VMware NSX Brownfield Deployment Guide

TECHNICAL WHITE PAPER

Table of Contents

Intended Audience

This document is targeted toward virtualization and network architects interested in deploying VMware® NSX network virtualization solution in a vSphere brownfield environment.

Overview

VMware's Software Defined Data Center (SDDC) architecture extends virtualization technologies across the entire physical data center infrastructure. VMware NSX, the network virtualization platform, is a key product in the SDDC architecture.

Having the ability to be deployed on any IP network, including both existing traditional networking models and next generation fabric architectures from any vendor, NSX is a completely non-disruptive solution. In fact, with NSX, the physical network infrastructure and services you already have are all you need to deploy a software-defined data center. In brownfield environments NSX can provide programmatic connectivity via L2 bridging or L3 routing and service insertion of existing physical services such as Load-Balancers and Firewalls, so such legacy services can be integrated and consumed by NSX Logical Networks.

Note: in the rest of this paper the terms "NSX" and "NSX-v" are used interchangeably and they always refer to the deployment of NSX with vSphere.

Deploying NSX on a Brownfield DC Network

Throughout the rest of this paper we will analyze different customer migration scenarios where NSX functionality can be seamlessly introduced in existing brownfield networks with minimal disruption and incremental functionality for the Data Center design.

In the first deployment scenario the goal is to introduce NSX on top of the brownfield network infrastructure; the assumption is that no major modifications are required in the network infrastructure itself. The specific scenario considered for this discussion is a traditional 3 layers DC network (access, aggregation, core), as highlighted in Figure 1:

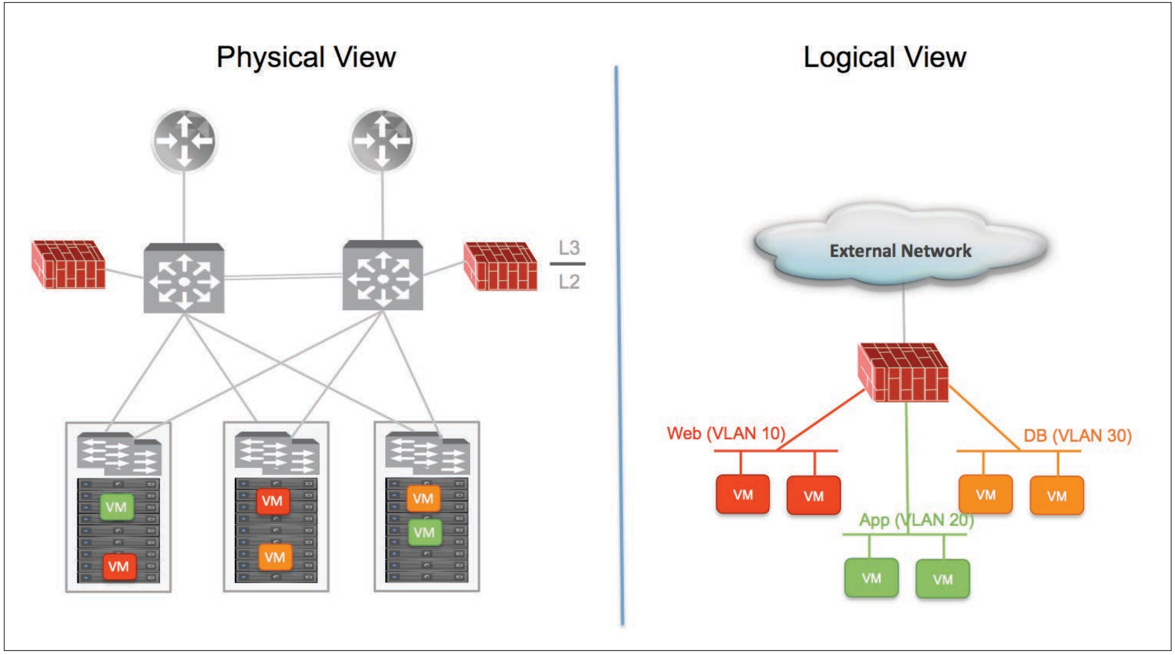


Figure 1: Classic 3 Layers DC Network Design

The main characteristics of such network are the following:

- The boundary between L2 and L3 network domains is positioned at the aggregation layer; more specifically, the physical FWs connected to the aggregation layer devices represent the default-gateway for the DC subnets (Web, App and DB tiers in this example).
- The access layer switches are functioning as pure L2 devices, switching traffic locally and to other access layer devices.
- Virtualized compute resources are connected to VLAN backed port-groups. The VLANs are spanned across the different access layer switches to enable L2 mobility for the deployed virtual machines.
- In this specific example we consider the deployment of a single POD (a POD is represented by the pair of aggregation switched and the connected access layer devices). Large DC deployments usually leverage multiple PODs interconnected at L3 via the Core routers.

Two specific migration use cases for this brownfield scenario are considered in this paper: micro-segmentation without overlays and micro-segmentation with overlays.

Scenario 1: Micro-segmentation without Overlays

The first migration scenario focuses on changing the way security policies are enforced for VMs connected to VLAN segments. The introduction of the Distributed Firewalling (DFW) capabilities offered with NSX ensures the optimal application of security policies for east-west (E-W) communication (zero-trust security policies) without the need of performing any change to the physical network infrastructure.

Multiple options can be offered for the application of security policies for north-south (N-S) communication, as it will be discussed in more detail in the “Detailed Migration Procedure” section. In all cases, one of the end-goals is removing the default gateway from the physical FW and positioning it on the aggregation layer device, so to be able to also optimize the handling of E-W communication.

Pre-Migration Considerations

- Initially security policies for all traffic flows are deployed on the centralized FW appliances. The FW is usually deployed as the default-gateway for the VMs, so that every IP subnet is segmented from the others leveraging security policies configured on the centralized FW. With this model, the physical FW is performing both routing and policy enforcement functionalities.

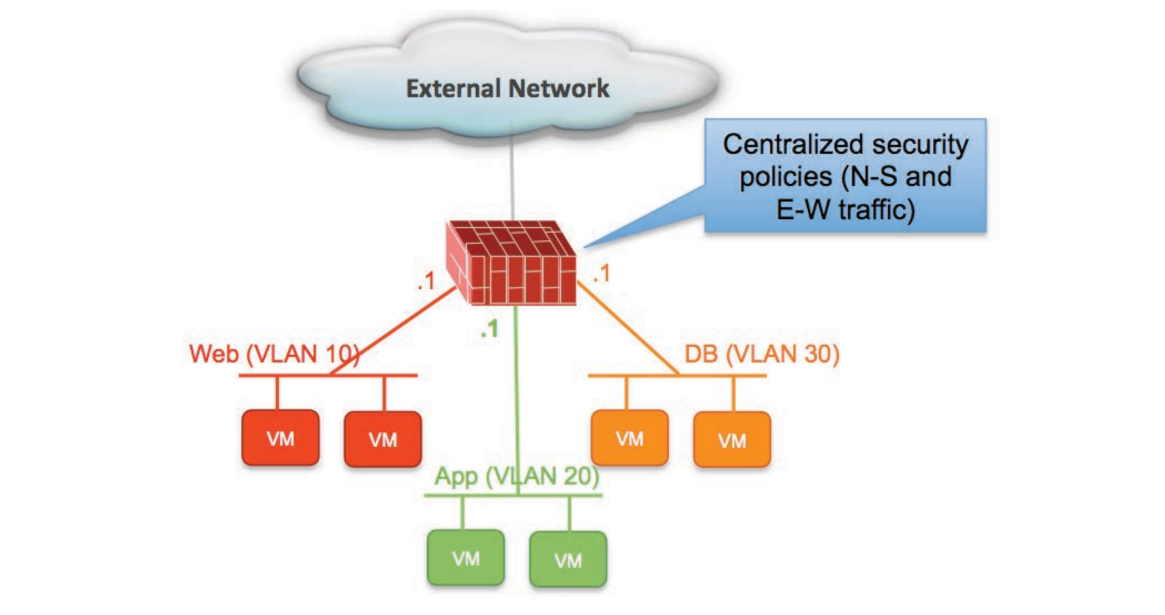


Figure 2: Centralized Security Policies on the Physical Firewall

Notice how intra-tier communications between VMs connected to the same VLAN segment are not subject to any policy enforcement, unless HW specific functionalities (as L2 ACLs or Private VLANs) are introduced in the physical switching infrastructure.

- One of the main requirement to configure NSX is for all the compute hosts to be already connected to the vSphere Distributed Switch (VDS). If that is not the case and the hosts are instead connected to the local vNetwork Standard Switch (VSS), it is first needed to migrate them to the VDS. Assuming that each ESXi host is equipped with at least two physical uplinks (vmnics) initially connected to the VSS, it is possible to ensure that the migration from VSS to VDS can be completed in a not disruptive fashion for the applications.

For more information on the recommended procedure for the VSS-VDS migration, please refer to the link below:

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

Similar considerations apply if the ESXi hosts are initially connected to a Nexus 1000v vSwitch and a migration to VDS will also be needed in that case as the first step of the migration process.

- ESXi Hosts requirements:

➢ Management cluster: the assumption is that a vCenter is already deployed to manage the compute clusters. The introduction of NSX to enable the DFW functionality does not require the provisioning of NSX Controllers, so it is only necessary to deploy the NSX Manager on the already existing Mgmt cluster.

➢ Compute clusters: there is no need to migrate VMs outside the existing compute clusters; the only requirement is to push the NSX VIBs to the Compute clusters to enable the DFW functionality. Additional ESXi hosts can later be added to existing Compute clusters or used to create new Compute clusters. In the former case, the VIBs will be automatically pushed to the hosts once they join the cluster; in the latter case, it is required to prepare new clusters for NSX from the NSX Manager UI.

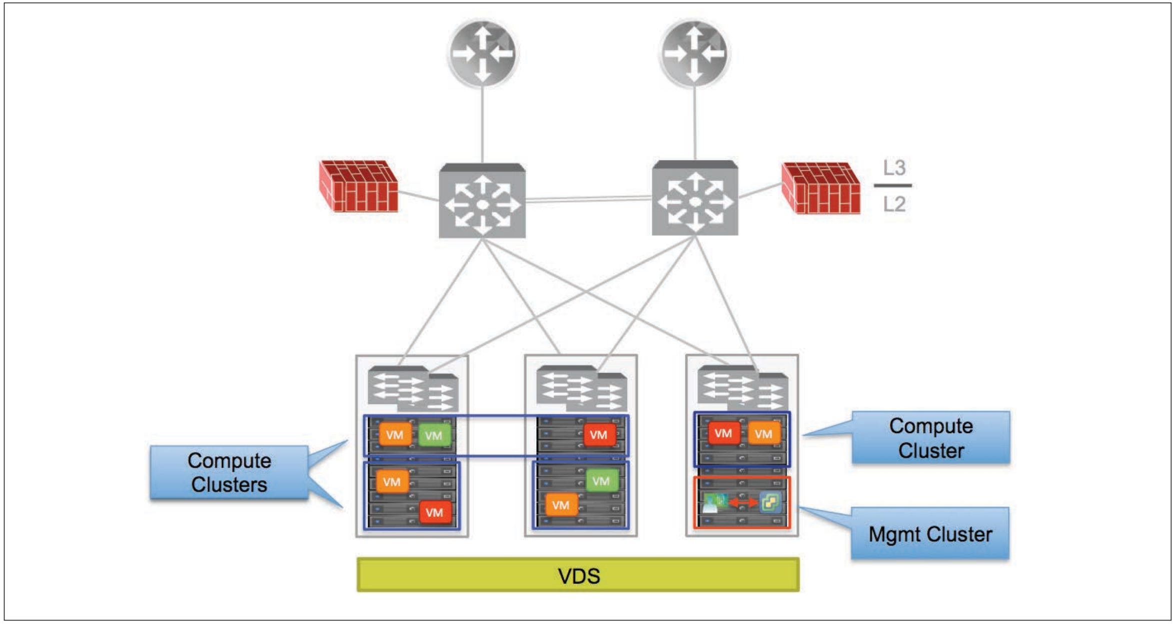


Figure 3: Compute Clusters, Mgmt Cluster and VDS

- Goals for the migration procedure:

➢ Adopting NSX DFW for optimal security policy enforcement for east-west communication (micro-segmentation or zero-trust security).

Note: the DFW is a micro-perimeter FW, which can apply security policies to all the flows originated or destined to each given VM, independently from the fact that the communication is confined internally to the DC (i.e. east-west flows) or established with the external network (i.e. north-south). Despite this clarification, in the context of this paper we focus specifically on the DFW security filtering applied to east-west communication, given the fact that a separate FW device (logical or physical) is usually deployed for north-south policy enforcement.

➢Moving the VM default gateway to the aggregation layer devices to remove this duty from the physical FW and increase east-west routing scalability.

➢Offer several alternative deployment options for handling north-south security policy enforcement, depending on the desire to utilize or not the pre-existing physical FW.

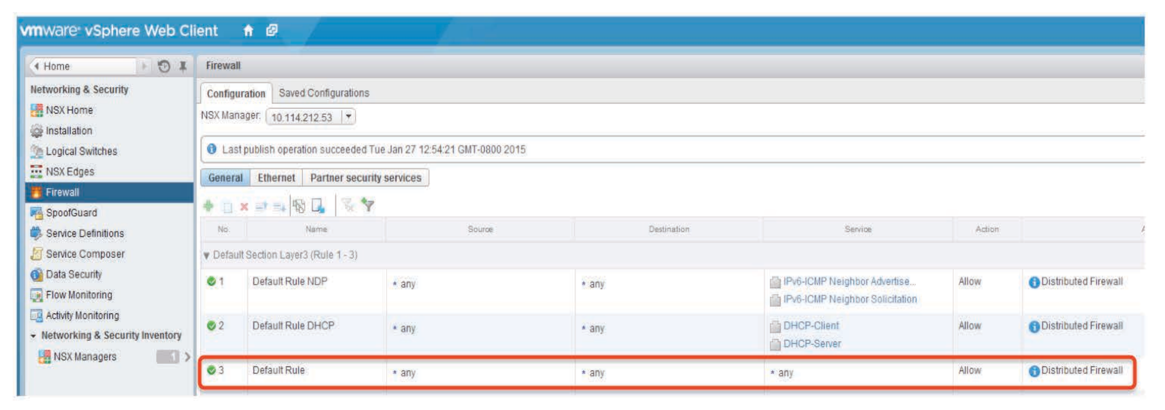
Detailed Migration Procedure

The step-by-step procedure to integrate NSX in this environment is described below:

Verify all the ESXi hosts that need to be prepared for NSX are connected to the VDS. If not, perform the required migration procedure (VSS to VDS or N1Kv to VDS).

b. Deploy the NSX Manager on the existing Mgmt cluster and link the NSX Manager to the vCenter already used to manage the existing management and compute clusters. As a reminder, there must always be a 1:1 relationship between the vCenter server that manages the compute cluster resources and the NSX Manager.

c. Prepare the compute clusters for NSX (by pushing the NSX VIBs to them): this is not expected to be a disruptive operation for existing application connected to the original VLAN environment; the DFW security policy enforcement is in fact applied as soon as the VIBs are pushed to the cluster, but there is a default rule permitting all communications (Figure 4).



NSX Manager: 10.114.212.53

Last publish operation succeeded Tue Jan 27 12:54:21 GMT-0800 2015

No.	Name	Source	Destination	Service	Action	
▾ Default Section Layer3 (Rule 1 - 3)						
✓ 1	Default Rule NDP	* any	* any	IPv6-ICMP Neighbor Advertise... IPv6-ICMP Neighbor Solicitation	Allow	Distributed Firewall
✓ 2	Default Rule DHCP	* any	* any	DHCP-Client DHCP-Server	Allow	Distributed Firewall
✓ 3	Default Rule	* any	* any	* any	Allow	Distributed Firewall

Figure 4: Default DFW Security Rules

Note: currently, a data-plane outage between 4-10 seconds is experienced when pushing the VIBs, which seems to be happening only when pushing the VIBs to ESXi hosts with Intel NIC cards (due to the current automatic enablement of Receive Side Scaling – RSS). Behavior is being investigated and could be corrected in future releases.

d. Deploy the distributed DFW policies. 3rd party solutions (like AlgoSec for example) will in future be integrated with NSX to be able to collect the FW rules from the physical FW and apply them directly to the DFW configuration. This will be possible on a per application basis, which makes sense since migrations are expected to be performed one application at the time. That said, the use of DFW allows creating security rules leveraging advanced constructs like Security Groups and vCenter objects, in addition to traditional IP or MAC addresses. The recommendation is hence to utilize those whenever possible and this implies a need to re-create the security policies on a per application basis (this represents also an opportunity to clean up and review the existing FW policies).

At this point, the security policies are still enforced in two places: the DFW and the physical FW. However, traffic not allowed by the policies configured on the DFW will be dropped directly at the HV level, already reducing the amount of traffic sent into the physical network infrastructure. This intermediate step also allows verifying that the DFW policies are working as designed; communication that may be allowed by mistake by a misconfigured DFW security policy is still dropped at the physical FW level, avoiding the creation of any security vulnerabilities

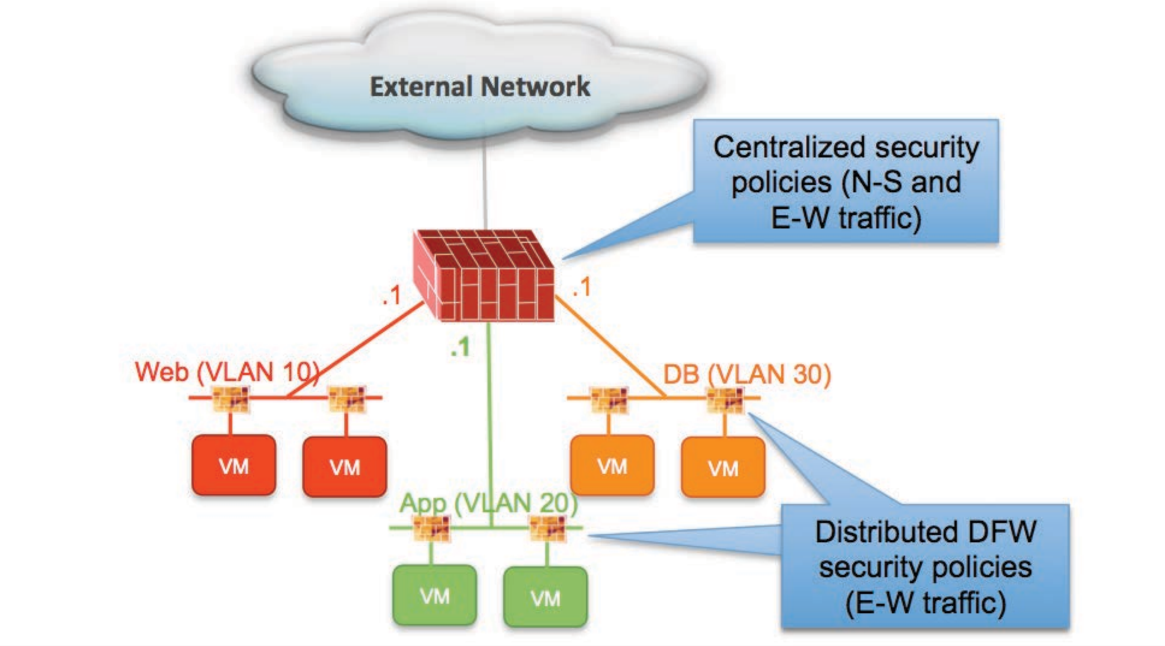

External Network
Centralized security policies (N-S and E-W traffic)
.1
.1
.1
Web (VLAN 10)
DB (VLAN 30)
VM
VM
VM
VM
App (VLAN 20)
VM
VM
Distributed DFW security policies (E-W traffic)

Figure 5: Centralized and Distributed Firewall Policies

e. Once the DFW functionality has been successfully verified, it is optionally possible to remove the policies for E-W traffic from the physical FW (the physical FW should anyway stop getting any deny hits, as traffic is now dropped by DFW). From this point on, all the policy enforcement for E-W communication is only applied at the DFW level.

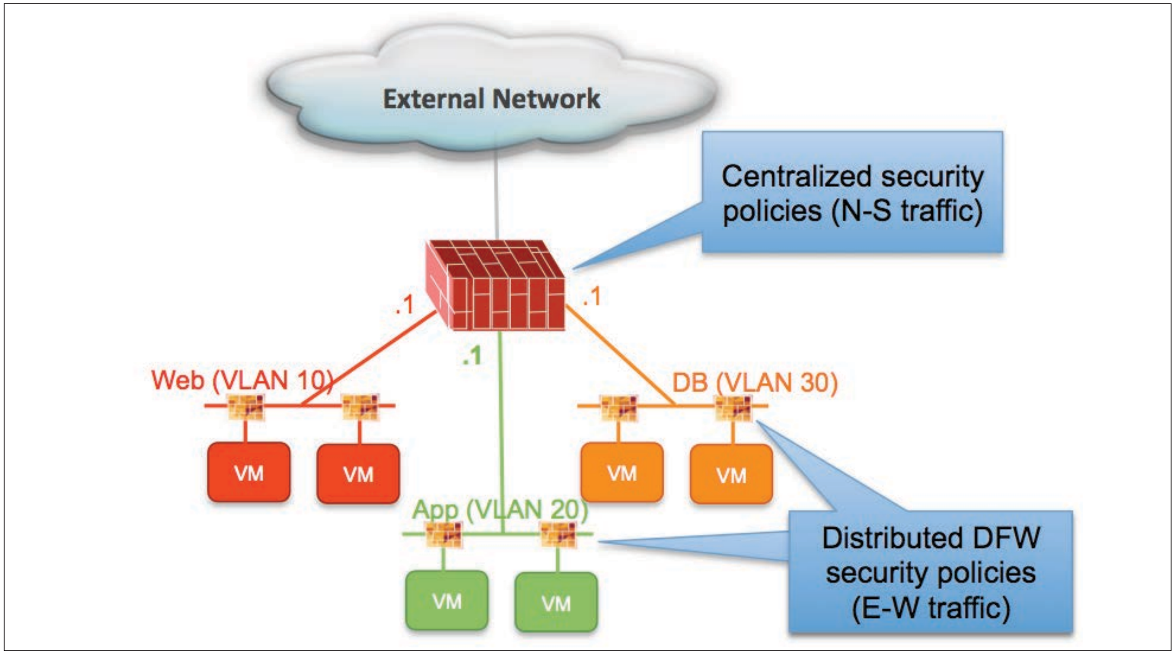

External Network
Centralized security policies (N-S traffic)
.1
.1
.1
Web (VLAN 10)
DB (VLAN 30)
VM
VM
VM
VM
App (VLAN 20)
VM
VM
Distributed DFW security policies (E-W traffic)

Figure 6: Physical FW Used only for N-S Security Policies

f. Configure the default gateway at the aggregation layer to be used for all routed E-W communication. This usually entails configuring VLAN Interfaces (SVIs) with the default gateway IP address for the Web, App and DB subnets (.1 is used in the example in Figure 6). At this point, the SVIs should be still kept in "shutdown" mode, since the physical FW is still performing routing functions.

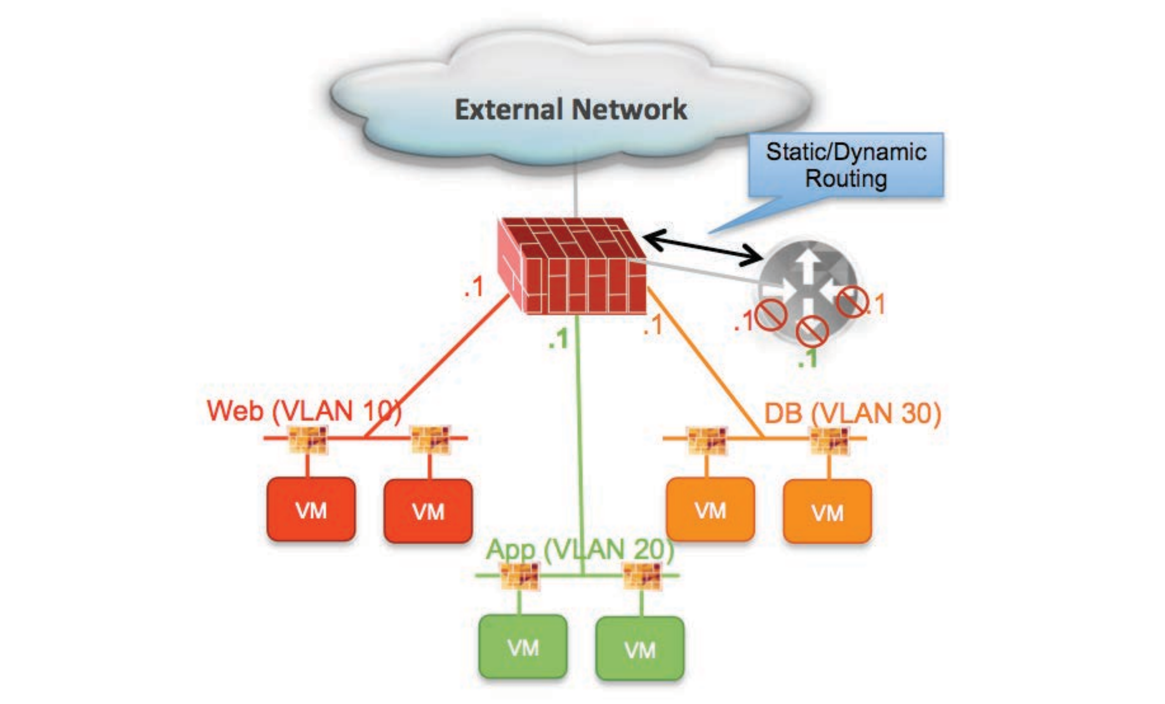


Figure 7: Configuration of Default Gateway on the Aggregation Layer Devices

Static or dynamic routing can also be established between the aggregation layer devices and the physical FW; this requires the definition of another SVI interface on the aggregation switches that will be used to establish the L3 static/dynamic relationship with the FW. This additional VLAN (VLAN 100 in this example) can be added to the other VLANs originally carried on the L2 trunk connection between the aggregation switch and the FW, as shown in Figure 8.

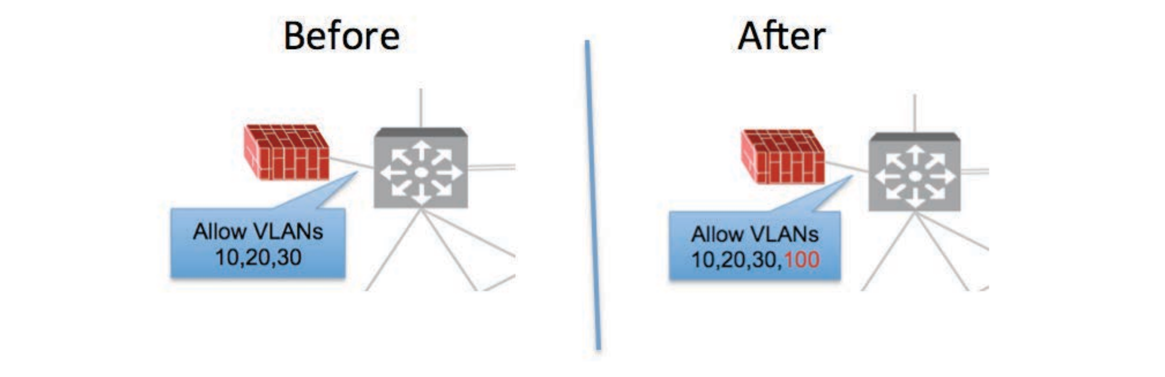


Figure 8: Adding L3 Peering VLAN on the L2 Trunk

This new VLAN represents the path that will be used for north-south communication once the next default gateway migration step is completed.

g. Switch the default gateway for the VMs subnets from the physical FW to the aggregation layer. This is a disruptive operation that should be performed during a maintenance window, as it requires the VMs to update their local ARP caches with the new MAC address of the default gateway.

Two alternative deployment models are discussed in this paper, depending on where security policing for N-S communication is performed.

1. The physical FW is still kept operational and it is connected inline on the north side of the aggregation layer devices.

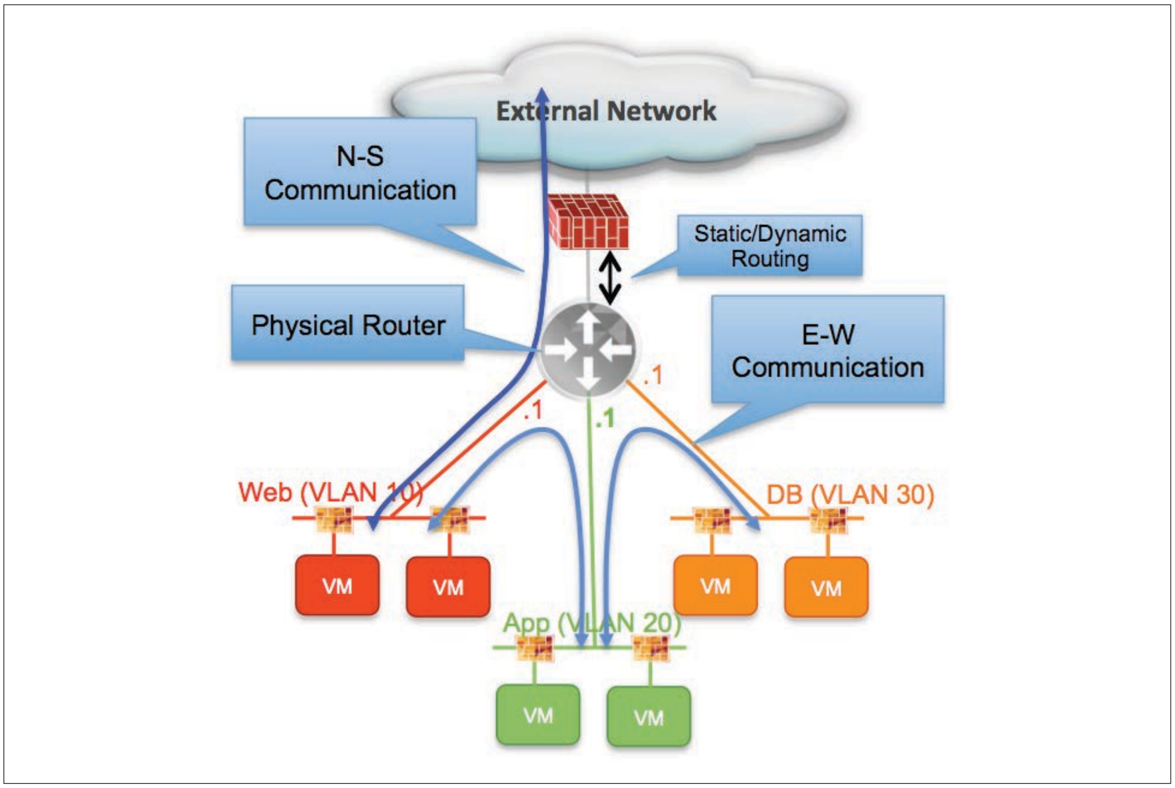


Figure 9: Physical FW Deployed Inline

This is an easy model where the physical FW is still used to apply security policies to the N-S communication, whereas the DFW is used for E-W traffic flows (and allows at the same time to enforce also intra-tier policy enforcement for VMs connected to the same VLAN).

Before the migration started, it's likely that the physical FW already had a static or dynamic routing peering with the next-hop device on the northbound side. This implies that the only required step is to establish the static/dynamic routing peering with the aggregation switch to start exchanging reachability information into the DC IP subnets, as already described in the previous migration step. Also, it is now possible to prune the VMs data VLANs (VLANs 10, 20 and 30) from the L2 trunks connecting the aggregation layer devices and the physical FW (shown in Figure 8).

Note: in this deployment model the physical FW may also perform NAT functionalities for the DC subnets leveraging private IP address space.

2. The physical FW is completely removed from the design.

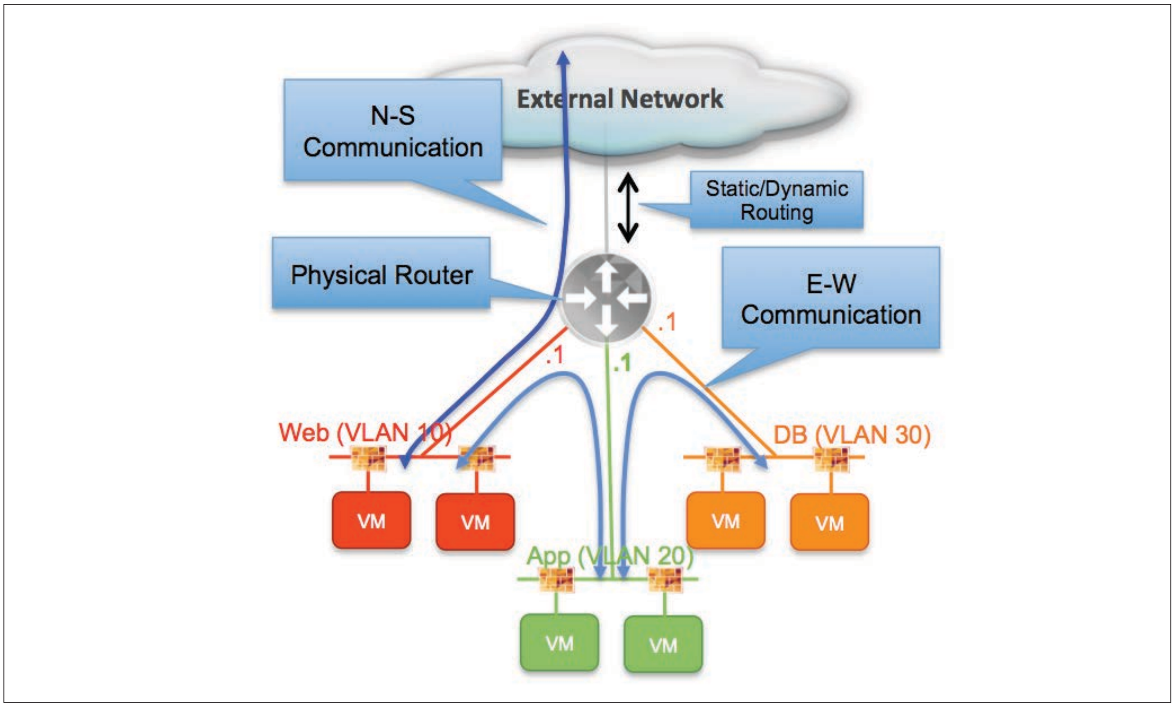


Figure 10: Use of DFW for Enforcing also N-S Policies

In this case, N-S policy enforcement can be applied at the DFW level or on the physical perimeter FW usually protecting the DC from the public Internet.

In any case, the physical FW originally connected at the aggregation layer can be decommissioned, which represent a saving both from OPEX and CAPEX perspectives.

Scenario 2: Micro-segmentation with Overlays

In addition to the FW policy changes discussed in the previous model, this scenario introduces the deployment of VXLAN based logical networks. The immediate advantage is the decoupling of connectivity in logical space from the physical underlay configuration: workloads can be connected to the same L2 domain (represented by a given VXLAN segment) independently from the specific configuration of the physical network (i.e. this can be achieved also in deployment s where the workloads are connected in parts of the network that are not L2 adjacent).

An additional requirement is introduced in this use case: the need to allow communication between VMs connected in logical space and bare-metal servers that cannot be virtualized. This is a common requirement and Figure 11 shows the typical example of a three tiers application deployment (Web, App and DB), where workloads belonging to the DB tier are a mix of VMs and bare-metal servers.

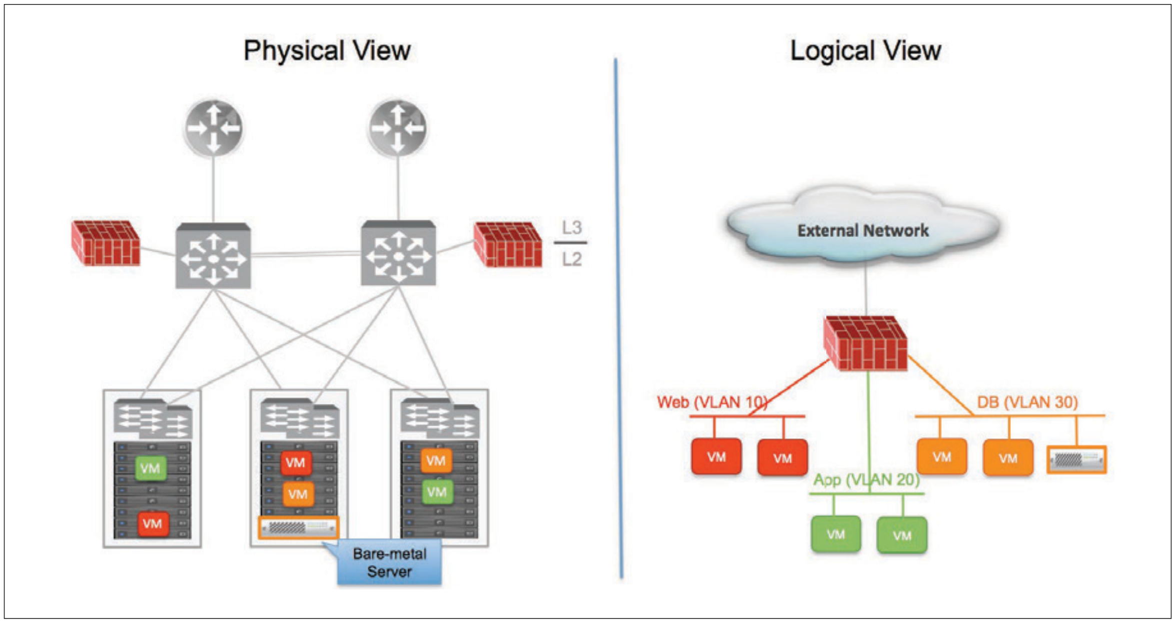


Figure 11: Three Layers DC Network with Virtualized and Bare-metal Servers

Deployment Considerations

- VDS considerations: similarly to what discussed in the previous use case, also the deployment of VXLAN logical networks mandates that the ESXi hosts are connected to the VDS switch. This implies that existing compute clusters leveraging VSS (or another virtual switch, as for example the Cisco Nexus 1000v) must be first migrated to VDS.

- Server requirements:

➢ Mgmt cluster: in addition to the deployment of the NSX Manager, a cluster of three NSX Controllers is also required to be able to configure logical networks and leverage the NSX advanced ARP suppression capabilities. The best practice is to deploy each Controller node on a separate ESXi host, so it may be needed to add ESXi hosts to the Mgmt cluster (just to ensure the NSX Controllers cluster can remain operational even under an ESXi host failure scenario).

➢ Compute clusters: may reuse existing ones or deploy/add new ones to be used for VMs connected to logical switches, as already discussed for the previous use case.

➢ Edge cluster: it is best practice recommendation to deploy a new Edge cluster to host the Edge Services Gateways and the DLR Control VMs. The Edge cluster could also be used to host the NSX Controllers, in deployments where the Mgmt cluster does not offer enough capacity.

Note: the Control VMs may instead be deployed as part of the Compute clusters. This is specifically recommended when deploying the ESG in ECMP mode, as it is best practice avoiding the deployment of an ESG node and a DLR Control VM on the same ESXi host. For more information please refer to the NSX Reference Design Guide:

https://communities.vmware.com/docs/DOC-27683

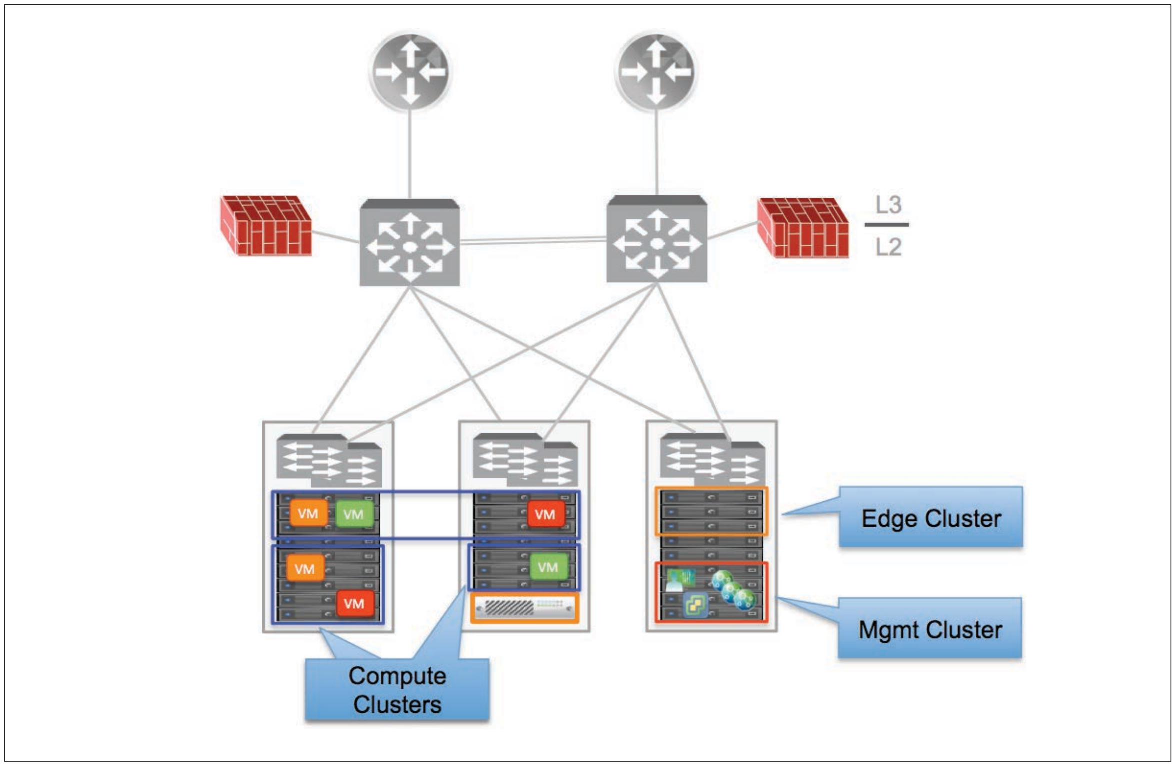


Figure 12: Compute, Edge and Mgmt Clusters

- Goals for the migration procedure:

➢ Migrate VMs to logical networks (VXLAN segments) to benefit of distributed network services in logical space (switching, routing, firewalling, load-balancing).

➢ As a side effect of the previous point, remove routing and the east-west security policy enforcement duties from the physical FW. Optionally, the physical FW could be decommissioned for CAPEX and OPEX savings (as discussed in the first migration use case).

➢ Preserve L2 communication between VMs connected to logical switches and bare-metal servers connected to VLANs.

➢ Simplify the configuration in the physical network: removing VM data VLANs that are not required anymore, default gateway configuration, etc. At the same time, few new VLANs will be introduced in the physical network:

o A new VLAN (the VXLAN transport VLAN) will be deployed in order to allow east-west communication between VMs belonging to different ESXi hosts.

o Another VLAN will also be needed between the NSX Edge Services Gateways and the physical network infrastructure to enable north-south communication.

The simplification of configuration in the physical network and the deployment of network services in logical space represent significant OPEX savings. Additionally, and more importantly, the introduction of NSX with its capability of decoupling logical and physical network connectivity (overlay vs. underlay) brings agility to the overall DC deployment, providing the ability to deploy applications almost in a real time manner as opposed to having to wait for change windows and removing the need to modify the configuration of the physical network.

Detailed Migration Procedure

The step-by-step procedure to integrate NSX in this environment is detailed below:

a. Verify all the ESXi hosts that need to be prepared for NSX are connected to the VDS. If not, perform the required migration procedure (VSS to VDS or N1Kv to VDS).

b. Deploy the NSX Manager on the existing Mgmt cluster and link the NSX Manager to the existing vCenter Server (already used to manage the Mgmt and Compute clusters).

c. Deploy the 3 NSX Controllers as part of the Mgmt cluster. Optionally, the NSX Controllers could be deployed as part of the Edge cluster. Recommendation is to have at least 3 ESXi hosts in either cluster to ensure the Controllers can be placed on separate hosts.

d. Prepare the compute and edge clusters for NSX (by pushing the NSX VIBs to them): in most cases this is a non-disruptive function for existing workloads connected to the original VLAN environment.

e. Deploy the DFW policies for E-W communication, as previously discussed in Migration Scenario 1. At this point, the policies are still enforced in two places: the DFW and the centralized FW. However, traffic not allowed by the policies will be dropped directly at the HV level, reducing already the amount of traffic sent into the physical network infrastructure.

f. Optionally remove the policies for E-W traffic from the centralized FW. At this point, all the E-W policy enforcement is only applied at the DFW level, the physical FW is still the default-gateway performing routing functionalities for all the VMs connected to VLAN backed port-groups.

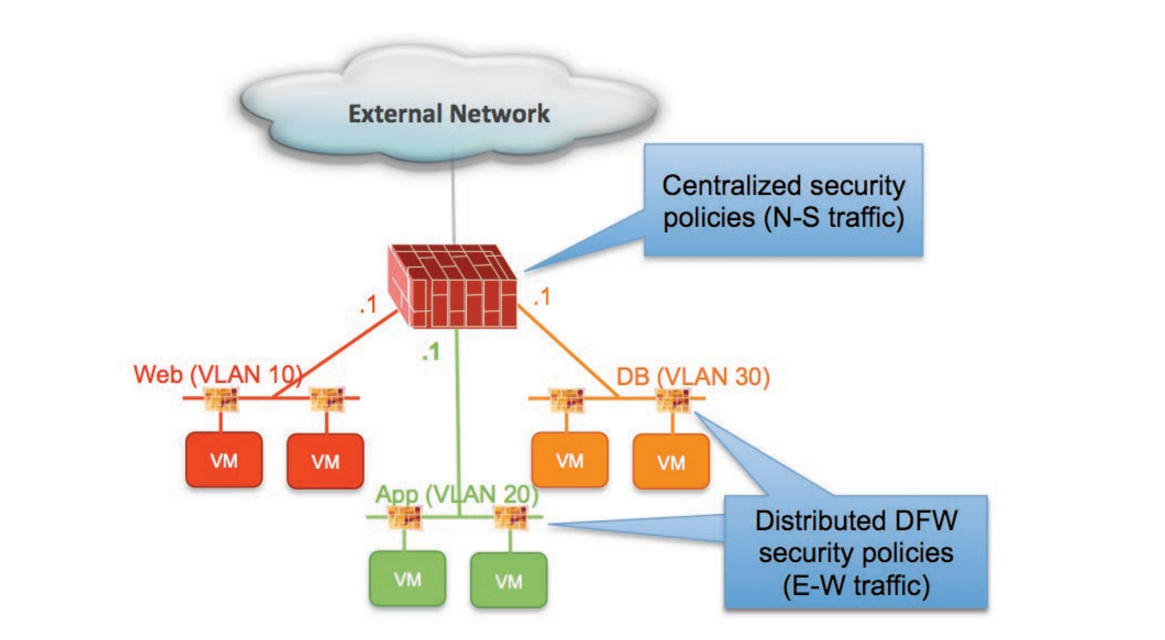

External Network
Centralized security policies (N-S traffic)
.1
.1
.1
Web (VLAN 10)
DB (VLAN 30)
App (VLAN 20)
VM
VM
VM
VM
VM
VM
Distributed DFW security policies (E-W traffic)

Figure 13: DFW for East-West Security Policies

As highlighted in Figure 13, the DFW capabilities are applied to the vnic interfaces of the virtualized workloads. However, configuration of these security policies allows also controlling communication from/to the bare-metal servers. This is true both for intra-subnet communication (i.e. between VMs and physical servers belonging to the same VLAN 30) and inter-subnet communication.

g. Configure VXLAN on all the compute and edge clusters: this operation typically is non disruptive for existing application connected to the original VLAN environment. Notice that this also requires adding to the network a VLAN used to carry the VXLAN encapsulated traffic originated by the VTEP VMkernel interfaces defined on each ESXi host. The default gateway for that VLAN can be configured on the aggregation layer switches.

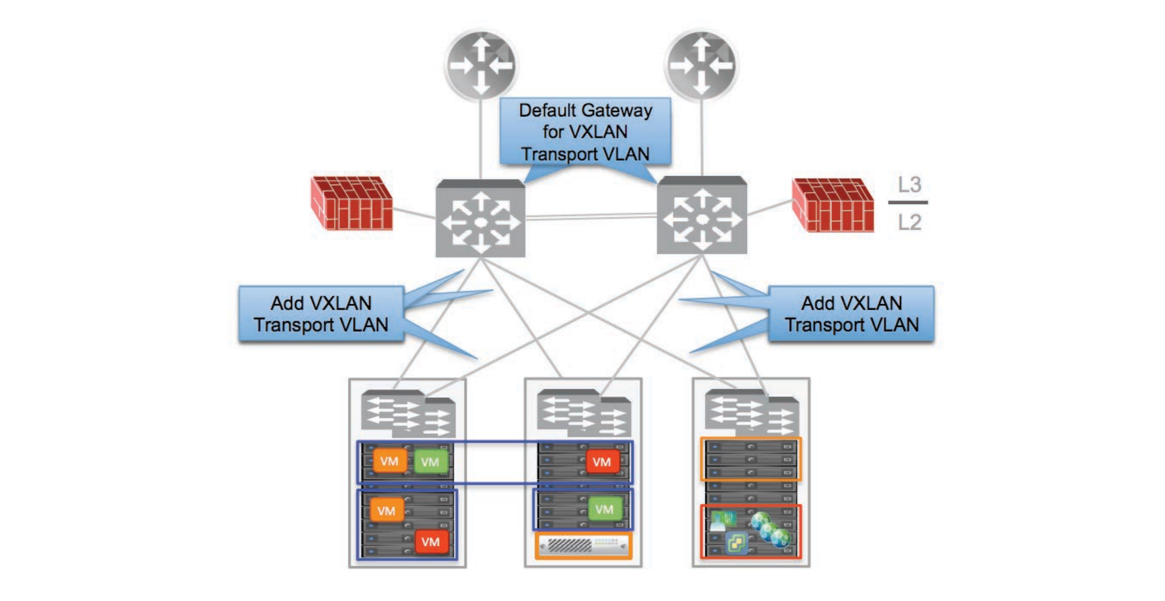


Figure 14: Configure VXLAN Transport VLAN in the Physical Network

h. Increase the MTU in the network infrastructure: since VXLAN traffic needs to be supported across the network, it is mandatory to increase the MTU to at least 1600B value on the following interfaces:

- L2 interfaces on the ToR devices: this configuration applies to the interfaces facing the ESXi hosts, the ones facing the aggregation layer switches and the interfaces between the ToRs.
- L2 and L3 interfaces on the aggregation switches: MTU must be increased on the L2 interfaces facing the various ToR switches, on the L2 trunk between the aggregation switches and on the L3 SVI interfaces relative to the VXLAN Transport VLAN. The last configuration is required in cases where VXLAN traffic must be routed between VTEP IP addresses belonging to different IP subnets.

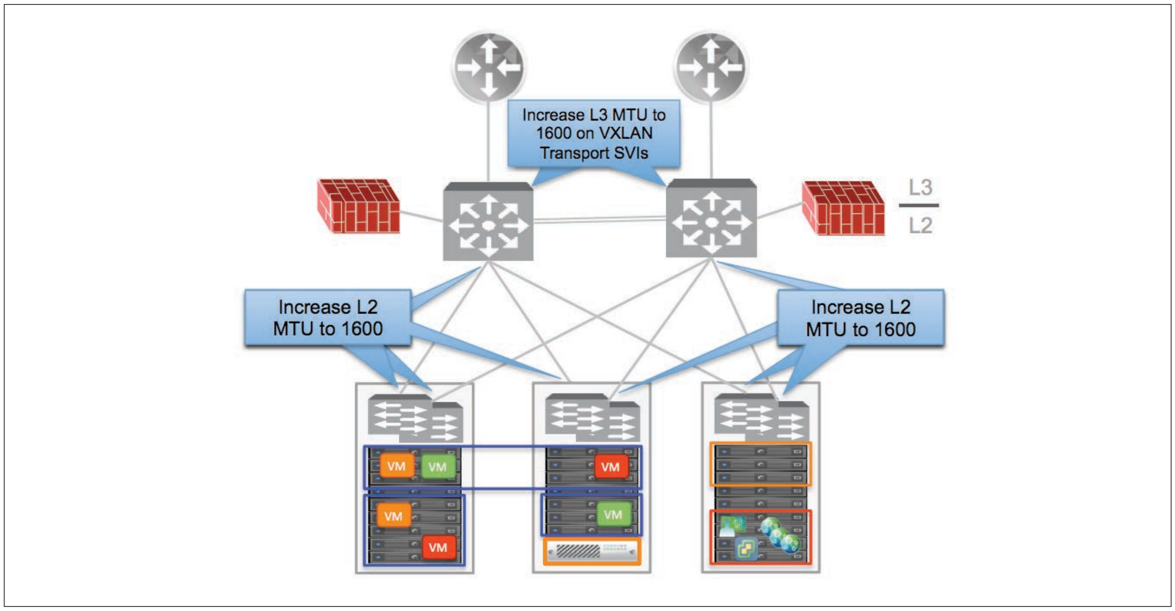


Figure 15: Increasing MTU on L2 and L3 Interfaces

In the scenario depicted in Figure 15 the VXLAN transport VLAN can easily be extended across all the server racks. It is hence likely that all the VTEP IP addresses for the ESXi hosts belonging to the Compute and Edge clusters are part of the same IP subnet, given the fact the L2/L3 boundary for the VXLAN Transport VLAN is positioned at the aggregation layer device. The IP subnet used for the VXLAN transport network should be assigned from the pool used for the underlay network infrastructure.

Note: changing the MTU on L2 interfaces may be disruptive for traffic originally flowing on that physical link (this mostly depends on the physical switch vendor). As a consequence, it is recommended to perform this operation one link at the time, so to ensure traffic can be recovered on the redundant available paths.

i. Create the Logical Switches (VXLAN segments) that will host the VMs after the migration is completed. This operation can be performed via the NSX Manager UI or leveraging REST API calls. Notice how at this point there are no VMs connected to those Logical Switches and at the same time the Logical Switches are not connected to any logical router interface yet.

j. Deploy NSX L2 Bridging that provides the VXLAN-to-VLAN communication for all the VMs that need to be migrated in logical space. Since, as previously mentioned, there is no problem in extending VLANs across the access layer devices deployed in separate racks, NSX L2 Bridging can be freely deployed as part of the Edge cluster or as part of a Compute cluster. A single NSX L2 Bridging instance can be active at any given time to perform bridging for up to 512 VXLAN/VLAN pairs. However, NSX offers a scale-out model allowing for the deployment of more than one bridging instance for multiple VXLAN-VLAN pairs (3 separate bridging instances are deployed in the example in Figure 16 below as part of the Compute clusters resources).

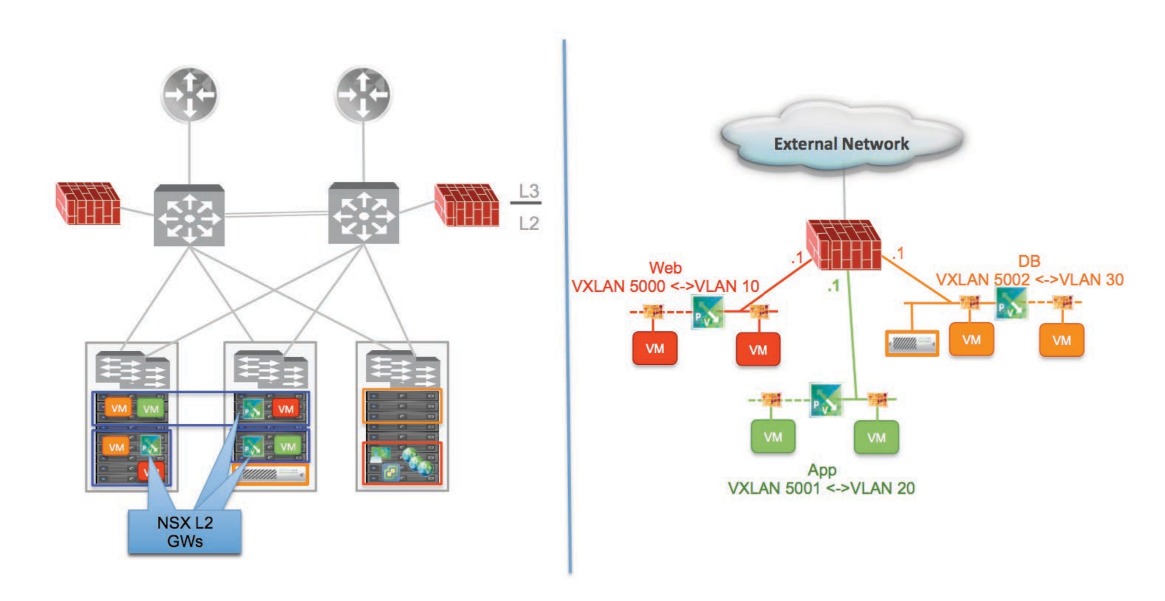


Figure 16: Deployment of NSX L2 Gateways as Part of the Compute Clusters

k. Start migrating VMs belonging to each VLAN into logical space. This migration process for each VM simply consists in changing the port-group for the VM's vnic from a VLAN backed port-group to a VXLAN backed one. This operation can be completed without impacting the data-plane traffic sourced and received by the VM.

During this intermediate migration step:

- The VMs belonging to the same L2 domain and connected to VLAN and VXLAN port-groups would communicate with each other leveraging NSX L2 Bridging. Same mechanism is used by the VMs connected to logical networks to communicate with bare-metal servers.
- The VMs migrated to the VXLAN segments still leverage the gateway in physical space on the FW.
- Inter-subnet routing is still performed by the FW (see right side of Figure 16).

l. Complete the VM migration and deploy the DLR Control VMs and the NSX Edges as part of the Edge cluster.

Note: it is best practice to define a dedicated DLR instance for routing, which is separate from the DLR instances previously introduced to enable the NSX L2 Bridging functionality.

A dynamic routing protocol can be configured on the DLR and on the Edge to start peering with each other. The DLR does not communicate any routing information yet, since no Logical Switches are connected to it at this point. The Edge can also establish static or dynamic routing with the physical FW to program in its forwarding table a valid path that will be used for N-S communication. This statically programmed or dynamically learned routing information is then communicated to the DLR leveraging the dynamic routing protocol.

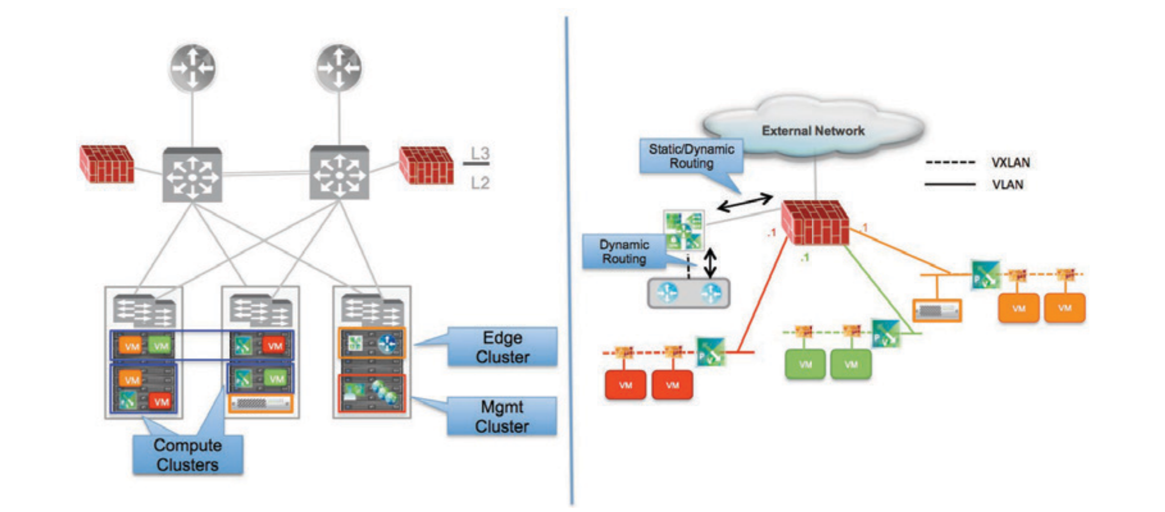


Figure 17: Deployment of DLR Control VMs and NSX Edges

As shown in Figure 17, at this point all the VMs have been migrated to the Logical Switches and still leverage NSX L2 Bridging to communicate with the default gateway positioned on the physical FW.

m. The next migration step is disruptive to traffic forwarding and should hence be performed during a maintenance window. The goal is to move the default gateway from the physical FW into logical space. In order to achieve this, it is required to connect the logical switches to the DLR, disconnect the physical FW from the VLAN segments and decommission the NSX L2 Bridging instances that are not required anymore (with the exception of the ones providing connectivity to the bare-metal servers).

From this point on, E-W communication is optimized by the DLR. For the handling of N-S communication and the application of security policies to those traffic flows, two different models can be proposed:

1. Keep the physical FW deployed inline on the north side of the NSX Edge.

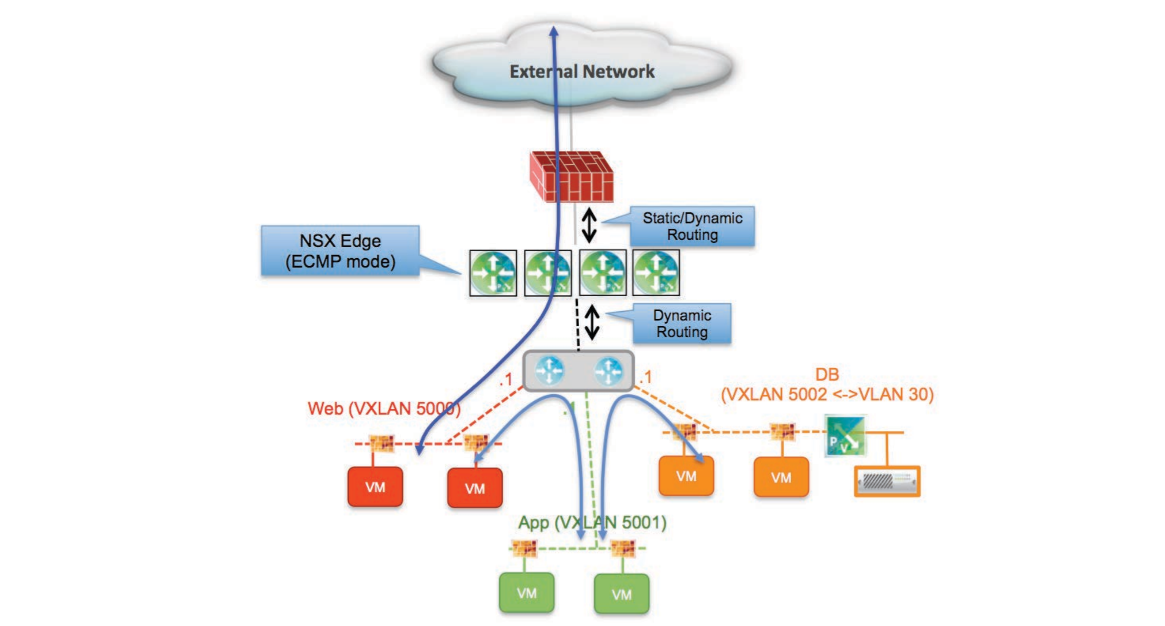


Figure 18: Physical FW Deployed Inline

In this case the security policies are enforced on the FW, together with NAT functions (if required). This allows deploying the NSX Edge in ECMP mode, for improved throughput and faster convergence.

2. Remove the physical FW and deploy inline FW functionality on the NSX Edge.

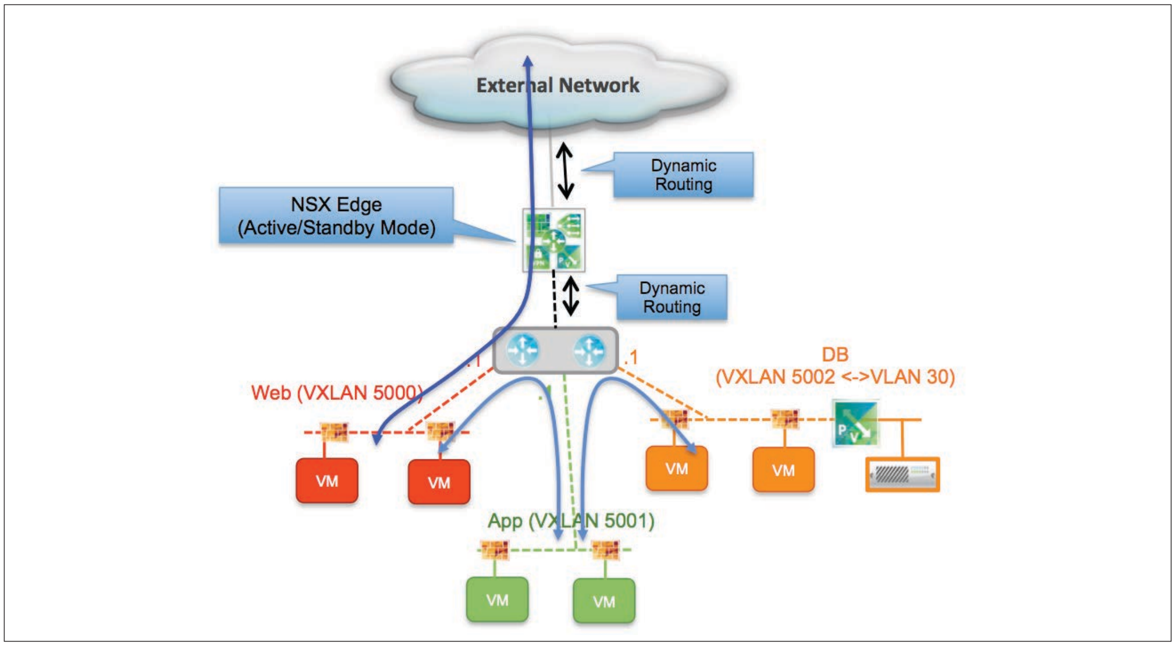


Figure 19: FW Functionality on the NSX Edge

The NSX Edge is here utilized also for FW and NAT, so it must be deployed in Active/Standby mode.

Note: the deployment models shown in Figure 18 and Figure 19, where the VXLAN segment 5002 is connected to the DLR and at the same time bridged to VLAN 30 can only be supported from NSX release 6.2. In deployments leveraging earlier releases it would be needed to connect VXLAN 5002 directly to the NSX Edge, which could hence be deployed in Active/Standby mode only. This current deployment option is shown in Figure 20.

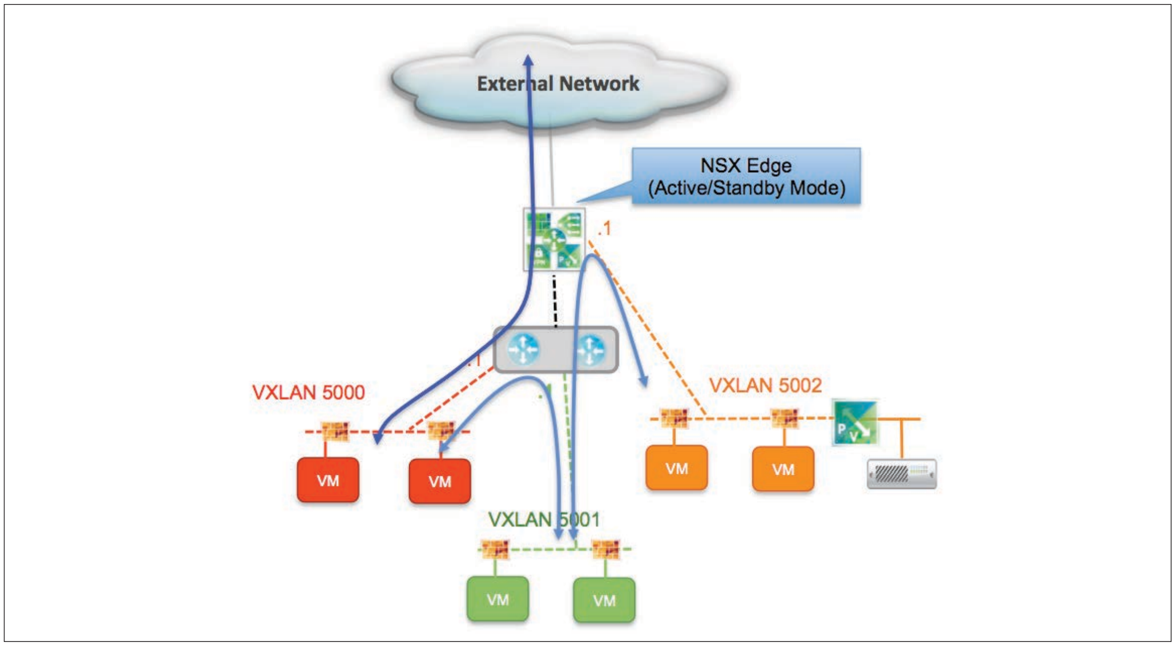


Figure 20: VXLAN Bridged Segment Connected to NSX Edge (pre-6.2 NSX Releases)

n. The final migration step consists of cleaning up the configuration on the network side (optionally including the physical FW) to remove the VLANs that originally where used only for the VMs (VLANs with bare-metal server obviously remain in service). The VLAN backed port-groups associated to those VLANs can also be removed from the vCenter server.principle is carried in the document and repeated to maintain ease of user readability.

Scenario 3: Deploying NSX on a Greenfield DC Network

A different scenario is the one where a customer is planning to build a greenfield infrastructure to deploy NSX, with the consequent need to migrate applications from the brownfield network to the new environment (Figure 21).

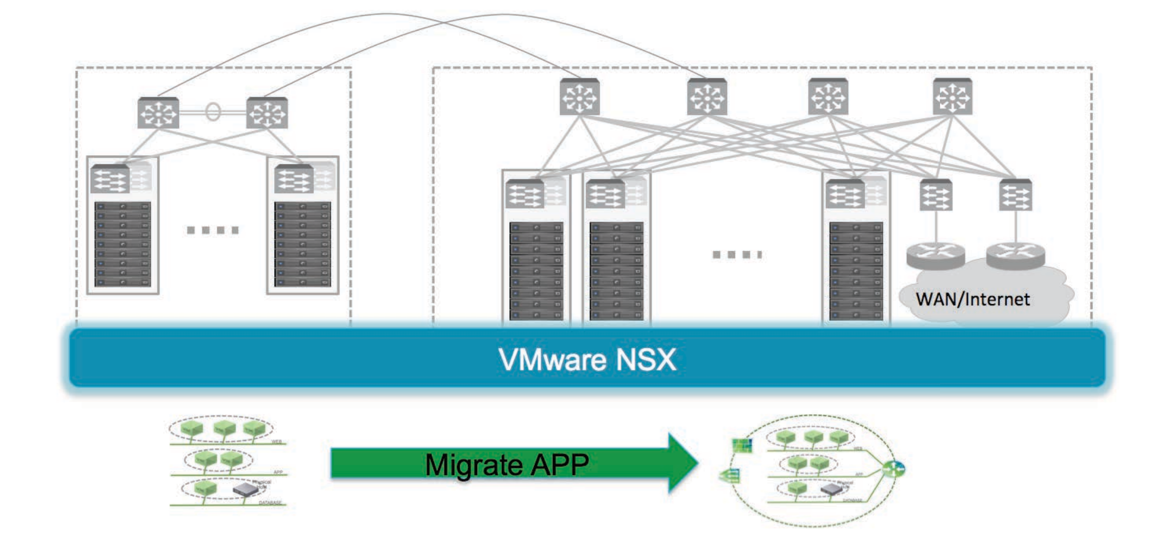


Figure 21: Migrating Apps between Brownfield and Greenfield Networks

Deployment Considerations

- The basic assumption is that the NSX deployment on the greenfield side can follow the best practice design recommendations discussed in the NSX Reference Design Guide:

  https://communities.vmware.com/docs/DOC-27683

  This implies deploying new Compute and Edge clusters, with best practice VDS configuration (separate VDS for Compute and Edge clusters). Those clusters are normally connected to a routed fabric network, deployed in a spine-leaf topology (the L2/L3 boundary is hence pushed down to the leaf devices).

- No assumptions are made for the Compute resources deployed in the brownfield site; this means they may be running an old vSphere version and being managed by an old vCenter release. The intent is to provide a migration procedure that does not require any upgrade on the existing gear connected to the brownfield network.

- L2 Gateways can be introduced as “anchor points” for VXLAN connectivity between the brownfield and greenfield networks, as highlighted in Figure 22.

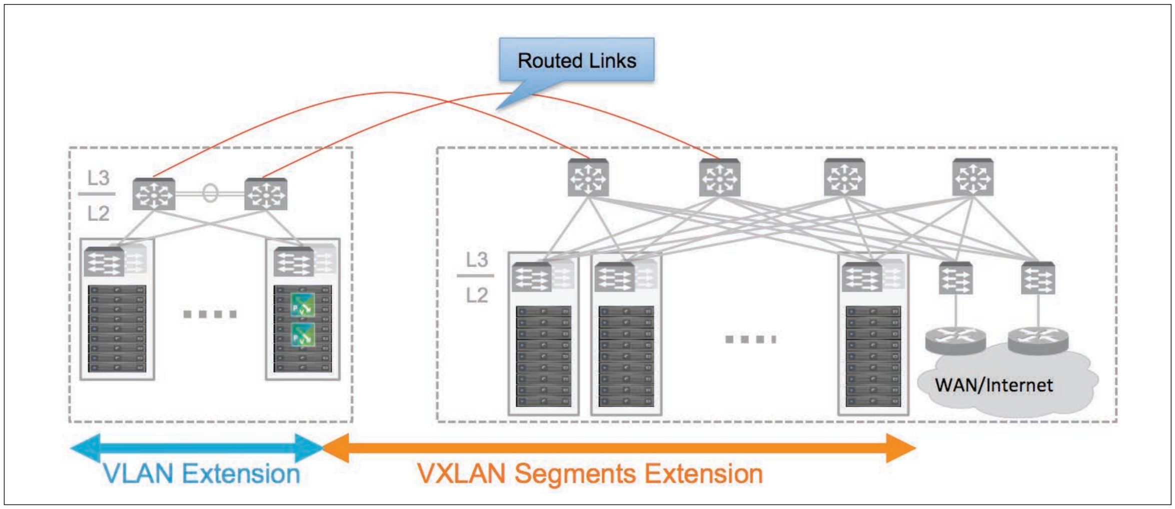


Figure 22: Use of NSX L2 Gateways for VM Migration

The main advantage of this approach is that L2 connectivity can be established between VMs migrated to the greenfield network and VMs (or bare-metal servers) that remain connected to VLANs in the brownfield side, without requiring any L2 extension (i.e. VLAN stretching) between the two networks. This L2 connectivity is likely required during the VMs migration phase (which can also last several months); at the same time, bare-metal hosts may remain connected to the brownfield network until the next refresh cycle (think for example Oracle RAC deployments, etc.).

Figure 22 shows how the use of VXLAN allows establishing L2 communications across routed connections. Also, since the greenfield infrastructure is usually deployed in closed proximity of the brownfield one, the use of VXLAN allows to fully leverage the available bandwidth between the two networks.

Note: the L3 links could terminate on the spine devices in the greenfield network, or optionally on a pair of leaf switches. The latter option (shown in Figure 23) is recommended to simplify the configuration and functionalities enabled on the spine switches, which essentially function as the “backplane” of the fabric.

For what concerns the deployment of L2 Gateways, a couple of options are applicable:

1. Deploy NSX L2 Bridging instances in the brownfield site: a couple of ESXi hosts (running vSphere release 5.5 or later) can be installed in the brownfield site; those hosts are managed by the greenfield vCenter server and become part of the greenfield NSX domain. Brownfield VLANs that need to be connected to the new VXLAN segments should just be carried to the rack where those bridging instances are deployed. As previously mentioned, this is usually not a big challenge in multi-layer network design where VLANs can be easily stretched across the access layer devices.
2. Deploy HW VTEP ToR switches in the brownfield network. Upcoming NSX release 6.2 will introduce full control and data plane integration between HW VTEPs and NSX Controllers (leveraging the OVSDB control plane), so that the deployment of those ToRs and the bridging configuration can be centrally managed from the NSX domain.

- Clusters deployment: the assumption is that a vCenter server is going to be deployed (with the other NSX components, like the NSX Manager and the NSX Controllers) in a new Mgmt cluster in the Greenfield network to manage the local Compute and Edge clusters. This is different from the vCenter Server already deployed in a Mgmt cluster in the brownfield side to manage the local Compute resources (Figure 23).

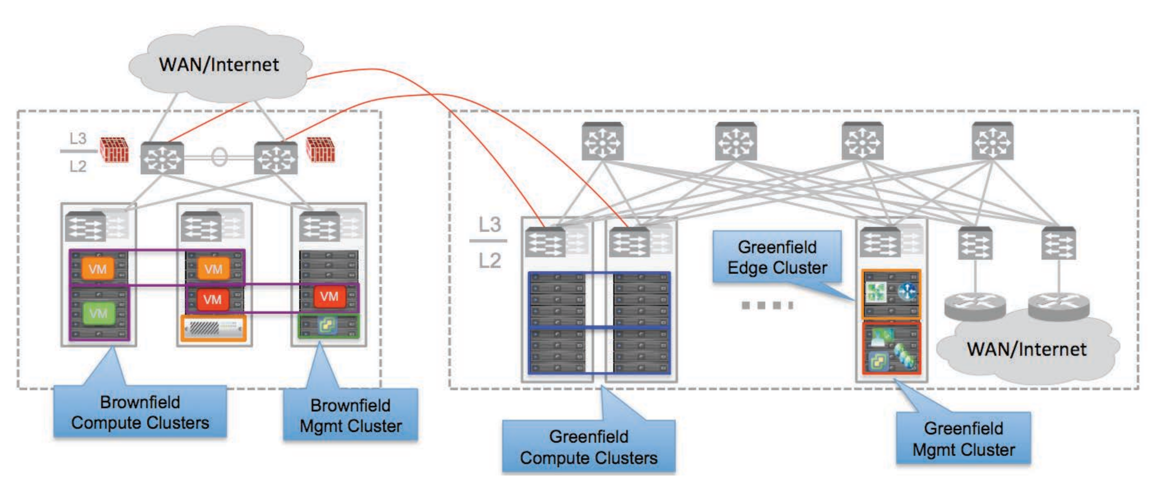


Figure 23: Separate vCenter Servers Deployed in Brownfield and Greenfield

- Goals of the migration:

➢ Onboarding applications originally running in the brownfield network to the greenfield infrastructure.

➢ Ensuring network and security services can be offered in logical space to those migrated applications.

➢ Maintaining connectivity to bare-metal servers that may remain deployed in the brownfield network.

➢ Ensuring optimized access to the application at the end of the migration to the greenfield network, removing the need for hair-pinning the communication across the links to the brownfield infrastructure.

Detailed Migration Procedure

The initial assumption before starting the migration procedure is that NSX with all its functional components has been deployed in the greenfield network. This implies also the creation of logical switches that will host the migrated VMs and of the logical routing components (DLR and NSX Edges) for establishing local communication to the external network. The logical switches can also be already connected to the DLR instance, but the DLR interfaces should be initially kept in disabled state, as shown in Figure 24.

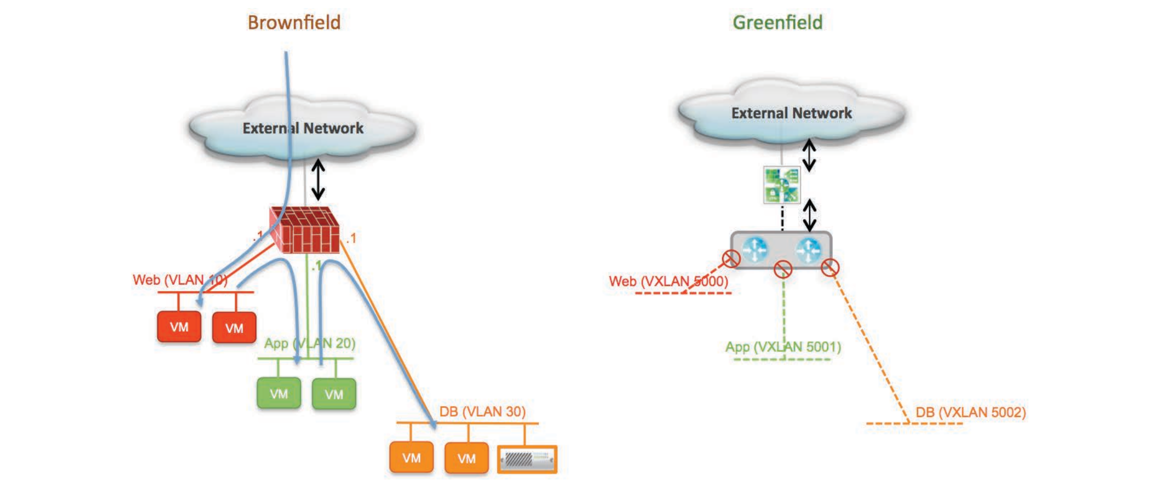


Figure 24: - Logical Views of Brownfield and Greenfield Networks

All the security enforcement and routing happens on the physical FW deployed in the brownfield network, and north-south communication also happens via the local connection to the external network.

a. The first step of the migration process consists in enabling L2 bridging between the VLANs used in the brownfield network and the VXLANs segments deployed in the NSX domain. In this example we consider the deployment of NSX L2 Bridging instances on ESXi hosts that are connected to the brownfield network. Those hosts must be managed by the vCenter server deployed in the greenfield network, since L2 Bridging is a logical functionality enabled in the NSX domain.

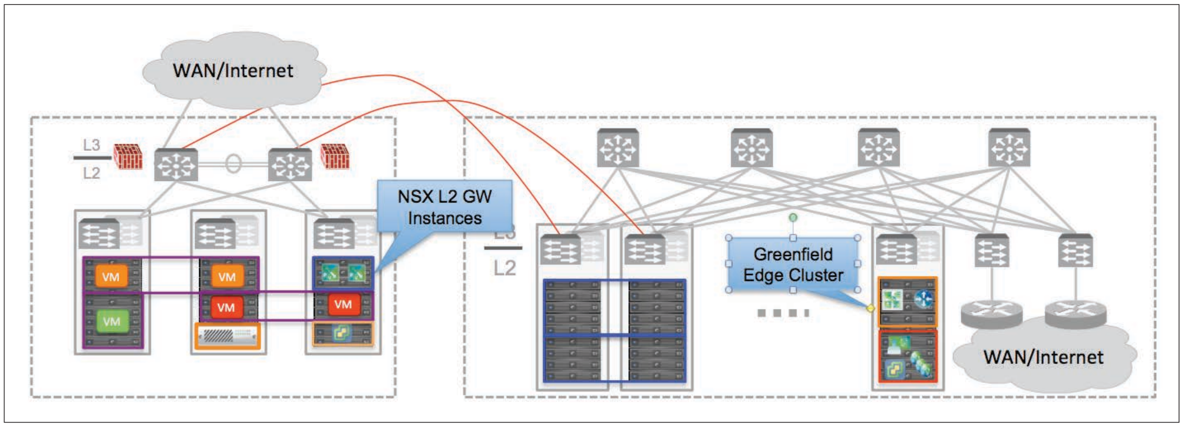


Figure 25: Deployment of NSX L2 Bridging Instances

Note: as previously mentioned, an alternative approach consists in connecting VXLAN capable switches (HW-VTEPs) to the brownfield network in order to perform the VXLAN-VLAN bridging function. Given the fact that full control-plane integration is not available until NSX release 6.2, this option is not covered in this initial version of the paper.

b. Configure the DFW rules that will be used to secure east-west communication once the VMs are migrated to the greenfield logical switches.

c. Configure the FW rules on the NSX Edge that will be used to secure north-south communication

d. Start migrating the VMs from the VLAN in the brownfield to the VXLAN in the greenfield. Typically this entails shutting down the VM in the brownfield side, copying the VM folder to the target datastore and importing it into the greenfield vCenter.

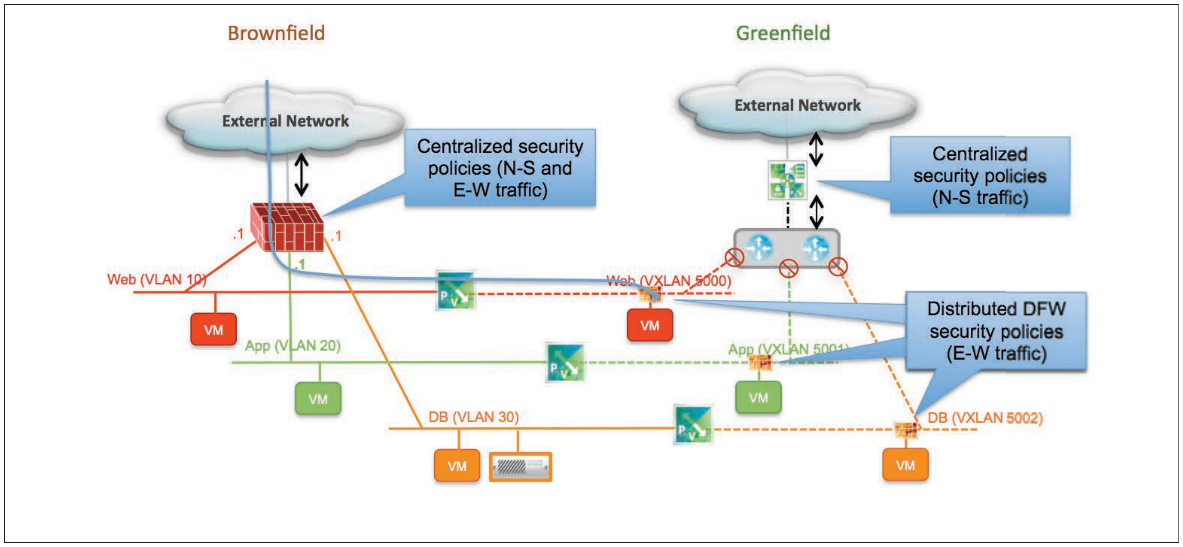


Figure 26: Migrating VMs between VLANs and VXLAN Segments

The physical firewall on the brownfield side still represents the default gateway for all the application tiers and keeps enforcing security policies for east-west and north-south communication. At the same time, the DFW starts filtering east-west traffic between each individual greenfield VM and everything else. As previously mentioned, this intermediate step allows validating that the DFW rules have been properly configured, without the risk of opening undesired security vulnerabilities.

Traffic flows between workloads not yet migrated and the greenfield networks are VXLAN encapsulated by the NSX L2 Bridging instances and sent via the L3 links connecting the brownfield and the greenfield networks.

e. Once the majority of VMs have been migrated to the greenfield network, it may make sense to move the default-gateway functionality for the different application tiers on the DLR. This is a disruptive step, as all the workloads (virtual and physical) must refresh their ARP cache with updated MAC information for the default-gateway.

At the same time it is also possible to take the physical FW out of the data path (for the specific IP subnets associated with the tiers of the application being migrated), so that north-south traffic starts also to flow locally in the greenfield network (Figure 27).

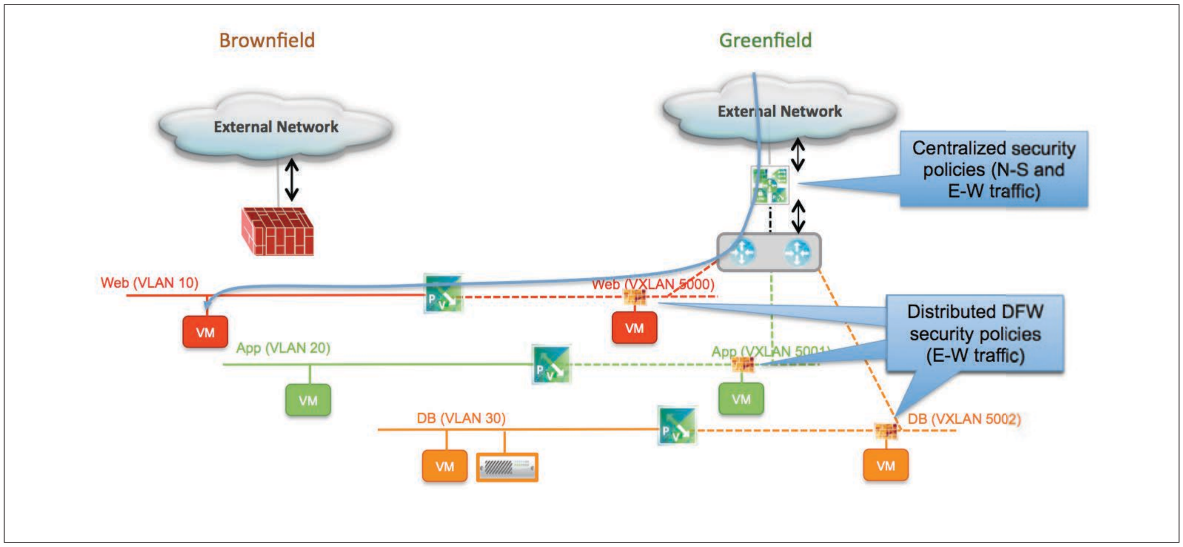
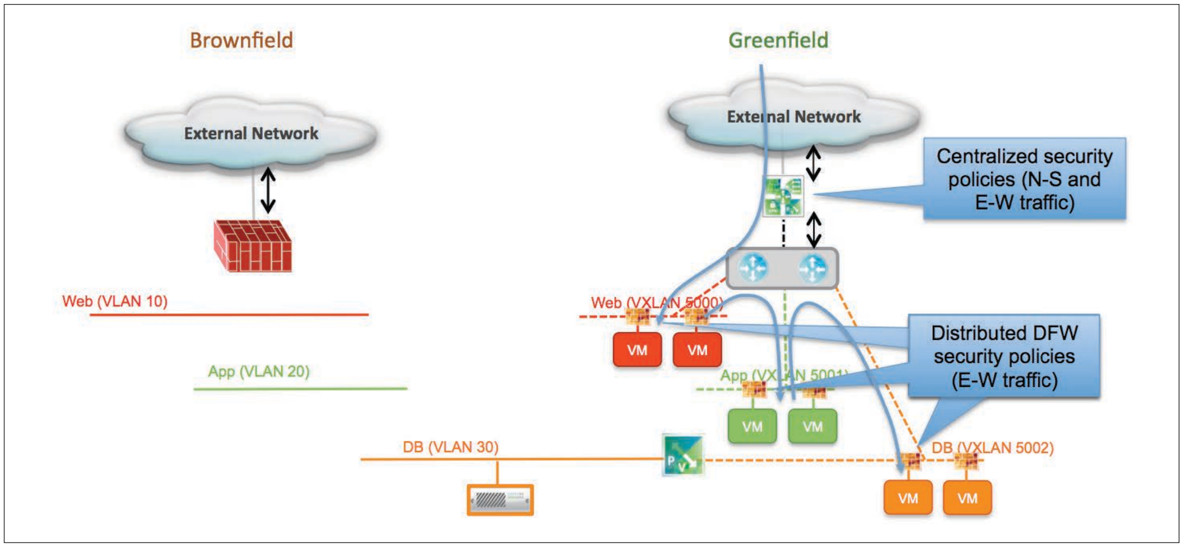


Figure 27: Enabling the Default Gateway Function in the Greenfield Network

f. Complete the onboarding process of the VMs to the greenfield logical switches. At this point it is also possible to decommission the NSX L2 Bridging instances that are not utilized anymore. The only bridging instances that remain active are the ones allowing connectivity to the bare-metal servers still connected to the brownfield network.



Centralized security policies (N-S and E-W traffic)

Distributed DFW security policies (E-W traffic)

Figure 28: Completion of the Migration Process

Conclusion

VMware NSX-v network virtualization solution addresses current challenges with physical network infrastructure and brings flexibility, agility and scale through VXLAN-based logical networks. Along with the ability to create on-demand logical networks using VXLAN, the NSX Edge Services Gateway helps users deploy various logical network services such as firewall, DHCP, NAT and load balancing on these networks. This is possible due to its ability to decouple the virtual network from the physical network and then reproduce the properties and services in the virtual environment.

In brownfield environments NSX can provide programmatic connectivity via L2 bridging or L3 routing and service insertion of existing physical services such as Load-Balancers and Firewalls, so such legacy services can be integrated and consumed by NSX Logical Networks.

As described in this document, in such brownfield environments NSX functionality can be incrementally introduced to meet different customer requirements and use-cases.

References

[1] VMware® NSX for vSphere (NSX-V) Network Virtualization Design Guide
https://communities.vmware.com/docs/DOC-27683

[2] VMware® vNetwork Distributed Switch: Migration and Configuration
http://www.vmware.com/files/pdf/vsphere-vnetwork-ds-migration-configuration-wp.pdf

[3] What's New in VMware vSphere 5.5
http://www.vmware.com/files/pdf/vsphere/VMware-vSphere-Platform-Whats-New.pdf

[4] vSphere 5.5 Configuration Maximums
http://www.vmware.com/pdf/vsphere5/r55/vsphere-55-configuration-maximums.pdf