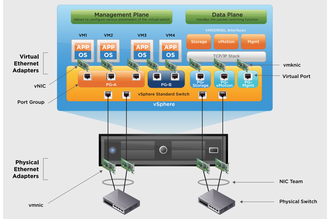
VMware vCloud Networking

vSphere Standard Switch (VSS)

Components, Terminology and Data Plane Details

VSPHERE NETWORKING COMPONENTS VSS 1



VSPHERE NETWORK TERMINOLOGY VSS 2

Virtual Ethernet Adapters
- Virtual network interface card (vnic) – virtual machine's interface to the network
- Virtual network kernel network interface card (vmknic) – vSphere hypervisor's interface to network (Mgmt, NFS, iSCSI, vMotion, FT)

Physical Ethernet Adapter
- Physical network interface card (pnic) – Helps communicate with entities outside ESXi host. This is also called as vmnic.

vSphere Standard Switch (VSS)
- Forwards packets between vnics, vmknics, and vmnics

Port Group (PG)
- Group of ports sharing the same configuration (e.g. vlan)

Virtual Port: These are the ports where virtual machines or vmknics are connected to the virtual switch. Type of port binding available – No binding, Also called as Ephemeral port binding.

Uplinks: Connections to physical switches

NIC Team: A group of vmnics connected to the same physical network

VIRTUAL ETHERNET ADAPTER TYPES VSS 3

vlance – Emulates AMD 79C970 PCnet32 LANCE 10 Mbps NIC

vmxnet – virtual adapter with performance optimization, requires VMware Tools

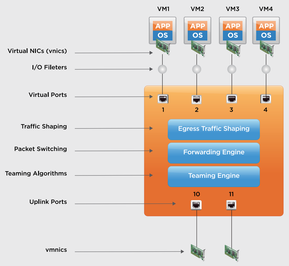
Flexible – Behaves as either vlance or vmxnet. If VMware Tools are installed it will act as vmxnet adapter

e1000 – Emulates Intel 82545EM Gigabit Ethernet NIC

vmxnet2 (enhanced) – Based on vmxnet, with jumbo frame and hardware offloads

vmxnet3 – New generation of virtual adapter, supports RSS, IPv6 offloads, and MSI/MS-X interrupts. Available on a limited set of guest operating systems

VSPHERE STANDARD SWITCH – DATA PLANE DETAILS VSS 4



VSS Configuration Parameters

VSS PARAMETERS - VLAN VSS 5

VLAN Configuration are done at port group level. The Following options are available as part of the VLAN ID (optional) field:
- None (0) – No VLAN tagging
  - Use this option when you don't want any VLAN tagging (default option). VLAN tagging is left to external switch. Also called as External Switch Tagging – EST
- All (4095) – Trunk configuration
  - Use 4095 when letting the Guest virtual machines to do the VLAN tagging. This is equivalent to the trunk port configuration on the physical switch port. Also called as Virtual Guest Tagging – VGT
- Type in any number as VLAN ID in the field – available numbers 1 to 4094
  - In this mode, based on the VLAN number configured on the port group the traffic going out of the virtual switch will be tagged accordingly. Also called as Virtual Switch Tagging – VST

Best Practices
- Use VLANs to isolate different traffic types. i.e use Virtual Switch Tagging (VST)
- While using VST, make sure the physical switch ports, where the host's vmnics are connected, are configured as trunk ports. These trunk ports then carry the tagged traffic from virtual switch.

VSS PARAMETERS - SECURITY VSS 6

Security Configuration at Port Group level
- Promiscuous mode
  - Reject – Only forwards traffic that is destined for the VM
  - Accept – All packets received on a particular VLAN of the port group are forwarded to all the VMs connected to the port group
- MAC address Change
  - Reject – If the VM changes it's MAC address from the one that was configured in .vmx file, the inbound traffic with new MAC is dropped

Accept – Any change in MAC address is accepted and frames are continued to receive
- Forged Transmit
  - Reject – Outbound traffic is dropped if sent with different MAC address
  - Accept – No check for MAC address is performed

Best practices
- Select Promiscuous mode – Reject (default)
- Select MAC address Change – Reject (default)
- Forged Transmit – Reject (default)

VSS PARAMETERS - TRAFFIC SHAPING VSS 7

Traffic Shaping Policy (Egress traffic management only)
- A virtual machine's network bandwidth can be controlled by enabling the network traffic shaper.

The following three parameters dictates the traffic-shaping policy on a virtual port
- Average Bandwidth – Specifies the allowed average bandwidth in kbps
- Peak Bandwidth – The maximum bandwidth allowed in kbps
- Burst Size – Establishes the maximum number of kilo bytes to allow in a burst

Best Practices
- Use this feature in the following scenarios
  - You want to control vMotion traffic going out of one physical interface such that it doesn't impact virtual machine traffic flowing through the same interface
  - You want to control a tenant's virtual machine bandwidth usage.

VSS PARAMETERS - TEAMING VSS 8

Teaming Configuration on a Port Group depends on the way uplinks of a virtual switch are connected to physical switches.

Based on the physical switch connection option (shown in VSS/VDS 3) choose the load balancing type
- Load balancing types
  - Route based on originating virtual port (Port ID Hash)
  - Route based on source MAC hash (MAC Hash)
  - Route based on IP hash (IP Hash)
  - Use Explicit Failover order (Explicit Failover)
- After choosing the Load balancing types
  - Make sure that the Active, Standby and Unused adapters are identified for the port group

Best Practices
- Always connect two or more vmnics to a virtual switch and configure teaming for redundancy and higher bandwidth.
- If possible terminate the vmnics on two separate physical access switch

VSS PARAMETERS – FAILOVER DETECTION VSS 9

Network Failover Configuration
- Link Status Only (default)
  - Use when supported by physical switches. For Example – Cisco switches.
- Beacon probing
  - Use when no link state tracking support is available on physical switches and you don't have redundant connection between access to distribution physical switches.

Best Practices
- Cisco switches provides Link State tracking feature. By enabling this feature any link failure between access and distribution (upstream) switch is indicated to the ESXi host such that the traffic gets moved to another working uplink. This feature should be used if you don't have redundant connection from the access switch to the distribution switches.
- Beacon probing should be used when you have more than 2 uplinks in a team. This software approach of detecting failure between the upstream links is very useful when there is no link state tracking support on the physical switches.

VSS PARAMETERS – FAILBACK OPTION AFTER LINK COMES BACK VSS 10

Failback Configuration
- Use this feature when using Explicit Failover (EF) teaming configuration
- Configure "No" when selecting EF teaming
- Default Failback is "yes" – That means after the link comes back after a failure the traffic will be moved back on the link.

Best Practices
- Failback configuration works with Active-Standby teaming option and not Active-Active teaming. So make sure when configuring Explicit failover the failback option is changed from default "yes" to "no". This change makes sure that you don't keep moving the traffic from one link to another if one link is flapping.

vSphere Switch (VSS/VDS) Operation and Connection Options

VSPHERE SWITCH OPERATION–MAC LEARNING VSS/VDS 1

Virtual switches (VSS/VDS) are not learning switches. They authoritatively configure the MAC forwarding table:

Forwarding Table Example

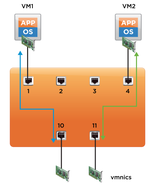
MAC	Port
VM1 MAC	1
VM2 MAC	4

Teaming configuration decides which VM traffic will be sent over which uplink

For Example: Port ID Hash Teaming

Based on the Port ID hash teaming algorithm VM1 traffic will be sent over uplink1 (virtual port 10)

Based on the Port ID hash teaming algorithm VM2 traffic will be sent over uplink2 (virtual port 11)

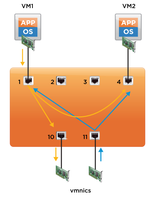


VSPHERE SWITCH OPERATION–HANDLING OF BROADCAST PACKETS VSS/VDS 2

VSS/VDS doesn't support Spanning Tree Protocol (STP) because they don't create loops

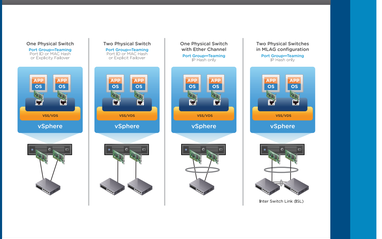
When VM1 sends a broadcast packet that packet is sent to all virtual ports and only one uplink port

When a broadcast packet is received from uplink port that packet is copied only to the VMs and not sent back to the other uplink port



This unique handling of Broadcast packets along with the Authoritative learning prevents loop in the network.

VIRTUAL SWITCH TO PHYSICAL SWITCH CONNECTION OPTIONS VSS/VDS 3



VSS TO VDS MIGRATION VSS/VDS 4

Three methods of Migration
- Using VDS wizards available in vCenter Server
- Using combination of the VDS and Host Profiles
- Using PowerCLI or vCLI commands

Key Steps while migrating using VDS wizards in vCenter Server
- Create a VDS first. Don't add the hosts to the VDS yet
- Next create the distributed virtual port groups to match the VSS port group configurations
- Add hosts to VDS
  - Select the hosts you wish to migrate, along with physical adapters per host. You can either decide to migrate all physical adapters to VDS at once or few at a time
  - Next step is to migrate the vmkernel virtual adapters from the existing VSS port groups to the VDS distributed port groups
  - After migrating vmkernel interfaces you can choose to migrate virtual machine networking from VSS to VDS

Best Practices
- If using Ether Channel configuration on physical switches make sure to disable it before going through migration steps

Migration Guide
- http://www.vmware.com/files/pdf/vsphere-vnetwork-ds-migration-configuration-wp.pdf

vSphere Distributed Switch (VDS)

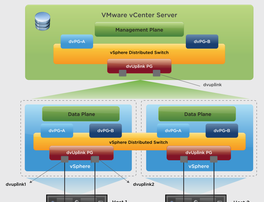
Components, Terminology and Data Plane Details

VSPHERE DISTRIBUTED SWITCH (VDS) COMPONENTS VDS 1

Key differences between VSS and VDS
- Management Plane and Data Plane are part of the host in VSS (as shown in VSS 1)
- In VDS, the data plane remains local to each host, but the management plane is centralized with vCenter Server acting as the central control point for all parameter configurations and virtual network management (as shown in VDS 1)

Key advantages of VDS
- Centralized management minimizes the configuration errors that could happen while managing VSS on individual hosts with separate management plane.
- Each vCenter Server instance can support up to 128 VDSs and each VDS can connect up to 500 hosts.
- Along with centralized and simplified management, VDS provides advanced virtual network capabilities.

One Management Plane : Allows to configure various parameters of the distributed switch
Data Plane : Handles the packet switching function on individual hosts



VDS SPECIFIC TERMINOLOGY VDS 2

dvUplinks
dvUplinks provide a level of abstraction for the physical NICs (vmnics) on each host. NIC teaming, load balancing, and failover policies on distributed port groups are applied to the dvUplinks and not to the vmnics on individual hosts.

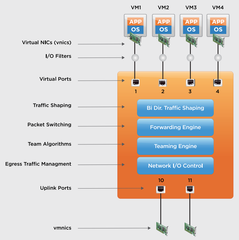
dvUplink Port group
While creating a new distributed switch the number of dvUplinks are defined as part of the dvUplink port group configuration.

Distributed Port Group (dvPG)
Distributed Port Group are port groups associated with a VDS and specify configuration that is common across a group of distributed virtual ports. Distributed Port Groups also defines how a connection is made through the VDS to the Network.

Distributed Virtual Port (dvPort)
These are the ports where virtual machines or vmknics are connected to the virtual switch. Type of port binding is Static, Ephemeral and Dynamic (deprecated option and won't be available in future release).

VSPHERE DISTRIBUTED SWITCH – DATA PLANE DETAILS VDS 3

When Compared with VSS data plane (VSS 4) VDS has one additional component in Egress traffic management also called as NIOC. Also, the Traffic shaping layer provides bi directional traffic shaping instead on only egress traffic shaping



VDS Configuration Parameters

VDS PARAMETERS - DVUPLINK VDS 4

dvUplink Configuration is done at dvUplink port group level

Depending on how many physical NICs are on hosts you can determine how many dvUplinks you should configure.

The following are some examples of dvUplink configurations
- If there are 4 hosts with 4 NICs each – Configure dvUplink as 4 (default)
- In a Heterogeneous environment where there are 2 hosts with 6 NICs and 2 other hosts with 8 NICs– Configure dvUplink as 8 (Highest common denominator)

Best Practices
- dvUplinks have to be mapped to vmnics on the hosts. This mapping process happens when the hosts are added to the distributed switch. You should do consistent mapping of dvuplink to vmnic across all the hosts.
- In case of host that does not have enough vmnics when compared to uplinks, you can leave the higher order dvUplinks not mapped. In this deployment you should make sure that the dvportgroups used on the host don't have teaming property that includes unmapped dvUplinks.

VDS PARAMETERS - DVPORT VDS 5

dvPort configuration is done at dvPortgroup level. You have following three options under the general dvport group properties
- Port binding types
  - Static binding (Default)
  - Dynamic binding (deprecated and won't be supported in future release)
  - Ephemeral – No binding (Equivalent to Standard switch option)
- Port allocation
  - Elastic (Default)
  - Fixed
- Number of ports – Required if using Fixed port allocation option

Best Practices
- Use Static binding and Elastic port allocation - The static port-binding configuration on a dvPortgroup helps users to do stateful monitoring of dvPorts. This is not possible with ephemeral and dynamic port-binding configurations, where users lose the visibility and troubleshooting capability.
- With elastic port allocation you don't have to manually manage the number of dvPorts.

VDS PARAMETERS - VLAN VDS 6

VLAN Configuration is done at distributed port group level
- The following four options are available
  - None – No VLAN tagging
  - VLAN – Allows you to specify the VLAN to use for tagging. This is also called as Virtual Switch Tagging (VST) mode
  - VLAN trunking – VLAN trunk range is configure from 0–4094. In this mode Guest virtual machines are allowed to do the tagging and those tags are carried through VDS. This is also called as Virtual Guest Tagging (VGT)
  - PVLAN – If PVLANs are configured at the VDS level this option allows to choose the PVLAN for the port group.

Best Practices
- Use VLANs to isolate different traffic types. i.e use Virtual Switch Tagging (VST)
- Make sure on the physical switch port, where the host is connected, all the configured VLANs are trunked
- PVLAN configuration can be used in creating Demilitarized Zones (DMZ) deployments.

VDS PARAMETERS - TEAMING VDS 7

Teaming Configuration on a distributed Port Group depends on the way uplinks of a virtual switch are connected to the physical switch.

Based on the physical switch connection option (shown in VSS/VDS 3) choose the load balancing type
- Load balancing types
  - Route based on originating virtual port (Port ID Hash)
  - Route based on source MAC hash (MAC Hash)
  - Route based on IP hash (IP Hash)
  - Use Explicit Failover order (Explicit Failover)
  - Route based on physical NIC load (LBT)
- After choosing the Load balancing types
  - Make sure that the Active, Standby and Unused adapters are identified for the port group

Best Practices
- Route based on physical NIC load (LBT) is the recommended option for teaming
  - It doesn't need any special configuration on the physical switch side and utilitzes the vmnics efficiently

vSphere Distributed Switch Features

VDS FEATURE - PRIVATE VLAN (PVLAN) VDS 8

Private VLAN feature provides isolation between the ports in the same broadcast domain. If you have limited VLANs or want to provide further segmentation in a broadcast domain then you can make use of this feature.

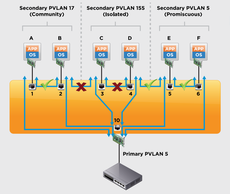
To use PVLAN feature on VDS, the physical switch needs to be PVLAN-capable and configured with the same VLAN IDs

PVLAN Basics – The basic concept behind private VLANs is to divide an existing VLAN, referred to as the primary VLAN, into one or more separate VLANs, called secondary VLANs. For Example:
- First a Private VLAN is identified by a primary VLAN ID 5
- Second this Primary VLAN ID can have multiple secondary VLAN IDs associated with it. VLAN ID 17 and VLAN ID 155 are secondary VLANs

Primary VLANs are promiscuous – virtual machines on primary VLAN can communicate with all ports in the same primary VLAN

Secondary VLANs can be either isolated or community
- Virtual machines in an isolated PVLAN can talk to no virtual machines except those in promiscuous PVLAN or also called primary VLAN
- Virtual machines in a community PVLAN can talk to each other and to the virtual machines in the promiscuous VLAN

The following diagram shows an example of PVLAN configuration with virtual machines in community, isolated and promiscuous zone. The communication table on the right shows whether communication between virtual machines is allowed or not.



Communication Table	B	D	F
A	Y	N	Y
C	N	N	Y
E	Y	Y	Y

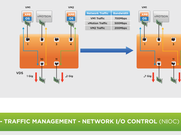
VDS FEATURE – LOAD BASED TEAMING VDS 9

Load Based Teaming is the advanced teaming option that takes into account the utilization of the physical NICs and rebalances the traffic to achieve even bandwidth utilization across all NICs in the team.

LBT Algorithm
- Every 30 sec. each NIC in the team is checked for BW utilization
- If any NIC's BW utilization goes above 75% the rebalance algorithm is kicked off
- Rebalance algorithm does its number crunching and identifies virtual machines or infrastructure traffic (NFS, vMotion, Management) that can be moved to the NIC that has enough unused BW.
- After the move is performed a reverse ARP process ensures that physical switches know about this change.

This Choice of load balancing doesn't require any specific configuration on the physical switches as is the case with IP hash teaming.

The following diagram shows how vMotion traffic is moved from the highly utilized uplink on the left to the other uplink by the LBT algorithm.

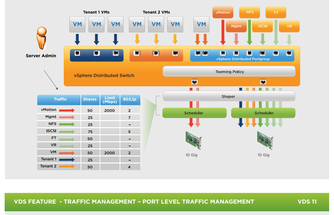


VDS FEATURE – TRAFFIC MANAGEMENT – NETWORK I/O CONTROL (NIOC) VDS 10

NIOC allows administrator to allocate I/O shares to different traffic types similar to allocating CPU and memory resources to a virtual machine
- Following traffic types are supported : Fault Tolerance, iSCSI, vMotion, Management, vSphere Replication (VR), NFS and virtual machine traffic. You can also create custom traffic types.
- The outgoing (egress) bandwidth for each traffic type can be controlled by setting the Shares and Limit parameters
- Shares parameter specifies the relative importance of a traffic type over other traffic and provides a guaranteed minimum when the other traffic competes for a particular network adapter. Shares apply at vmnic level.
- Limits specify the absolute maximum bandwidth for a traffic type. Limits apply across team.

Best Practices
- Make use of Shares parameter rather than Limit because Limit parameter puts hard limit on traffic type. So even when BW is available the traffic type can't utilize it
- Tag the important traffic with 802.1p tag to provide end to end Quality of Service (QoS)

The following diagram shows vMotion and Management traffic through 10 Gig NIC on the left.
- vMotion – 50 shares, Mgmt – 25 shares, Total = (50 + 25) = 75 shares
  vMotion BW = (50/75) * 10 Gig = 6.67 Gig, Mgmt. BW = (25/75) * 10 gig = 3.33 Gig
  In this example vMotion can't use more than 2 Gig BW because of the Limit parameter configured.



VDS FEATURE – TRAFFIC MANAGEMENT – PORT LEVEL TRAFFIC MANAGEMENT VDS 11

This feature is similar to the Egress traffic shaping feature in the standard switch. In VDS the Traffic shaping can be done on the ingress traffic as well.

A virtual machine's ingress and Egress network bandwidth can be controlled by enabling the network traffic shaper.

The following three parameters dictates the traffic-shaping policy on a virtual port
- Average Bandwidth – Specifies the allowed average bandwidth in kbps
- Peak Bandwidth – The maximum bandwidth allowed in kbps
- Burst Size – Establishes the maximum number of kilo bytes to allow in a burst

Best Practices
- Use this feature in the following scenarios
  - When you want more granular control over traffic type that can't be achieved through the NIOC feature.
  - You want to limit input traffic to a virtual machine or to a vmkernel nic.

VDS FEATURE – NETFLOW VDS 12

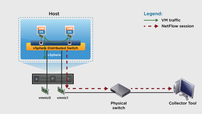
NetFlow is a networking protocol that collects IP traffic information as records and sends them to a collector tools for traffic flow analysis.

NetFlow features on VDS allows you to monitor the virtual machine to virtual machine flows that are not seen on the physical switch ports.

VDS support NetFlow v10 also called as IPFIX

Net Flow Use cases
- NetFlow helps monitor application flows and measures flow performance over time. It also helps in network capacity planning exercises.
- Make use of NetFlow to measure the bandwidth requirement for different traffic types and then use that information to configure the NIOC shares and limits parameter.

The following diagram shows an example how VDS sends the network flow information across to a Collector Tool placed centrally in the network.



VDS FEATURE – PORT MIRRORING VDS 13

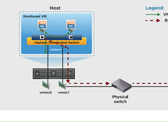
Port mirroring is the capability on a network switch to send a copy of network packets seen on a switch port to a network monitoring device connected to another switch port. Port mirroring is also referred to as Switch Port Analyzer (SPAN) on Cisco switches

The following are standard port mirroring options available to perform the network troubleshooting and monitoring activities
- Switched Port Analyzer (SPAN)
- Remote Switched Port Analyzer (RSPAN)
- Encapsulated Remote Switched Port Analyzer (ERSPAN)

VDS supports all of the above port mirroring options. Administrators now have the visibility into the virtual infrastructure traffic that never hits physical switch ports.

Best Practices
- Choose the port mirroring session type based on where the monitoring device is placed in the network. For example
  - RSPAN and ERSPAN makes sense when Analyzer is placed centrally in the network
  - SPAN works when you want to monitor traffic on a vSphere host. You can send the mirror traffic to a virtual machine running on the same host.

The following diagram shows an example how a monitored VM's port traffic gets mirrored to the analyzer through the vmnic1 interface.



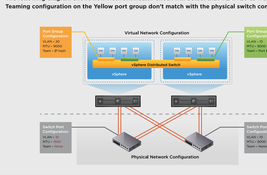
VDS FEATURE - NETWORK HEALTH CHECK (5.1 FEATURE) VDS 14

Network Health Check feature detects the common configuration errors across virtual and physical switch parameters. Following three parameters are checked
- Mismatched VLAN trunks between virtual switch and physical switch
- Mismatched MTU setting between vNIC, virtual switch, physical adapter, and physical switch ports.
- Mismatched Teaming Configurations

Network Health Check Use cases
- vSphere administrators now can provide network related warning information to the Network admins, and facilitate speedy resolution to configuration issues. Specially useful in Siloed organization with specific responsibilities.

Best Practices
- Default this feature is disabled. You should enable it during the network configuration process.
- Use it while debugging any network issue.

The following diagram shows how Network Health Check will detect that VLAN 20 as well as MTU and Teaming configuration on the Yellow port group don't match with the physical switch configurations



VDS FEATURE - ROLLBACK AND RECOVERY (5.1 FEATURE) VDS 15

Rollback and Recovery feature provides two options to recover from misconfiguration of Management Network
- Automatic Rollback happens when accidental misconfiguration of management port group is detected. This is default enabled.
- Recover Management network on VDS through Direct Console User Interface (DCUI) of hosts. This allows you to configure the management network parameters on VDS.

Best Practices
- Keep the automatic rollback feature on
- In case when the changes are made to the external network and hosts loose connectivity to the vCenter Server, you can recover the host management network by connecting through the DCUI of the host.

VDS FEATURE - CONFIGURATION BACKUP AND RESTORE (5.1 FEATURE) VDS 16

Configuration backup and restore feature helps take snapshots of working VDS configuration.

These network configuration snapshots can be used for new deployments, rollback, and sharing purposes

The Following operations are supported as part of this feature
- Backup VDS/dvPortgroup configuration on disk (Export)
- Restore VDS/dvPortgroup from a backup (Import)
- Create new entity from backup (Import)
- Revert to the previous dvPortgroup configuration after changes are made

Snapshot Use cases
- Recover your virtual network configuration quickly after any vCenter Server database corruption scenario or any other catastrophic failures
- Create VDS configuration templates that can be deployed in new environments
- Create revision control system of VDS configurations and keep track of the virtual network configuration changes.

RESOURCES VDS 17

vSphere Networking Community
http://communities.vmware.com/community/vmtn/server/networking

vSphere Networking Standard Virtual Switch Video
www.youtube.com/watch?v=seUXJ6Uy4h8

vSphere Blog Networking
http://blogs.vmware.com/vsphere/networking/

Technical Resources for Virtual Networking
http://www.vmware.com/technical-resources/virtual-networking/resources.html

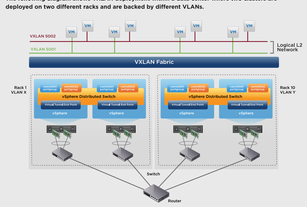
VDS Best Practices paper
http://www.vmware.com/resources/techresources/10250/

Virtual Extensible Local Area Network (VXLAN)

Components, Terminology and Data Plane Details

VXLAN VXLAN 1

The following diagram shows VXLAN deployment within a data center where two clusters are deployed on two different racks and are backed by different VLANs.



VXLAN SPECIFIC TERMINOLOGY VXLAN 2

Overlay Network
It is a network built over another network. For Example, VXLAN network is built over IP network.

Encapsulation and Decapsulation
Process of adding and removing packet header is called as Encap and Decap respectively

VTEP
"Virtual Tunnel Endpoint" refers to the encapsulation/de-encapsulation endpoints of a VXLAN tunnel. In the vSphere environ- ment the vmkernel module of VDS on each host acts as VTEP.

Virtual Wire or VXLAN Segment ID
A unique logical layer 2 network identified by a unique 24 bit segment ID.

MTU
"Maximum Transmission Unit" of a communications protocol of a layer is the size (in bytes) of the largest protocol data unit that the layer can pass onwards.

IGMP
"Internet Group Management Protocol" communications protocol is used by hosts and adjacent routers to establish multicast group memberships. It is analogous to ICMP for unicast connections.

IGMP Snooping
A network switch keeps a map/table of IGMP conversation between hosts and routers to allow filtering of unneeded Multicasts for particular links.

LACP
"Link Aggregation Control Protocol" provides a method to control the bundling of several physical ports together to form a single channel.

VXLAN PACKET HEADER DETAILS VXLAN 3

Ethernet in IP – overlay network
- Entire L2 frame encapsulated in UDP frame
- 50 bytes of overhead

Include 24 bit VXLAN Identifier
- 2^24 = 16,777,216 logical networks

VXLAN can cross Layer 3

Tunnels between vSphere hosts
- VM's do NOT see VXLAN ID
- External switch don't see VMs IP and MAC address

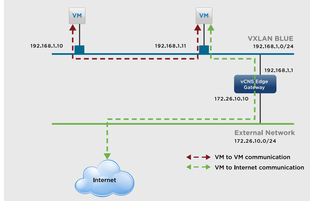
IP multicast used for L2 broadcast, unknown unicast

Technology submitted to IETF for standardization
- With Cisco, Citrix, Red Hat, Broadcom, Arista and Others

VXLAN Traffic Flows

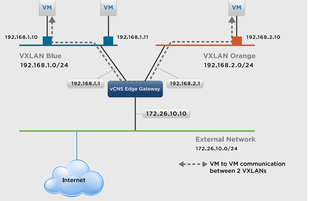
VXLAN COMMUNICATION PATHS – TWO VMS ON SAME VXLAN VIRTUAL WIRE VXLAN 4

The following diagram shows VXLAN traffic flow between virtual machines and to the external world



VXLAN COMMUNICATION PATHS – TWO VMS ON DIFFERENT VXLAN VIRTUAL WIRES VXLAN 5

The following diagram shows VXLAN traffic flow between virtual machines on two different virtual wires



VXLAN ADVANTAGES VXLAN 6

Provides the ability to provision on-demand logical layer 2 isolated networks
With 24 bit VXLAN identifier you can have the ability to create 16 million isolated networks

Allows customer the flexibility to provision compute resources across layer 2 boundaries.
For example, If there is a rack in the datacenter that has run out of compute capacity and there is another rack that is backed by different VLAN with some free compute resources, you can now easily increase the compute resources for your application by extending your logical layer 2 network over the rack where compute capacity is available.

On demand networks without physical network re-configuration
If customers want to deploy a new application they don't have to request the network admins to plumb new VLANs through their physical network infrastructure.

VXLAN INFRASTRUCTURE PREREQUISITES VXLAN 7

IP Multicast forwarding is required
- More multicast groups are better
- Multiple segments can be mapped to a single multicast group
- If VXLAN transport is contained in a single VLAN, "IGMP Querier" must be enabled for that VLAN on the switch
- If VXLAN transport is traversing routers, multicast routing must be enabled.

Increased MTU needed to accommodate VXLAN encapsulation overhead
- Physical infrastructure must carry 50 bytes more than the VM VNIC MTU size.
- E.g. 1500 MTU on VNIC + 50 = 1550 MTU on switches and routers.

Leverage 5-tuple hash distribution for uplink and inter-switch LACP
- Encapsulation will generate a source UDP port based on a hash of the inner packet 5-tuple (header information)

If VXLAN traffic is traversing a router, proxy ARP must be enabled on first hop router