Course and Labs

Design Network in AWS

Mohanraj Shanmugam

Course and Labs

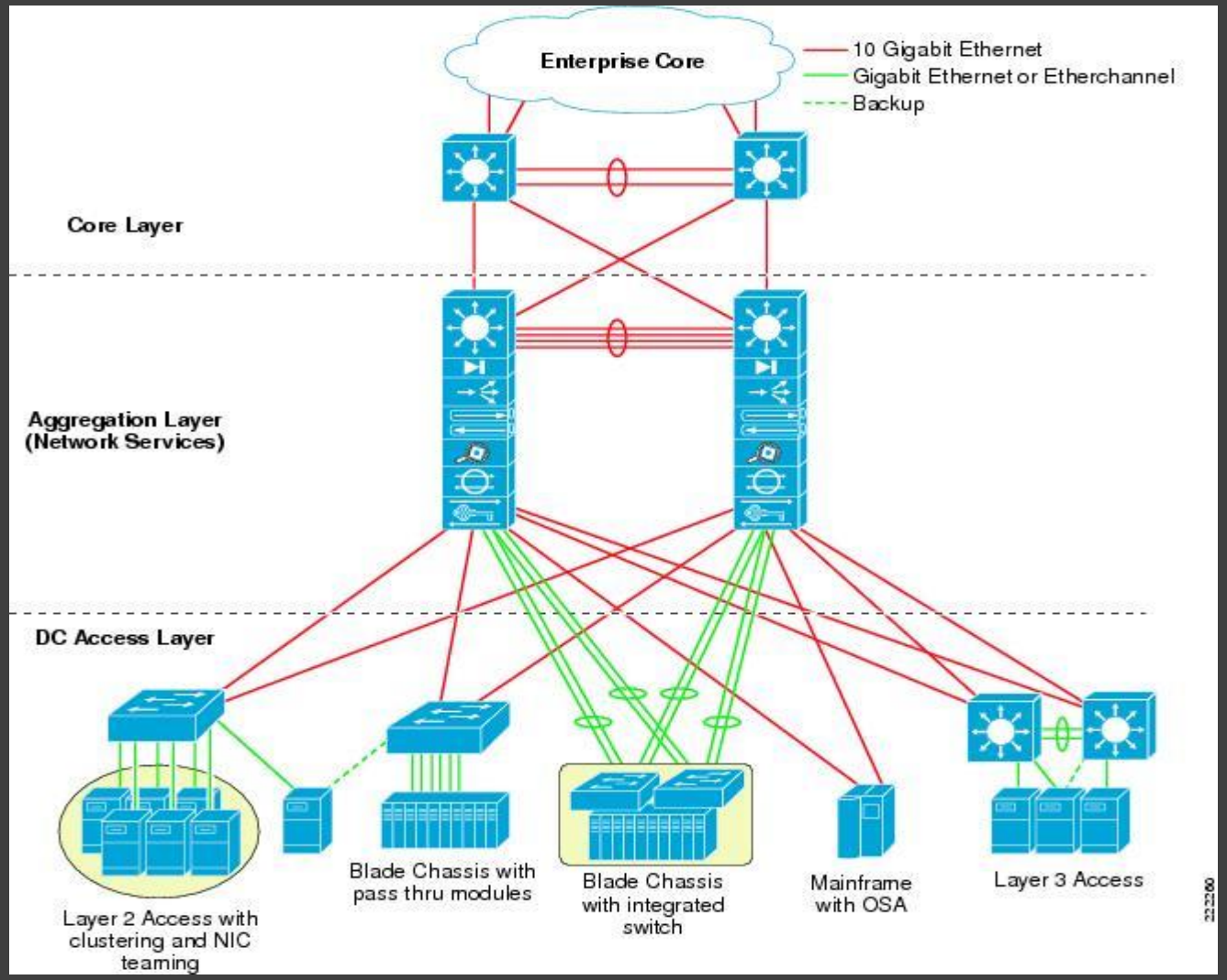
Topic 1: Datacenter Networking and SDN

Mohanraj Shanmugam

AWS Networking

- AWS Networking will give a overview of how we network the AWS compute (EC2), Storage, Database and Other Services.
- All Amazon network are virtualized and It works same way as our Virtual network in our datacenter.
- Amazon Virtual Network is as scalable as other AWS Resources
- Lets understand how to setup a virtual Enterprise network in a Datacenter before starting the AWS Networking

Traditional Physical Datacenter Network



- All Servers are connected to Access layer switch
- Each server will have multiple NIC cards, Team or bond Two or more NIC Cards for High Availability and Link aggregation is configured at Access switch level to support Team or Bond.
- Each server requires multiple networks like Production, Backup , Clustering and Management Networks.
- Each network connected to different NIC cards or Bonds and wired separately
- Each network is Segmented by VLAN so that it can communicate within the segmented server
- All access layer switches are connected to Aggregation layer switch where all segmentation and routing takes place across networks
- All Network Functions like Load balancer, Firewall, VPN, Proxy and Intrusion prevention system is connected to aggregation layer
- All external networks like Internet and leased line and WAN network Terminate at Aggregation Layer
- Core or Metro layer connects between Data centers and Provide Layer 3 and Layer 2 high speed routing between Datacenters

Virtualize Enterprise Network

- The network team is being bombarded with configuration requests that can take days or weeks to handle, There are emerging Network Technologies which will automate and move towards the Programmable Network to provide as a service
- The Three Main Categories of Virtual Enterprise Network are:
  - Network Virtualization
  - Network Function Virtualization
  - Software Defined Networking
- Lets understand each of this in detail

Network Virtualization

- Network virtualization is the process of combining hardware network resources and software network resources into a single administrative unit.
- The goal of network virtualization is to provide systems and users with efficient, controlled, and secure sharing of the networking resources.
- There are two types of Network Virtualization
  - Internal virtualization
  - External virtualization

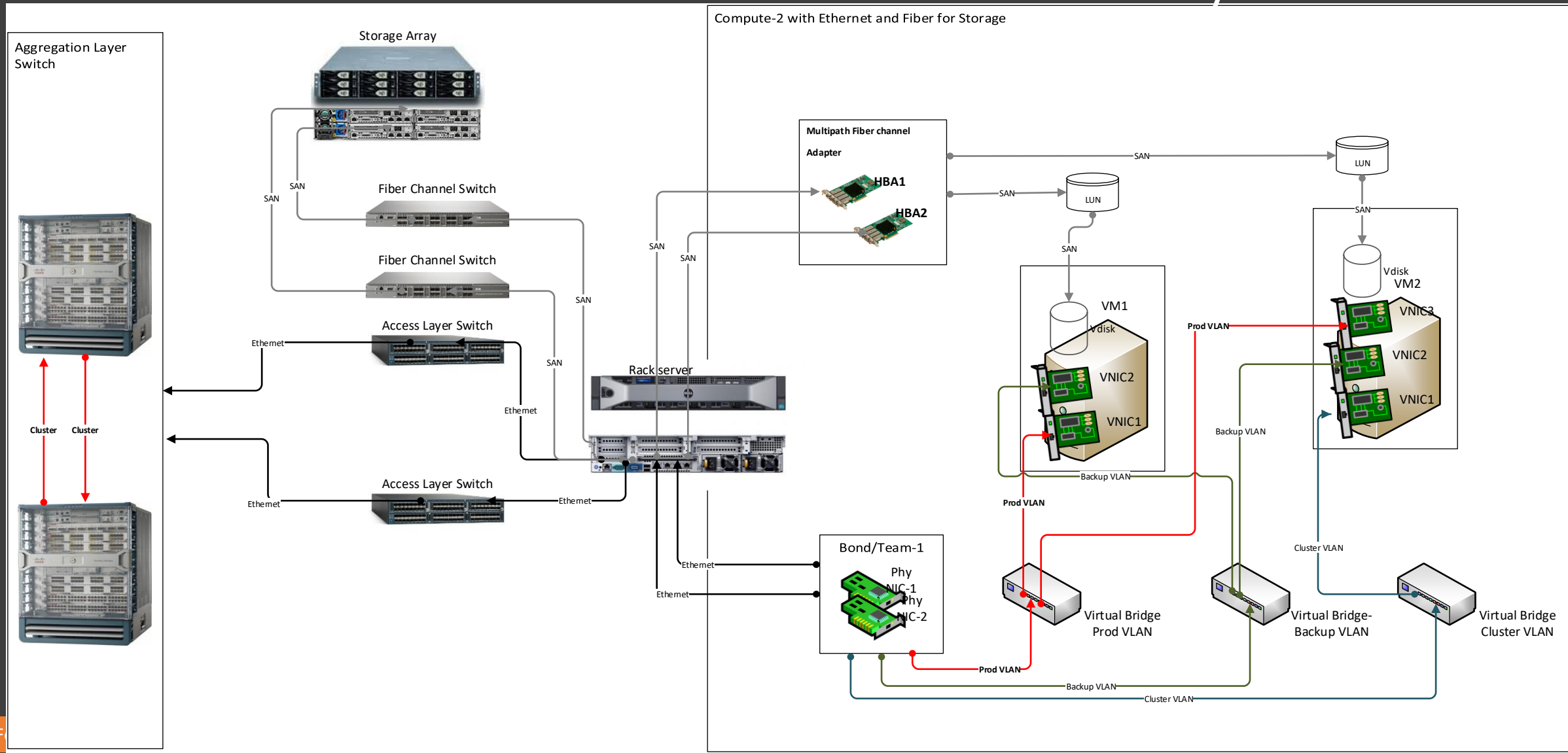
External virtualization

- External network virtualization combines or subdivides one or more local area networks (LANs) into virtual networks to improve a large network's or data center's efficiency.
- External virtual networks are administered by software as a single entity.
- Examples of external virtual networks include large corporate networks and data centers.
- Components of External Virtualization
  - Access Switches
  - Fiber Switches
  - Convergent Switches
  - Aggregate switches
  - VLAN
  - Physical Server Adapters

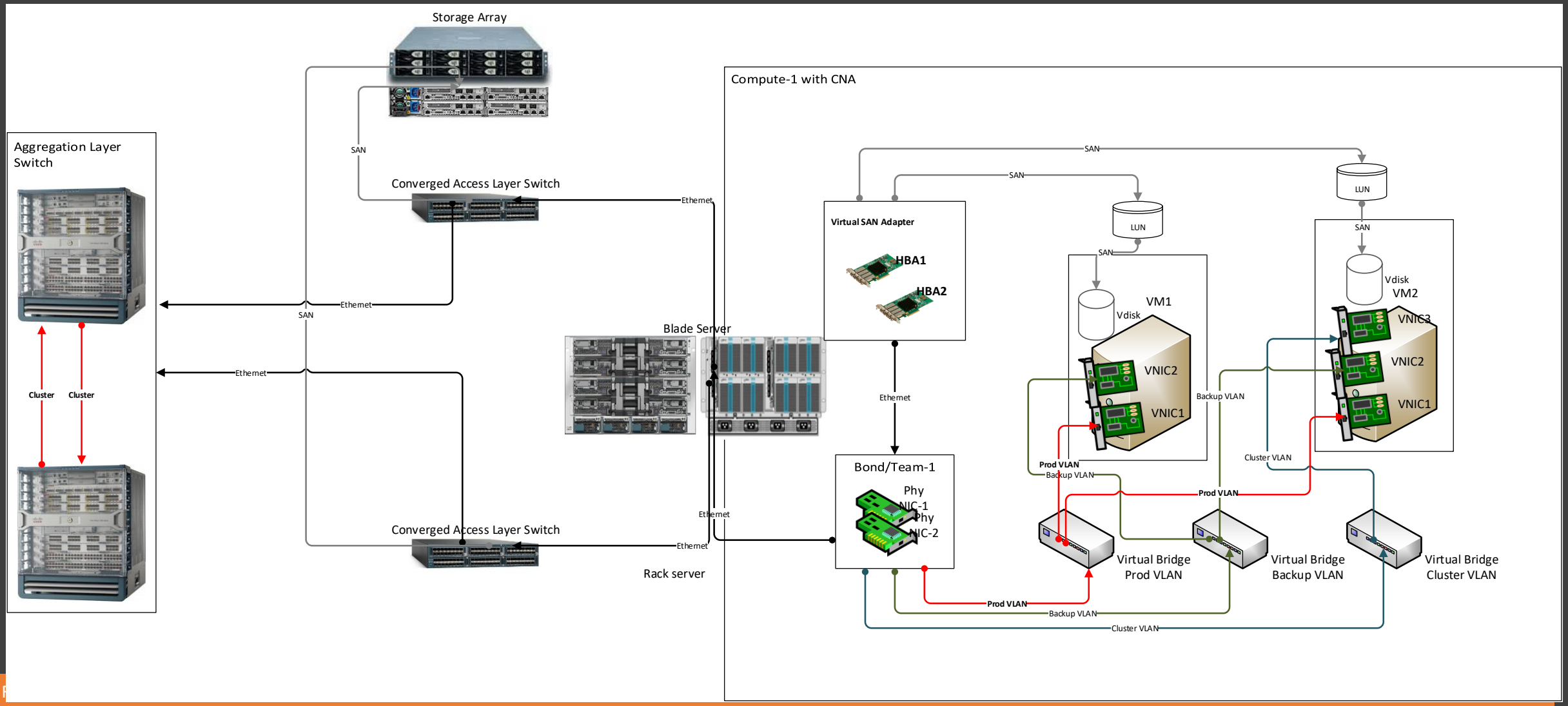
Internal Virtualization

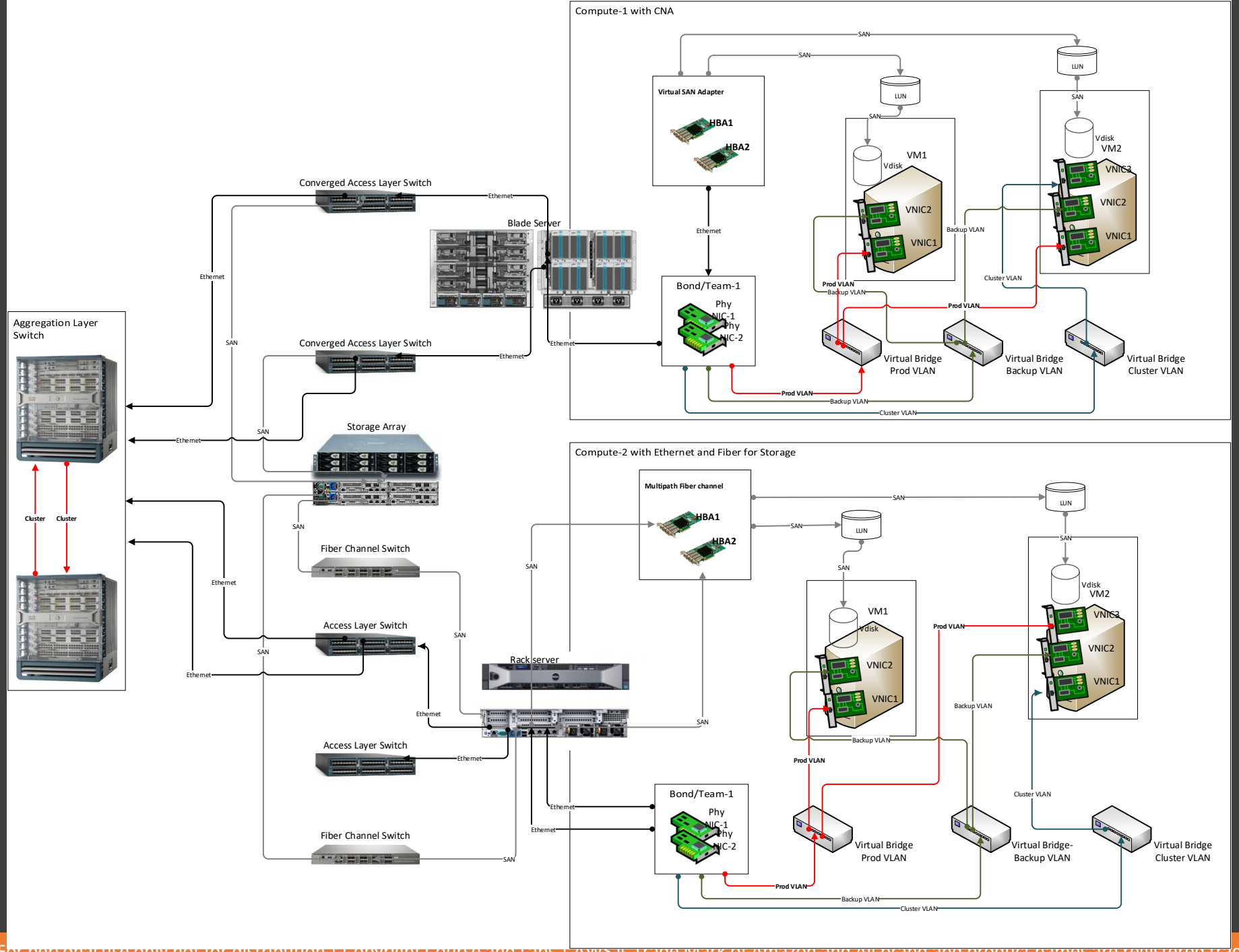
- Internal Network Virtualization provides network functionality purely based on software.
- An example of is the network topology used by common virtualization produces such as KVM or VMWARE ESX.
- In these you use existing network in your environment and present it to the virtual machines using a simple bridged or NAT based networking.
- Components of Internal Networking
  - Vitual NIC
  - Virtual SAN Adapter
  - Virtual Bridge or Switch
  - Physical NIC

Network Virtualization – Traditional Physical Server



Network Virtualization- Blade Server and CNA



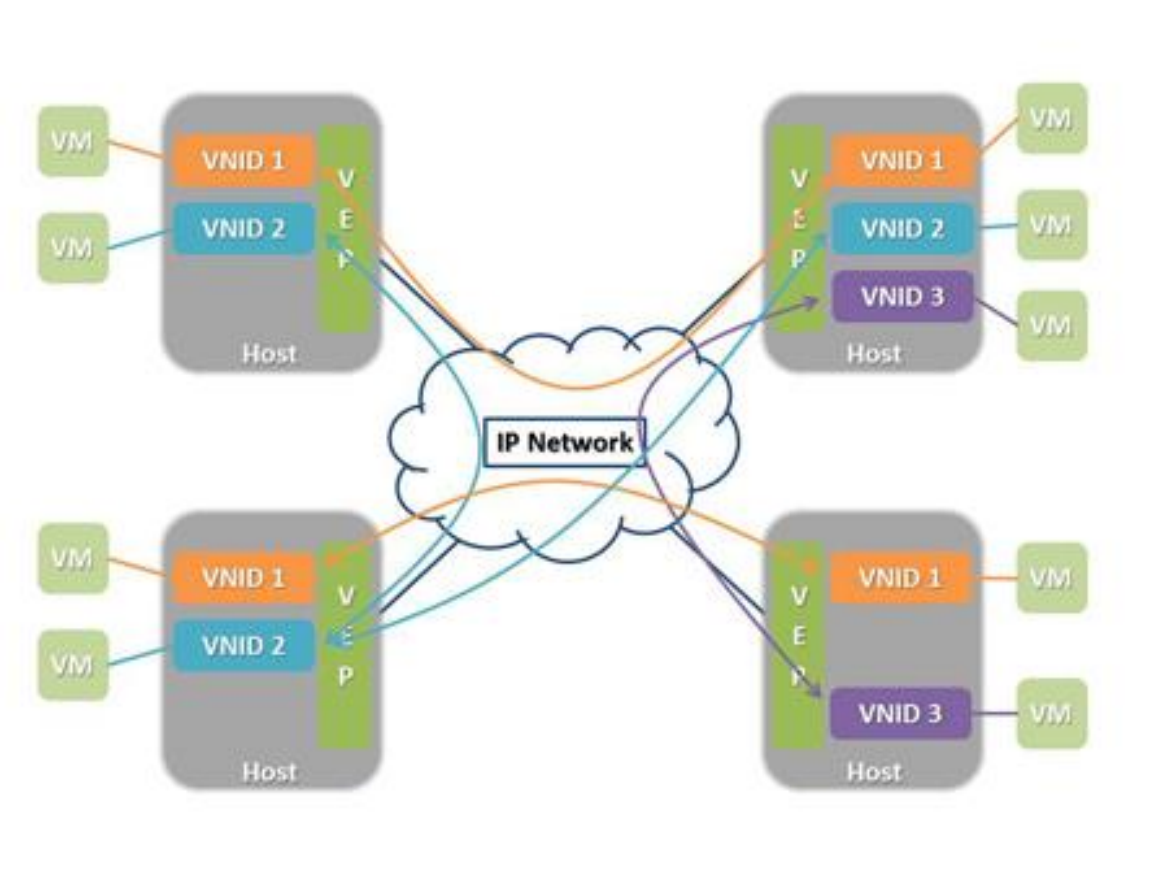


Network Virtualization

Network Overlays

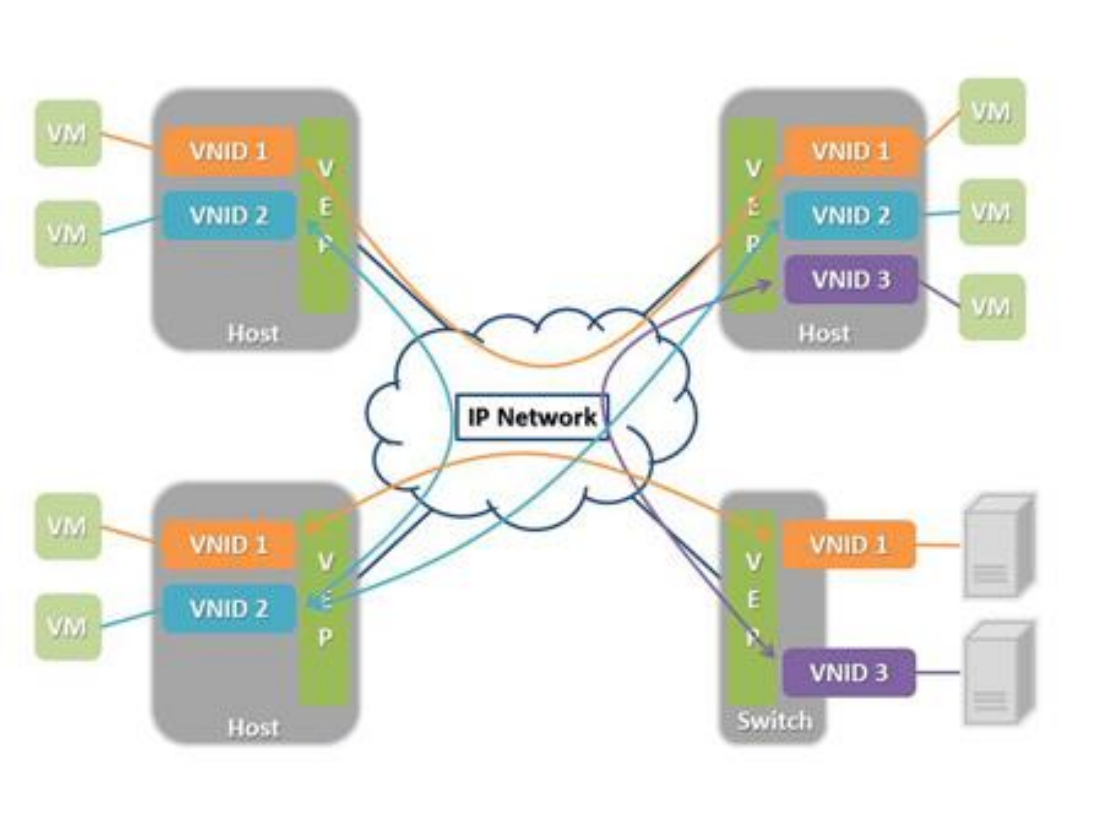
- The 802.1q standard defines the VLAN tag as a 12-bit space, providing for a max of 4,096 VLANs .
- This is an easily reachable ceiling in multitenant environments where multiple internal or external customers will request multiple subnets.
- A physical server now has multiple Virtual Machines (VMs) each with its own Media Access Control (MAC) address. This requires larger MAC address tables in the switched Ethernet network due to potential attachment of and communication among hundreds of thousands of VMs.
- To Overcome that we use Overlay, overlay is used to carry the MAC traffic from the individual VMs in an encapsulated format over a logical "tunnel".
- The popular Standards of Overlay are:
  Virtual Extensible LAN (VXLAN),
  Network Virtualization using Generic Routing Encapsulation(NVGRE),

Network Overlays



- Endpoints are assigned to a virtual network via a Virtual Network ID (VNID).
- These endpoints will belong to that virtual network regardless of their location on the underlying physical IP network.
- In diagram 1 there are four virtual hosts connected via an IP network.
- Each host contains a Virtual End Point (VEP), which is a virtual switch capable of acting as the encapsulation/de-encapsulation point for the virtual networks (VNIDs.)
- Each host has two or more VNIDs operating on it and each workload assigned to a given VNID can communicate with other workloads in the same VNID, while maintaining separation from workloads in other VNIDs on the same or other hosts.

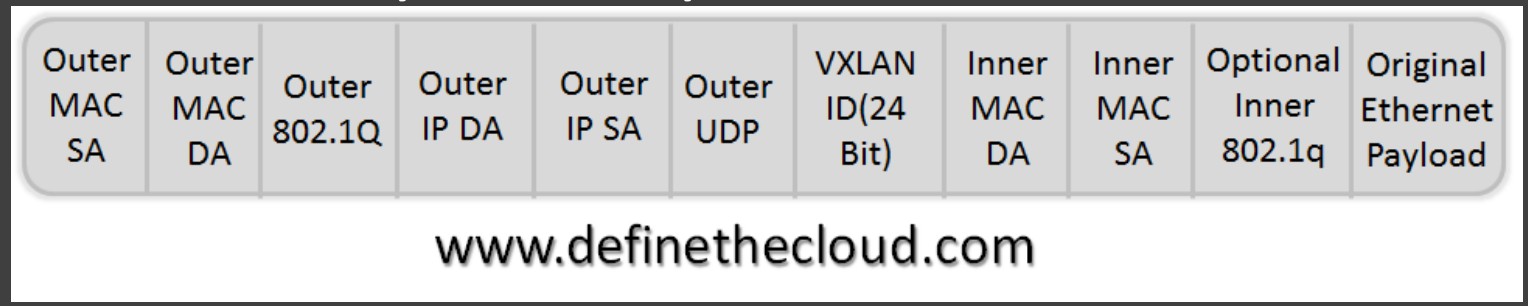
Network Overlays



- Depending on the chosen encapsulation and configuration method, hosts that do not contain a given VNID will either never see packets destined for that VNID, or will see them and drop them at ingress. This ensures the separation of tenant traffic.
- The same concept would apply if using a physical switch with the VEP functionality. This would allow physical devices to be connected to the overlay network
- With a physical switch capable of acting as the tunnel end-point, you can add both physical servers and appliances (firewalls, load balancers, and so on) to the overlay.
- This model is key to a cohesive deployment in mixed workload environments common in today's data centers.

Virtual Extensible LAN (VXLAN),

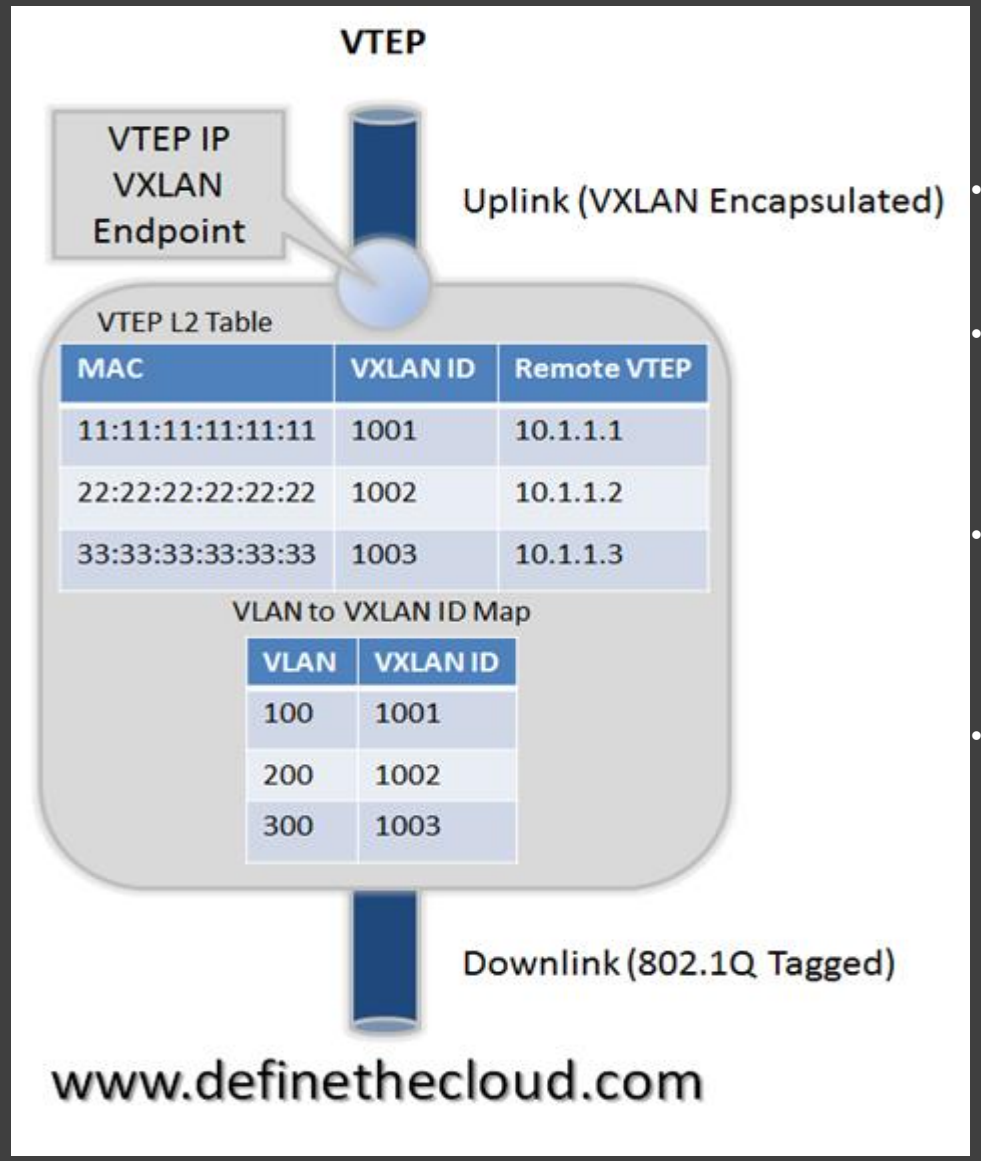
- **Virtual Extensible LAN** (**VXLAN**) is majorly backed up by Cisco and Vmware
- The standard for VXLAN is under the scope of the IETF NVO3 working group.
- VXLAN's goal is allowing dynamic large scale isolated virtual L2 networks to be created for virtualized and multi-tenant environments.
- It does this by encapsulating frames in VXLAN packets.
- VXLAN utilizes a 24-bit VXLAN header, shown in the diagram, to identify virtual networks.
- This header provides for up to 16 million virtual L2 networks.
- Frame encapsulation is done by an entity known as a VXLAN Tunnel Endpoint (VTEP.)



Virtual Extensible LAN (VXLAN),

- A VTEP has two logical interfaces: an uplink and a downlink.
- The uplink is responsible for receiving VXLAN frames and acts as a tunnel endpoint with an IP address used for routing VXLAN encapsulated frames.
- These IP addresses are infrastructure addresses and are separate from the tenant IP addressing for the nodes using the VXLAN fabric.
- VTEP functionality can be implemented in software such as a virtual switch or in the form a physical switch.
- VXLAN frames are sent to the IP address assigned to the destination VTEP; this IP is placed in the Outer IP DA.
- The IP of the VTEP sending the frame resides in the Outer IP SA.
- Packets received on the uplink are mapped from the VXLAN ID to a VLAN and the Ethernet frame payload is sent as an 802.1Q Ethernet frame on the downlink.
- During this process the inner MAC SA and VXLAN ID is learned in a local table. Packets received on the downlink are mapped to a VXLAN ID using the VLAN of the frame.
- A lookup is then performed within the VTEP L2 table using the VXLAN ID and destination MAC; this lookup provides the IP address of the destination VTEP. The frame is then encapsulated and sent out the uplink interface.

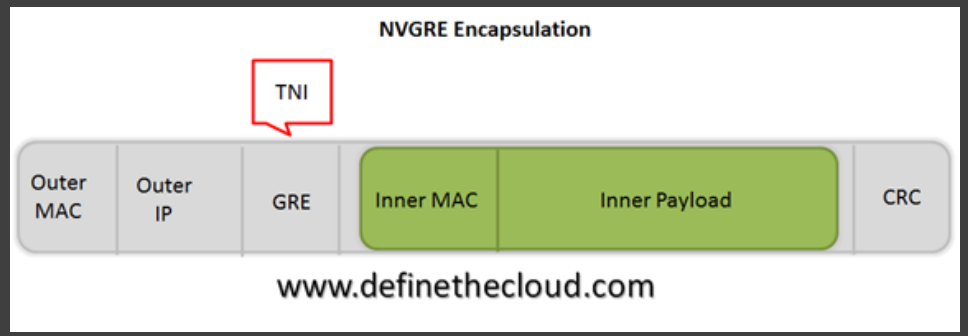
Virtual Extensible LAN (VXLAN),



- Using the diagram above for reference a frame entering the downlink on VLAN 100 with a destination MAC of 11:11:11:11:11:11 will be encapsulated in a VXLAN packet with an outer destination address of 10.1.1.1. The outer source address will be the IP of this VTEP (not shown) and the VXLAN ID will be 1001.
- In a traditional L2 switch a behavior known as flood and learn is used for unknown destinations (i.e. a MAC not stored in the MAC table. This means that if there is a miss when looking up the MAC the frame is flooded out all ports except the one on which it was received. When a response is sent the MAC is then learned and written to the table. The next frame for the same MAC will not incur a miss because the table will reflect the port it exists on. VXLAN preserves this behavior over an IP network using IP multicast groups.
- Each VXLAN ID has an assigned IP multicast group to use for traffic flooding (the same multicast group can be shared across VXLAN IDs.) When a frame is received on the downlink bound for an unknown destination it is encapsulated using the IP of the assigned multicast group as the Outer DA; it's then sent out the uplink. Any VTEP with nodes on that VXLAN ID will have joined the multicast group and therefore receive the frame. This maintains the traditional Ethernet flood and learn behavior.
- VTEPs are designed to be implemented as a logical device on an L2 switch. The L2 switch connects to the VTEP via a logical 802.1Q VLAN trunk. This trunk contains an VXLAN infrastructure VLAN in addition to the production VLANs. The infrastructure VLAN is used to carry VXLAN encapsulated traffic to the VXLAN fabric. The only member interfaces of this VLAN will be VTEP's logical connection to the bridge itself and the uplink to the VXLAN fabric. This interface is the 'uplink' described above, while the logical 802.1Q trunk is the downlink.

Network Virtualization using Generic Routing Encapsulation(NVGRE)

- NVGRE uses Generic Routing Encapsulation (GRE) to tunnel layer 2 packets over layer 3 networks.
- Its principal backer is Microsoft. Other companies supporting the development of NVGRE include Chelsio Communications, F5 Networks, AristaNetworks, Mellanox, Broadcom, Dell, Emulex, Intel a nd Hewlett Packard.
- It uses the lower 24 bits of the GRE header to represent the Tenant Network Identifier (TNI.) Like VXLAN this 24 bit space allows for 16 million virtual networks.



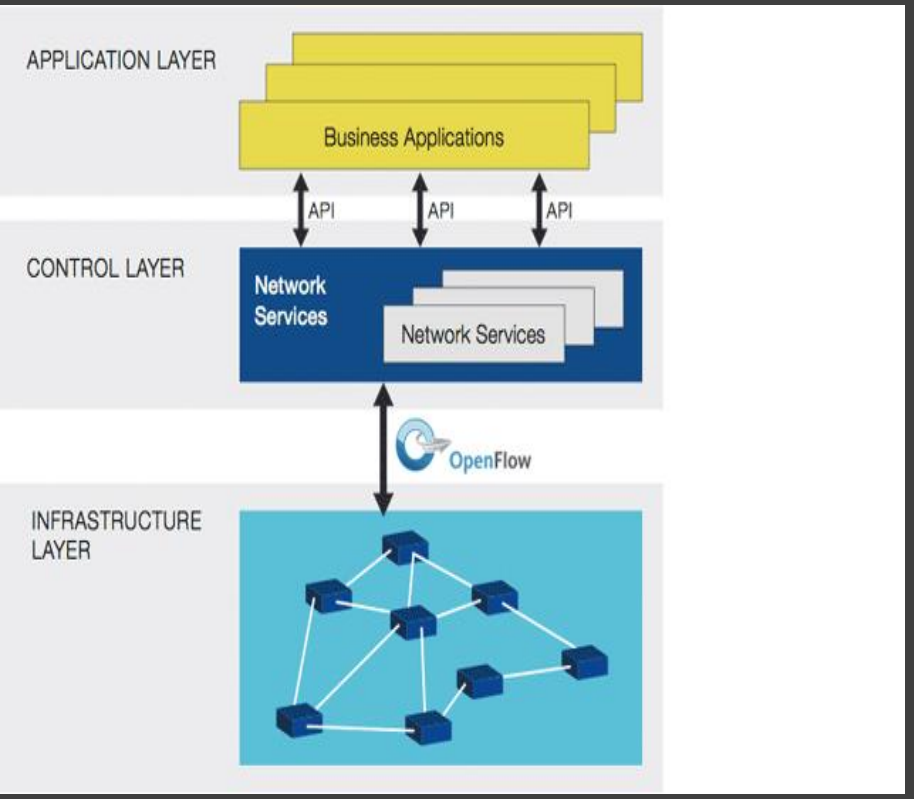
VXLAN Vs NVGRE

- NVGRE provides optional support for broadcast via IP multi-cast
- It supports multi-pathing capabilities.
- It Supports Jumbo frames

Software Defined Network (SDN)

- Software-Defined Networking (SDN) is an emerging architecture that is dynamic, manageable, cost-effective, and adaptable, making it ideal for the high-bandwidth, dynamic nature of today's applications.
- This architecture decouples the network control (Control Pane) and forwarding functions (Data Pane)
- IT enables the network control to become directly programmable and the underlying infrastructure to be abstracted for applications and network services.
- The OpenFlow® protocol is a foundational element for building SDN solutions.

Software Defined Network (SDN)



The SDN architecture is:

Directly programmable: Network control is directly programmable because it is decoupled from forwarding functions.

Agile: Abstracting control from forwarding lets administrators dynamically adjust network-wide traffic flow to meet changing needs.

Centrally managed: Network intelligence is (logically) centralized in software-based SDN controllers that maintain a global view of the network, which appears to applications and policy engines as a single, logical switch.

Programmatically configured: SDN lets network managers configure, manage, secure, and optimize network resources very quickly via dynamic, automated SDN programs, which they can write themselves because the programs do not depend on proprietary software.

Open standards-based and vendor-neutral: When implemented through open standards, SDN simplifies network design and operation because instructions are provided by SDN controllers instead of multiple, vendor-specific devices and protocols.

Software Defined Network (SDN)

- SDN addresses the fact that the static architecture of conventional networks is ill-suited to the dynamic computing and storage needs of today's data centers, campuses, and carrier environments. The key computing trends driving the need for a new network paradigm include:
- **Changing traffic patterns**: Applications that commonly access geographically distributed databases and servers through public and private clouds require extremely flexible traffic management and access to bandwidth on demand.
- **The "consumerization of IT"**: The Bring Your Own Device (BYOD) trend requires networks that are both flexible and secure.
- **The rise of cloud services**: Users expect on-demand access to applications, infrastructure, and other IT resources.
- **"Big data" means more bandwidth**: Handling today's mega datasets requires massive parallel processing that is fueling a constant demand for additional capacity and any-to-any connectivity.

OpenFlow

- OpenFlow® is the first standard communications interface defined between the control and forwarding layers of an SDN architecture.
- OpenFlow® allows direct access to and manipulation of the forwarding plane of network devices such as switches and routers, both physical and virtual (hypervisor-based).
- OpenFlow-based SDN technologies enable IT to address the high-bandwidth, dynamic nature of today's applications, adapt the network to ever-changing business needs, and significantly reduce operations and management complexity.

OpenFlow

Programmability

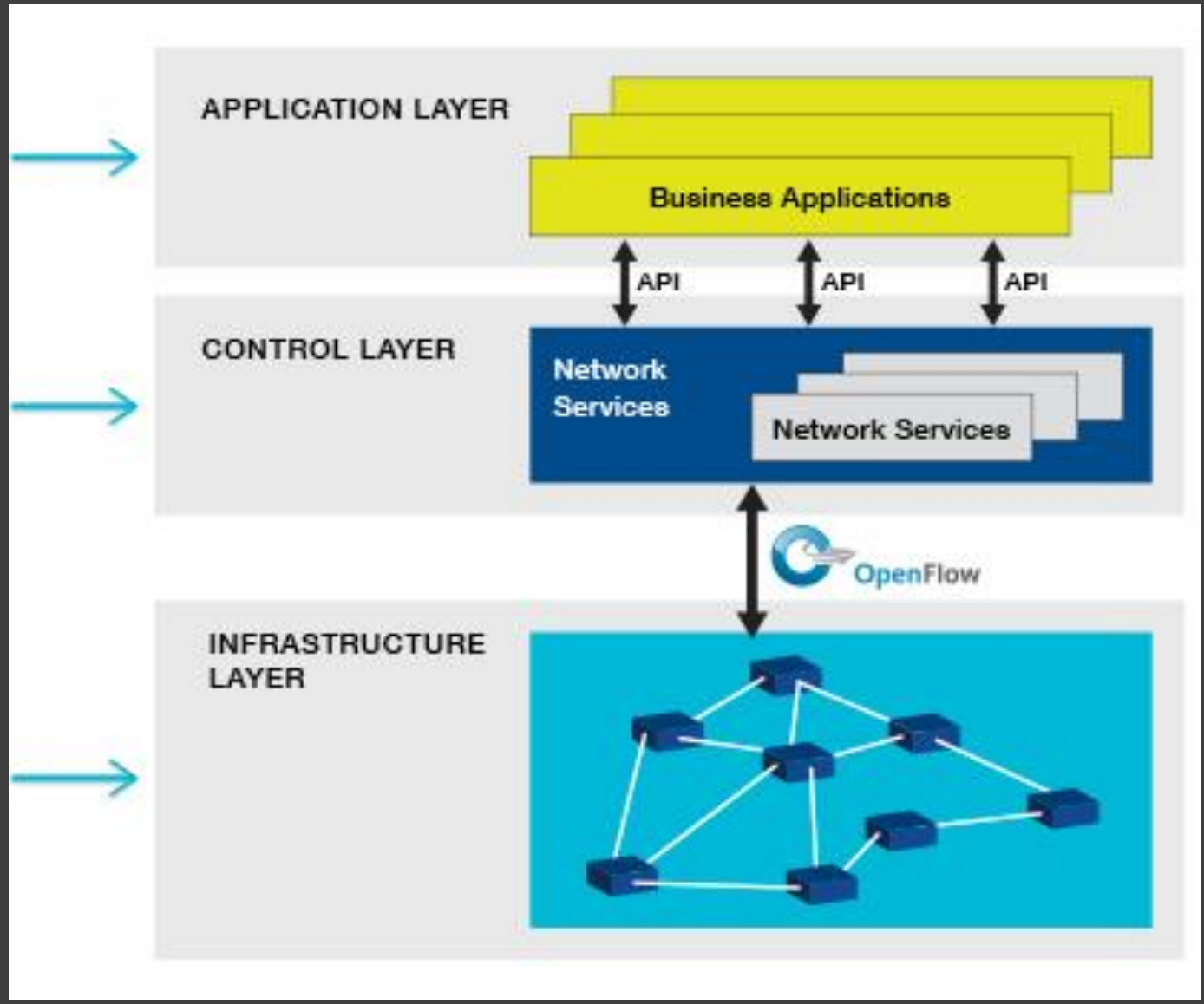
- Enable innovation/differentiation
- Accelerate new features and services introduction

Centralized Intelligence

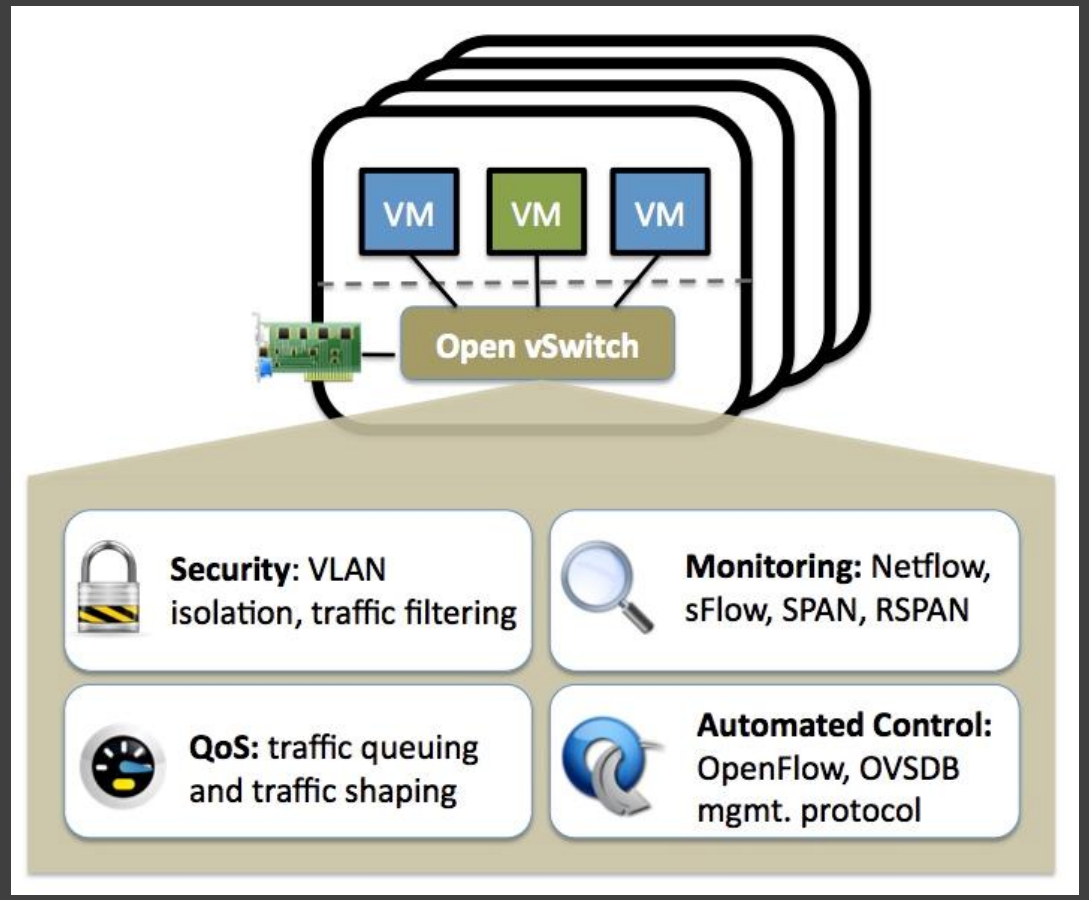
- Simplify provisioning
- Optimize performance
- Granular policy management

Abstraction

- Decouple:
  - Hardware & Software
  - Control plane & forwarding
  - Physical & logical config.



Open VSwitch



- Open vSwitch is a production quality, multilayer virtual switch licensed under the open source **Apache 2.0** license.
- It is designed to enable massive network automation through programmatic extension, while still supporting standard management interfaces and protocols (e.g. NetFlow, sFlow, IPFIX, RSPAN, CLI, LACP, 802.1ag).
- In addition, it is designed to support distribution across multiple physical servers similar to VMware's vNetwork distributed vswitch or Cisco's Nexus 1000V.

Network Function Virtualization

- **Network-Function Virtualization** (NFV) is a network architecture concept that uses the technologies of IT virtualization to virtualize entire classes of network node functions into building blocks that may connect, or chain together, to create communication services.
- NFV is virtualizing Layer 4-7 functions such as virtualized load balancers, firewalls, intrusion detection devices and WAN accelerators, Routers, Proxies provided by a software by a inside a VM or a group of VMs.
- For Example providing a Load balancer service using HAProxy in a Linux VM