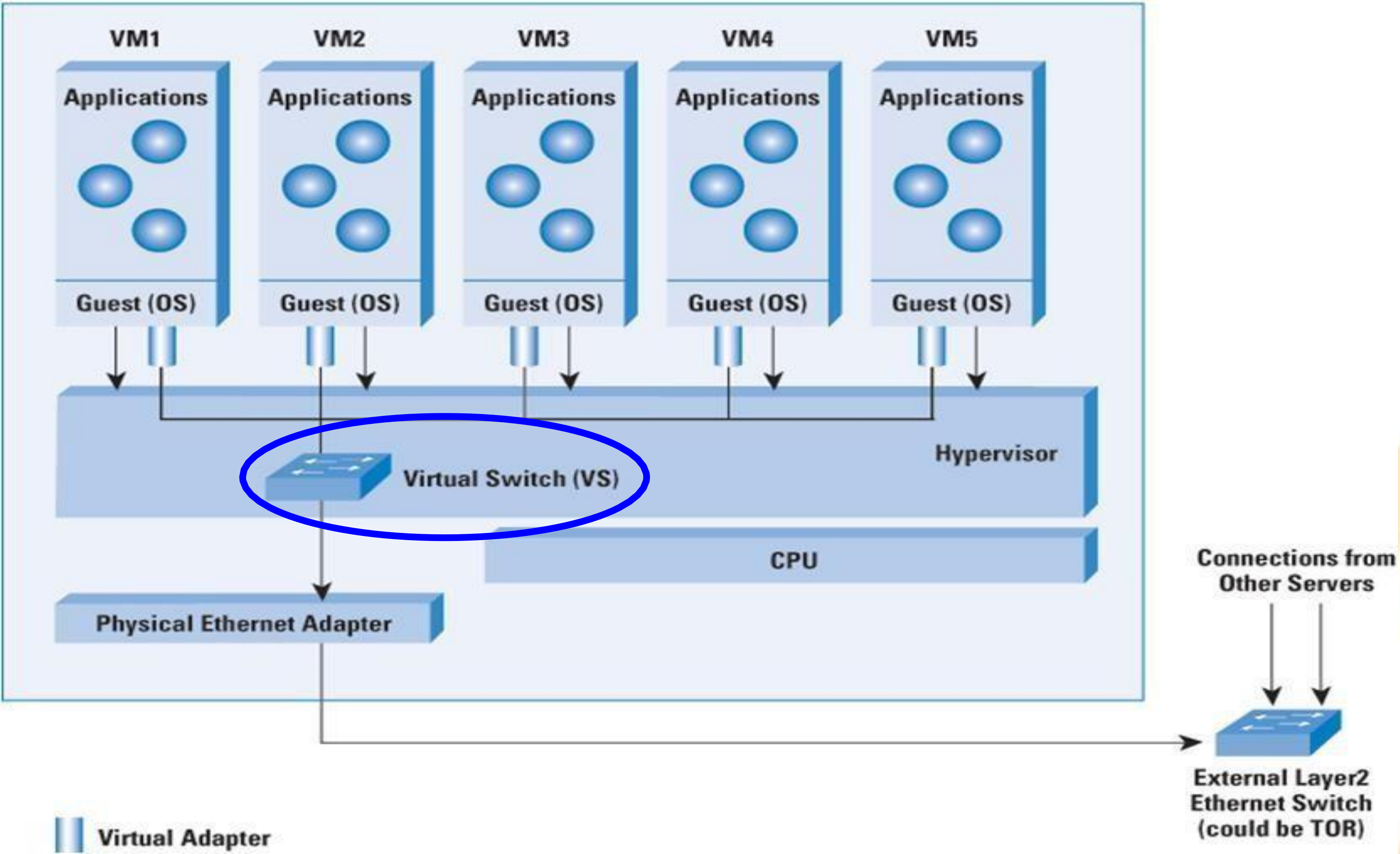
VM are mainly used as client workstations

Host-only Connection

In host-only mode, the VM can communicate with each other and with the host machine but can not communicate with any other machine on the physical network. Used when we want set up an “isolated” private network. Internal packets will not leak to the physical network



Virtual Ethernet Switch


VM1
VM2
VM3
VM4
VM5
Applications
Guest (OS)
Hypervisor
Virtual Switch (VS)
CPU
Physical Ethernet Adapter
Connections from Other Servers
External Layer2 Ethernet Switch (could be TOR)
Virtual Adapter

A Virtual Switch is similar to a Physical Switch

- Both maintain a MAC-Port forwarding Table and performs the following
  - Look at the destination MAC address of an arriving frame
  - Forward the frame to one or more ports
  - Avoid unnecessary delivery (unlike the Hub)
- Both supports port-based VLANs. Each port of a vSwitch (or a physical switch) can be configured as an
  - Access Port: Belong to a single VLAN
  - Trunk Port: Belong to multiple VLANs (Tagging is needed)
- Port Mirroring: Both supports copying frames to a mirror port allowing network Admin the ability to analyze, test and debug using a sniffing tool. He can also monitor applications such as IDS (Intrusion Detection)

A Virtual Switch is different to a Physical Switch

- The Hypervisor provides a direct channel from the vNIC for configuration information and hence there is no need for "Learning" unicast addresses or perform IGMP snooping (to learn MG membership)
- The Hypervisor enforces a single tier networking topology, i.e., there is no way to interconnect multiple virtual switches ⇒ No loops ⇒ No need for SPT algorithm
- Network traffic can not flow directly from one vSwitch to another within the same host. Virtual switches provide all ports you need in one switch. Furthermore, Virtual switches can NOT share the same PNIC.
- No entries in a virtual switch table that points to a port on another virtual switch. Every destination the switch looks up matches ports on the same vSwitch as the port where the frame came from.

Virtual Ports

- The ports of a virtual switch provide logical connections among VMs and between a VM and the physical network It works much like a physical Ethernet switch.
- The vSwitch detects which virtual machines are logically connected to each of its virtual ports and uses that information to forward traffic to the correct virtual machines.
- A vSwitch can be connected to physical switches by using physical Ethernet adapters, also referred to as uplink adapters, to join virtual networks with physical networks.

Up-Link Ports

- Uplink ports are ports associated with the PNIC. They provide connection between the virtual network and the physical network
- Some Virtual switches may not connect to the physical network and hence they have no up-link ports. An example is a virtual switch providing connectivity from a virtual firewall device to the VM it is protecting.
- VNIC connects to virtual ports when the VM is powered. THE VNIC updates the virtual switch port with its MAC address at initialization and when it changes its MAC address.
- If a Virtual switch port is configured as a VLAN trunk, its uplink port (on physical switch) must all be configured as a Trunk port

Virtualization Demands on Switching

- Consider a data center with 100 servers, each with 16 virtual machines but with one physical 10-Gbps Ethernet connection to the external switch from each physical machine. If we were to carry forward the model where each physical server is replaced with its virtual equivalent but still needs to be addressable (MAC/IP Addresses layer)
- We would need 16 MAC and IP addresses for the virtual servers that now reside "on top" of the single physical link, for a total of 1600 addresses across all servers!!! This problem is exacerbated when you increase the number of VMs per server
  - Switching between MAC addresses belonging to the virtual machines is done by the virtual switch inside the server.

Virtual Switch

- The virtual switch treats the physical link as an uplink to the external physical switch.
- This intra-machine Virtual Machine (VM) switch with an uplink to the external switch is completely in line with access and aggregation switch topologies where the access layer is subsumed inside the server.
- Note that each physical host can have more than one virtual switch to support greater logical segmentation. In such cases, it is common for each of the virtual switches to have its own physical uplink to the external Ethernet switch.

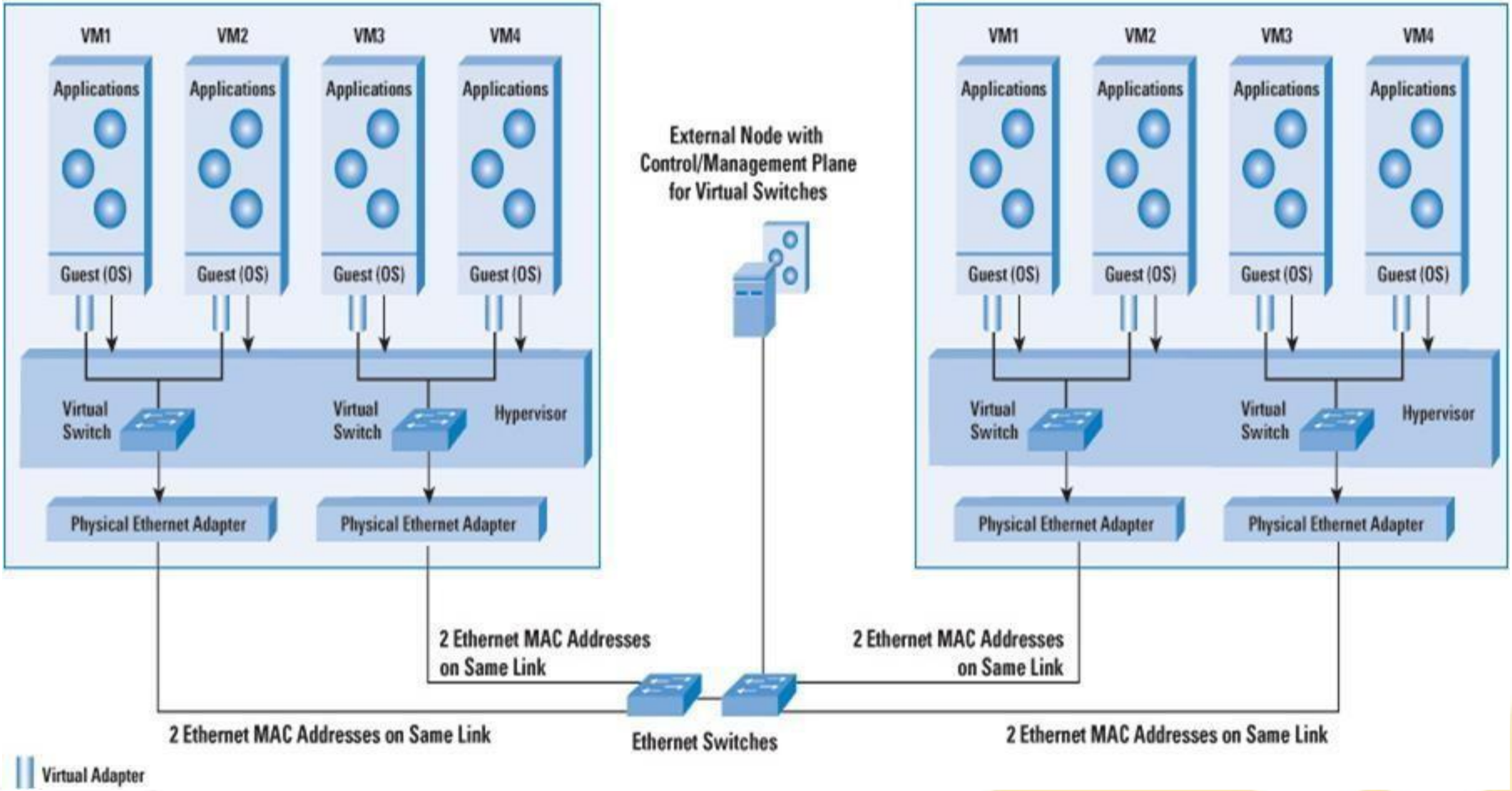
Virtual Switch (Continued)

- The virtual switch does not need to learn MAC addresses like a traditional switch. It assumes that all destination-unknown frames should be forwarded over the physical link (or uplink to the physical switch).
- In addition, it switches traffic between the intra-machine VMs according to policy. For example, you could prohibit two VMs on the same machine from communicating with each other by configuring an access control list on the virtual switch.
- The VMs may all be on the same or on different VLANs. Broadcasts and intra-VLAN traffic are forwarded according to the rules for each VLAN. In effect, the virtual switch is a simple function that is used for aggregation and access control within a physical server containing VMs.

Inter-VM Traffic

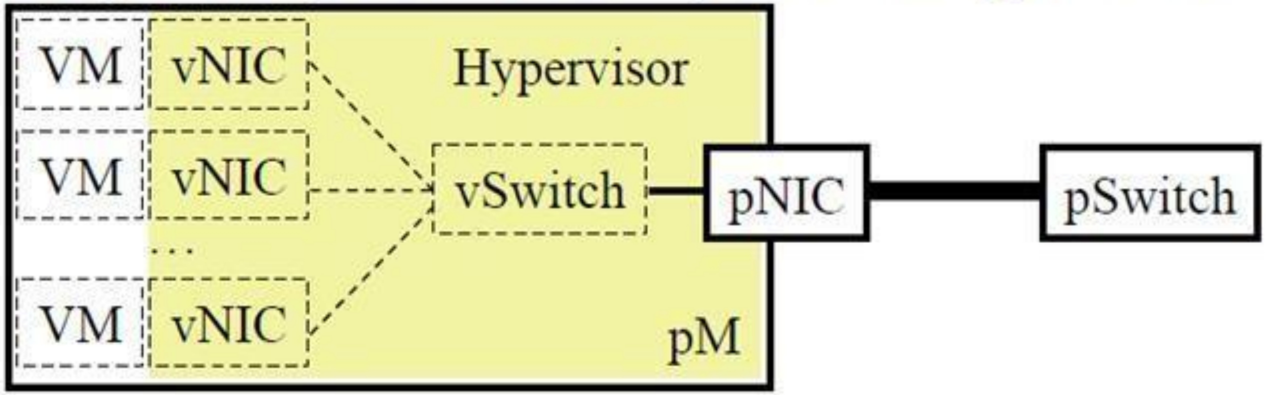
- Inter-VM traffic within the same physical machine is not visible to the network and cannot be subjected to appropriate monitoring by network administrators.
- The IEEE is discussing approaches to providing external network switches the visibility into the intra-VM traffic.
- One option includes "hair pinning," where inter-VM traffic would still be carried over to an external switch and brought back to the same physical server.

Virtual Switch Aggregation &Management


VM1
VM2
VM3
VM4
Applications
Guest (OS)
Virtual Switch
Hypervisor
Physical Ethernet Adapter
External Node with Control/Management Plane for Virtual Switches
2 Ethernet MAC Addresses on Same Link
Ethernet Switches
Virtual Adapter

vSwitch

- **Problem**: Multiple VMs on a server need to use one physical network interface card (pNIC)
- **Solution**: Hypervisor creates multiple vNICs connected via a virtual switch (vSwitch)
- pNIC is controlled by hypervisor and not by any individual VM
- **Notation**: From now on prefixes p and v refer to physical and virtual, respectively. For VMs only, we use upper case V.



Virtual Edge Bridging

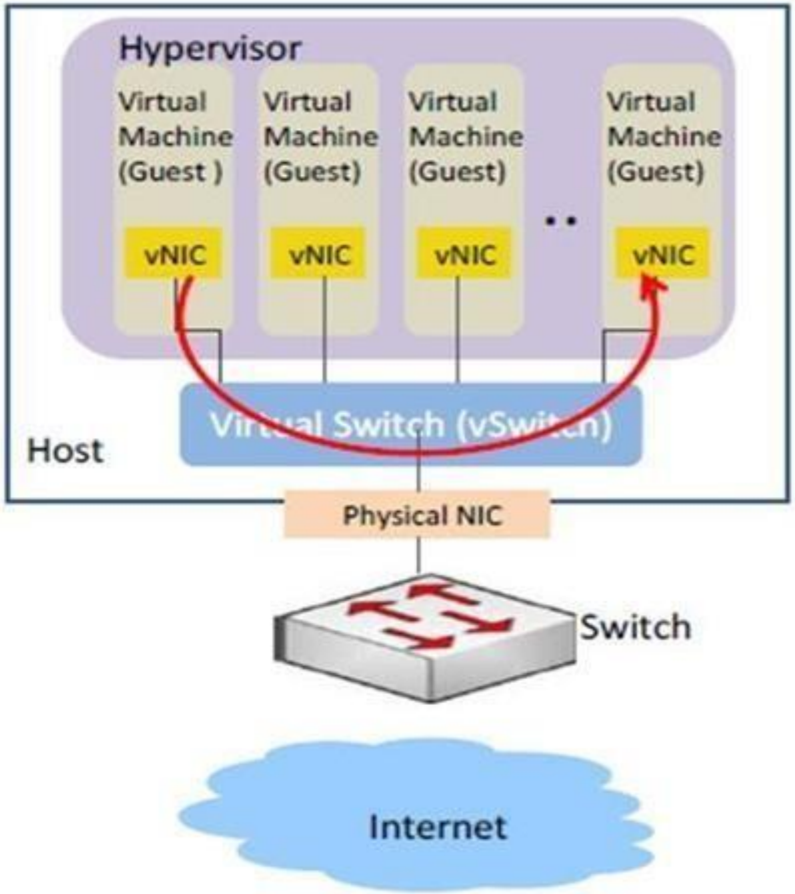


Figure 1. Virtual switch and physical switch

- The hypervisors implement vSwitch
- Each VM has at least one virtual network interface cards (vNICs) and shared physical network interface cards (pNICs) on the physical host through vSwitch
- Administrators don't have effective solution to separate packets from different VM users
- For VMs reside in the same physical machine, their traffic visibility is a big issue

Virtual Ethernet Port Aggregator

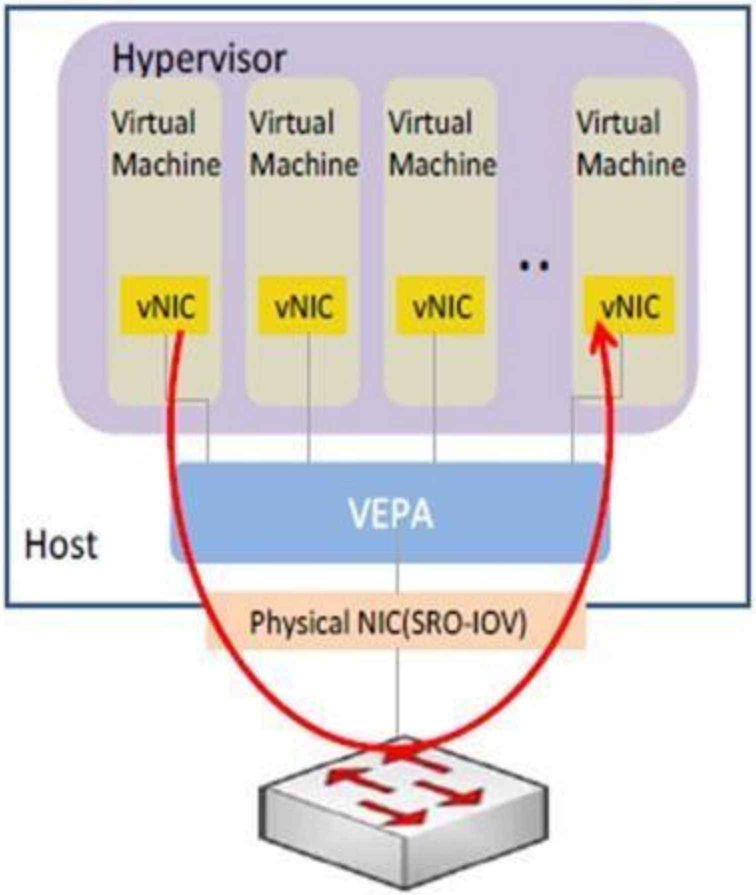
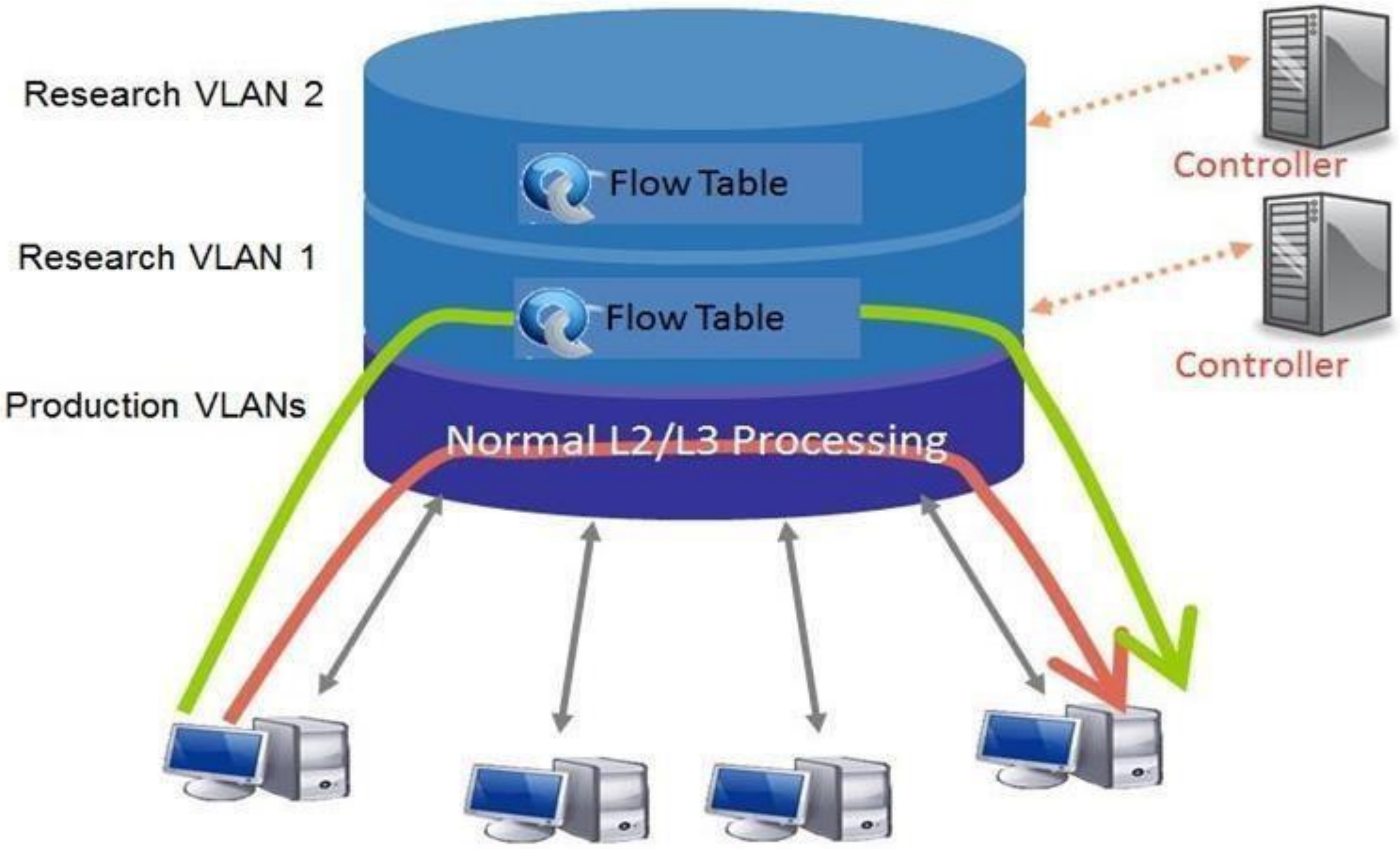


Figure 2. Virtual Ethernet Port Aggregator (VEPA)

- VEPA software update is required for host servers in order to force packets to be transmitted to external switches
- An external VEPA enabled switch is required for communications between VMs in the same server
- VEPA supports "hairpin" mode which allows traffic to "hairpin" back out the same port it just received it from--- requires firmware update to existing switches

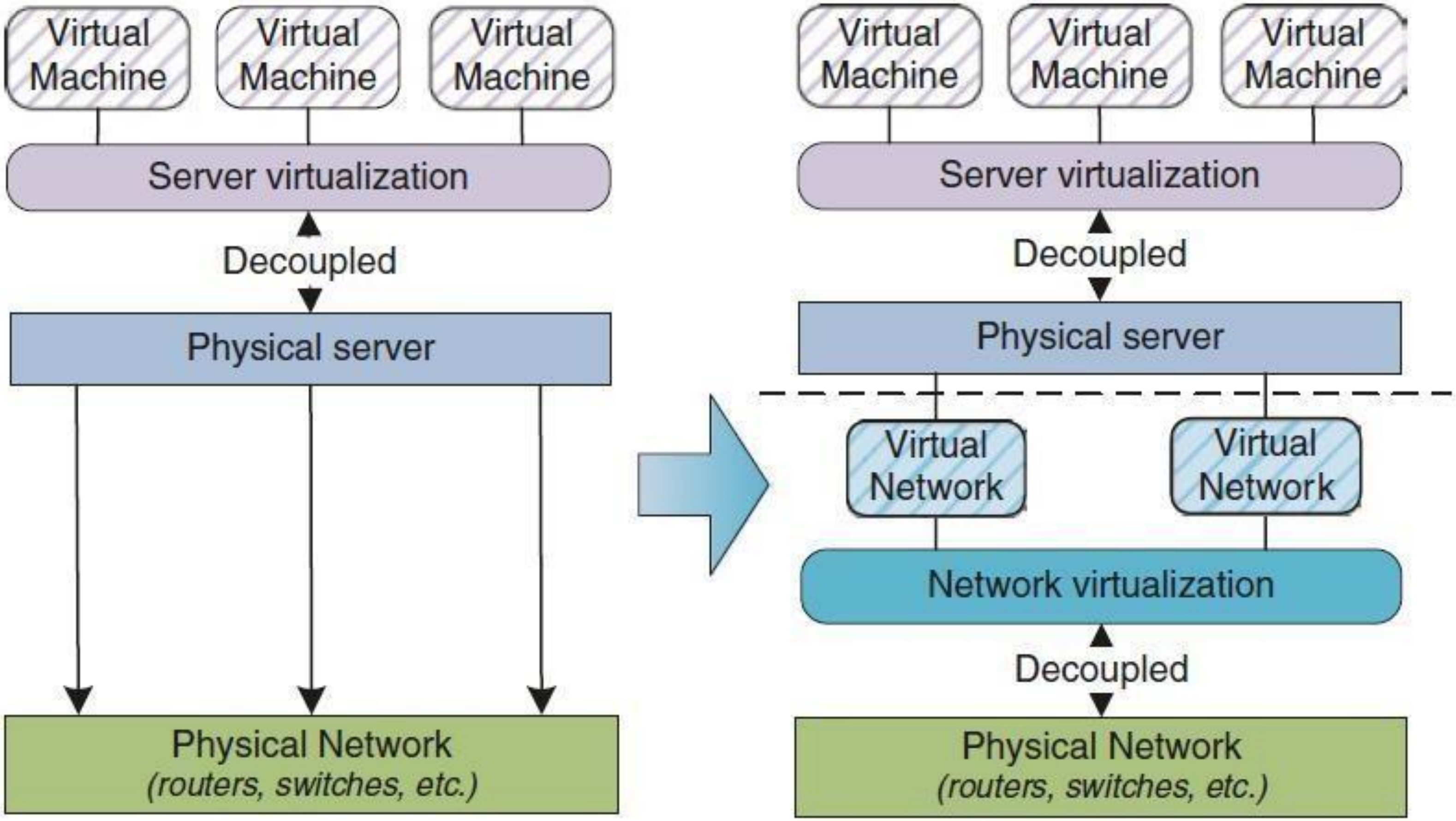
Switch-based Virtualization



Network Virtualization

- **Decoupling the services provided by a network from the physical infrastructure**
- **Virtual network is a "container" of network services, provisioned by software**
- **Faithful reproduction of services provided by a physical network**
- **Main Idea: Sharing the network**
  - Different controllers for different users/traffic
  - Isolation (bandwidth, table space, flow space)
- **Main Idea: Abstracting the topology**
  - One big virtual switch
  - Many virtual switches to one physical switch
  - Arbitrary network topologies

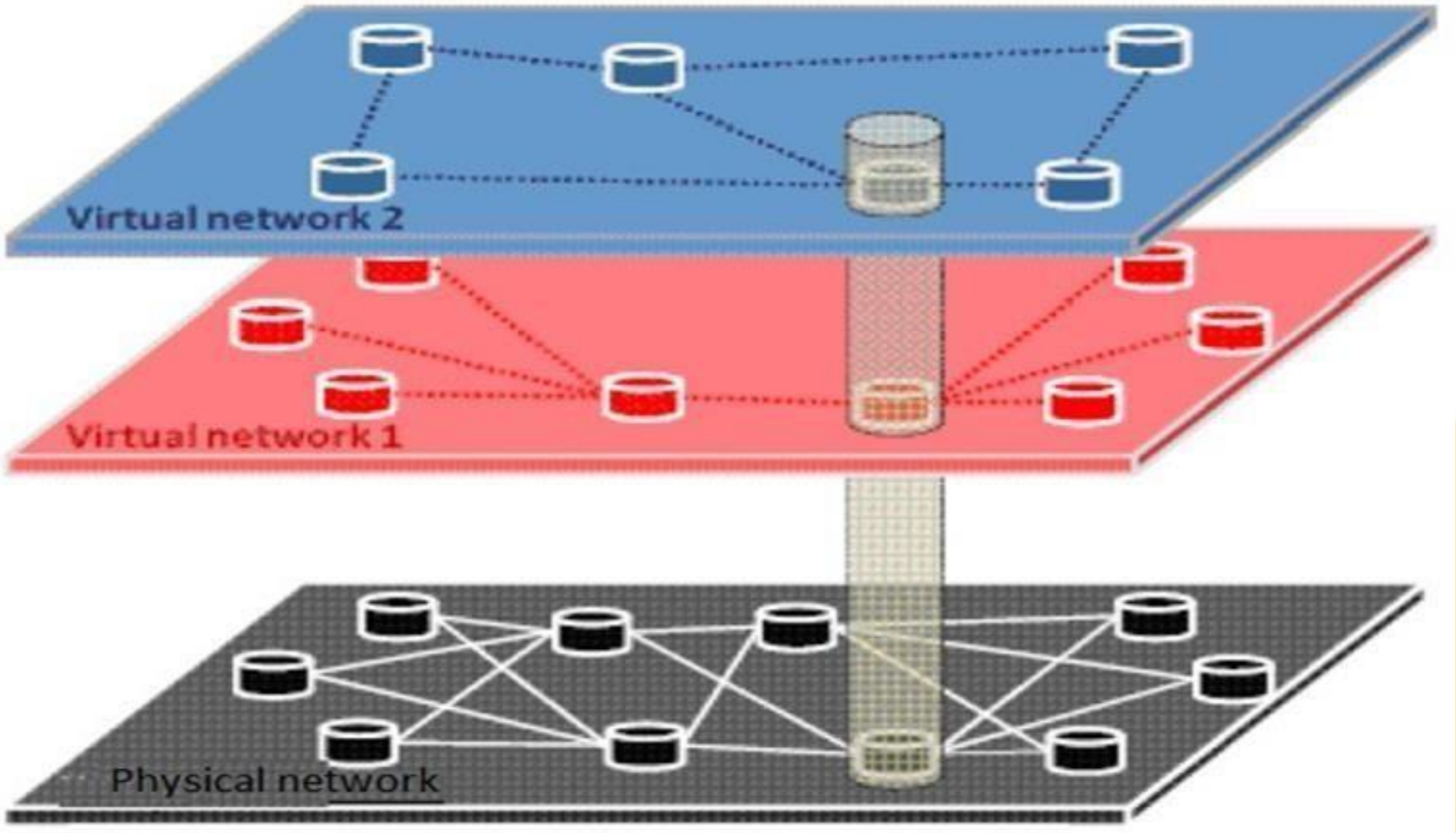
Server vs. Network Virtualization



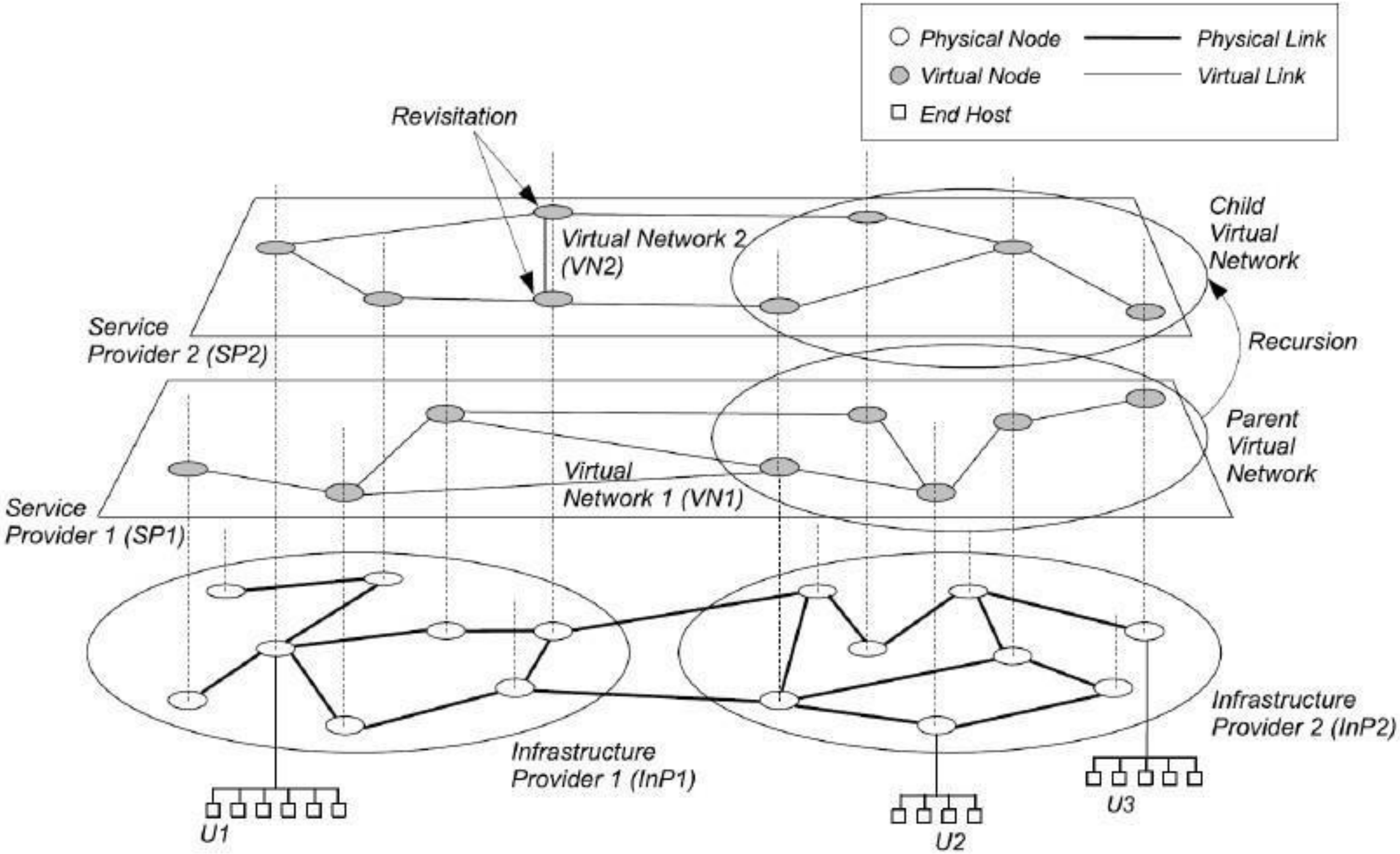
Reasons for Network Virtualization

- **Multiple administrative groups**
  - Different departments on a campus
- **Multiple customers**
  - Tenants in a shared data center
  - Researchers on a shared infrastructure
- **Experiments vs. operational network**
  - Support research without breaking real services
- **Expanding a network’s footprint**
  - Lease components in another carrier’s network
- **Multiple services or applications in one domain**

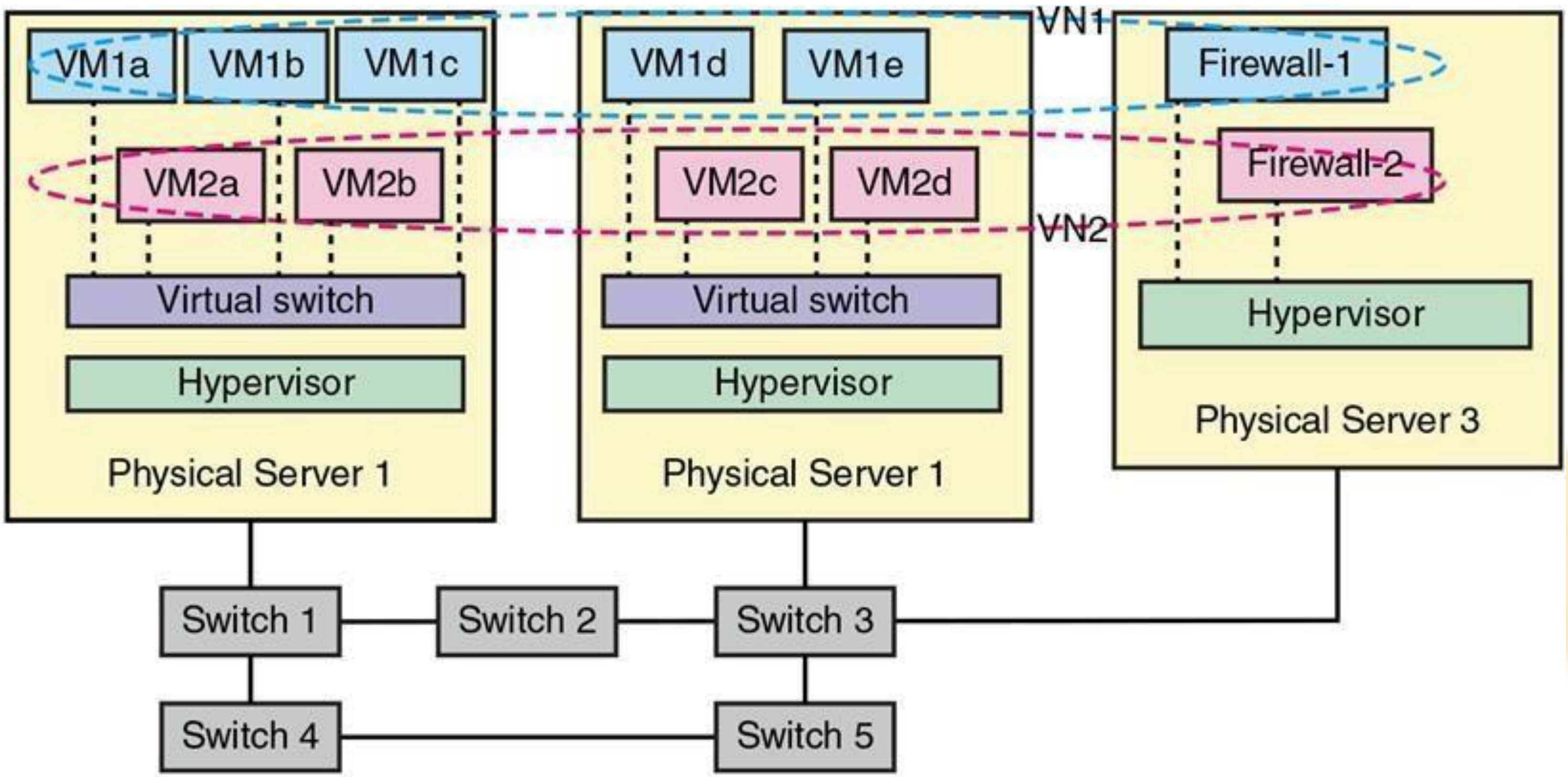
Network Slicing Model



Example of Network Virtualization

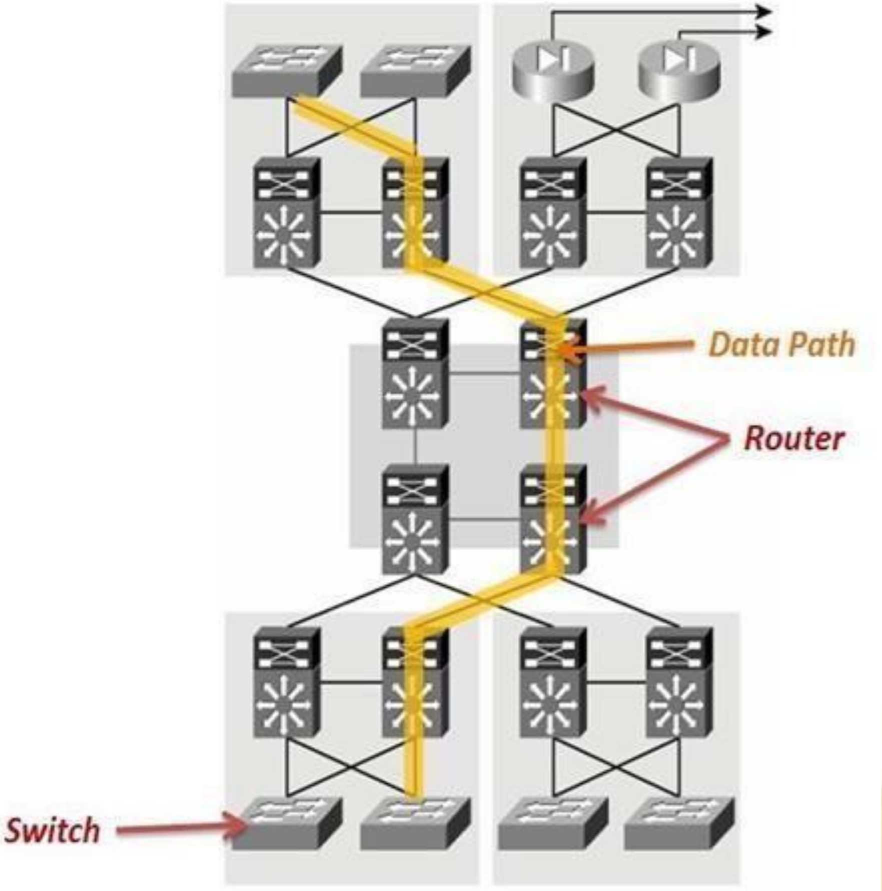


Example of Network Virtualization



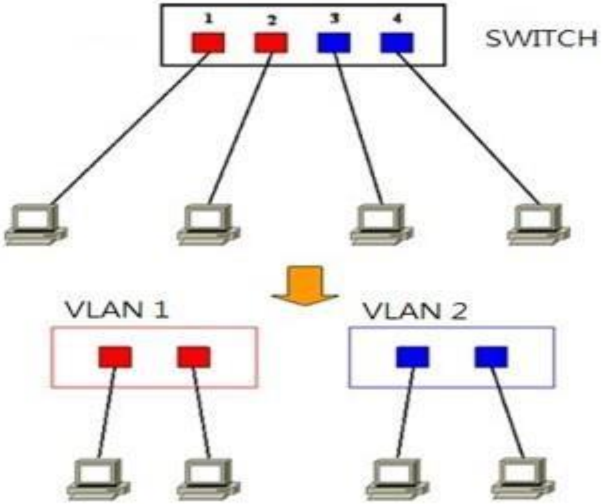
Components of Network Virtualization

- Two Virtualization components :
  - Device Virtualization
    - Virtualize physical devices in the network
  - Data Path Virtualization
    - Virtualize communication path between network access points

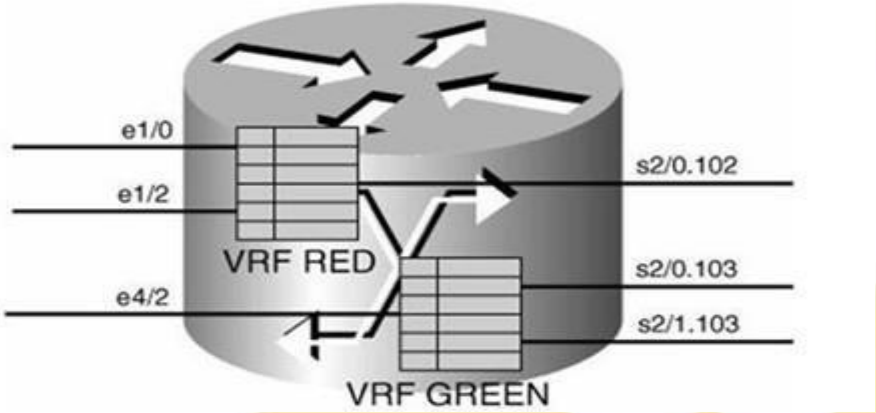


Device Virtualization

- Device Virtualization
  - Layer 2 solution
    - Divide physical switch into multiple logical switches.

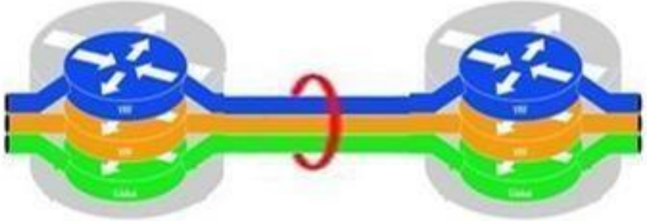


  - Layer 3 solution
    - VRF technique (Virtual Routing and Forwarding)
    - Emulate isolated routing tables within one physical router.

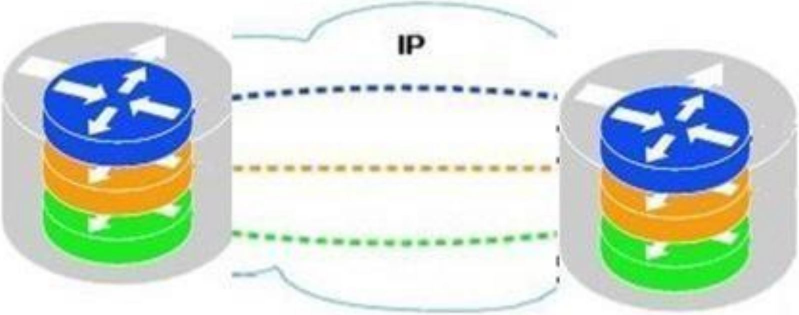


Data Path Virtualization

- Data Path Virtualization
  - Hop-to-hop case
    - Consider the virtualization applied on a single hop data-path.
  - Hop-to-cloud case
    - Consider the virtualization tunnels allow multi-hop data-paths.

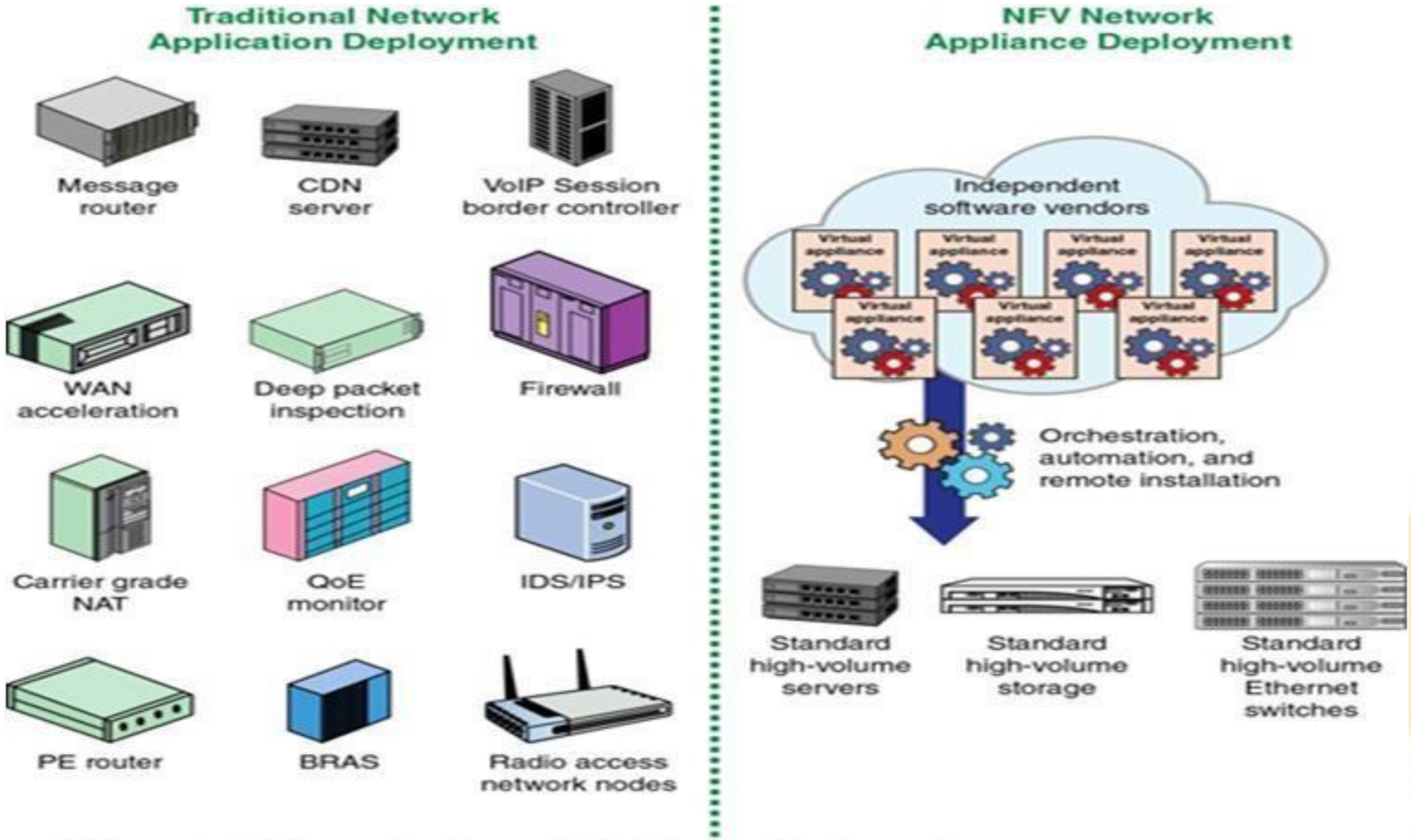


L2 based labeling allows single hop data path virtualization



Tunnels allow multi-hop data path virtualization

Network Functions Virtualization


Traditional Network
Application Deployment
Message router
CDN server
VoIP Session border controller
WAN acceleration
Deep packet inspection
Firewall
Carrier grade NAT
QoE monitor
IDS/IPS
PE router
BRAS
Radio access network nodes
NFV Network
Appliance Deployment
Independent software vendors
Virtual appliance
Orchestration, automation, and remote installation
Standard high-volume servers
Standard high-volume storage
Standard high-volume Ethernet switches

CDN = content delivery network
WAN = wide area network
NAT = network address translation
QoE = quality of experience
VoIP = voice over Internet Protocol

IDS = intrusion detection system
IPS = intrusion prevention system
PE = provider edge router
BRAS = broadband remote access server