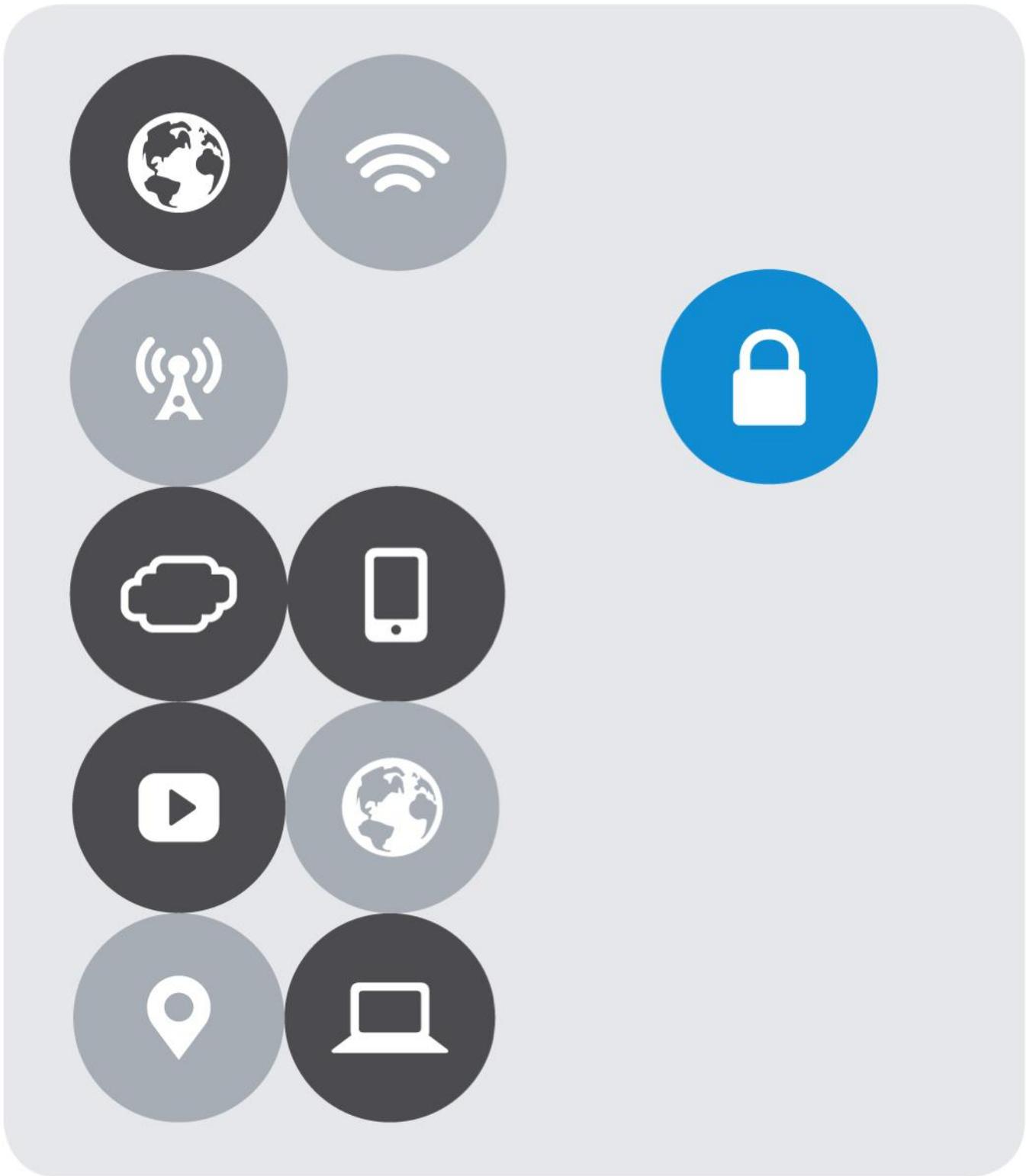




DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP



Version History

Date	Version	Author	Description
December 2019	1.0	Ulises Alonso Camaró, F5 Networks	Initial public version.
June 2020	2.0	Ulises Alonso Camaró F5 Networks	<p>Validated with NSX-T 3.0.</p> <p>Updated all screenshots and configuration flows to match NSX-T 3.0.</p> <p>Changed network addressing to allow a lab with all topologies simultaneously.</p> <p>Changed Logical Router/LR nomenclature for Gateway following new NSX-T convention.</p> <p>Modified topologies B and D so they are now more generic and can take advantage that a single BIG-IP Virtual Server can listen to multiple IPs. This also avoids L3 routing hops in Topology D.</p> <p>Extended information on BIG-IP version compatibility with a dedicated section.</p>
September 2020	2.1	Ulises Alonso Camaró F5 Networks	<p>Extended topology suitability matrix based on flow's direction with the inter-tenant E-W flows case.</p> <p>Added MAC masquerading information.</p> <p>Added VMC on AWS section.</p> <p>Added section on Hybrid and Multi-Cloud design considerations.</p> <p>Renamed from "Integration guide" to "deployment guide"</p>

INTRODUCTION	6
NSX-T VERSIONS CONSIDERED IN THIS GUIDE	7
BIG-IP VERSIONS CONSIDERED IN THIS GUIDE	7
DEPLOYMENT GUIDE OVERVIEW	7
INLINE TOPOLOGIES OVERVIEW	8
PARALLEL TOPOLOGIES OVERVIEW	9
TOPOLOGIES' MAIN CHARACTERISTICS SUMMARY	10
ADDITIONAL DEPLOYMENT POSSIBILITIES	12
NSX-T DESIGN CONSIDERATIONS	13
Design consideration: Layer 2 networking.....	13
Design consideration: NAT.....	13
Design consideration: Use of dynamic routing (BGP) with upstream networks.....	13
Design considerations: NSX-T's distributed functions.....	14
Design consideration: Centralized management.....	14
TOPOLOGY A: BIG-IPS INLINE-CONNECTED TO NSX-T'S TIER-0 GATEWAY	15
Implementation: Active/Standby NSX-T Edge with static routing.....	17
Implementation: Active/Standby NSX-T Edge using BGP.....	25
Implementation: Active/Active NSX-T Edge using BGP ECMP.....	37
TOPOLOGY B: BIG-IPS INLINE – CONNECTED LIKE AN NSX-T'S TIER-1 GATEWAY	49
Implementation: BIG-IPs inline-connected like an NSX-T's Tier-1 Gateway.....	50
TOPOLOGY C: BIG-IPS PARALLEL-CONNECTED TO NSX-T'S TIER-0 GATEWAY	59
Implementation: BIG-IPs parallel-connected to NSX-T's Tier-0 Gateway.....	61
TOPOLOGY D: BIG-IPS PARALLEL-CONNECTED TO NSX-T'S TIER-1 GATEWAY	71
Implementation: BIG-IPs parallel-connected to NSX-T's Tier-1 Gateway.....	76

VMWARE CLOUD ON AWS	81
Introduction	81
Sample topology D for VMC on AWS – VMC configuration.....	82
Sample topology D for VMC on AWS – BIG-IP configuration.....	85
Alternative topologies for BIG-IP in VMC on AWS.....	87
HYBRID AND MULTI-CLOUD DESIGN CONSIDERATIONS	89
Introduction and Use Cases	89
Overall approach	89
SaaS Security and multi-cloud	90
Generic Public Cloud and VMC on AWS connectivity options	90
VMware HCX – Hybrid Cloud Extension	91
Design Guidelines – VMC on AWS with local VPC workloads.....	92
Design Guidelines – multi-cloud	93
Cloud Bursting with multi-cloud	95
Design Guidelines – single site with cloud bursting.....	96
GENERAL NOTES	98
BGP configuration details.....	98
Best practices for BIG-IP in VMware NSX-T	99
MAC Masquerade in NSX-T	102
VMC on AWS	103
Considerations for container platforms.....	103
General guidelines	103
Exposing container services.....	103
Relevant configuration settings when using Red Hat OpenShift.....	105
Relevant configuration settings when using Pivotal PKS.....	105
VERIFYING THE DEPLOYMENT	107
Basic testing	107
Dynamic routing testing.....	108
End to End testing: test egress forwarding connectivity through the BIG-IP.....	110
End to End testing: test egress forwarding connectivity without the BIG-IP.....	112
End to End testing: test Ingress connectivity through the BIG-IP.	112

Introduction

The Software-Defined Data Center (SDDC) is characterized by server virtualization, storage virtualization, and network virtualization. Server virtualization has already proved the value of SDDC architectures in reducing costs and complexity of the compute infrastructure. VMware NSX network virtualization provides the third critical pillar of the SDDC. It extends the same benefits to the data center network to accelerate network service provisioning, simplify network operations, and improve network economics.

This guide provides configuration guidance and best practices for the topologies in most common scenarios ensuring compatibility and minimal disruption to the existing environments. Unlike with NSX-V, F5 BIG-IP does not participate in the control plane of the overlay networking. This is due to NSX-T's lack of a publicly documented API. The integration is based on routing within the overlay networks. This has the following implications:

- For North-South traffic flows this is not an issue because the number of networks to which the F5 BIG-IP has to be connected is small and is not expected to change often.
- For East-West traffic this inhibits the possibility of using F5 BIG-IP hardware. Also, the number of network segments to which the F5 BIG-IP is expected to be connected for this use case is very high, but the VMware hypervisor only allows the VMs to be connected with up to 10 vNICs¹ with one network segment per vNIC. In this guide this VMware limitation is overcome by creating multiple clusters of BIG-IPs. This allows higher distribution of the traffic and CPU utilization across the VMware cluster.

Using F5 BIG-IP ADC instead of NSX-T's load balancer provides the following benefits:

- F5 BIG-IPs can be connected to either Tier-0 (internally or externally) and to Tier-1 distributed routers while NSX-T's load balancer can only be placed logically connected to Tier-1 Gateways.
- NSX-T's load balancer is not a distributed function and runs centralized on NSX-T Edge's nodes, which can represent a bottleneck. F5 BIG-IP can run in multiple hypervisors concurrently by either running Active-Active F5 Scale-N clusters or multiple F5 BIG-IP clusters.
- F5 BIG-IP provides proven, scalable, and world-class performance for ADC, NAT and Firewall capabilities, and provides additional functionalities such as Advanced WAF, SSL-VPN, Anti-DDoS protection, Secure Web Gateway with Identity Management and many other solutions with a unified management & visibility with F5 BIG-IQ.

¹ For checking vSphere's limits consult the link <https://configmax.vmware.com/quest?vmwareproduct=vSphere&release=vSphere%206.7&categories=1-0> and search "Networking Virtual Devices" or "Virtual NICs per virtual machine".

NSX-T versions considered in this guide

This guide considers NSX-T versions 2.4-3.0 but given that the F5 BIG-IP integration is transparent from NSX-T point of view² this documentation should apply to upcoming NSX-T releases as well.

BIG-IP versions considered in this guide

Any BIG-IP Virtual Edition version is supported as long as the hypervisor is supported. Please check the page [BIG-IP VE Supported Platforms](#) in clouddocs.f5.com for the most up to date information. When using BIG-IP Hardware platforms any BIG-IP version is supported.

Additionally, when using BIG-IP (either Hardware or Virtual Edition) north of the NSX-T Edge nodes this arrangement typically uses BGP (specially for Active-Active deployments) in which case BIG-IP will require the Advanced Routing module to be provisioned. See [K46129932: How to verify Advance Routing Module is provisioned](#) for more details.

Deployment guide overview

The document includes the 4 most common topologies:

- **Inline topologies:**

- Topology A: BIG-IPs inline-connected to NSX-T's Tier-0 Gateway.
- Topology B: BIG-IPs inline-connected like NSX-T's Tier-1 Gateways.

- **Parallel topologies (these require SNAT):**

- Topology C: BIG-IPs parallel-connected to NSX-T's Tier-0 Gateway.
- Topology D: BIG-IPs parallel-connected to NSX-T's Tier-1 Gateway.

There is a section with implementation details for each topology, and for Topology A there are three implementation options. This is followed by a section containing details common to all topologies and best practices when deploying F5 in VMware. Then, a section for configuring and testing a service with F5 BIG-IP. Finally, there is a section with considerations for container platforms, Red Hat OpenShift and other Kubernetes based options.

² To be precise, in some topologies BIG-IP is connected to NSX-T Edge using eBGP but BGP is an Internet standard, not NSX-T specific.

Inline topologies overview

A main characteristic of inline topologies is they do not require the use of SNAT (Secure Network Address Translation), keeping the client IP address unchanged. Another benefit is that traffic flows are easier to understand and troubleshoot.

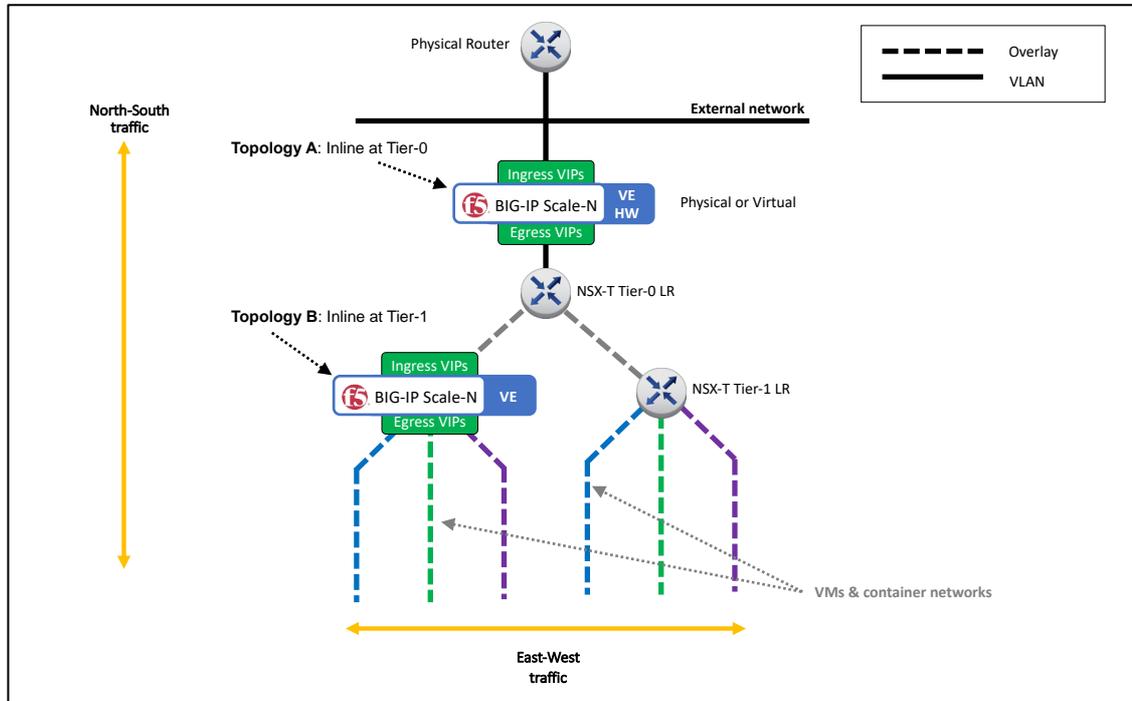


Figure 1 - BIG-IPs in inline-connected topologies A and B.

- Topology A – BIG-IPs inline-connected to NSX-T’s Tier-0 Gateway.

This topology allows the use of either BIG-IP hardware or Virtual Editions (VE). In this topology the F5 BIG-IP is placed in a special vantage point for all tenants where security-related services can be enforced easily (for example WAF, Firewall and anti-DDoS) and also NAT if needed.

For this topology three possible configurations are described:

- NSX-T Edge cluster in Active-Standby mode using a static routing.
- NSX-T Edge cluster in Active-Standby mode using a dynamic routing with BGP.
- NSX-T Edge cluster in Active-Active mode using dynamic routing with BGP ECMP.

- Topology B – BIG-IPs inline-connected like an NSX-T’s Tier-1 Gateway.

This topology is similar to Topology A but allows per-Tenant BIG-IP clusters, hence providing isolation between tenants. In this topology it is proposed eliminating NSX-T’s Tier-1 Gateways to keep a 2-tier routing model while keeping BIG-IPs inline to the traffic path (there is more information in the Topology B section). This topology only uses BIG-IP Virtual Editions.

Parallel topologies overview

In these topologies, the paths for plain forwarding traffic and the traffic handled by BIG-IP services are different:

- The BIG-IPs are not inline for plain forwarding traffic and hence this traffic doesn't need SNAT.
- For BIG-IP services, the traffic goes through the BIG-IPs through a parallel path and SNAT is required in order to keep traffic symmetric. See the Design considerations section for more information when using NAT.

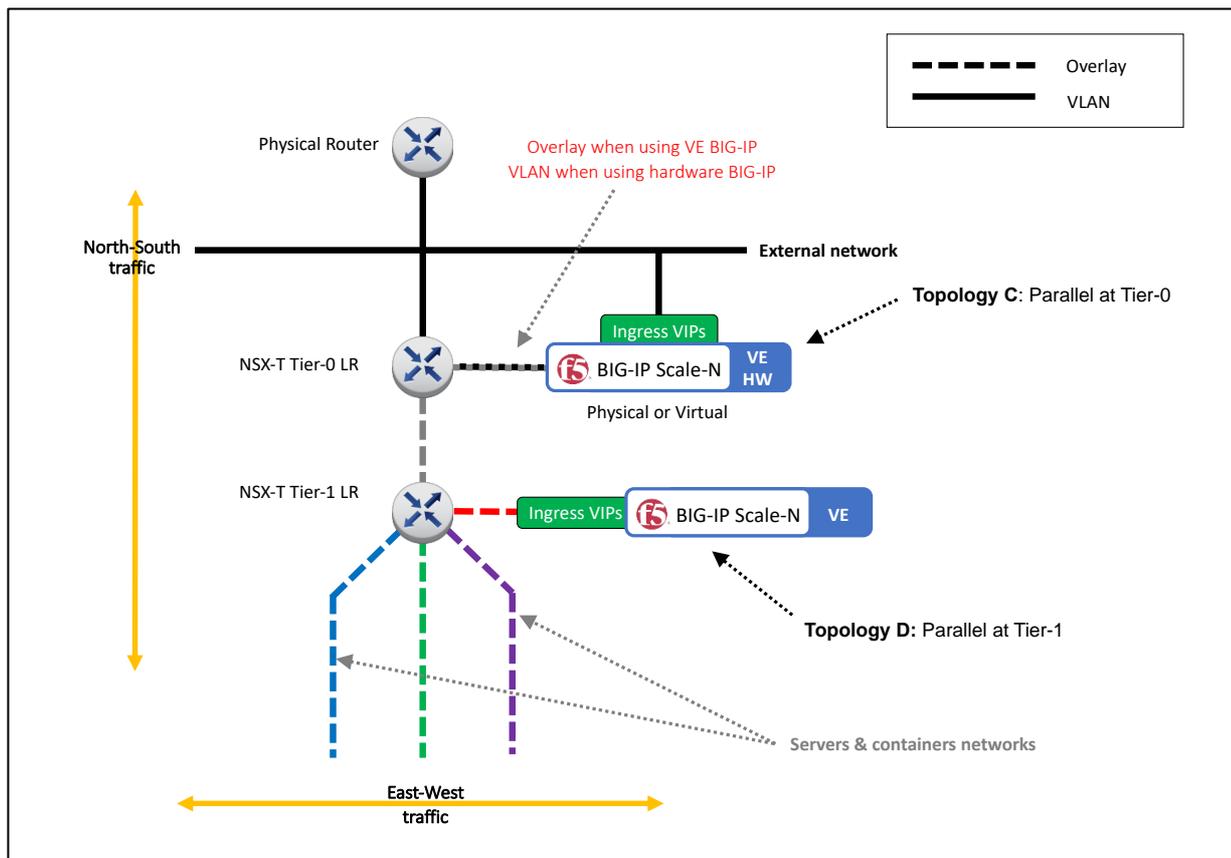


Figure 2 - BIG-IPs in parallel-connected topologies C and D.

- Topology C – BIG-IPs parallel-connected to NSX-T's Tier-0 Gateway.

Like Topology A which is also connected to a Tier-0 Gateway, this topology allows the use of either BIG-IP hardware or Virtual Editions. Other than the requirement of using SNAT, the main difference from Topology A is that each tenant can have their own BIG-IPs instances with complete isolation. This can be achieved either using BIG-IP hardware instantiating vCMP guests or using F5 BIG-IP Virtual Edition instances for each tenant.

- Topology D – BIG-IPs parallel-connected to NSX-T's Tier-1 Gateway.

This topology is similar to Topology C but with the BIG-IPs attached to the Tier-1 routers and would allow that Edge services could be applied at the NSX-T boundary for all traffic

flows without any traffic bypassing these Edge services. This is equivalent to the topology used by NSX-T Load Balancers.

Although this topology can be used for both North-South and East-West services traffic, it can be useful combining Topology D for East-West traffic with Topology A for North-South traffic. This combined A & D Topology is especially useful when high performance is required, and NSX-T Edges operate in Active-Active mode with ECMP. In this case, the F5 BIG-IP has to take over NSX-T Edge's stateful functions. The BIG-IP can also perform additional single-point control functionalities such as WAF, anti-DDoS, or SSL-VPN, which are not available in NSX-T Edge.

Topologies' main characteristics summary

The next tables show a summary of the characteristics of each topology. A description of the characteristics is at the bottom each table. Some of the characteristics are direct consequence of the topology type and tier placement, this is the case of being able to keep the client address or being an enforcement point where all flows go through the BIG-IP.

Note that both topologies that are applied to Tier-0 allow multi-tenancy with either software partitions or virtualization partitions (vCMP).

Topology	Type	Tier	VE	HW	Keeps client address	Inter-tenant distributed forwarding path	Enforcement point	Allows per-tenant VE clusters
A	Inline	Tier-0	✓	✓	✓	Not applicable	✓ (for all tenants)	
B	Inline	Tier-1	✓		✓		✓ (per tenant)	✓
C	Parallel	Tier-0	✓	✓		Not applicable		✓
D	Parallel	Tier-1	✓			✓		✓

Topology: the name of the topology used in this guide.

Type: If all the traffic goes through the BIG-IPs (Inline) or not (Parallel). When a topology is inline implies that the BIG-IPs are able to be an enforcement point for all traffic and it is guaranteed no traffic will by-pass BIG-IP's topologies.

Tier: If the BIG-IPs are attached to a Tier-0 or Tier-1 NSX-T Gateway. In the case of Topology C the proposed topology actually replaces NSX-T's Tier-1 Gateway. See topology's section for more details.

VE: the topology allows the use BIG-IP Virtual Edition.

HW: the topology allows for hardware appliances or chassis. Hardware platforms with vCMP technology is recommended. This allows hard resource isolation between tenants.

Keeps source address: Ingress traffic doesn't need to translate the source IP address of the clients. This avoids the need of using the X-Forwarded-For HTTP header.

Inter-tenant distributed forwarding path: when using plain routing between tenant workloads the processing path is fully distributed by only using NSX-T's networking. In other words, this scenario is a path between Tier-1 workload to another Tier-1 workload and not using BIG-IP services. Note that when using NSX-T's native LB the processing is done centralized in the NSX-T Edge nodes.

Enforcement point: this is characteristic of being an Inline topology type as described above.

Allows per-tenant VE clusters: the topology allows creating separate BIG-IP VE clusters for each tenant where these do not interfere with each other.

Topology	Suitable for North-South	Suitable for intra-tenant East-West	Suitable for inter-tenant East-West
A	✓	NA	✓ (If VIPs are not in tenant's segments)
B	✓	✓ (BIG-IP is a tenant enforcement point)	✓ (BIG-IP is a tenant enforcement point)
C	✓ (for non-routing traffic)	NA	✓ (If VIPs are not in tenant's segments)
D	✓ (for non-routing traffic)	✓ (for non-routing traffic)	✓ (for non-routing traffic)

Suitable for North-South: North-South flows is traffic that goes in and out of the NSX-T deployment. In the case of topologies C and D the routed traffic doesn't get any BIG-IP service applied.

Suitable for intra-tenant East-West: traffic that doesn't use a Tier-0 Gateway. BIG-IP at Tier-0 (topologies A and C) don't affect East-West traffic flows. Topology B or D should be chosen depending on if it is required that the BIG-IP be a tenant enforcement point. Although Topology D doesn't allow the BIG-IP to be an enforcement point it allows distributed L3 forwarding by using only Tier-1 gateways for these flows.

Suitable for inter-tenant East-West: traffic that uses Tier-0 Gateway. When routed these flows typically take advantage of distributed processing and traffic goes directly from VM to VM. BIG-IP at Tier-0 can deal with these flows if the VIPs are not in tenant's segments. Note that when using BIG-IP for these flows it doesn't incur in more node hops than native NSX-T LB because the native NSX-T LB is implemented in the Edge nodes and also represent a node hop. For topologies B and D it is the same situation as for intra-tenant East West flows.

Additional deployment possibilities

There are many other topology possibilities; the following examples have specific use cases:

- BIG-IP Service scaling group (SSG) for CPU-intensive workloads such as Advanced WAF in large scale deployments.
- Per-App VE which provides DevOps teams with an ADC and a WAF to deliver services and security just for the application they are developing.

For more information on these please consult [BIG-IP Cloud Edition Solution Guide](#).

NSX-T design considerations

Design consideration: Layer 2 networking

This guide doesn't suggest any specific Layer 2 configuration. The Layer 2 configuration depends on the overall vCenter and more predominantly the NSX-T configuration. Because of this, the configuration examples in this guide start at Layer 3. It is a pre-requisite of the examples to have Layer 2 previously configured.

In general, it is recommended to use redundancy at all Network Layers. In the case of Layer 2 networking this is typically achieved by using LACP³ which is supported in the ESXi/vSphere hypervisor and in the NSX-T Transport and Edge nodes. In the case of BIG-IP hardware platforms LACP is supported. The VMs in ESXi/vSphere do not receive the LACP frames from the hypervisor hence the network appliances such as BIG-IP VE cannot implement LACP and this must be configured instead at the hypervisor level. In other words, LACP should be configured in the NSX-T transport nodes or ESXi/vSphere and this will be transparent to the BIG-IP VE.

Design consideration: NAT

When using BIG-IP for North-South traffic workloads (VM or containers) it is important that the F5 BIG-IP has direct visibility of the IP addresses of these VMs or containers, otherwise health-checking probes do not have visibility of the actual service, especially when 1:1 NAT mapping is not applied.

If NAT is required, it can be performed by the F5 BIG-IPs, which has the added value of offloading this functionality from NSX-T Edge. This in turn allows NSX-T Edge nodes to run in Active-Active mode with ECMP without restrictions - NAT in Tier-0 can only run in Active-Active when using Reflexive (stateless) mode⁴.

In many instances, services need to be aware of the client's IP address. In these cases, and when the F5 BIG-IP performs NAT, the client IP address can be added in the HTTP payload using the `X-Forwarded-For` header for unencrypted and encrypted traffic by performing SSL/TLS termination in the F5 BIG-IP. This capability of always being able to insert the `X-Forwarded-For` header is an important reason for choosing F5 BIG-IP for NAT functionality.

Design consideration: Use of dynamic routing (BGP) with upstream networks

³ LACP - Link Aggregation Control Protocol is an IEEE standard.

⁴ Reflexive NAT - <https://docs.vmware.com/en/VMware-NSX-T-Data-Center/2.5/administration/GUID-46900DFB-58EE-4E84-9873-357D91EFC854.html>

NSX-T Edge's Tier-0 routers exchange routes with upstream devices by means of eBGP. It is recommended the use of dynamic routing in the following use cases:

- When using NSX-T Edge in Active-Active mode.
- When the NSX-T deployment doesn't have a contiguous address space with a single prefix.
- When IP addresses can migrate to other deployments.
- When NSX-T Edges are connected using several subnets to the upstream networks.

Design considerations: NSX-T's distributed functions

NSX-T provides distributed processing for switching, routing, firewall, and NAT depending on the configuration. F5 Networks recommends taking advantage of NSX-T's distributed processing whenever possible. Other features and configurations such as stateful NAT, LB, Edge Firewall are not compatible with distributed processing or Active-Active Tier-0 routers. When these functions cannot be run in a distributed manner, F5 recommends running these in F5 BIG-IP.

Design consideration: Centralized management

It is recommended to consider BIG-IQ which provides the following functionality:

- Centralized Management including self-service app-owner portal, application templates with security policies.
- Per-app analytics, performance metrics and dashboards.
- Multi-cloud capable and enabler for centralized CI/CD integrations.
- Fine grained RBAC where demarcation between the network, security, and app teams can be well established with their own specific views of a deployment.

Topology A: BIG-IPs inline-connected to NSX-T's Tier-0 Gateway.

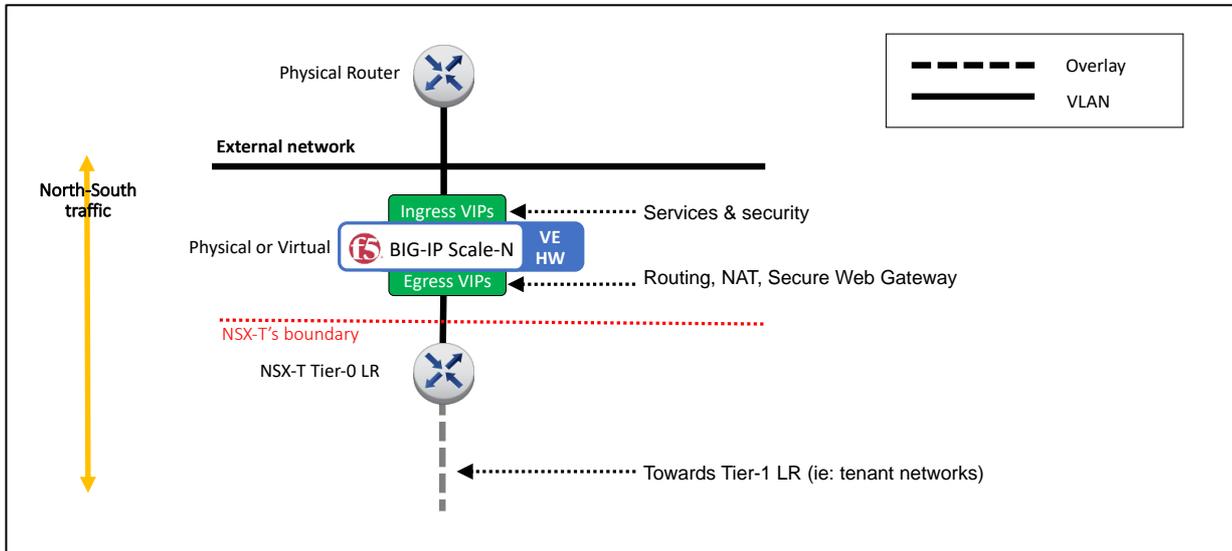


Figure 3 – Topology A overview (simplified view without HA components).

The main feature of this topology is that the F5 BIG-IP can easily be an enforcement point for North-South traffic. In this scenario, F5 can be either deployed as hardware or as a Virtual Edition. When using a Virtual Edition, multi-tenancy can be achieved by using separate logical partitions. When using BIG-IP hardware, multi-tenancy can also be achieved with full isolation by using vCMP.

When NSX-T Edge is running in Active-Active mode with ECMP, it is not able to run stateful services (ie: edge firewall, load balancing, or NAT with the exception of Reflexive NAT). In this high-performance use case, this functionality can be off-loaded to the F5 BIG-IP (hardware platforms are recommended, using chassis for ultimate scalability without reconfiguration).

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

When using this logical topology there are two alternatives for the physical topology. These can be seen in the next figure.

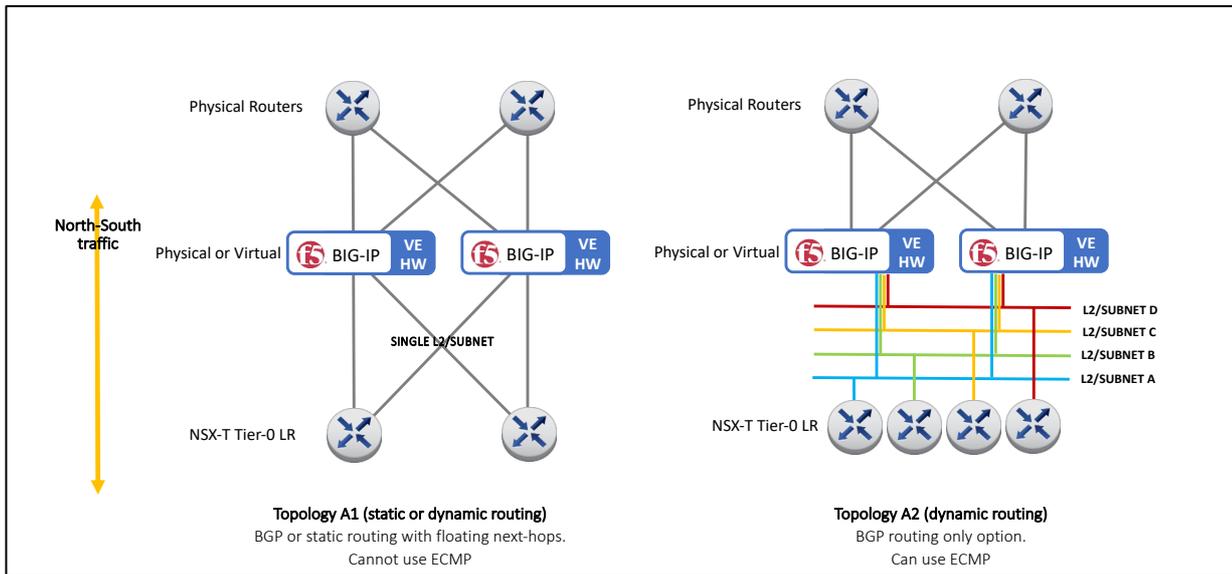


Figure 4 - L1/L2 options for Topology A.

Implementation: Active/Standby NSX-T Edge with static routing

The next figure shows the configuration which will be implemented in this section.

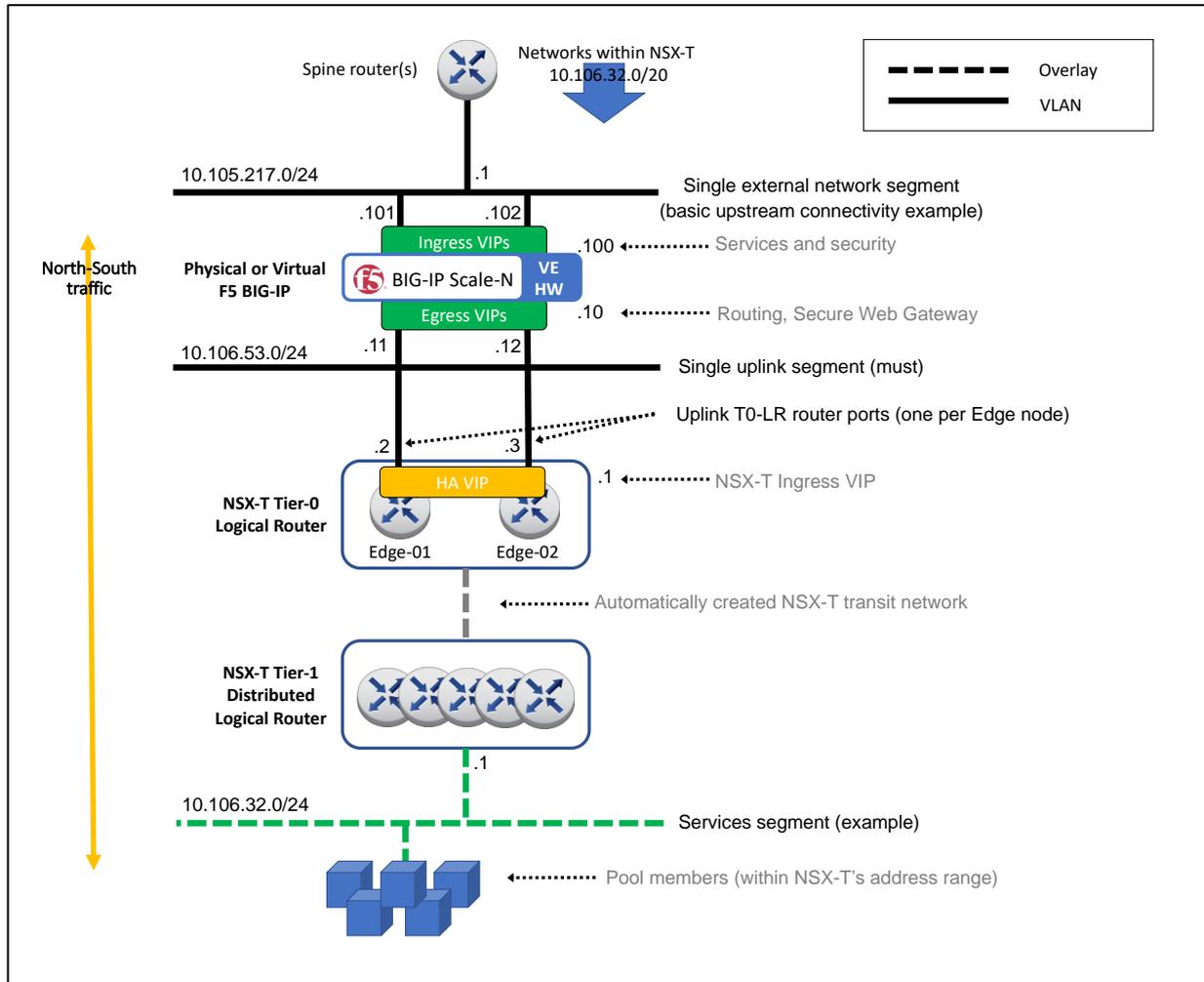


Figure 5 – Example of topology A using static routing used through this section.

Given the many possibilities of configuring NSX-T Edge nodes and their logical switch uplink ports, it is assumed that these have been already created. This guide is focused in the

configuration for the Layer 3 and higher layers that are specific to this topology. See section Design consideration: Layer 2 networking for details.

1. Create the Tier-0 configuration.

1.1. Create a Tier-0 Gateway in Active-Standby mode.

In NSX-T manager, go to `Networking > Tier-0 Gateways > Add Gateway > Tier-0` as shown in the next figure.

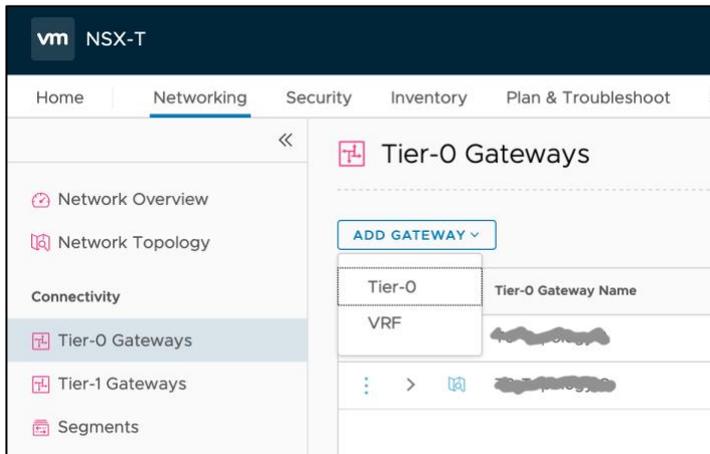


Figure 6 - Adding a Tier-0 Gateway/Gateway.

In the New Tier-0 Router dialog, complete the following:

- Name: T0-topology A in this example.
- Edge Cluster: Select the existing Edge cluster.
- High Availability Mode: Active-Standby.
- Failover Mode: Non-Preemptive (to avoid double failover once the failed unit recovers).

The screenshot shows the 'Tier-0 Gateways' configuration interface. At the top left, there is a red icon and the title 'Tier-0 Gateways'. Below the title is a blue button labeled 'ADD GATEWAY'. The main configuration area is a light blue panel with a table-like structure. The table has three columns: 'Tier-0 Gateway Name', 'HA Mode', and 'Linked Tier-1 Gateways'. The first row contains the following values: 'T0-Topology A' (with a red asterisk), 'Active Standby' (with a red asterisk), and an empty field. Below the table, there are several settings: 'Fail Over' set to 'Non Preemptive', 'Edge Cluster' set to 'nsx-edge-cluster-topology-a', and 'Tags' with two dropdowns: 'Tag (Required)' and 'Scope (Optional)'. Below the tags, there is a note: 'Max 30 allowed. Click (+) to save.' At the bottom of the configuration area, there are three buttons: 'SAVE', 'CANCEL', and 'Unsaved Changes'. Below the configuration area, there are several expandable sections: 'INTERFACES', 'ROUTING', 'MULTICAST', 'BGP', and 'ROUTE RE-DISTRIBUTION'.

Figure 7 - Filling the details of a Tier-0 Gateway/Gateway.

1.2. Create an Interface for each Edge Node used by the Tier-0 Gateway/Gateway.

Select the router created (T0-Topology-A in our example) and create two interfaces in the UI by first selecting the Edit option in the T0 Gateway, then scrolling down to the

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

Interfaces section clicking in the `Set` option of `External` and `Service` Interfaces. Enter the following parameters for each interface:

- **Name:** In this example, `edge-1-uplink-red` is used for the first router port and `edge-2-uplink-red` for the second (we will use `edge-*-uplink-blue` in the BGP+ECMP scenarios).
- **Type:** `External`
- **Edge Node:** This will be `edge-1-topology-a` and `edge-2-topology-a` for each external interface respectively.
- **MTU:** use external network's MTU, which should be the same on the BIG-IP.
- **URPF Mode:** `Strict` is a good practice providing security with no expected performance impact. `Strict` should be used unless asymmetric paths are used.
- **Segment:** This is the L2 network to which the interface is attached to. It is a prerequisite to have this previously created. See section `Design consideration: Layer 2 networking` for details.
- **IP Address/mask:** this is the IP address assigned to the address port in the shared segment between the NSX-T Edge nodes and the F5 BIG-IPs. In this example, `10.106.53.1/24` is used for router port in `edge-01` and `10.106.53.2/24` in `edge-02`.
- Click `Add`.

Name	Type	IP Address / Mask	Connected To(Segment)	Status
edge-1-uplink-red	External	10.106.53.1/24	vlan-353	

Edge Node: edge-1-topology-a

MTU: Enter MTU (Minimum 64)

PIM: Disabled

ND Profile: default

Tags: Tag (Rec), Scope (C)

URPF Mode: Strict

Figure 8 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.

Name	Type	IP Address / Mask	Connected To(Segment)	Status
edge-1-uplink-red	External	10.106.53.1/24	vlan-353	Success
edge-2-uplink-red	External	10.106.53.2/24	vlan-353	Success

Figure 9 – Final Gateway Port configuration of the Tier-0 Gateway.

1.3. Create an HA VIP for the Tier-0 Gateway.

The HA VIP is an IP address that will be shared by the two Edge Nodes used for the Tier-0 Gateway just created and it will be used as the ingress IP to the NSX-T networks.

Select the Gateway just created (T0-Topology A in our example), and create an HA VIP in the UI by selecting `Edit > HA VIP Configuration > Set` and entering the following parameters:

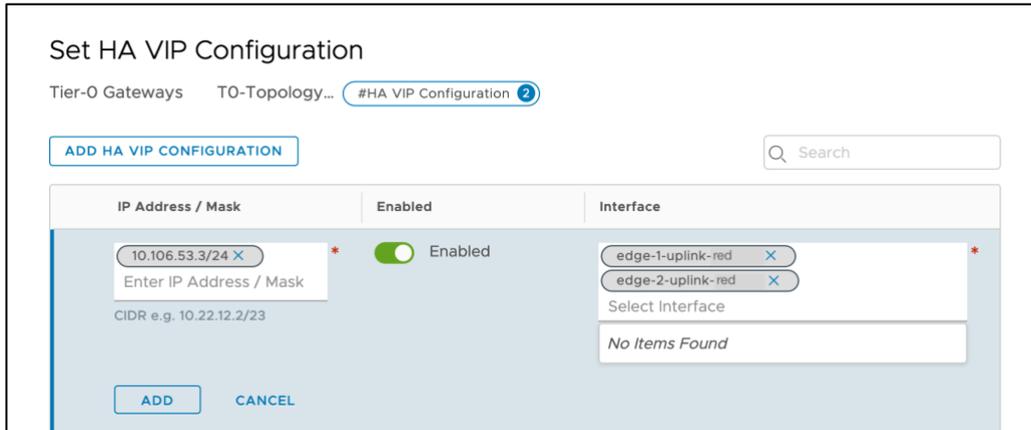


Figure 10 - Adding an HA VIP to NSX-T's T0 Gateway.

Selecting the two external interfaces just created.

1.4. Add a default route in the Tier-0 Gateway towards the BIG-IP cluster floating Self IP address.

In our example, the BIG-IP cluster floating address to use as the next hop is 10.106.53.10. Select the T0-Topology A Gateway created and then create a static routing in the UI by selecting `Routing > Static Routes > Set` as follows and entering as Next Hop BIG-IP's floating-IP, in this example 10.106.53.10:

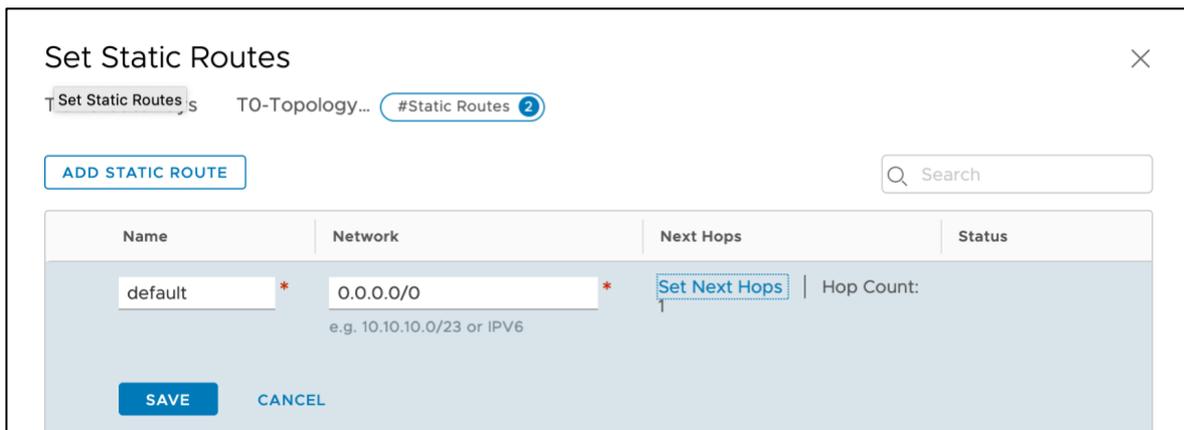


Figure 11 – Adding Tier-0 Gateway's default route.

2. Create a Tier-1 Gateway.

This will be used later to instantiate a VM and perform a verification of the deployment.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

In NSX-T manager, select `Networking > Tier-1 Gateways > Add Tier-1 Gateway > Tier-1 Router` filling the following parameters:

- Name: In this example, `T1-Topology A`.
- Tier-0 Router: Select the Tier-0 router (`T0-Topology A` in our example).
- Edge Cluster: The name of the Edge Cluster of the NSX-T Edge nodes being used.
- Failover Mode: `Non-Preemptive` (to avoid double failover once the failed unit recovers).
- Route Advertisement: at least “All Connected Segments [...]” should be enabled.
- Click `Add`.

Figure 12 – Filling the properties when creating a Tier-1 Gateway.

The next step is to create a network attached to this Tier-1 Gateway. In the UI, select `Networking > Segments > Add Segment` and enter the following parameters:

- Segment Name: in this example, `segment-332`.
- Connectivity: the Tier-1 Gateway, in this case `T1-Topology A`.
- Subnets: this really indicates both the subnet and the IP address of the Tier-1 Gateway in this segment, in this case `10.106.32.1/24`

This configuration can be seen in the next figure:

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

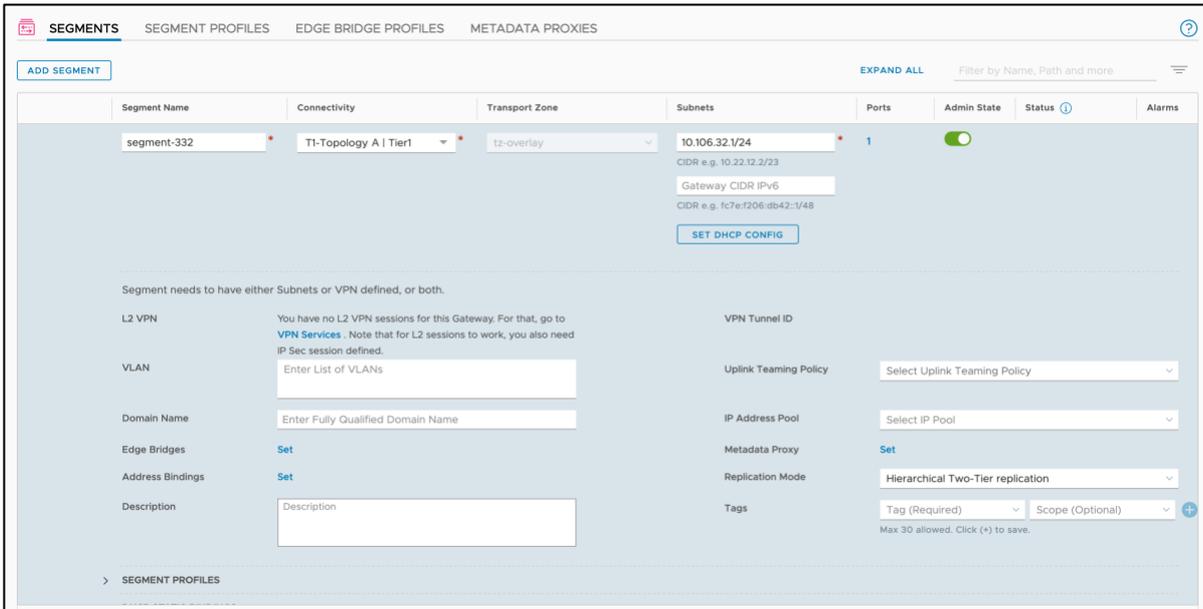


Figure 13 - Adding a segment to the T1 Gateway.

3. Create the Layer 3 configuration in the BIG-IP.

First, create the Self IPs and floating Self IPs towards the spine routers (north-bound) and towards the NSX-T Tier-0 Gateway (south-bound). These do not require any special configuration. An example of the first BIG-IP unit is shown next.

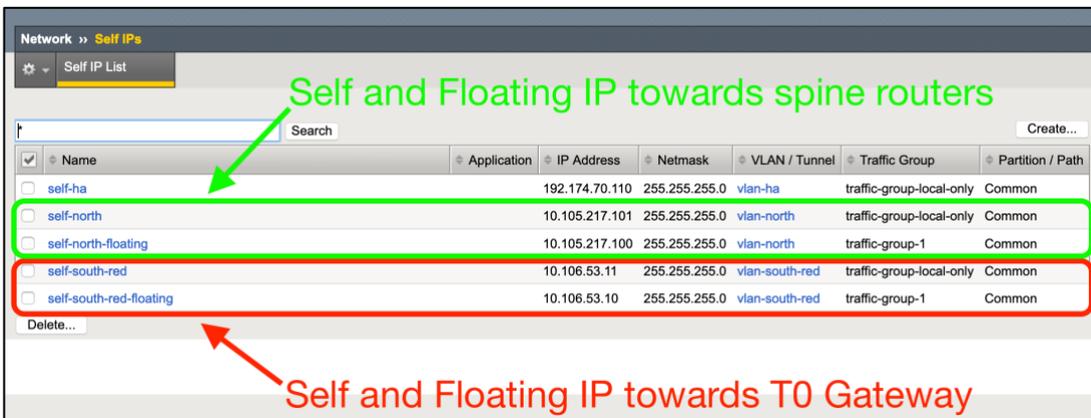


Figure 14 – Self IPs and floating Self IPs required (shown in BIG-IP unit 1).

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

Note: the non-floating Self IPs are per BIG-IP unit, while the floating Self IPs are synchronized across the BIG-IP units.

The next step is to configure the static routing in the BIG-IP. Typically, these involve two routes:

- A default route using spine router as gateway.
- A route towards the NSX-T IP address range using NSX-T's Tier-0 HA VIP as gateway.

These routes can be shown in the next figure and should be configured in both BIG-IP units (this configuration is not synchronized automatically across BIG-IPs).

<input checked="" type="checkbox"/>	Name	Application	Destination	Netmask	Route Domain	Resource Type	Resource	Partition / Path
<input type="checkbox"/>	default		Default IPv4		Partition Default Route Domain	Gateway	10.105.217.1	Common
<input type="checkbox"/>	nsxt		10.106.32.0	255.255.240.0	Partition Default Route Domain	Gateway	10.106.53.3	Common

Figure 15 – Static routes required in the BIG-IP units.

At this point, follow the testing steps described in the Verifying the deployment section.

Implementation: Active/Standby NSX-T Edge using BGP

The next figure shows the configuration implemented in this section. This topology differs from the previous Topology A implementation, which used static routing, in the next-hops used by the BIG-IP and the Tier-0 Gateways.

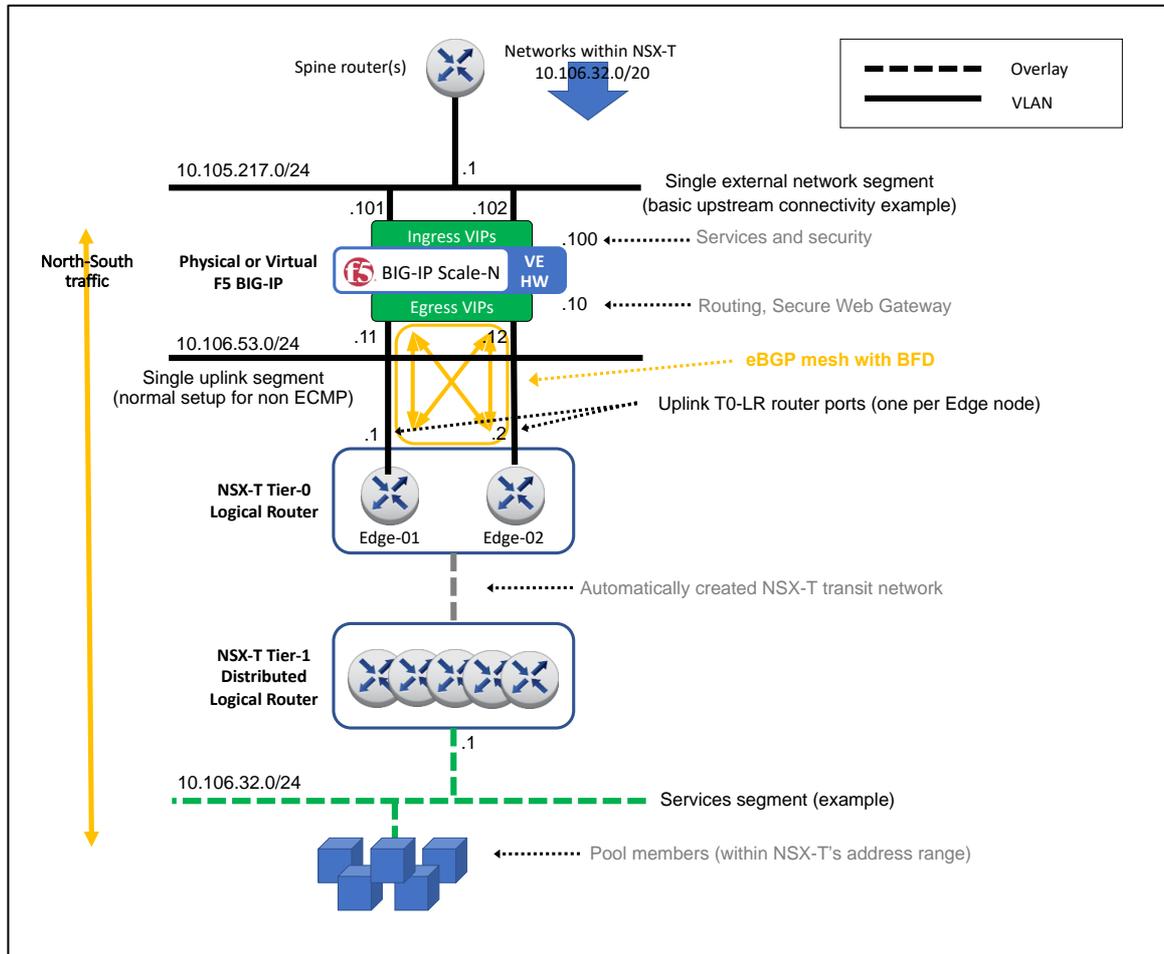


Figure 16 – Example of topology A using BGP routing used through this section

Given the many possibilities of configuring NSX-T Edge nodes and their logical switch uplink ports, it is assumed that these have been already created. This guide is focused in the

configuration for the Layer 3 and higher layers that are specific to this topology. See section Design consideration: Layer 2 networking for details.

1. Create the Tier-0 configuration.

1.1. Create a Tier-0 Gateway in Active-Standby mode.

In NSX-T manager, go to `Networking > Tier-0 Gateways > Add Gateway > Tier-0` as shown in the next figure.

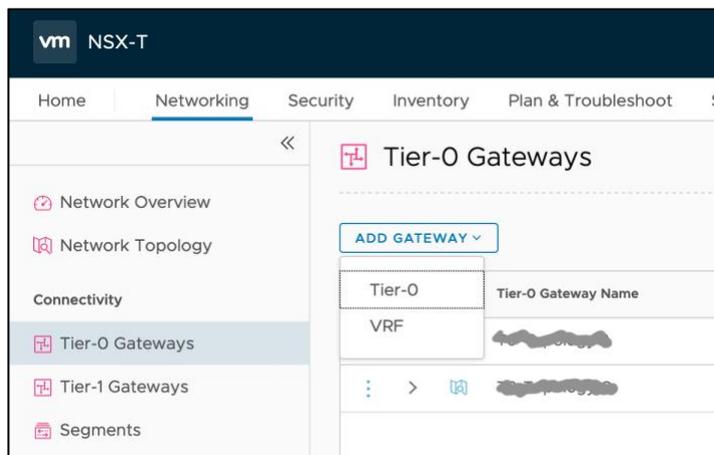


Figure 17 - Adding a Tier-0 Gateway.

In the New Tier-0 Router dialog, complete the following:

- Name: T0-topology A in this example.
- Edge Cluster: Select the existing Edge cluster.
- High Availability Mode: Active-Standby.
- Failover Mode: Non-Preemptive (to avoid double failover once the failed unit recovers).

The screenshot shows the 'Tier-0 Gateways' configuration interface. At the top left is a red icon with a white 'T' and 'L' and the text 'Tier-0 Gateways'. Below this is a blue button labeled 'ADD GATEWAY'. The main area is a light blue form with a table-like structure. The first row has three columns: 'Tier-0 Gateway Name' with the value 'T0-Topology A' and a red asterisk, 'HA Mode' with a dropdown menu set to 'Active Standby' and a red asterisk, and 'Linked Tier-1 Gateways' which is empty. Below this are several rows of settings: 'Fail Over' with a dropdown set to 'Non Preemptive'; 'Edge Cluster' with a dropdown set to 'nsx-edge-cluster-topology-a' and a blue 'x' icon; 'Additional Settings', 'Route Distinguisher for VRF Gateways', and 'EVPN Settings' each with a right-pointing chevron icon. The 'Tags' section has two dropdown menus: 'Tag (Required)' and 'Scope (Optional)', followed by a blue '+' icon and the text 'Max 30 allowed. Click (+) to save.' At the bottom of the form are three buttons: 'SAVE' (blue), 'CANCEL' (grey), and 'Unssaved Changes' (grey). Below the form are several expandable sections: 'INTERFACES', 'ROUTING', 'MULTICAST', 'BGP', and 'ROUTE RE-DISTRIBUTION', each with a right-pointing chevron icon.

Figure 18 - Filling the details of a Tier-0 Gateway/Gateway.

1.2. Create an Interface for each Edge Node used by the Tier-0 Gateway/Gateway.

Select the router created (T0-Topology-A in our example) and create two interfaces in the UI by first selecting the Edit option in the T0 Gateway, then scrolling down to the

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

Interfaces section clicking in the Set option of External and Service Interfaces. Enter the following parameters for each interface:

- Name: In this example, `edge-1-uplink-red` is used for the first router port and `edge-2-uplink-red` for the second (we will use `edge-*-uplink-blue` in the BGP+ECMP scenarios).
- Type: External
- Edge Node: This will be `edge-1-topology-a` and `edge-2-topology-a` for each external interface respectively.
- MTU: use external network's MTU, which should be the same on the BIG-IP.
- URPF Mode: `Strict` is a good practice providing security with no expected performance impact. `Strict` should be used unless asymmetric paths are used.
- Segment: This is the L2 network to which the interface is attached to. It is a prerequisite to have this previously created. See section Design consideration: Layer 2 networking for details.
- IP Address/mask: this is the IP address assigned to the address port in the shared segment between the NSX-T Edge nodes and the F5 BIG-IPs. In this example, `10.106.53.1/24` is used for router port in `edge-01` and `10.106.53.2/24` in `edge-02`.
- Click Add.

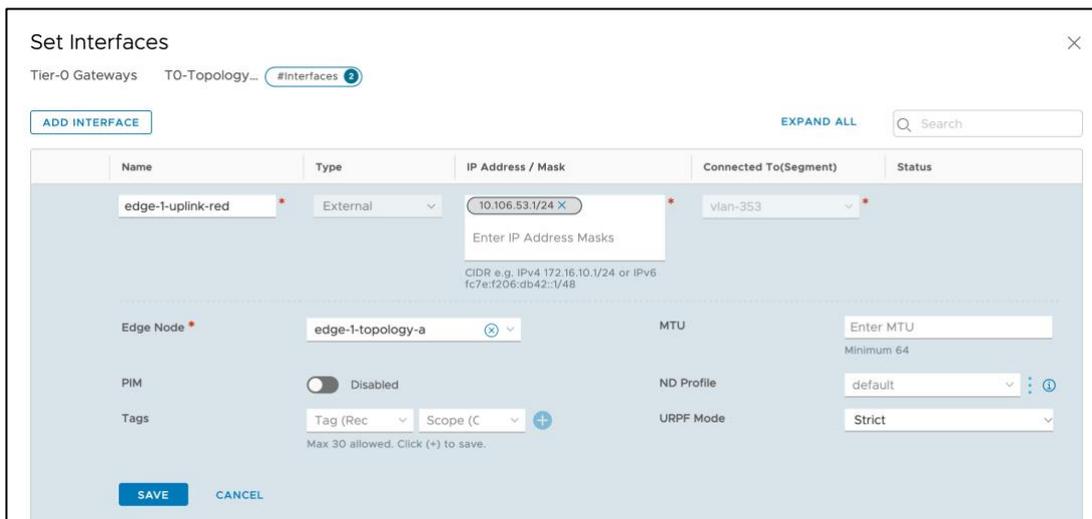


Figure 19 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.



Figure 20 – Final Gateway Port configuration of the Tier-0 Gateway.

1.3. In the Tier-0 Gateway, configure a BGP peering mesh with the F5 BIG-IPs.

In this section, it is described a BGP configuration (eBGP to be more precise) where both the NSX-T Edge cluster and the F5 BIG-IP cluster have an Active-Standby configuration. The steps involved are:

- Enable BGP in the Tier-0 Gateway.
- Configure a BGP peering mesh with the F5 BIG-IPs.
- Enable BFD in the BGP peerings.

These steps are described next.

1.3.1. Enable BGP in the Tier-0 Gateway.

In NSX-T manager, select the Tier-0 Gateway the UI by clicking `Networking > Routers` then follow the `Routing > BGP` dialogs of the router. Click the Edit button and set the values as follows:

- Local AS: This is typically within the private range 64.512 – 65.534.
- Graceful restart: Set to disable as per VMware’s best practice `NSXT-VI-SDN-038`.
- ECMP: Set to disable.

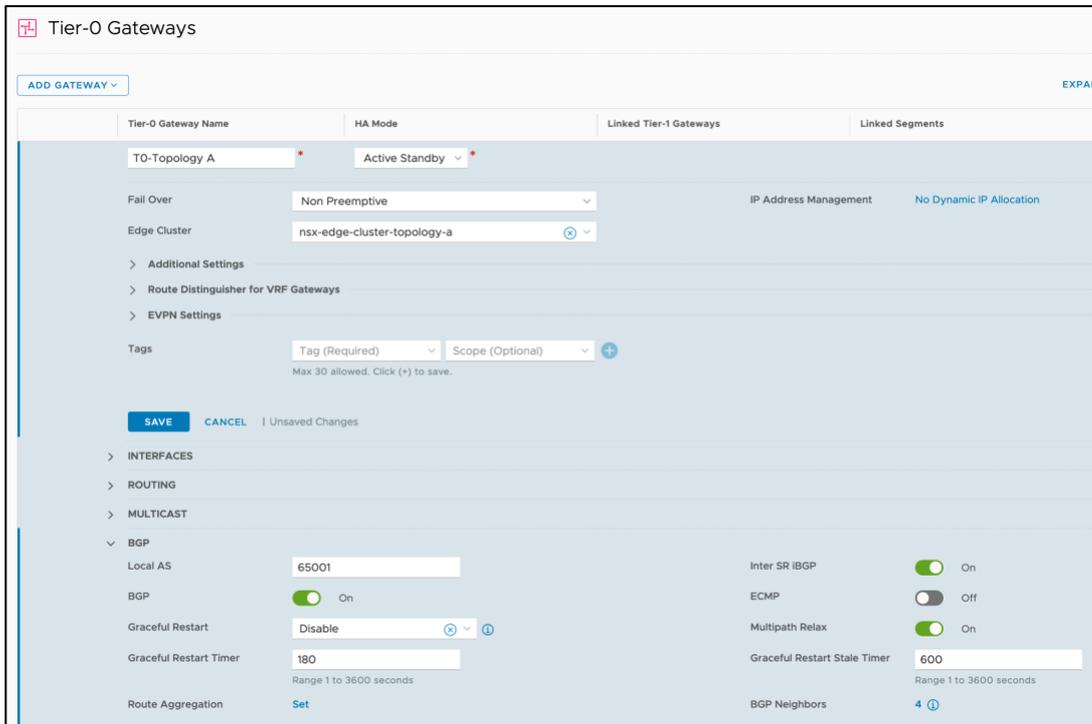


Figure 21 – Enable BGP in the Tier-0 Gateway in Active-Standby mode.

1.3.2. Configure a BGP peering mesh with the F5 BIG-IPs.

In the same BGP section, click the link `Set` in the BGP Neighbors field and complete the tabs: Neighbor, Local Address and BFD for the two BIG-IP Self IPs. In the next

figure, the peering configuration for the BIG-IP unit #1 is shown. The only configuration difference between BIG-IP unit #1 and unit #2 is the Neighbor Address.

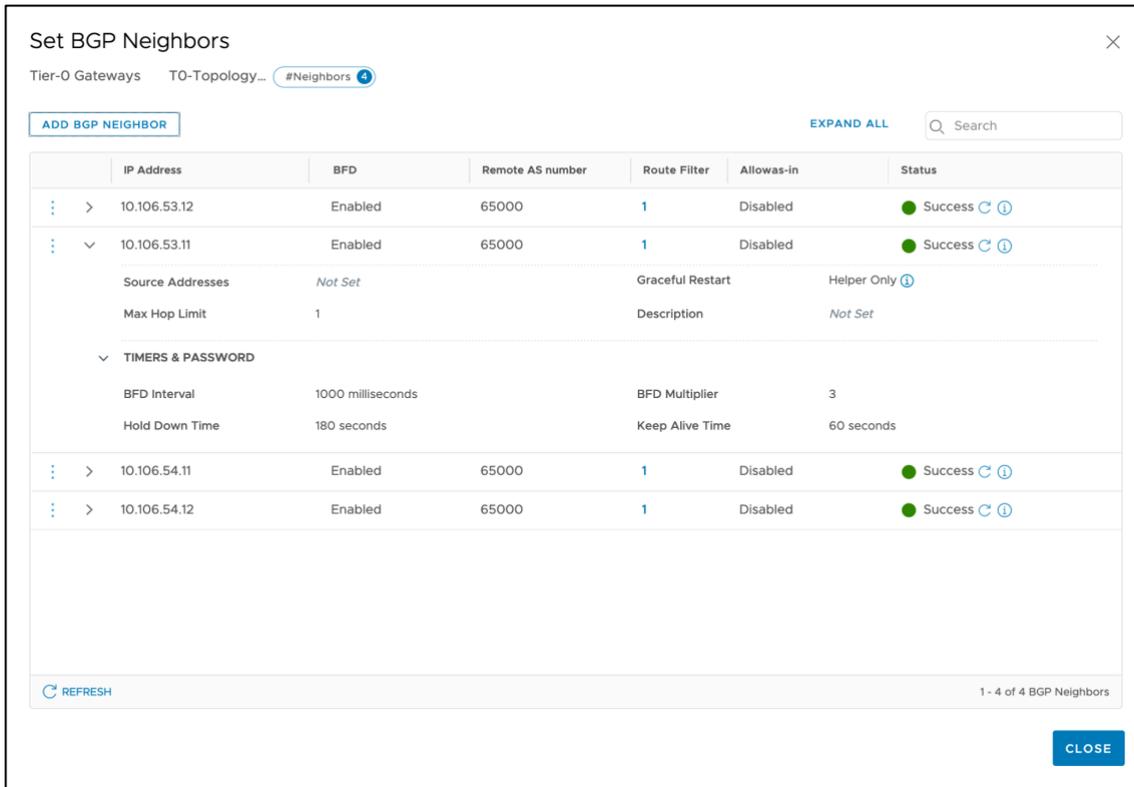


Figure 22 – Adding a BGP neighbor.

In this figure, the default values are used with the exception of the following fields:

- Neighbor Address: this is the (non-floating) Self IP of each F5 BIG-IP.
- Remote AS: typically, this is a value given by the network administrators within a private AS range.
- Password: this provides security to the peerings and avoids unwanted peerings.
- Source Address: by not specifying a source address, NSX-T will establish a BGP peering from each T0 Gateway’s uplink interface with each BIG-IP address. In this example this will establish two BGP peers for each entry.
- In the **BFD Configuration** section, the appropriate BFD settings depend if the BIG-IPs/NSX-T Edges are bare metal (timers set to 300ms) or virtual machines (timers set to 1000ms) as described in [BGP configuration details](#) within the [GENERAL NOTES](#) section.

The remaining step is to redistribute the NSX-T routes into NSX-T’s BGP which then will be announced to the BGP peers (in this case the F5 BIG-IPs). This is done at Tier-0 Gateway level in the section shown in the next figure.



Figure 23 - Enabling Route redistribution at T0 Gateway

Create a redistribution entry which includes NSX connected networks as it can be seen in the next figure.

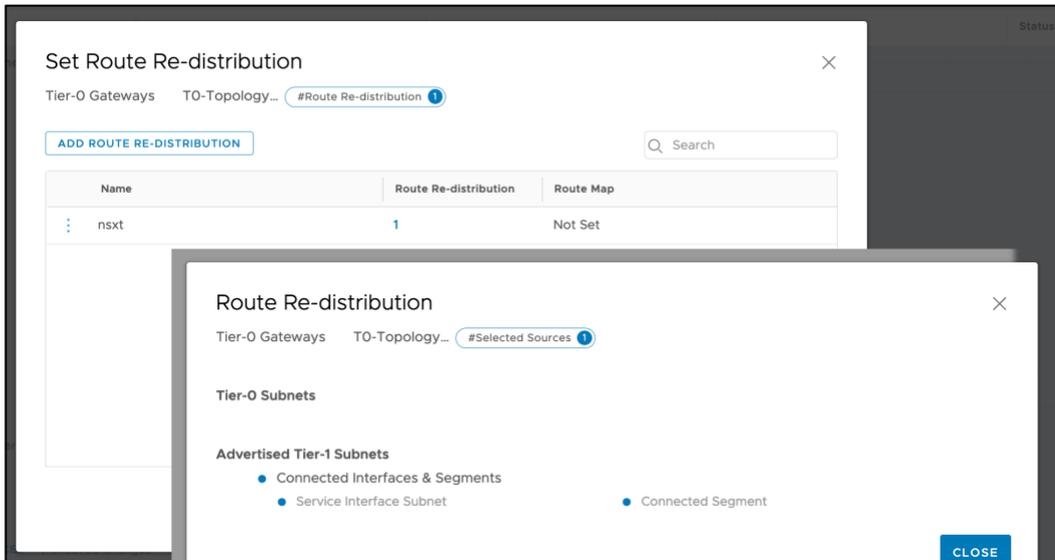


Figure 24 - Route redistribution settings at T0 Gateway

2. Create a Tier-1 Gateway.

This will be used later to instantiate a VM and perform a verification of the deployment.

In NSX-T manager, select **Networking > Tier-1 Gateways > Add Tier-1 Gateway > Tier-1 Router** filling the following parameters:

- **Name:** In this example, `T1-Topology A`.
- **Tier-0 Router:** Select the Tier-0 router (`T0-Topology A` in our example).
- **Edge Cluster:** The name of the Edge Cluster of the NSX-T Edge nodes being used.
- **Failover Mode:** `Non-Preemptive` (to avoid double failover once the failed unit recovers).
- **Route Advertisement:** at least “All Connected Segments [...]” should be enabled.
- Click **Add**.

Figure 25 – Filling the properties when creating a Tier-1 Gateway.

The next step is to create a network attached to this Tier-1 Gateway. In the UI, select **Networking > Segments > Add Segment** and enter the following parameters:

- **Segment Name:** in this example, `segment-332`.
- **Connectivity:** the Tier-1 Gateway, in this case `T1-Topology A`.
- **Subnets:** this really indicates both the subnet and the IP address of the Tier-1 Gateway in this segment, in this case `10.106.32.1/24`.

This configuration can be seen in the next figure:

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

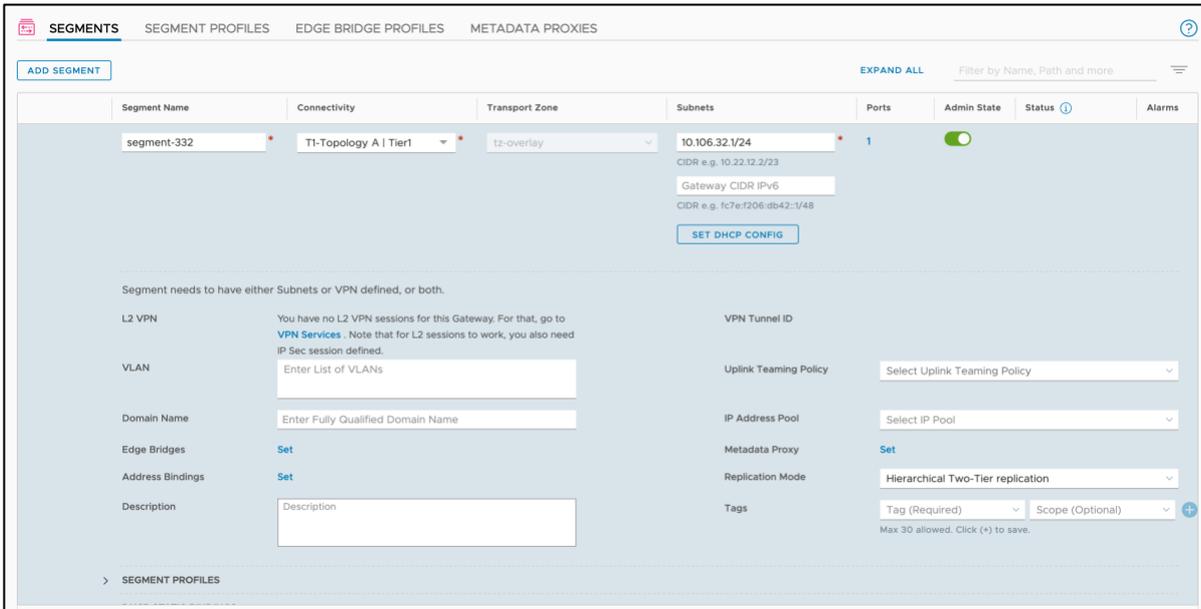


Figure 26 - Adding a segment to the T1 Gateway.

3. Create the Layer 3 configuration in the BIG-IP.

First, create the Self IPs and floating Self IPs towards the spine routers (north-bound) and towards the NSX-T Tier-0 Gateway (south-bound). These do not require any special configuration. An example of the first BIG-IP unit is shown next.

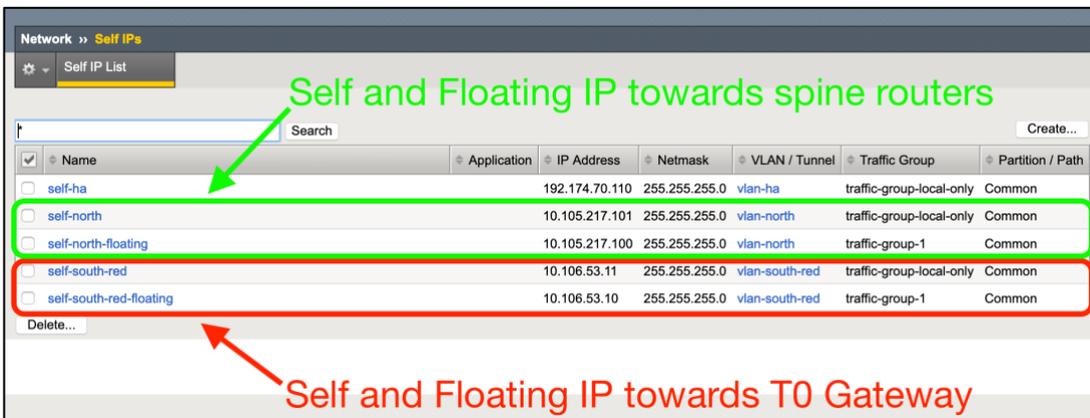


Figure 27 – Self IPs and floating Self IPs required (shown in BIG-IP unit 1).

The non-floating Self IPs need to allow TCP port 179 in order the BGP peering to be established. This is done by configuring the port lock down security feature of the Self IPs as shown in the next figure. BFD protocol will be automatically allowed.

The screenshot shows the configuration page for a Self IP named 'self-south-red'. The breadcrumb navigation is 'Network >> Self IPs >> self-south-red'. The 'Properties' tab is selected. The configuration is as follows:

Field	Value
Name	self-south-red
Partition / Path	Common
IP Address	10.106.53.11
Netmask	255.255.255.0
VLAN / Tunnel	vlan-south-red
Port Lockdown	Allow Custom
Custom List	<ul style="list-style-type: none"><input type="radio"/> TCP <input type="radio"/> UDP <input checked="" type="radio"/> Protocol: ICMP Add<input type="radio"/> All <input type="radio"/> None <input checked="" type="radio"/> Port: AddTCP: 179UDP: Protocol: ICMPDelete
Traffic Group	<input type="checkbox"/> Inherit traffic group from current partition / path traffic-group-local-only (non-floating)
Service Policy	None

Figure 28 – Allowing BGP in the non-floating Self IPs.

Note that the non-floating Self IPs are per BIG-IP unit whilst the floating Self IPs are synchronized across the BIG-IP units.

The next step is to configure the BGP routing in the BIG-IP. This involves two steps:

- Enabling BGP and BFD protocols in the routing domain used to connect to the NSX-T environment. This is done in the UI.
- Configuring BGP and BFD in the ZebOS cli (imish).

In order to enable BGP and routing protocols. Use the BIG-IPs UI and browse through Network > Route Domains > 0 (assuming that the default routing domain is the one being used). In this window enable BFD and BGP as seen in the next figure. Note that given this is

part of F5 BIG-IP's base config it is not synchronized and must be done in all the F5 BIG-IP units.

Network » Route Domains » 0

Properties

General Properties

Name	0
Partition / Path	Common
ID	0
Description	<input type="text"/>

Configuration

Strict Isolation	<input checked="" type="checkbox"/> Enabled																		
Parent Name	None																		
VLANs	<table border="0"> <tr> <td>Members:</td> <td></td> <td>Available:</td> </tr> <tr> <td>socks-tunnel</td> <td></td> <td></td> </tr> <tr> <td>vlan-ha</td> <td><<</td> <td></td> </tr> <tr> <td>vlan-north</td> <td></td> <td></td> </tr> <tr> <td>vlan-south-blue</td> <td>>></td> <td></td> </tr> <tr> <td>vlan-south-red</td> <td></td> <td></td> </tr> </table>	Members:		Available:	socks-tunnel			vlan-ha	<<		vlan-north			vlan-south-blue	>>		vlan-south-red		
Members:		Available:																	
socks-tunnel																			
vlan-ha	<<																		
vlan-north																			
vlan-south-blue	>>																		
vlan-south-red																			
Dynamic Routing Protocols	<table border="0"> <tr> <td>Enabled:</td> <td></td> <td>Available:</td> </tr> <tr> <td>BFD</td> <td></td> <td>IS-IS</td> </tr> <tr> <td>BGP</td> <td><<</td> <td>OSPFv2</td> </tr> <tr> <td></td> <td></td> <td>OSPFv3</td> </tr> <tr> <td></td> <td>>></td> <td>PIM</td> </tr> <tr> <td></td> <td></td> <td>RIP</td> </tr> </table>	Enabled:		Available:	BFD		IS-IS	BGP	<<	OSPFv2			OSPFv3		>>	PIM			RIP
Enabled:		Available:																	
BFD		IS-IS																	
BGP	<<	OSPFv2																	
		OSPFv3																	
	>>	PIM																	
		RIP																	
Bandwidth Controller	None																		
Connection Limit	0																		
Eviction Policy	None																		

Update Cancel

Figure 29 – Enabling BFD and BGP in F5 BIG-IP. This must be performed in all units.

The next step is to configure BFD and BGP itself. Log in through SSH into each BIG-IP unit and run the `imish` command which enters the ZebOS cli (ZebOS uses a typical router cli command set). The F5 BIG-IP must mimic NSX-T's BGP configuration. This is shown in the next figure with embedded comments.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

```
service password-encryption ← good security practice
!
interface VLAN196
  bfd interval 1000 minrx 1000 multiplier 3 ← matches Tier-0 config
!
router bgp 65000
  bgp router-id 10.105.196.11 ← per BIG-IP value
  redistribute kernel ← redistributes BIG-IP configured routes into BGP
  neighbor 10.106.53.1 remote-as 65001
  neighbor 10.106.53.1 password ***enter password in clear, it will be encrypted***
  neighbor 10.106.53.1 timers 60 180 ← matches Tier-0 config
  neighbor 10.106.53.1 fall-over bfd
  no neighbor 10.106.53.1 capability graceful-restart ← as per VMware's
  neighbor 10.106.53.1 route-map default-route recommendation NSXT-
  neighbor 10.106.53.2 remote-as 65001 VI-SDN-038
  neighbor 10.106.53.2 password ***enter password in clear, it will be encrypted***
  neighbor 10.106.53.2 timers 60 180
  neighbor 10.106.53.2 fall-over bfd
  no neighbor 10.106.53.2 capability graceful-restart
  neighbor 10.106.53.2 route-map default-route
!
bfd gtism enable ← safety feature enabled by default
!
ip prefix-list default-route seq 5 permit 0.0.0.0/0
!
route-map default-route permit 5 ← route-map to set the next-hop to the floating-IP
  match ip address prefix-list default-route
  set ip next-hop 10.105.53.10 primary
!
```

Figure 30 – ZebOS BGP without ECMP configuration in the BIG-IP.

At this point, follow the testing steps described in the Verifying the deployment section.

Implementation: Active/Active NSX-T Edge using BGP ECMP

For large / high performance deployments, NSX-T Edge nodes are typically configured in Active-Active. In this deployment guide it is assumed that when using NSX-T Active-Active the most likely scenario is that NSX-T Edge nodes are bare metal servers and the BIG-IPs are implemented in hardware. When using Active/Active NSX-T Edge it is likely to be used with ECMP⁵ which provides additional L3 load sharing paths. This scenario is outlined in the next figure for two NSX-T Edge nodes with two uplink Layer 3 paths. We will use a different Layer 2 segment for each Layer 3 path for additional isolation and bandwidth.

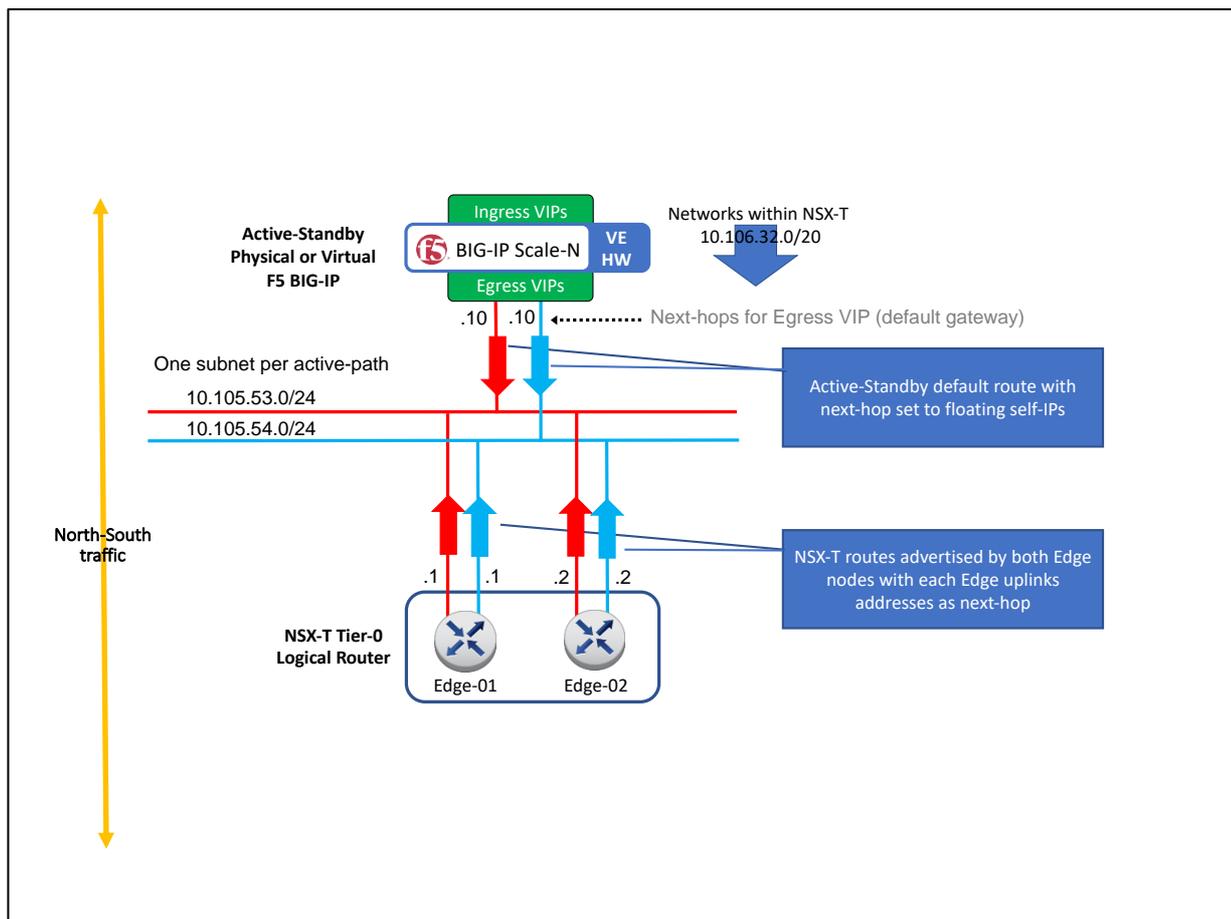


Figure 31 – Active-Active NSX-T Edge with two ECMP uplinks and BIG-IP in Active-Standby.

In this scenario the NSX-T Edge nodes are not able to process traffic in a stateful manner. The F5 BIG-IPs in Active-Standby will implement the services that require processing the traffic in a stateful manner. Given that it is highly likely that BIG-IP hardware is used, an F5 BIG-IP Active-Active setup is not required in this scenario.

An F5 BIG-IP Active-Active setup in this scenario would require a more complex configuration

⁵ Please note that NSX-T Edge Active-Active doesn't imply the use ECMP or vice versa.

in order to keep the traffic symmetry outside the NSX-T environment. Instead, if ultimate scalability is required, the best option is adding blades with a chassis platform which provides ultimate scale-out performance without requiring any reconfiguration and keeps a simple architecture.

In this topology, each Edge node needs two uplinks which must be in different logical switches and different transport zones. The Edge nodes share the logical switches for each uplink subnet. Figure 32 shows the detail of the BGP peerings established between NSX-T edge nodes and the BIG-IPs. Note that although the Edge nodes have as next-hop the floating Self IPs of each subnet, the BGP peerings are setup with the non-floating Self IPs. In total 4 BGP peerings are created but unlike with the previous BGP configuration without ECMP, this time each peer uses a different Layer 3 network for each peering.

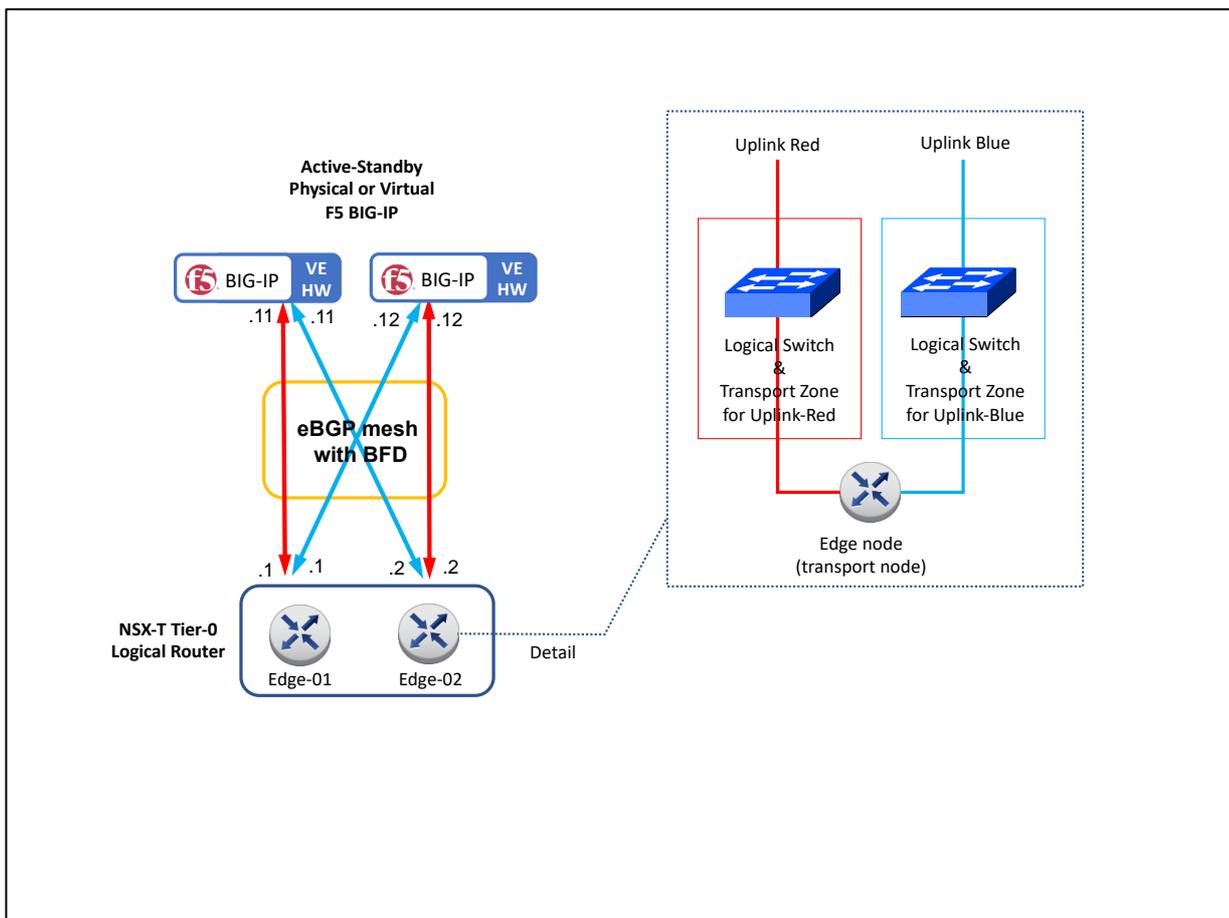


Figure 32 – BGP peering detail with two uplink Layer 3 paths & transport zones for ECMP.

Given the many possibilities of configuring NSX-T Edge nodes and their logical switch uplink ports, it is assumed that these have been already created. This guide is focused in the configuration for the Layer 3 and higher layers that are specific to this topology. See section Design consideration: Layer 2 networking for details.

1. Create a transport zone for each uplink

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

In NSX-T manager, create a separate transport zone of type VLAN and logical switches for each uplink subnet.

Ultimately there will be used 3 transport zones, one for each uplink (tz-vlan-uplink-red and tz-vlan-uplink-blue) and one for the overlay networking. All these are shown in the next figure.

Transport Zone	ID	Traffic Type	Transport Node Membe	Status	Where Used
tz-overlay	1b3a...963e	Overlay	8	Up	Where Used
tz-vlan	a95c...aeba	VLAN	2	Up	Where Used
tz-vlan-uplink-blue	015a...84c3	VLAN	2	Up	Where Used
tz-vlan-uplink-red	afa3...3857	VLAN	2	Up	Where Used

Figure 33 – Overall configuration of transport zones. The used ones by this topology are highlighted (red and blue for the uplinks).

2. Edit the Edge transport nodes to add the two uplink transport zones.

Go to `System > Fabric > Nodes > Edge Transport Nodes` and Edit each Edge transport node associated with the T0 Gateway, adding a switch (N-VDS switch) for each Uplink transport zone created in the previous steps. This is shown in the next figure.

Edit Edge Transport Node - edge-1-topology-a ×

Name *

Description

+ ADD SWITCH

> edge-nvds-overlay	DELETE
> edge-nvds-red	DELETE
> edge-nvds-blue	DELETE

Figure 34 – Adding the switches for each Uplink transport zone in each Edge transport nodes.

Besides each transport-zone, each associated N-VDS switch requires specific Uplink profile and Uplink interfaces. An example for Transport Zone `tz-vlan-uplink-red` is shown next.

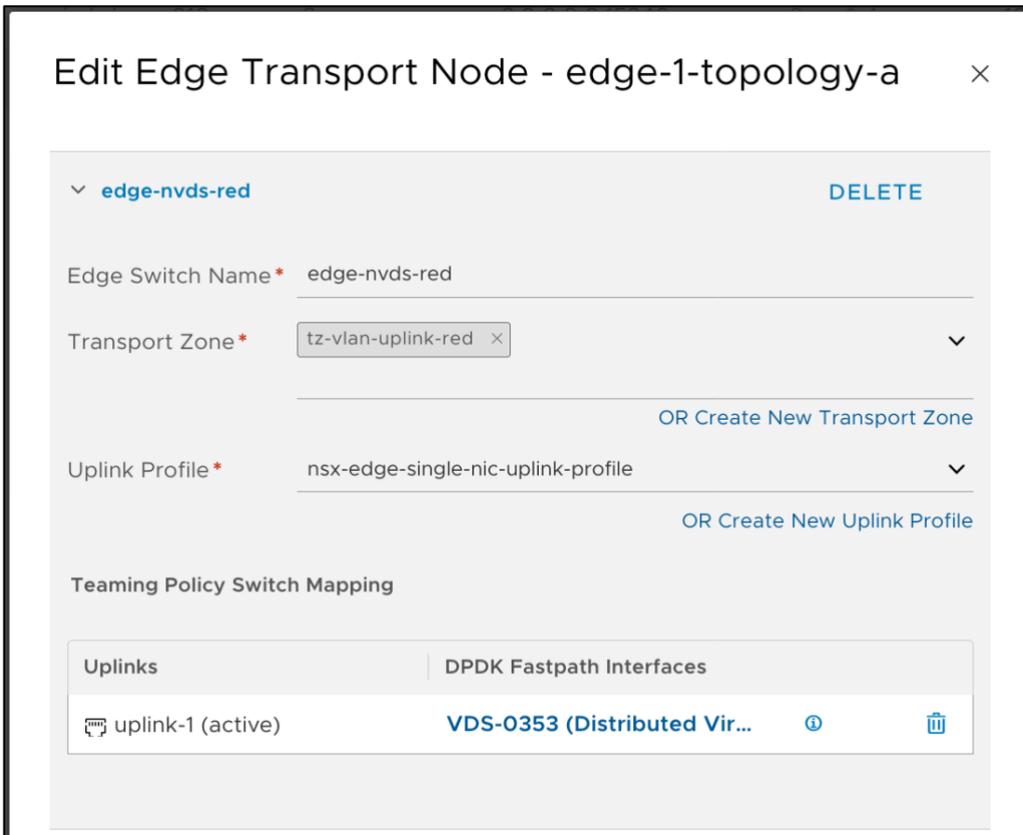


Figure 35 – N-VDS switch configuration for a sample Uplink transport zone.

3. Create a Tier-0 configuration.

3.1. Create a Tier-0 Gateway in Active-Active mode.

In NSX-T manager, go to `Networking > Tier-0 Gateways > Add Gateway > Tier-0` as shown in the next figure.

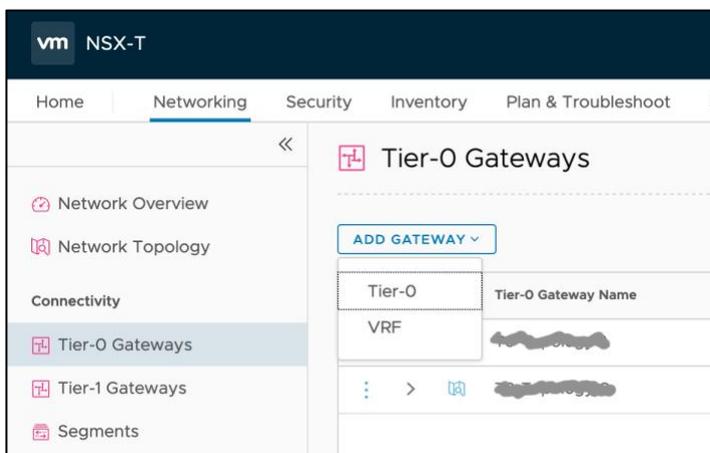


Figure 36 - Adding a Tier-0 Gateway.

In the New Tier-0 Router dialog, complete the following:

- Name: T0-topology A in this example.
- Edge Cluster: Select the existing Edge cluster.
- High Availability Mode: Active-Active.

The screenshot shows the 'Tier-0 Gateways' configuration page. At the top left is a red icon with a plus sign and the text 'Tier-0 Gateways'. Below this is a blue button labeled 'ADD GATEWAY'. The main area is a form with two columns: 'Tier-0 Gateway Name' and 'HA Mode'. The 'Tier-0 Gateway Name' field contains 'T0-Topology A' with a red asterisk. The 'HA Mode' dropdown is set to 'Active Active' with a red asterisk. Below these are sections for 'IP Address Management' (set to 'No Dynamic IP Allocation'), 'Edge Cluster' (set to 'nsx-edge-cluster-topology-a'), and 'Additional Settings' (with expandable sections for 'Route Distinguisher for VRF Gateways' and 'EVPN Settings'). There is also a 'Tags' section with 'Tag (Required)' and 'Scope (Optional)' dropdowns, and a note 'Max 30 allowed. Click (+) to save.' At the bottom are sections for 'INTERFACES', 'ROUTING', 'MULTICAST', 'BGP', and 'ROUTE RE-DISTRIBUTION', each with an expandable arrow.

Figure 37 - Filling the details of a Tier-0 Gateway in Active-Active mode.

3.2. Create a Router interface for each Edge Node used by the Tier-0 Gateway.

Select the just created Tier-0 Gateway and create 1 Gateway port for each peering address. This is one Gateway's interface for the combination of each subnet (two in this example) and NSX-T Edge nodes (two in this example). In total 4 Gateway interfaces will be created as shown next. It is very important to correctly assign the right Edge Transport

node and switch. The ports and their configuration used in this example are shown next. The settings for each Gateway's interfaces are analogous to the Active-Standby setup.

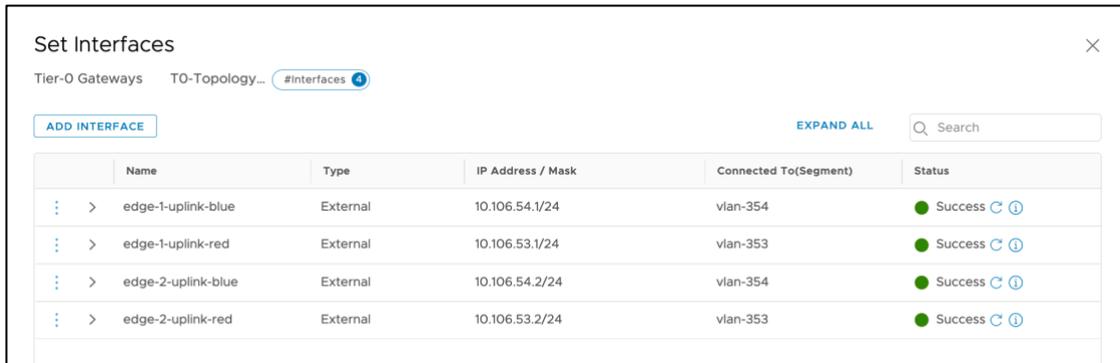


Figure 38 – Adding the Gateway's interfaces for the Uplink paths.

3.3. Enable BGP in the Tier-0 Gateway likewise the Active-Standby setup but in this case enabling ECMP.

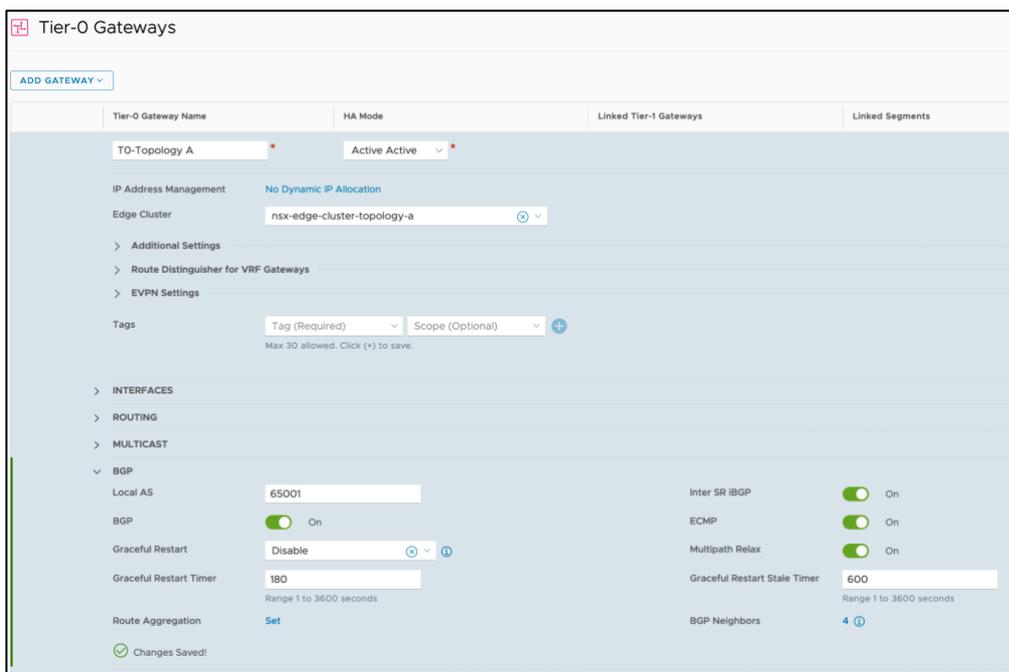


Figure 39 - Enable BGP with ECMP in the Tier-0 Gateway in Active-Active mode.

Configure a BGP peering mesh with the F5 BIG-IPs.

Unlike in the Active-Standby setup, in this case the source address for each peering will be specified. Overall the configuration settings to be used are shown next:

- **Neighbor Address:** this is the (non-floating) Self IP of each F5 BIG-IP.
- **Remote AS:** typically, this is a value given by the network administrators within a private AS range.
- **Password:** this provides security to the peerings and avoids unwanted peerings.
- **Source Address:** by not specifying a source address, NSX-T will establish a BGP peering from each T0 Gateway's uplink interface with each BIG-IP address. In this example this will establish two BGP peers for each entry.

- **BFD Configuration:** the appropriate BFD settings depend if the BIG-IPs/NSX-T Edges are bare metal (timers set to 300ms) or virtual machines (timers set to 1000s) as described in [BGP configuration details](#) within the [GENERAL NOTES](#) section.

Ultimately the configuration should be similar to the one in the following figure:

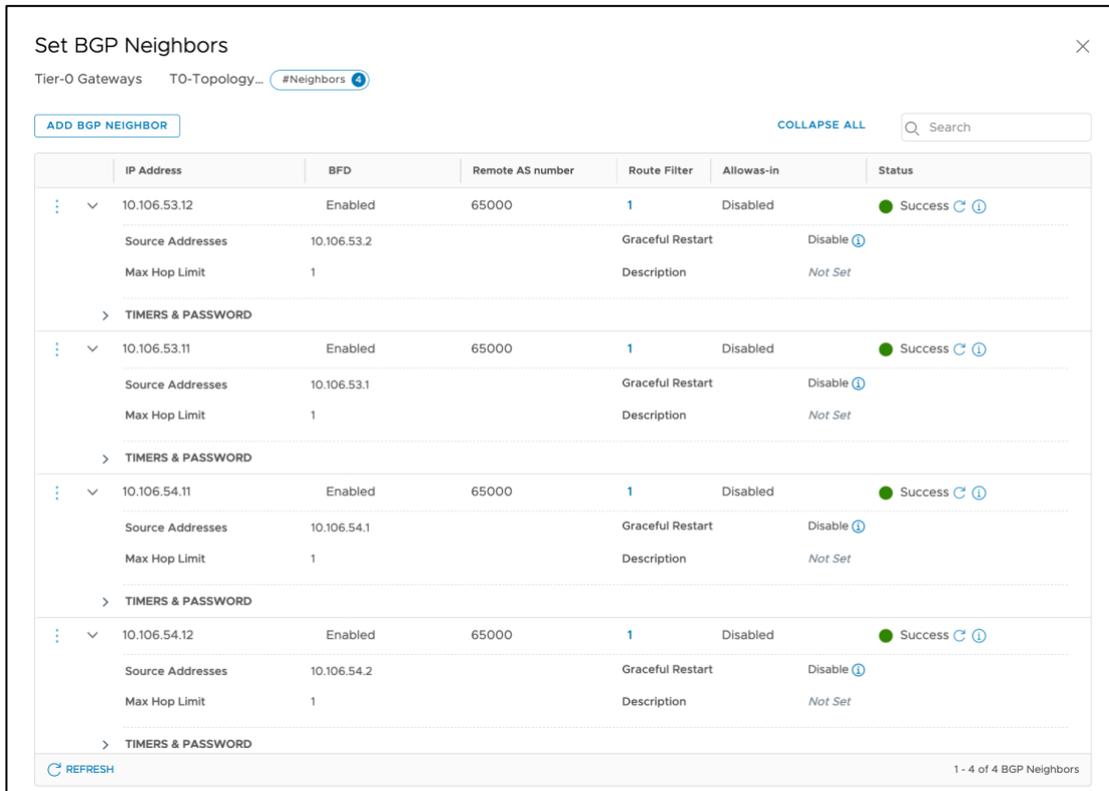


Figure 40 – BGP peerings for ECMP.

The remaining step is to redistribute the NSX-T routes into NSX-T's BGP which then will be announced to the BGP peers (in this case the F5 BIG-IPs). This is done at Tier-0 Gateway level in the section shown in the next figure.



Figure 41 - Enabling Route redistribution at T0 Gateway

Create a redistribution entry which includes NSX connected networks as it can be seen in the next figure.

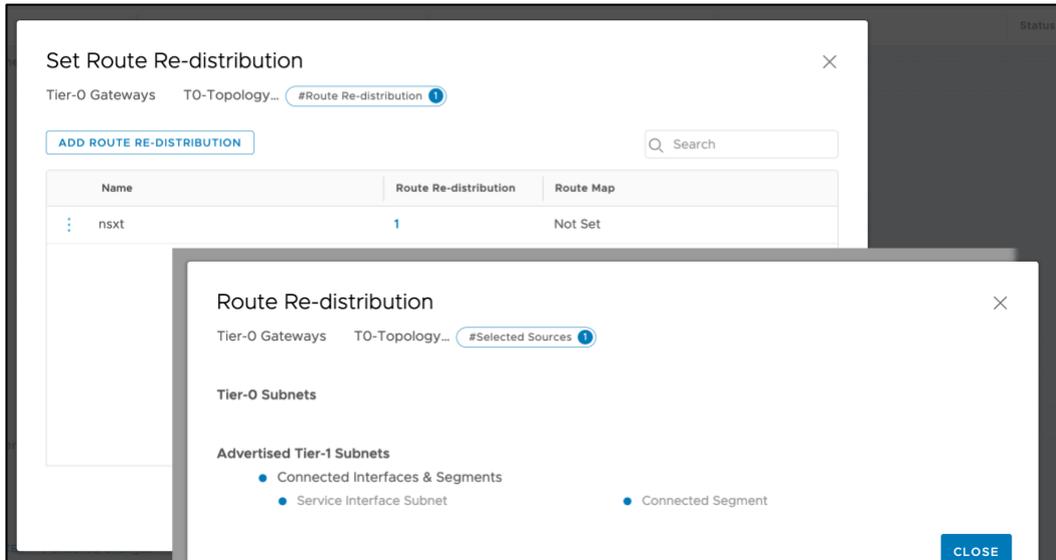


Figure 42 - Route redistribution settings at T0 Gateway

4. Create a Tier-1 Router. This step is the same as in the Active-Standby setup.
5. Create the Layer 3 configuration for the BIG-IP side.

Overall, the configuration of Self IPs is analogous to the Active-Standby setup but in this case, there are two segments (`vlan-south-blue` and `vlan-south-red`). The overall configuration for BIG-IP unit #1 is shown in the next figure.

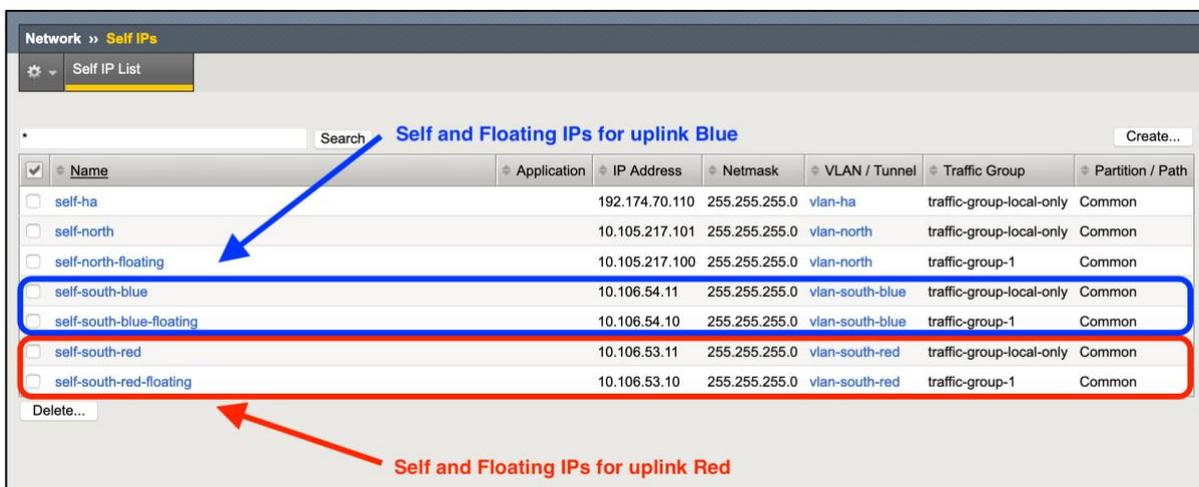


Figure 43 – Self IP in BIG-IP unit #1 for the NSX-T uplinks.

The Self IPs towards NSX-T's uplinks have the same configuration as in the Active-Standby configuration using BGP. Please check the Active-Standby implementation section for details on configuring these Self IPs.

The next step is to configure BFD and BGP itself. For this log in through SSH into each BIG-IP unit and run the `imish` command which enters the ZebOS cli (ZebOS uses a typical router cli command set). The F5 BIG-IP must mimic NSX-T's BGP configuration. This is shown in the

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

next figure with embedded comments. The differences between with the Active-Standby setup are shown in colors other than orange.

```
Service password-encryption ← good security practice

interface vlan-south-blue
  bfd interval 1000 minrx 1000 multiplier 3 ← matches Tier-0 config
!
interface vlan-south-red
  bfd interval 1000 minrx 1000 multiplier 3
!
router bgp 65000
  bgp router-id 192.174.70.111 ← per BIG-IP value
  max-paths ebgp 8 ← ECMP
  redistribute kernel ← redistributes BIG-IP configured routes into BGP
  neighbor 10.106.53.1 remote-as 65001
  neighbor 10.106.53.1 password ***enter password in clear, it will be encrypted***
  neighbor 10.106.53.1 timers 60 180 ← matches Tier-0 config
  neighbor 10.106.53.1 fall-over bfd
  no neighbor 10.106.53.1 capability graceful-restart ← as per VMware's
  neighbor 10.106.54.1 route-map default-route-uplink-red out recommendation NSXT-VI-SDN-038
  neighbor 10.106.54.1 remote-as 65001
  neighbor 10.106.54.1 password ***enter password in clear, it will be encrypted***
  neighbor 10.106.54.1 timers 60 180
  neighbor 10.106.54.1 fall-over bfd
  no neighbor 10.106.54.1 capability graceful-restart
  neighbor 10.106.54.1 route-map default-route-uplink-blue out
!
bfd gtsm enable ← safety feature enabled by default
!
ip prefix-list default-route seq 5 permit 0.0.0.0/0
!
route-map default-route-uplink-red permit 5 ← route-map to set the next-hop to the
  match ip address prefix-list default-route floating-IP, one per load sharing path.
  set ip next-hop 10.106.53.10 primary
!
route-map default-route-uplink-blue permit 5 ← route-map to set the next-hop to the
  match ip address prefix-list default-route floating-IP, one per load sharing path.
  set ip next-hop 10.106.54.10 primary
!
```

Figure 44 – ZebOS BGP ECMP configuration in the BIG-IP.

One key aspect of doing L3 path load sharing (in this case using BGP+ECMP) is that the BIG-IP can receive traffic for the same flow in different VLANs (asymmetric traffic) by default, as a security feature the BIG-IP doesn't allow such behavior blocking this traffic.

Asymmetric traffic is allowed in the BIG-IP by unsetting the parameter VLAN-Keyed Connections as shown in the next figure. This must be configured in all the BIG-IP units.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

System » Configuration : Local Traffic : General

Device Local Traffic OVSDB App IQ

Properties

Auto Last Hop	<input checked="" type="checkbox"/> Enabled
Maintenance Mode	<input type="checkbox"/>
VLAN-Keyed Connections	<input type="checkbox"/> Must be unset to allow asymmetric traffic
Path MTU Discovery	<input checked="" type="checkbox"/> Enabled
Reject Unmatched Packets	<input checked="" type="checkbox"/> Enabled
Eviction Policy	default- eviction-policy
Default Per Virtual Server SYN Check™ Threshold	0 Half-Open TCP connections
Global SYN Check™ Threshold	64000 Half-Open TCP connections
Layer 2 Cache Aging Time	300 seconds
Share Single MAC Address	<input type="checkbox"/>
SNAT Packet Forwarding	TCP and UDP Only
Hardware VLAN SYN Cookie Protection	<input checked="" type="checkbox"/> Enabled

Figure 45 – Configuration required for ECMP which might generate asymmetric traffic.

At this point, follow the testing steps described in the Verifying the deployment section.

Topology B: BIG-IPs inline – connected like an NSX-T's Tier-1 Gateway.

In the next figure it is shown an overview of this topology.

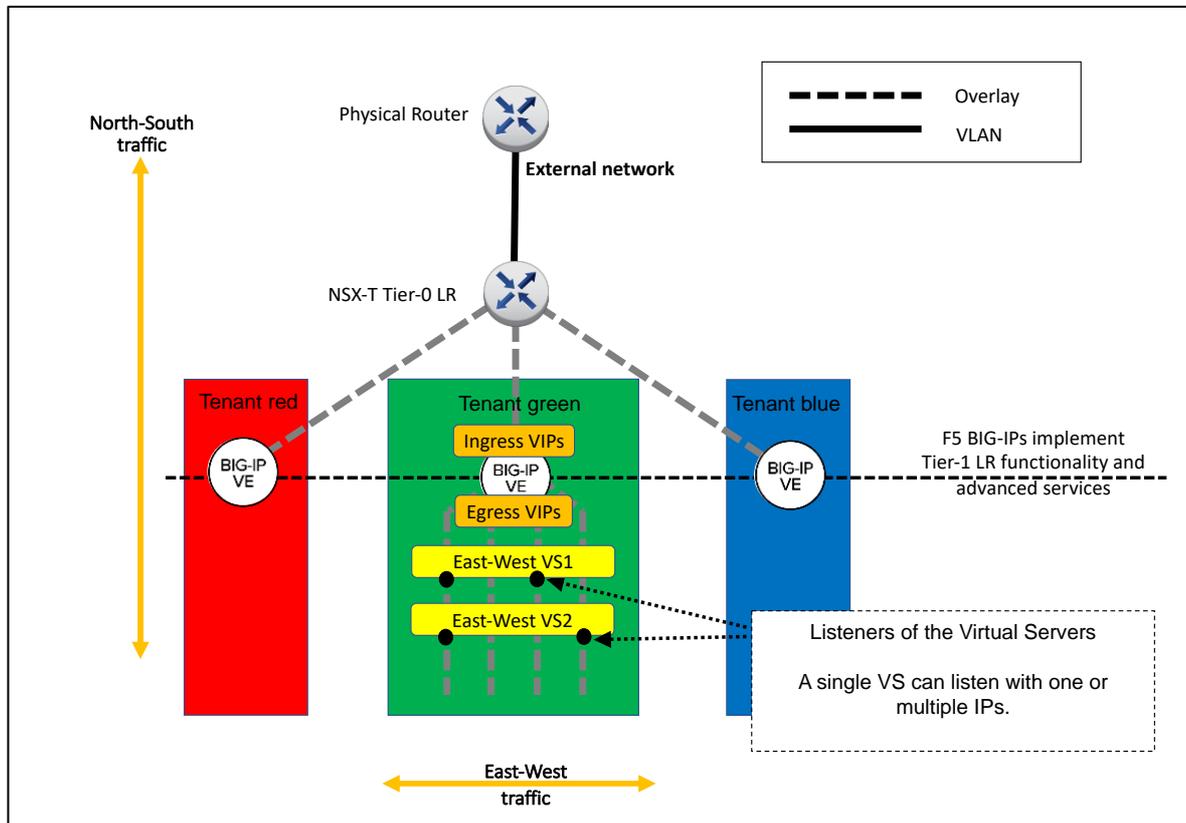


Figure 46 – Overview of BIG-IPs inline-connected like an NSX-T's Tier-1 Gateway.

The main characteristic of this topology is that NSX-T's Tier-1 Gateways are replaced by BIG-IPs. NSX-T's distributed firewall works normally, but this topology eliminates NSX-T's distributed routing between the segments at the Tier-1. This is not as performance impacting as it might seem. It only impacts performance when there is plain routing between the segments. If the services between the segments are implemented with load balancing (which is beneficial for availability of the services) there is no performance impact because load balancing is always implemented in a centralized manner (whether implementing it with NSX-T's LB or BIG-IP ADC or any other VM-based load balancer), unless using NSX-T's DLB which has very limited functionality.

Eliminating the NSX-T's Tier-1 Gateway keeps a simpler 2-tier routing and allows F5 BIG-IP Services to be implemented between the tenant segments. If it is expected to have a high volume of plain routing traffic between the tenant's segments, then NSX-T's distributed Gateway should be inserted south of tenant's BIG-IPs, creating a 3-tier routing where BIG-IP's routing tier would just be transit between NSX-T's top and bottom Gateways.

Unlike other LB implementations it is not necessary to dedicate a subnet for East-West VIPs. BIG-IP Virtual Servers can have one or more VIPs listening in one or more segments independently of the address of the VIP. This will be exemplified in the implementation section.

It is recommended to have BIG-IP clusters specific for each tenant. This is aligned with

VMware’s vision where the Tier-1’s domain can be managed by each tenant. The benefits of using BIG-IQ for centralized management and visibility are more relevant in this this topology. Additionally, having several BIG-IP clusters distributes the workload across the ESXi hypervisors unlike NSX-T’s LBs which might be more limited running in NSX-T Edge’s hosts only.

Implementation: BIG-IPs inline-connected like an NSX-T’s Tier-1 Gateway.

In the next figure, the configuration to be implemented is shown.

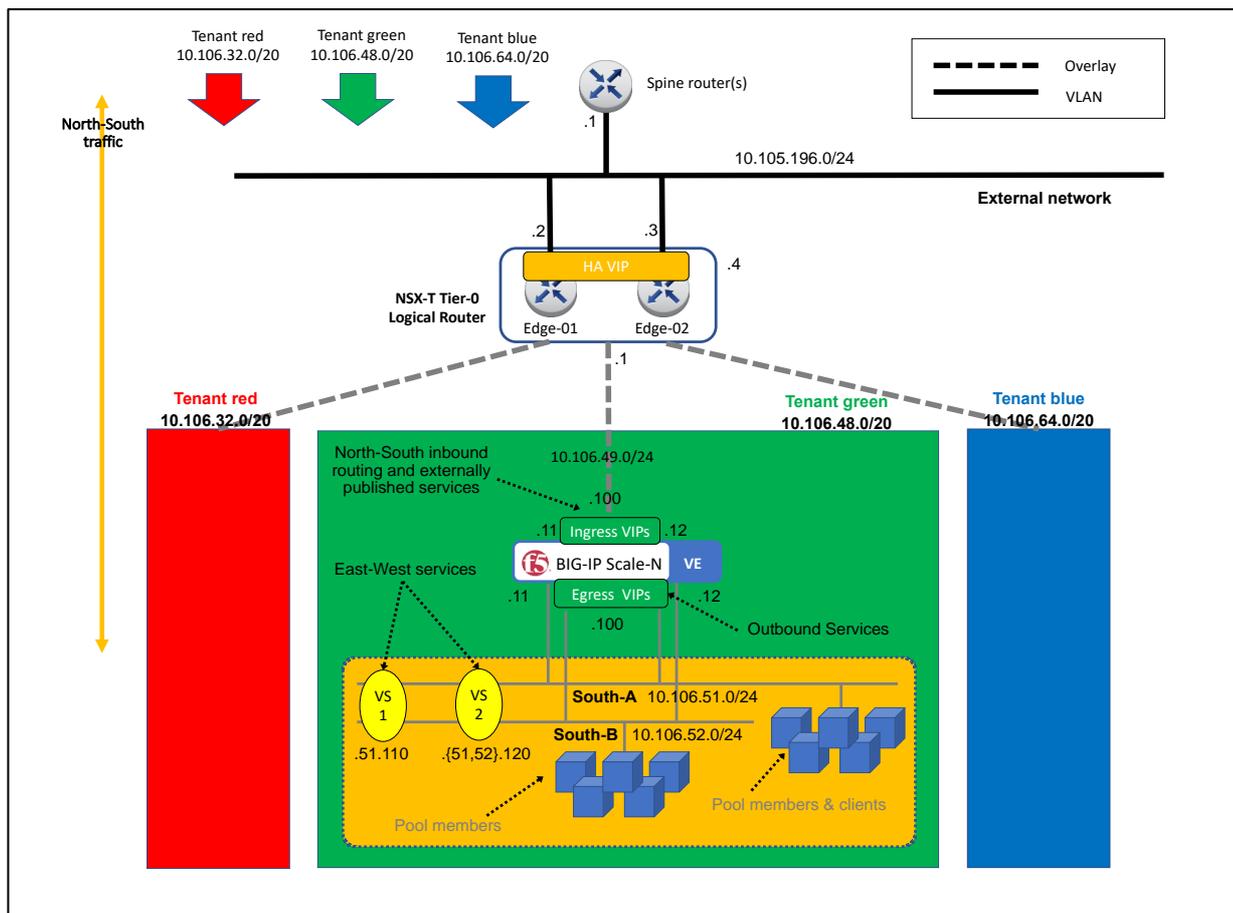


Figure 47 – Topology B example implementation.

In order to have a manageable network, contiguous networks are used for each tenant. In this example, /20 prefixes are used. This is especially relevant in this topology because NSX-T’s Gateways are not used. Only NSX-T Gateways can advertise routes within the whole NSX-T network. In the case of using BIG-IP as a Tier-1 Gateway replacement, it is needed to configure static routes in NSX-T’s Tier-0. By having contiguous networks for each tenant, it is only needed a single routing entry per tenant.

The transit network between the Tier-0 and the BIG-IPs uses a /24. Using a /24 prefix is larger than strictly necessary for an HA-pair (only 4 hosts address would be needed) but allows for more ingress VIP addresses and expanding the BIG-IP HA cluster into a Scale-N Active-Active cluster (up to 8 BIG-IPs per cluster) or multiple BIG-IP clusters.

From the figure above, it can be seen that this topology is only supported by BIG-IP VE. The configuration will be detailed next. As with all other topologies, this guide focuses in the configuration for the Layer 3 and higher layers that are specific to this topology.

1. Create the Tier-0 configuration.

1.1. Create a Tier-0 Gateway in Active-Standby mode.

In NSX-T manager, go to `Networking > Tier-0 Gateways > Add Gateway > Tier-0` as shown in the next figure.

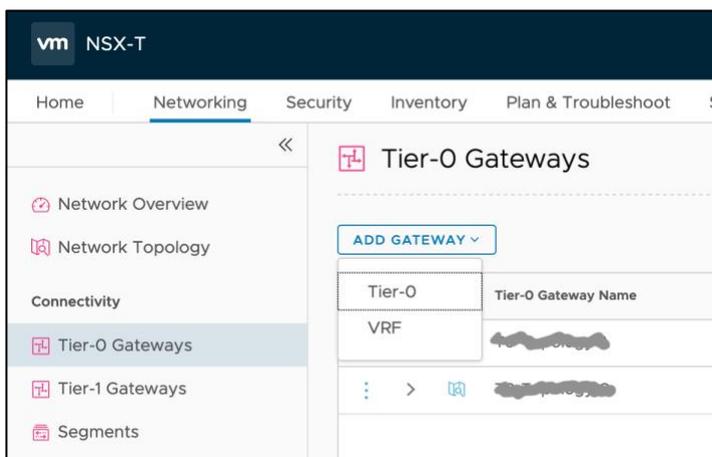


Figure 48 - Adding a Tier-0 Gateway.

In the New Tier-0 Router dialog, complete the following:

- Name: `T0-topology B` in this example.
- Edge Cluster: Select the existing Edge cluster.
- High Availability Mode: `Active-Standby`.
- Failover Mode: `Non-Preemptive` (to avoid double failover once the failed unit recovers).

Tier-0 Gateway Name	HA Mode	Linked Tier-1 Gateways
T0-Topology B *	Active Standby *	

Fail Over: Non Preemptive

Edge Cluster: nsx-edge-cluster-topology-b

Tags: Tag (Required) Scope (Optional) +
Max 30 allowed. Click (+) to save.

SAVE CANCEL | Unsaved Changes

INTERFACES

ROUTING

MULTICAST

BGP

ROUTE RE-DISTRIBUTION

Figure 49 – Filling the details of a Tier-0 Gateway.

1.2. Create an Interface for each Edge Node used by the Tier-0 Gateway.

Select the router created (`T0-Topology B` in our example) and create two interfaces in the UI by first selecting the Edit option in the T0 Gateway, then scrolling down to the

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

Interfaces section clicking in the `Set` option of `External` and `Service` Interfaces. Enter the following parameters for each interface:

- **Name:** In this example, `edge-1-uplink-red` is used for the first router port and `edge-2-uplink-red` for the second (we will use `edge-*-uplink-blue` in the BGP+ECMP scenarios).
- **Type:** `External`
- **Edge Node:** This will be `edge-1-topology-a` and `edge-2-topology-a` for each external interface respectively.
- **MTU:** use external network's MTU, which should be the same on the BIG-IP.
- **URPF Mode:** `Strict` is a good practice providing security with no expected performance impact. `Strict` should be used unless asymmetric paths are used.
- **Segment:** This is the L2 network to which the interface is attached to. It is a prerequisite to have this previously created. See section `Design consideration: Layer 2 networking` for details.
- **IP Address/mask:** this is the IP address assigned to the address port in the shared segment between the NSX-T Edge nodes and the F5 BIG-IPs. In this example, `10.106.53.1/24` is used for router port in `edge-01` and `10.106.53.2/24` in `edge-02`.
- Click `Add`.

Name	Type	IP Address / Mask	Connected To(Segment)	Status
edge-1-uplink-red	External	10.106.53.1/24	vlan-353	

Figure 50 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.

Name	Type	IP Address / Mask	Connected To(Segment)	Status
edge-1-uplink-red	External	10.106.53.1/24	vlan-353	Success
edge-2-uplink-red	External	10.106.53.2/24	vlan-353	Success

Figure 51 – Final Uplink interface configuration of the Tier-0 Gateway.

1.3. Create an HA VIP for the Tier-0 Gateway.

The HA VIP is an IP address that will be shared by the two Edge Nodes used for the Tier-0 Gateway created and will be used as the ingress IP to the NSX-T networks.

Select the Router created (T0-Topology A in our example), and create an HA VIP in the UI by selecting **Edit > HA VIP Configuration > Set** and entering the following parameters:

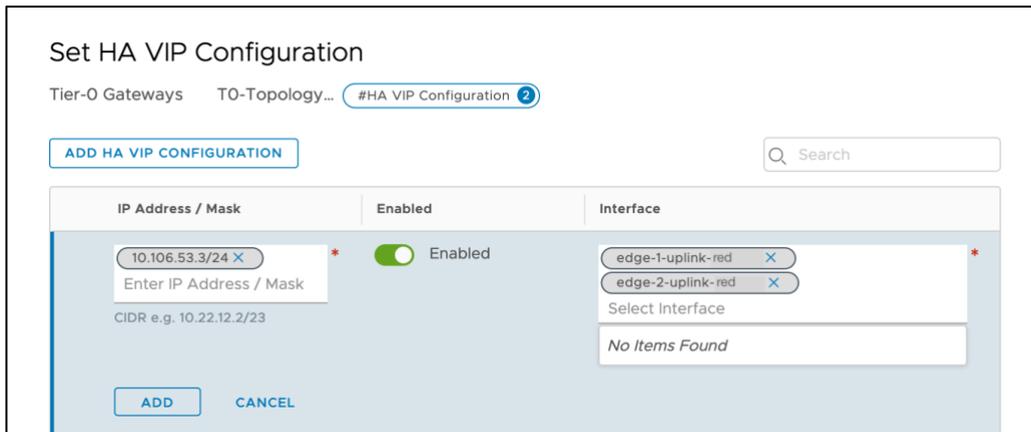


Figure 52 - Adding an HA VIP to NSX-T's T0 Gateway.

Selecting the two external interfaces just created.

Add a default route in the Tier-0 Gateway towards the BIG-IP cluster floating Self IP address.

In our example, the BIG-IP cluster floating address to use as the next hop is 10.106.53.10. Select the T0-Topology A Gateway created and then create a static routing in the UI by selecting **Routing > Static Routes > Set** as follows and entering as Next Hop BIG-IP's floating-IP, in this example (not shown in the figure) 10.106.53.10.

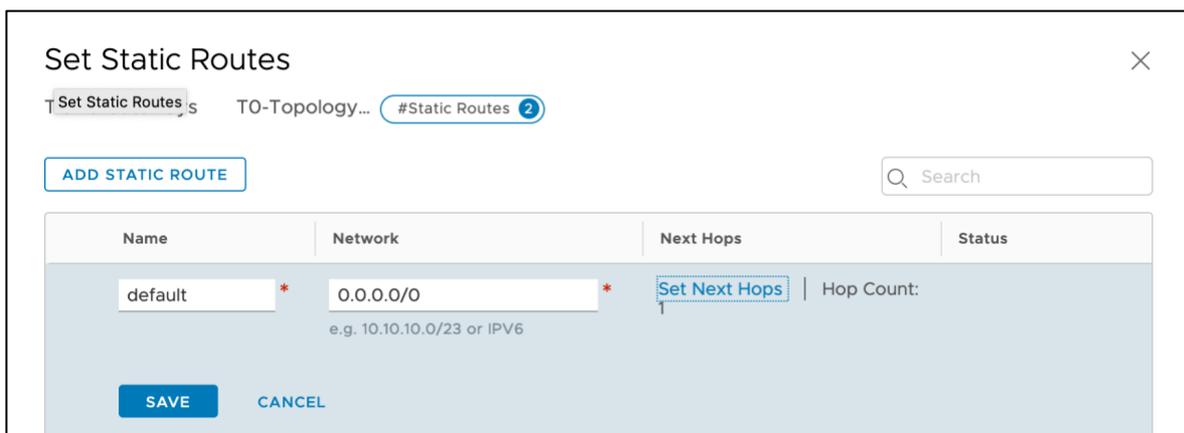


Figure 53 – Adding Tier-0 Gateway's default route.

2. Create a segment for the transit network between Tier-0 Gateway/Edges and the BIG-IPs.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

Go to **Networking > Segments > ADD SEGMENT** and create a Logical Switch within the overlay Transport Zone and attaching it to the Tier-0 Gateway as follows:

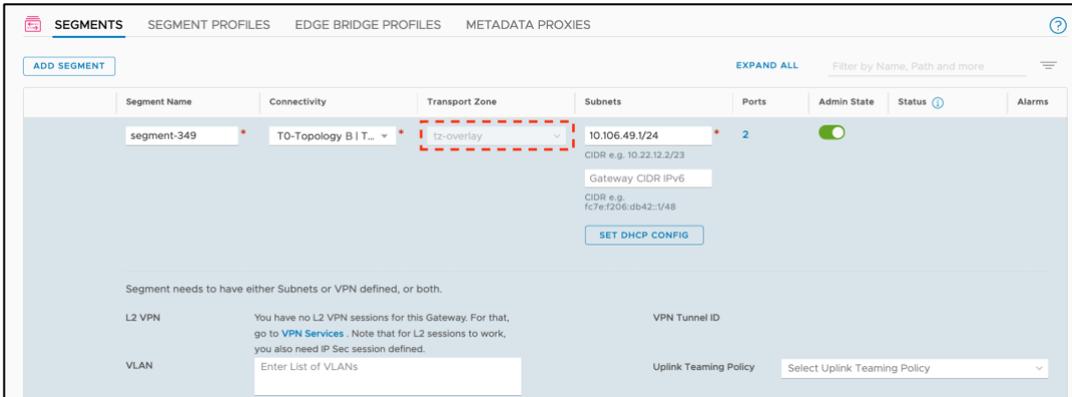


Figure 54 – Creating an overlay segment for the transit network between the Tier-0 Gateway and the BIG-IPs.

2.1. Add tenants' routes to Tier-0 Gateway.

By using a contiguous prefix per tenant it is only needed to add a single route to the existing routing table. Ultimately the routing table will look like Figure 55.

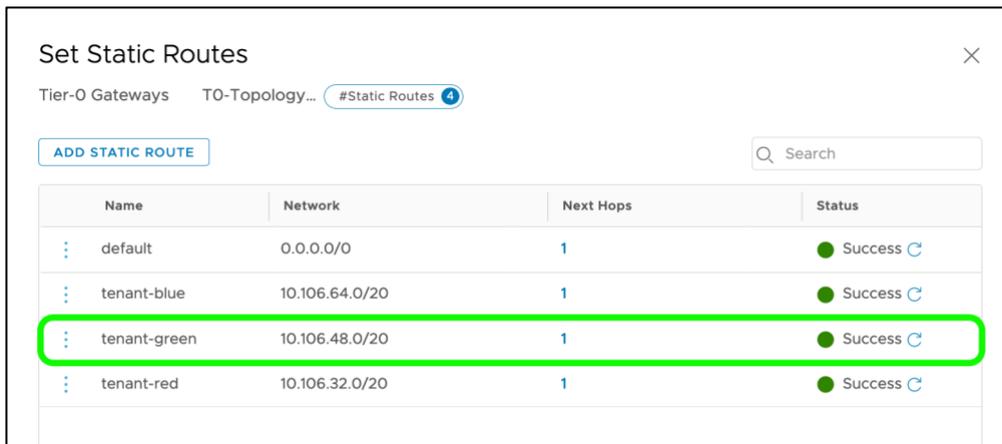


Figure 55 – Adding tenant's routing entries. Highlighted is the routing entry for tenant green for which BIG-IPs are configured in this section.

2.2. Create tenant's segments.

Follow the same steps as for creating the segment for the transit network, creating as many logical switches as networks are going to be used for the tenant. In this example we

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

will create only the ones for the tenant green, these will be:

- segment 349/transit network - 10.106.49.0/24
- segment 351/tenant network - 10.106.51.0/24
- segment 352/tenant network - 10.106.52.0/24

3. Create the Layer 3 configuration in the BIG-IP side.

Unlike in Topology A's implementations, in this topology the BIG-IPs will use NSX-T overlay segments for the data traffic. After creating the segments in the NSX manager, the BIG-IP VE can be attached to these segments just like a non NSX-T segment:

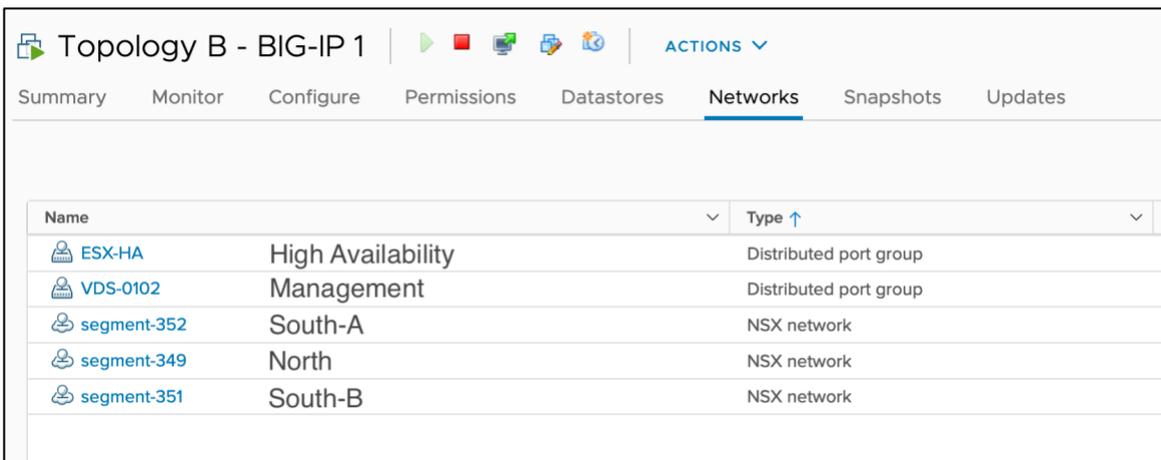


Figure 56 – Attaching the BIG-IP to NSX-T segments.

Notice the different types of Networks (NSX and regular/non-NSX). The BIG-IP will make use of all these networks just like any regular untagged VLAN as shown in the next figure:

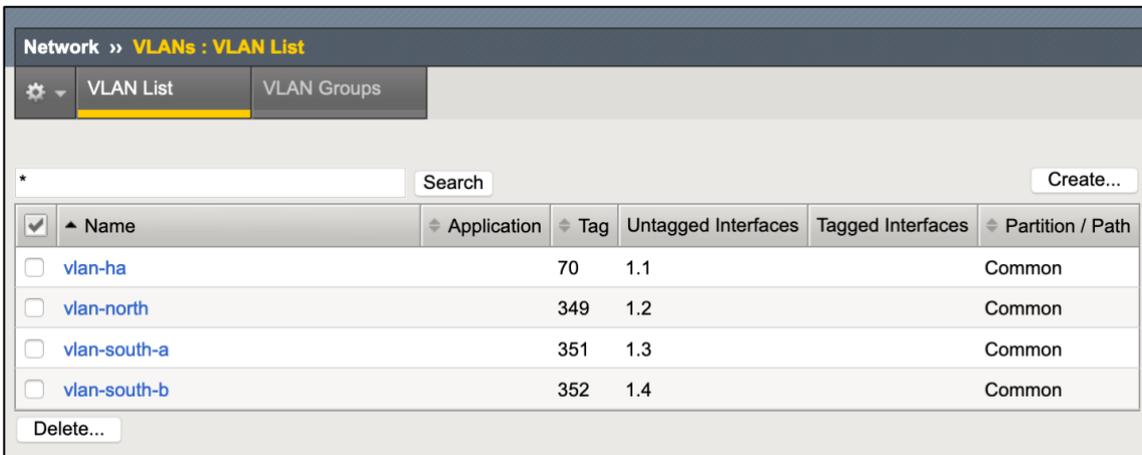


Figure 57 – Adding the NSX-T segment to the BIG-IP is just like a regular untagged VLAN.

Next, create the Self IPs and floating Self IPs towards the Tier-0 Gateways (north-bound) and for the tenants' networks (south-bound). None of these require any special configuration. An example of the first BIG-IP unit is shown next.

<input type="checkbox"/>	Name	Application	IP Address	Netmask	VLAN / Tunnel	Traffic Group	Partition / Path
<input type="checkbox"/>	self-ha		192.174.70.116	255.255.255.0	vlan-ha	traffic-group-local-only	Common
<input type="checkbox"/>	self-north		10.106.49.101	255.255.255.0	vlan-north	traffic-group-local-only	Common
<input type="checkbox"/>	self-north-floating		10.106.49.100	255.255.255.0	vlan-north	traffic-group-1	Common
<input type="checkbox"/>	self-south-a		10.106.51.101	255.255.255.0	vlan-south-a	traffic-group-local-only	Common
<input type="checkbox"/>	self-south-a-floating		10.106.51.100	255.255.255.0	vlan-south-a	traffic-group-1	Common
<input type="checkbox"/>	self-south-b		10.106.52.101	255.255.255.0	vlan-south-b	traffic-group-local-only	Common
<input type="checkbox"/>	self-south-b-floating		10.106.52.100	255.255.255.0	vlan-south-b	traffic-group-1	Common

Figure 58 – Self IPs and floating Self IPs required (shown in BIG-IP unit 1).

Please note that the non-floating Self IPs are per BIG-IP unit whilst the floating Self IPs are synchronized across the BIG-IP units.

The next step is to configure the static routing in the BIG-IP. In this case, it is only required a default route towards the Tier-0 Gateway because all other networks are directly connected. This is shown in the next figure and should be configured in both BIG-IP units (this configuration is not synchronized automatically across BIG-IPs).

<input type="checkbox"/>	Name	Application	Destination	Netmask	Route Domain	Resource Type	Resource	Partition / Path
<input checked="" type="checkbox"/>	default	Default IPv4	Partition Default Route Domain		Gateway	10.106.49.1	Common	

Figure 59 – Static route required in the BIG-IP units.

At this point follow the testing steps described in the Verifying the deployment section.

Details for East-West traffic flows.

As mentioned previously, it is not required to dedicate a subnet for East-West VIPs, in fact BIG-IP Virtual Servers can have one or more IP addresses listening in one or more segments independently of the address. This is exhibit in the implementation diagram where

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

there are shown two Virtual Servers:

- VS1 listens in two VLANs but has a single IP.
- VS2 listens in two VLANs but has two IPs.

These would be implemented as follows

The screenshot shows the VMware NSX-T interface for configuring Virtual Servers. The top part displays a table of Virtual Servers:

Status	Name	Description	Application	Destination	Service Port	Type	Resources	Partition / Path
<input type="checkbox"/>	web-ingress			10.106.49.110	80 (HTTP)	Standard	Edit...	Common
<input type="checkbox"/>	web-eastwest-singleIP	VS1		10.106.51.110	80 (HTTP)	Standard	Edit...	Common
<input type="checkbox"/>	web-eastwest-multipleIP	VS2		example-VS-multipleIP	80	Standard	Edit...	Common
<input type="checkbox"/>	router			Any IPv4	0 (Any)	Forwarding (IP)	Edit...	Common

Below the table, the configuration for the 'example-VS-multipleIP' address list is shown in a modal window:

Shared Objects » Address Lists » example-VS-multipleIP

Properties

Name: example-VS-multipleIP

Description: [Empty field]

Addresses: 10.106.51.120, 10.106.52.120

[Add an Address...] [Add]

[Update] [Delete]

Figure 60 - Example of East-West Virtual Servers with multiple IP addresses

It is important to differentiate the following Virtual Server Settings:

- The destination address of the Virtual server (which is shown in the figure above).
- The segments where the Virtual Server is going to listen (this is independent of the destination address) and it is configured in the BIG-IP by selecting the VLANs where the Virtual server will be enabled or disabled.
- The source address of the Virtual Server which is a set of prefixes which limit the application of the Virtual Server. The main use of this feature is to have a different Virtual Server for the same destination and VLAN combination, and the Virtual Server that applies will depend on the source of the request.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

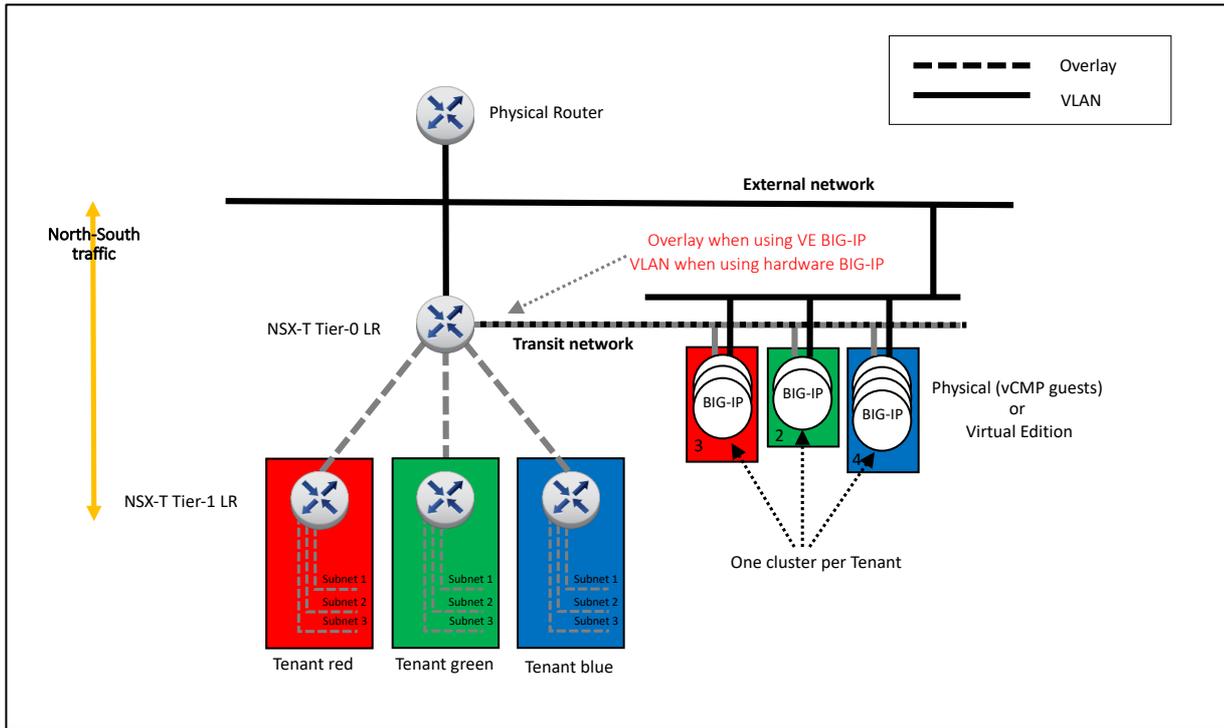


Figure 62 – Topology C with multiple tenants shown.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

This topology has the following benefits:

- Allows direct path to NSX-T which in turn allows NSX-T Edge to perform NAT at Tier-0 without eliminating direct IP visibility from the BIG-IP.
- Allows the deployment of a BIG-IP cluster for different tenants without impacting each other.
- Allows the use of either hardware or virtualized BIG-IPs.

On the other hand, it has the following drawbacks:

- It is a more complex topology, with different paths for the same endpoints.
- Requires SNAT, hiding client's IP addresses.

This topology is suitable for ADC, WAF & Identity management use cases but requires that the direct path is tightly controlled in NSX-T's firewall otherwise security functionalities would be bypassed.

Implementation: BIG-IPs parallel-connected to NSX-T's Tier-0 Gateway.

In the next figure, the configuration which will be implemented in this section is shown.

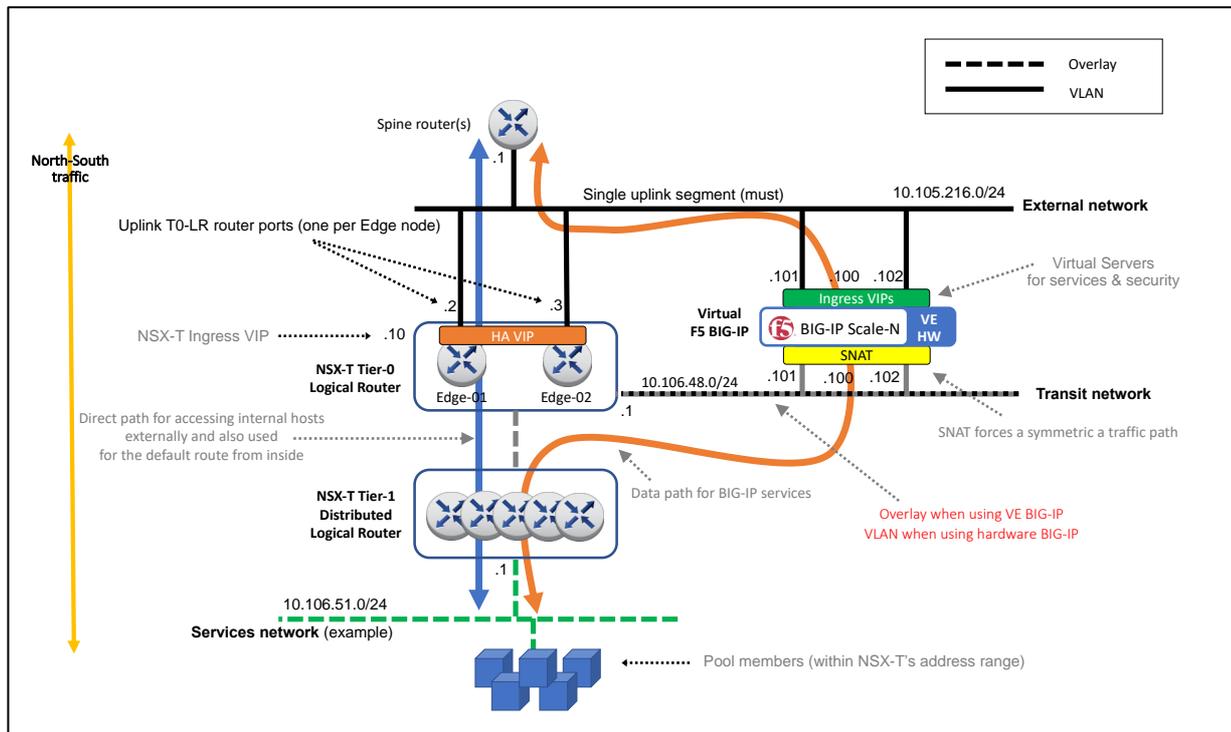


Figure 63 – Topology C example implementation.

In the example used for this topology BIG-IP VE is used which means that the segment between the BIG-IP and the Edge nodes uses the NSX-T overlay. This will be shown in the following configuration. Given the many possibilities of configuring NSX-T Edge nodes and their logical switch uplink ports, it is assumed that these have been already created. This guide is focused in the configuration for the Layer 3 and higher layers that are specific to this topology. See section Design consideration: Layer 2 networking for details.

1. Create the Tier-0 configuration.

1.1. Create a Tier-0 Gateway in Active-Standby mode.

In NSX-T manager, go to `Networking > Tier-0 Gateways > Add Gateway > Tier-0` as shown in the next figure.

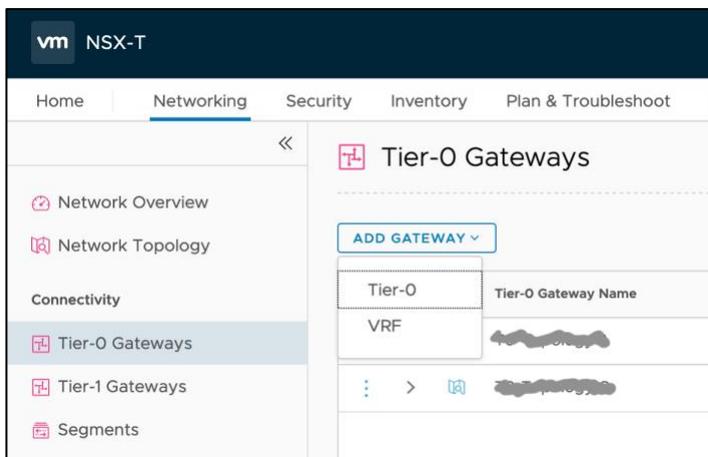


Figure 64 - Adding a Tier-0 Gateway.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

In the New Tier-0 Router dialog, complete the following:

- Name: T0-topology C in this example.
- Edge Cluster: Select the existing Edge cluster.
- High Availability Mode: Active-Standby.
- Failover Mode: Non-Preemptive (to avoid double failover once the failed unit recovers).

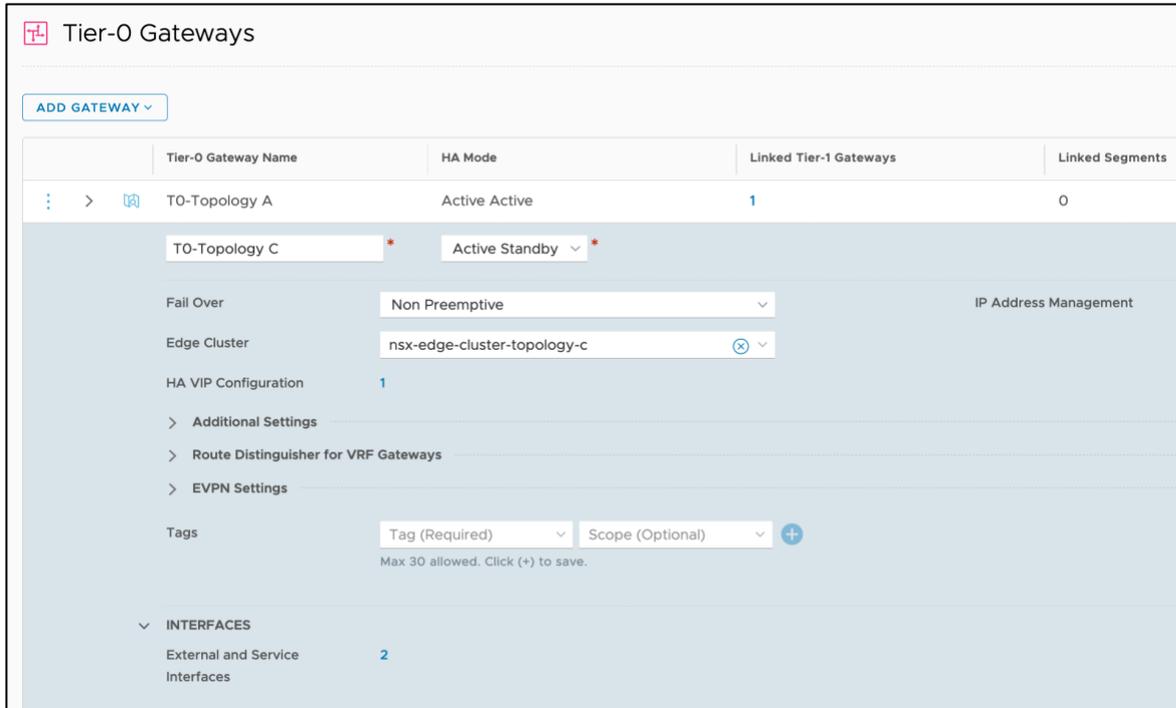


Figure 65 - Filling the details of a Tier-0 Gateway.

1.2. Create an Interface for each Edge Node used by the Tier-0 Gateway.

Select the router created (T0-Topology C in our example) and create two interfaces in the UI by first selecting the Edit option in the T0 Gateway, then scrolling down to the

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

Interfaces section clicking in the `Set` option of `External` and `Service` Interfaces. Enter the following parameters for each interface:

- **Name:** In this example, `edge-1-uplink-vlan216` is used for the first router port and `edge-2-uplink-vlan216` for the second.
- **Type:** `External`
- **Edge Node:** This will be `edge-1-topology-c` and `edge-2-topology-c` for each external interface respectively.
- **MTU:** use external network's MTU, which should be the same on the BIG-IP.
- **URPF Mode:** `Strict` is a good practice providing security with no expected performance impact. `Strict` should be used unless asymmetric paths are used.
- **Segment:** This is the L2 network to which the interface is attached to. It is a prerequisite to have this previously created. See section `Design consideration: Layer 2 networking` for details.
- **IP Address/mask:** this is the IP address assigned to the address port in the shared segment between the NSX-T Edge nodes and the F5 BIG-IPs. In this example, `10.106.53.1/24` is used for router port in `edge-01` and `10.106.53.2/24` in `edge-02`.
- Click `Add`.

Name	Type	IP Address / Mask	Connected To(Segment)	Status
edge-1-uplink-vlan216	External	10.105.216.11/24	vlan-216	

Edge Node: edge-1-topology-c

MTU: Enter MTU (Minimum 64)

PIM: Disabled

ND Profile: default

URPF Mode: Strict

Tags: Tag (Req), Scope (C) (Max 30 allowed. Click (+) to save.)

Buttons: SAVE, CANCEL

Figure 66 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.

Name	Type	IP Address / Mask	Connected To(Segment)	Status
edge-1-uplink-vlan216	External	10.105.216.11/24	vlan-216	Success
edge-2-uplink-vlan216	External	10.105.216.12/24	vlan-216	Success

Figure 67 – Final Uplink interface configuration of the Tier-0 Gateway.

1.3. Create an HA VIP for the Tier-0 Gateway.

The HA VIP is an IP address that will be shared by the two Edge Nodes used for the Tier-0 Gateway created and will be used as the ingress IP to the NSX-T networks.

Select the Router created (T0-Topology A in our example), and create an HA VIP in the UI by selecting `Edit > HA VIP Configuration > Set` and entering the following parameters:

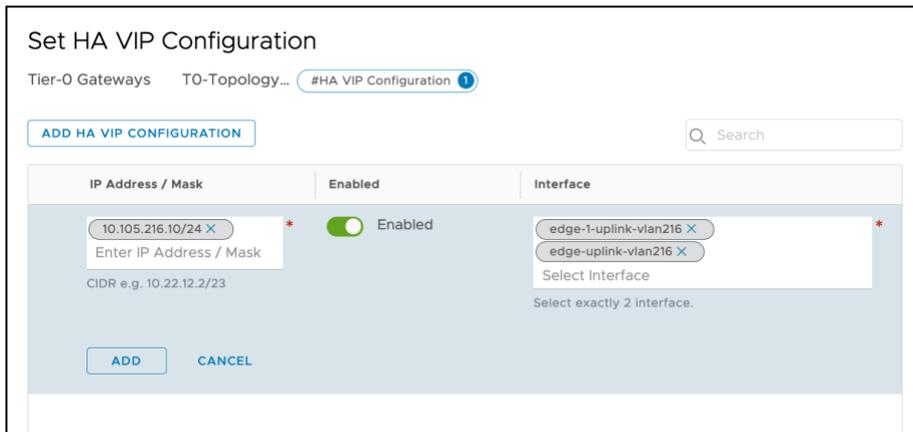


Figure 68 - Adding an HA VIP to NSX-T's T0 Gateway.

Select the two external interfaces just created.

Add a default route in the Tier-0 Gateway towards the BIG-IP cluster floating Self IP address.

In our example, the BIG-IP cluster floating address to use as the next hop is 10.106.53.10. Select the T0-Topology A Gateway created and then create a static routing in the UI by selecting `Routing > Static Routes > Set` as follows and entering as Next Hop BIG-IP's floating-IP, in this example 10.106.216.1:



Figure 69 – Adding Tier-0 Gateway's default route.

1.4. Create the transit network between the Tier-0 Gateway/Edges and the BIG-IP.

1.4.1. Create a segment for the transit network.

Go to `Networking > Segments > ADD SEGMENT` and create a Segment within the Overlay or a VLAN Transport Zone, this will mainly depend if the BIG-IP is a VE or hardware. In this case we are using a VE and the transit network will be in the overlay Transport

Zone. The segment (we use segment-348 in this example) must be attached to the Tier-0 Gateway previously created. This configuration is shown next.

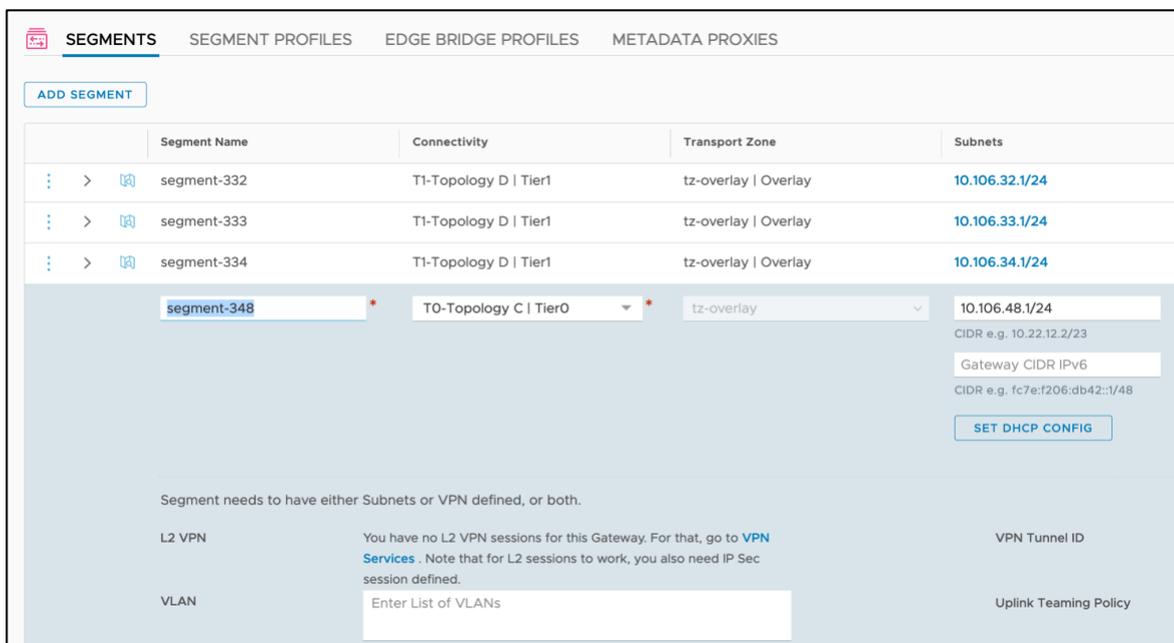


Figure 70 - Creating the Transit segment (segment-348) within the Overlay Transport Zone for a BIG-IP VE

2. Create a Tier-1 Gateway.

Although not part of this topology, this configuration be used later to instantiate a VM and perform a verification of the deployment.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

In NSX-T manager, select `Networking > Tier-1 Gateways > Add Tier-1 Gateway > Tier-1 Router` filling the following parameters:

- Name: In this example, `T1-Topology C`.
- Tier-0 Router: Select the Tier-0 router (`T0-Topology C` in our example).
- Edge Cluster: The name of the Edge Cluster of the NSX-T Edge nodes being used.
- Failover Mode: `Non-Preemptive` (to avoid double failover once the failed unit recovers).
- Route Advertisement: at least “All Connected Segments [...]” should be enabled.
- Click Add.

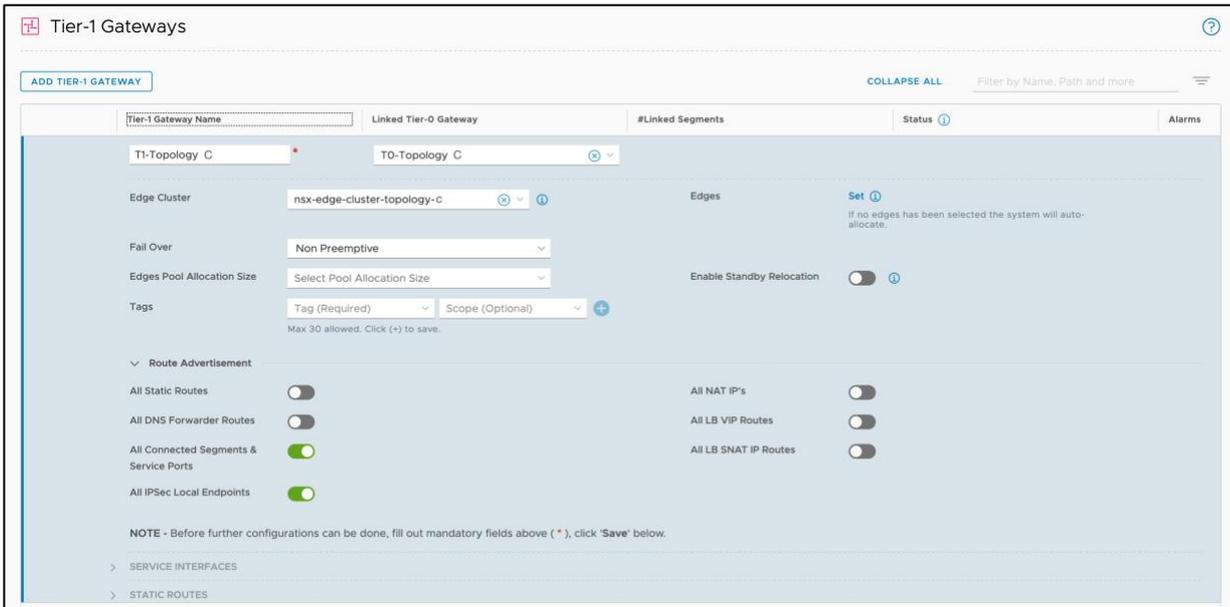


Figure 71 – Filling the properties when creating a Tier-1 Gateway.

The next step is to create a network attached to this Tier-1 Gateway. In the UI, select `Networking > Segments > Add Segment` and enter the following parameters:

- Segment Name: in this example, `segment-351`.
- Connectivity: the Tier-1 Gateway, in this case `T1-Topology C`.
- Subnets: this really indicates both the subnet and the IP address of the Tier-1 Gateway in this segment, in this case `10.106.51.1/24`

This configuration can be seen in the next figure:

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

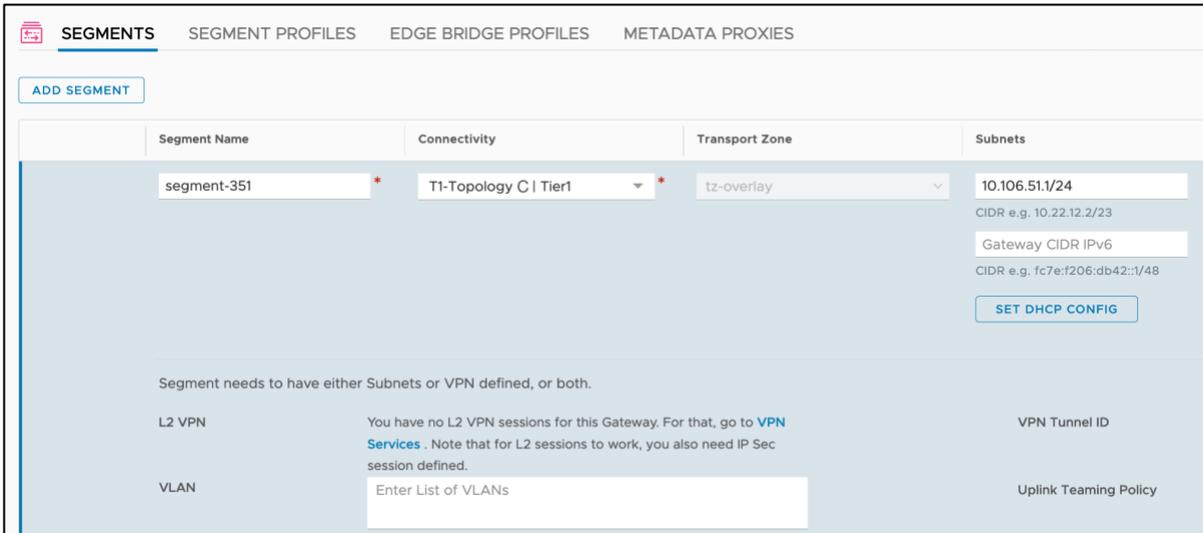


Figure 72 - Adding a segment to the T1 Gateway.

3. Create the Layer 3 configuration on the BIG-IP.

In this example, we are using BIG-IPs VE and for the transit network NSX-T overlay segments. The configuration used in this example is shown next:

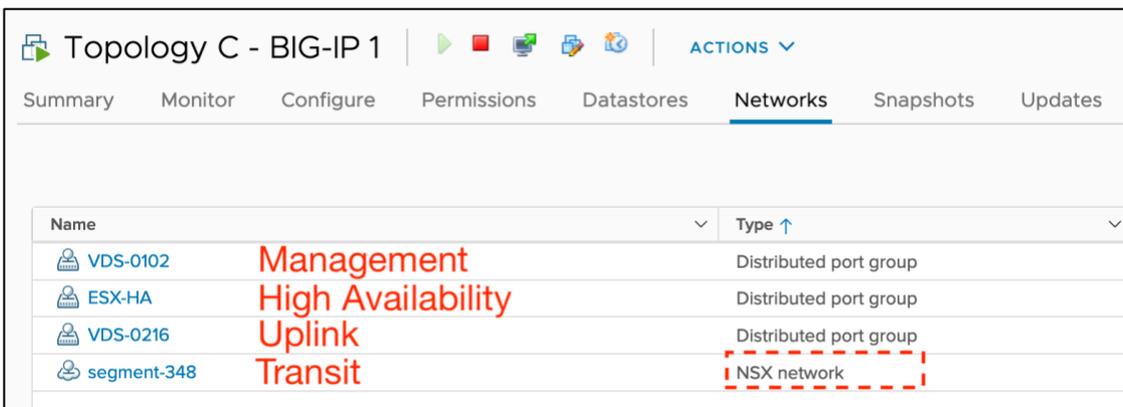


Figure 73 - Attaching the BIG-IP to an NSX-T overlay segment for the transit network.

The BIG-IP will make use of all these networks just like any regular untagged VLAN as shown in the next figure:

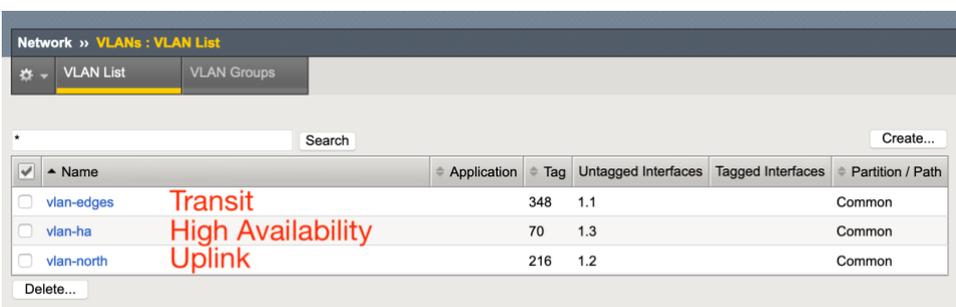


Figure 74 – Adding the Layer 2 networks to the BIG-IPs.

Next, create the Self IPs and floating Self IPs towards the spine routers (north-bound) and towards the NSX-T networks (south-bound) through the NSX-T Tier-0 Gateway's transit

network. These do not require any special configuration. An example of the first BIG-IP unit is shown next.

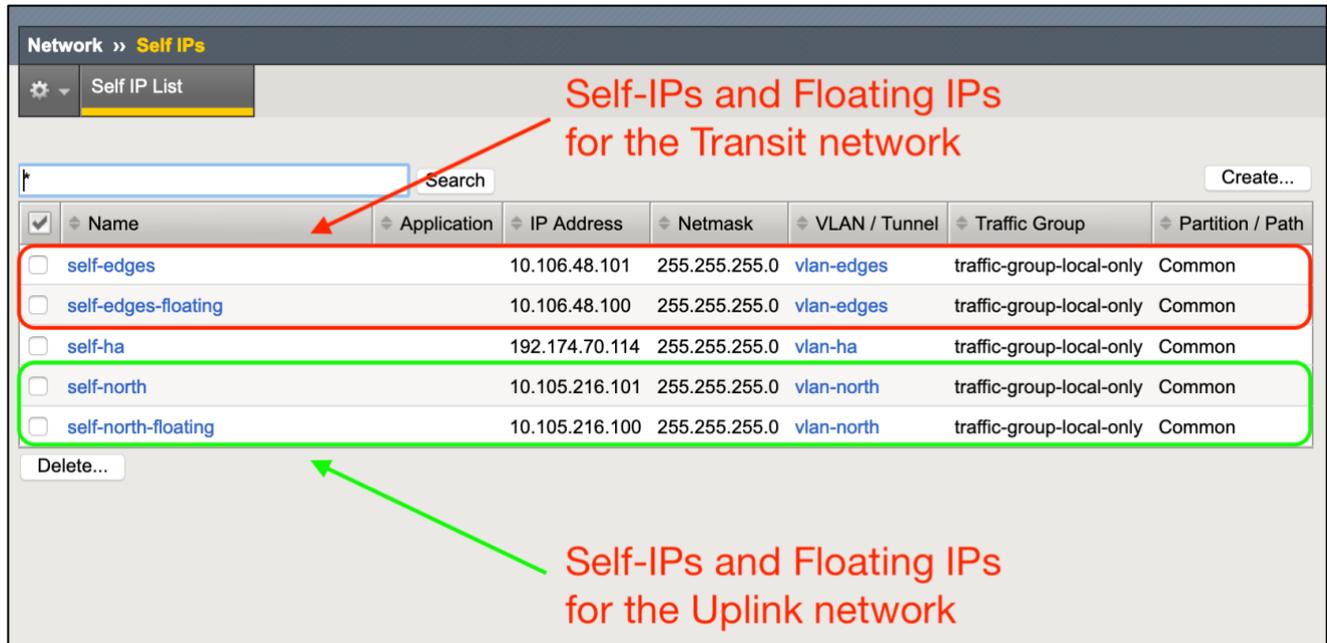


Figure 75 – Self IPs and floating Self IPs required (shown in BIG-IP unit 1).

Note that the non-floating Self IPs are per BIG-IP unit while the floating Self IPs are synchronized across the BIG-IP units.

The next step is to configure the static routing on the BIG-IP. Typically, these involve two routes:

- A default route using spine router as gateway.
- A route towards the NSX-T IP address range using the IP address of NSX-T’s Tier-0 transit network as gateway.

These routes can be shown in the next figure and should be configured in both BIG-IP units (this configuration is not synchronized automatically across BIG-IPs).

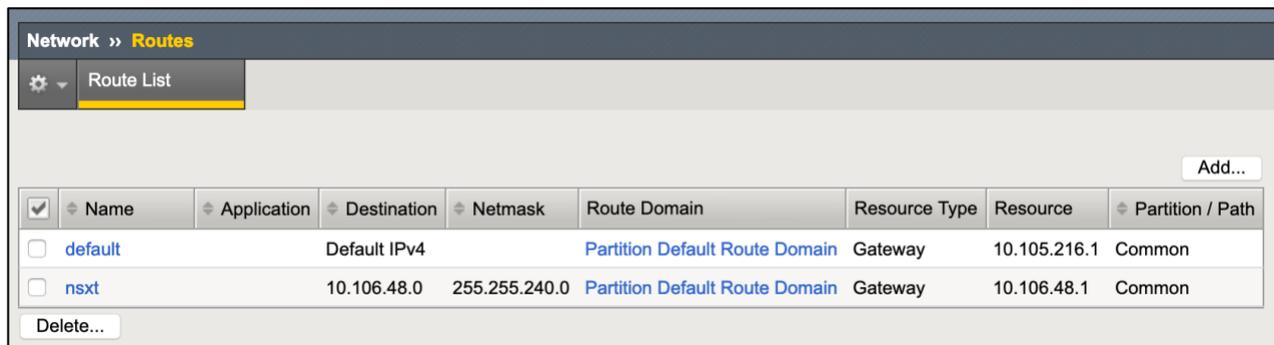


Figure 76 - Static routes required on the BIG-IP units.

At this point, follow the testing steps described in the Verifying the deployment section.

Topology D: BIG-IPs parallel-connected to NSX-T's Tier-1 Gateway.

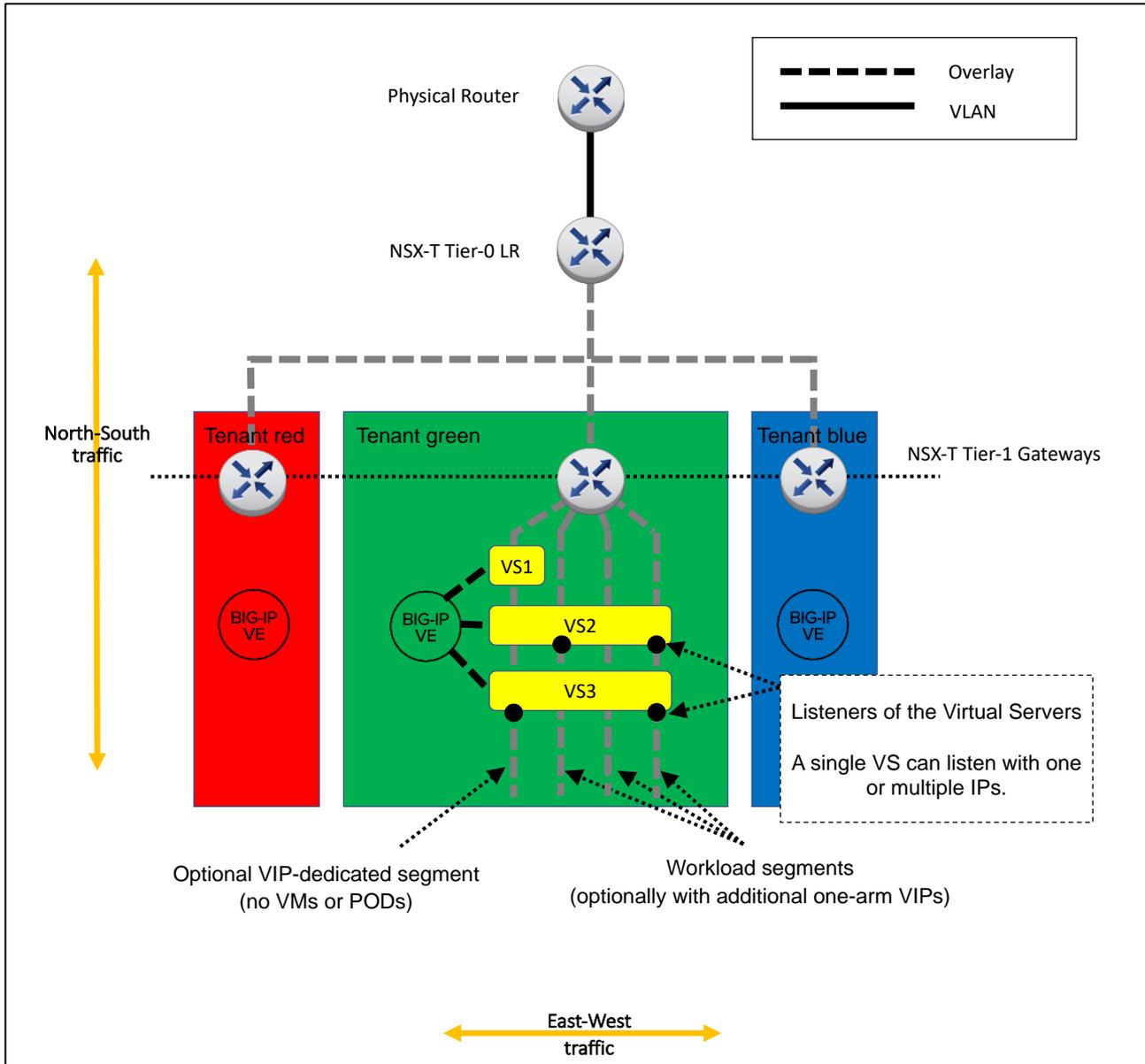


Figure 77 - Topology D overview (simplified view without HA components).

The ideal scenario to handle East-West traffic is to have a BIG-IP cluster for each tenant. This is aligned with VMware's vision where the Tier-1's domain can be managed by each tenant. The benefits of using BIG-IP for centralized management and visibility are more relevant in this topology. Additionally, having several BIG-IP clusters distributes the workload across the ESXi hypervisors unlike NSX-T's LBs, which might be more limited running in NSX-T Edge's hosts only.

In the next figure, an implementation example of this topology is shown, which describes the flows for North-South traffic:

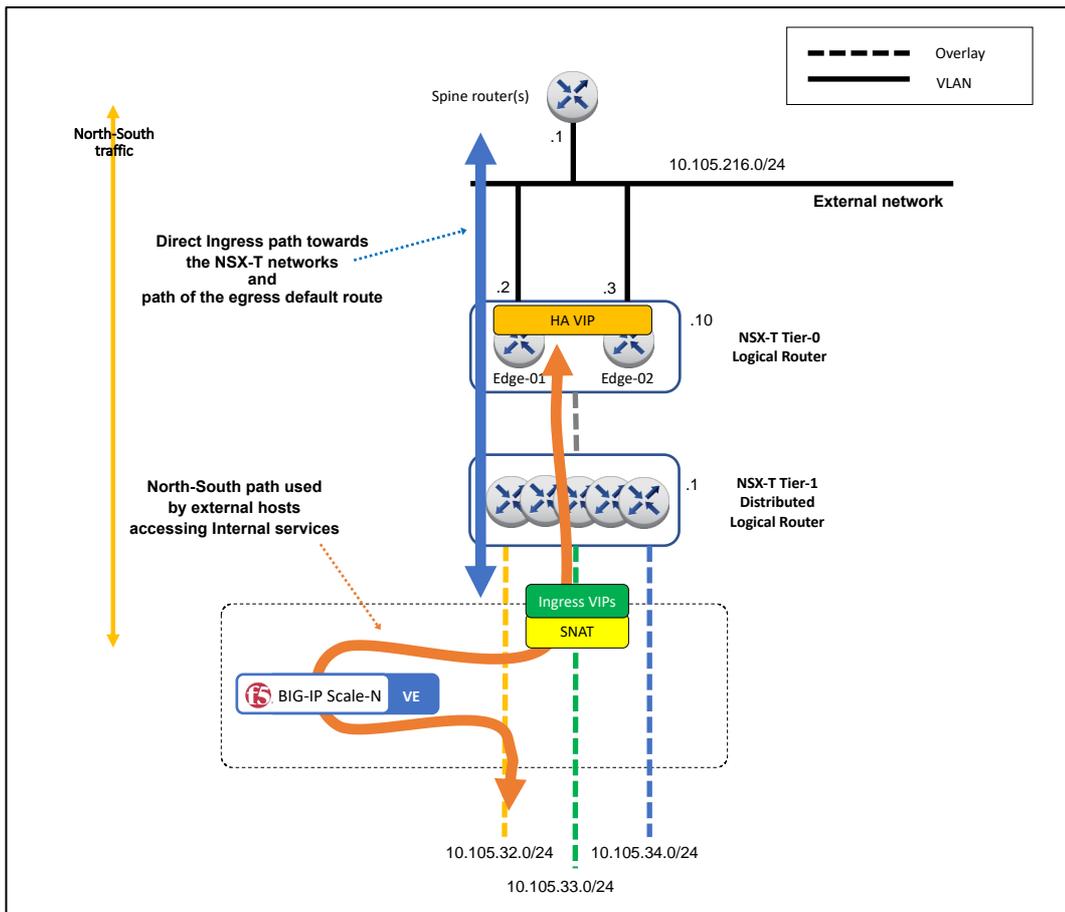


Figure 78 – Topology D example implementation – North/South traffic flows.

Two North-South traffic flows can be seen in the figure:

- Ingress traffic through the Tier-0 Gateway direct to the workload servers (blue color), either from outside the NSX-T environment (shown in the figure) or from another tenant (not shown). This traffic reaches the VMs directly, no LB or services are applied to it. No SNAT is required. Normally, these flows are not allowed freely and filtering rules are set in the NSX-T's firewall.
- Ingress traffic reaching tenant's services (orange color). The VIPs might be in a given subnet and the workload servers in any other subnet. The traffic doesn't go through the Tier-1 Gateway twice.

In the next figure, an implementation example of this topology is shown, this time describing the flows for East-West traffic:

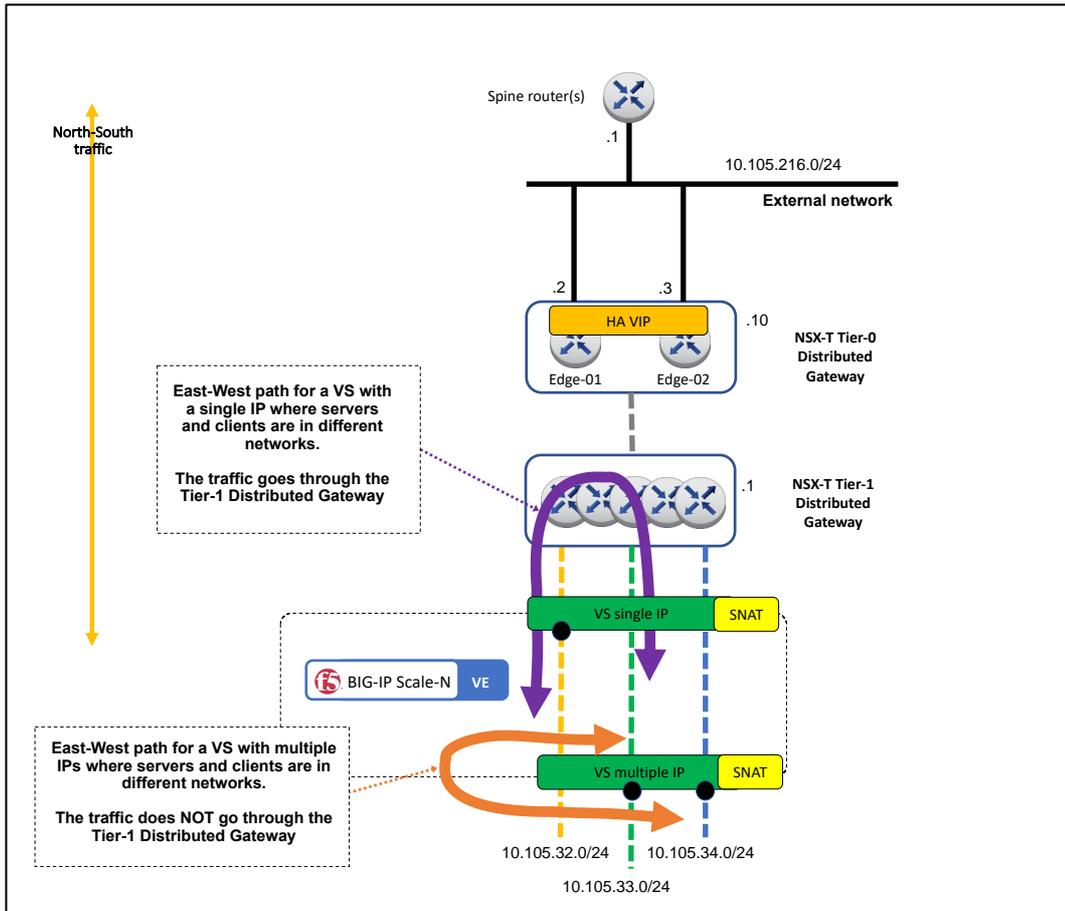


Figure 79 – Topology D example implementation – East/West traffic flows.

In the figure above we can differentiate two East-West flows within the same Tenant (within the routing scope of a Tier-1 Gateway):

- The purple flow shows a typical Virtual Server with a single IP address (VIP). The flow outlined is between segments orange and green. The VIP belongs to segment orange and the client is in the green segment. In order for the client to reach the VIP it has to go through the Tier-1 Gateway. This is an efficient path though because Layer 3 processing is distributed.
- The orange flow shows a Virtual server with two IP addresses (VIPs), one in segment green and another in segment blue. This arrangement allows that regardless the clients are in segment green or blue, they never have to go through the Tier-1 Gateway. This improves performance and simplifies the traffic flows.

Please note that in both Virtual Server configurations SNAT is required to avoid Direct Server Return (DSR) which would not allow for proxy based advanced services. DSR is out of scope of this guide.

Additionally different Virtual Servers with the same destination IP/port can be implemented by using the `Source Address` setting in the Virtual Servers.

Local Traffic » Virtual Servers : Virtual Server List » webservers

Properties Resources Security Statistics

General Properties

Name	webservers
Partition / Path	Common
Description	
Type	Standard
Source Address	<input checked="" type="radio"/> Host <input type="radio"/> Address List 0.0.0.0/0
Destination Address/Mask	<input checked="" type="radio"/> Host <input type="radio"/> Address List [Redacted]
Service Port	<input checked="" type="radio"/> Port <input type="radio"/> Port List 80 HTTP
Notify Status to Virtual Address	<input checked="" type="checkbox"/>
Availability	<input checked="" type="radio"/> Available (Enabled) - The virtual server is available

Figure 80 – Source Address setting to discriminate the prefixes to which the Virtual Server applies.

Although topology D can be used for both North-South and East-West traffic, it is important to note that this topology can be combined with Topology A. In such combined scenario Topology D would be used only for East-West traffic within a tenant (and could be managed by each

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

tenant) and Topology A could be used for North-South flows. An example of this combined topology is shown in Figure 81.

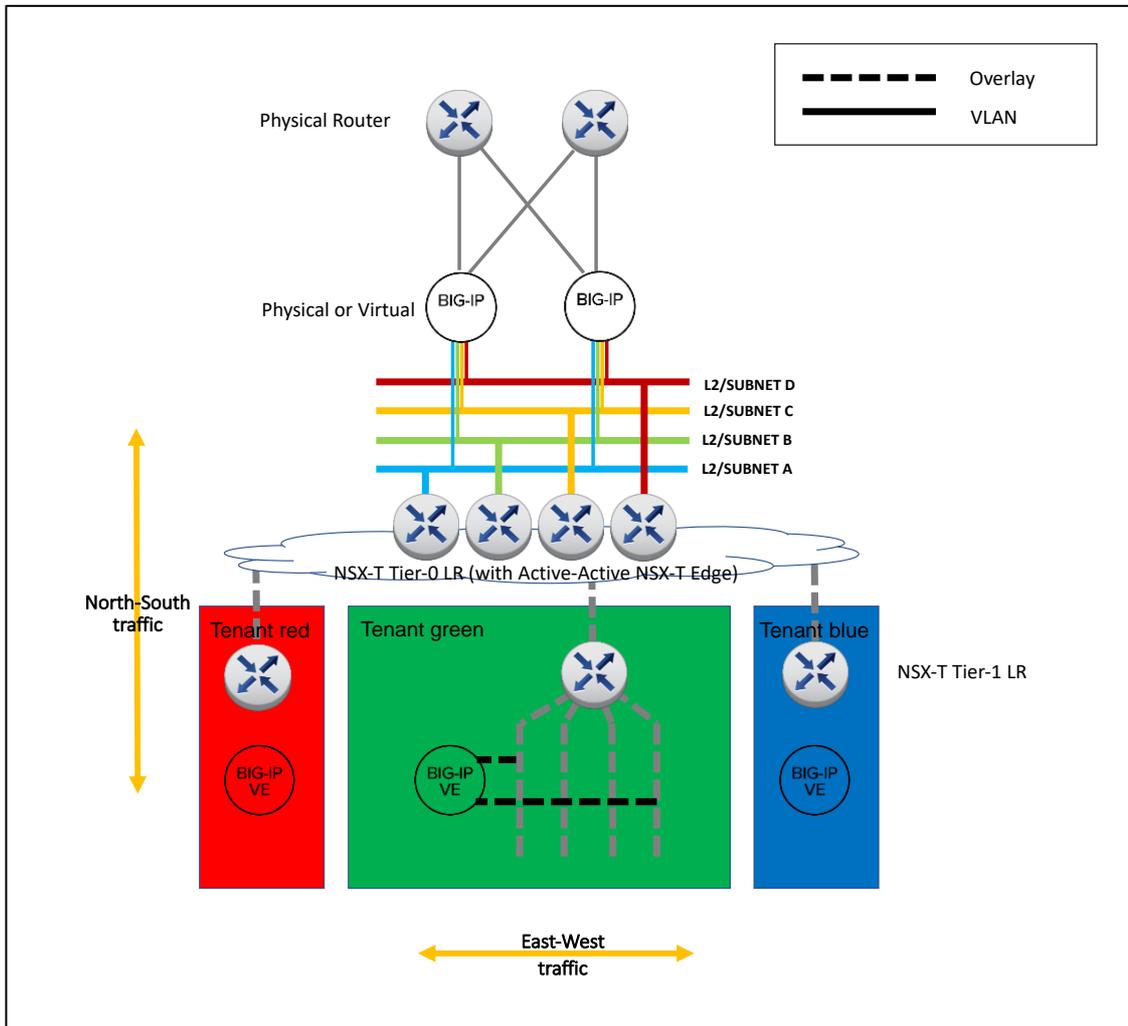


Figure 81 - Combined Topology A and D.

Implementation: BIG-IPs parallel-connected to NSX-T's Tier-1 Gateway.

Figure 82 shows the configuration which is implemented in this section.

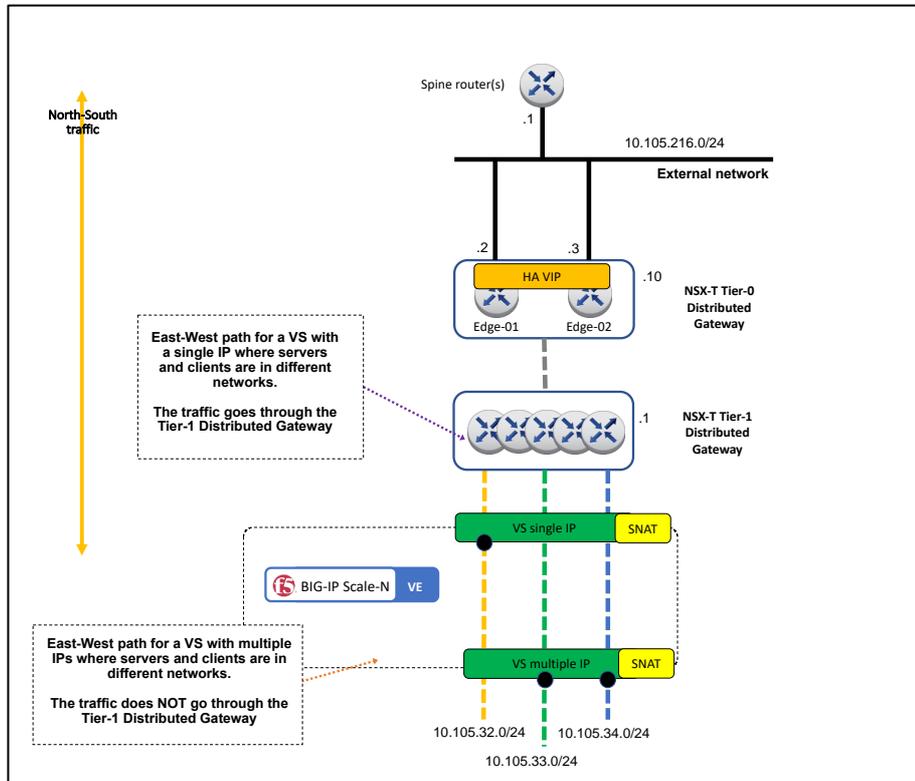


Figure 82 - Topology D implementation used through this section.

Note that in this example topology that there is no virtual server for the egress traffic. The outbound traffic from the internal hosts is routed directly to the Tier-1 Gateway. If the deployment requires an egress VIP to install advanced services such as Web Gateway this would be better using any of the inline topologies (topology A or C).

The configuration steps are described next and we start with the previously existing Tier-0 Gateway of topology A, to which we will attach the Tier-1 Gateway. There is no limitation in the Tier-0 Gateway chosen.

1. Create a Tier-1 Gateway.

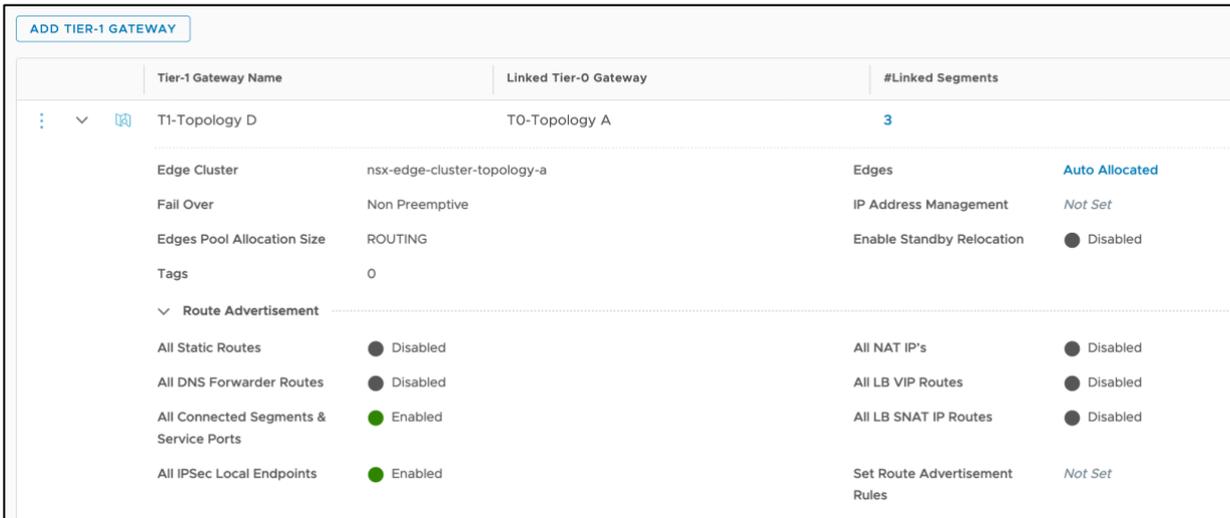
This Tier-1 Gateway will have a transit network towards Tier-0 (automatically created) and in this example 3 user segments in the overlay transport zone (orange, green and blue).

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

In NSX-T manager, select `Networking > Tier-1 Gateways > Add Tier-1 Gateway > Tier-1 Router` filling the following parameters:

- Name: In this example, `T1-Topology D`.
- Tier-0 Router: Select the Tier-0 router (`T0-Topology A` in our example).
- Edge Cluster: The name of the Edge Cluster of the NSX-T Edge nodes being used.
- Failover Mode: `Non-Preemptive` (to avoid double failover once the failed unit recovers).
- Route Advertisement: at least “All Connected Segments [...]” should be enabled.
- Click Add.



Tier-1 Gateway Name	Linked Tier-0 Gateway	#Linked Segments	
T1-Topology D	T0-Topology A	3	
Edge Cluster	nsx-edge-cluster-topology-a	Edges	Auto Allocated
Fail Over	Non Preemptive	IP Address Management	Not Set
Edges Pool Allocation Size	ROUTING	Enable Standby Relocation	Disabled
Tags	0		
Route Advertisement			
All Static Routes	Disabled	All NAT IP's	Disabled
All DNS Forwarder Routes	Disabled	All LB VIP Routes	Disabled
All Connected Segments & Service Ports	Enabled	All LB SNAT IP Routes	Disabled
All IPsec Local Endpoints	Enabled	Set Route Advertisement Rules	Not Set

Figure 83 – Filling the properties when creating a Tier-1 Gateway.

The next step is to create the orange, green and blue networks and attach them to this Tier-1 Gateway. In the UI, select `Networking > Segments > Add Segment` and enter the following parameters:

- Segment Name: in this example `segment-332`, `segment-333` and `segment-333` respectively.
- Connectivity: the Tier-1 Gateway, in this case `T1-Topology D`.
- Subnets: this really indicates both the subnet and the IP address of the Tier-1 Gateway in this segment, in this case `10.106.{32,33,34}.1/24`

This configuration can be seen in the next figure:

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

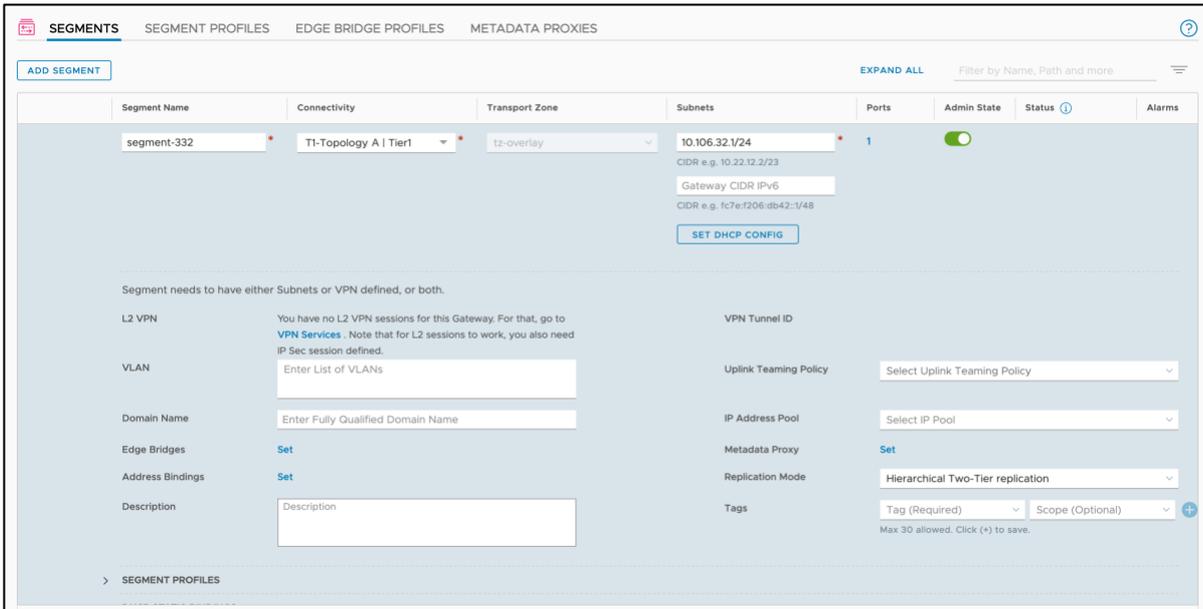


Figure 84 - Adding a segment to the T1 Gateway.

2. Create the Layer 3 configuration in the BIG-IP

First, create the Self IPs and floating Self IPs in the VIP segment that are attached to the Tier-1 Gateway. These do not require any special configuration. An example of the first BIG-IP unit is shown in Figure 85.

Network » Self IPs							
Self IP List							
* <input type="text"/> Search Create...							
<input checked="" type="checkbox"/>	Name	Application	IP Address	Netmask	VLAN / Tunnel	Traffic Group	Partition / Path
<input type="checkbox"/>	self-ha		192.174.70.112	255.255.255.0	vlan-ha	traffic-group-local-only	Common
<input type="checkbox"/>	self-south-a		10.106.31.11	255.255.255.0	vlan-south-a	traffic-group-local-only	Common
<input type="checkbox"/>	self-south-a-floating		10.106.31.10	255.255.255.0	vlan-south-a	traffic-group-1	Common
<input type="checkbox"/>	self-south-b		10.106.32.11	255.255.255.0	vlan-south-b	traffic-group-local-only	Common
<input type="checkbox"/>	self-south-b-floating		10.106.32.10	255.255.255.0	vlan-south-b	traffic-group-1	Common
<input type="checkbox"/>	self-south-c		10.106.33.11	255.255.255.0	vlan-south-c	traffic-group-local-only	Common
<input type="checkbox"/>	self-south-c-floating		10.106.33.10	255.255.255.0	vlan-south-c	traffic-group-1	Common

Figure 85 – Self IPs and floating Self IPs required (shown in BIG-IP unit 1).

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

Note that the non-floating Self IPs are per BIG-IP unit while the floating Self IPs are synchronized across the BIG-IP units.

We will use a default route to reach the non-directly connected networks. We will use the first self-IP to reach the Tier-1 Gateway. This is shown in Figure 86:

Properties	
Name	default
Description	
Destination	0.0.0.0
Netmask	0.0.0.0
Resource	Use Gateway...
Gateway Address	IP Address 10.106.32.1
MTU	

Cancel Repeat Finished

Figure 86 – Static routes required in the BIG-IP units.

At this point, follow the testing steps described in the Verifying the deployment section.

VMware Cloud on AWS

Introduction

VMware Cloud (VMC) on AWS provides NSX-T networking with several restrictions. Among these, one of the most relevant is that it constrains the users to using only one Tier-1 Gateway (Compute Gateway in VMC nomenclature) per Tier-0 Gateway⁶. Besides the limitations compared to a native NSX-T it provides the following advantages:

- It allows to deploy Data Centers on demand (SDDC – Software Defined Data Center) on AWS infrastructure.
- VMC is deployed within an AWS VPC (Virtual Private Cloud) which allows simple access to AWS services such as *Direct Connect* or additional user compute in EC2.
- Analogously to the previous item, the EC2 compute resources in the VPC can also make use of the VMC deployment. The VPC and the VMC deployment are connected using plain routing.

The next picture shows a scenario where we have two VMware deployments, one of them within VMC where we also make use of additional EC2 compute resources within the same VPC where the SDDC is.

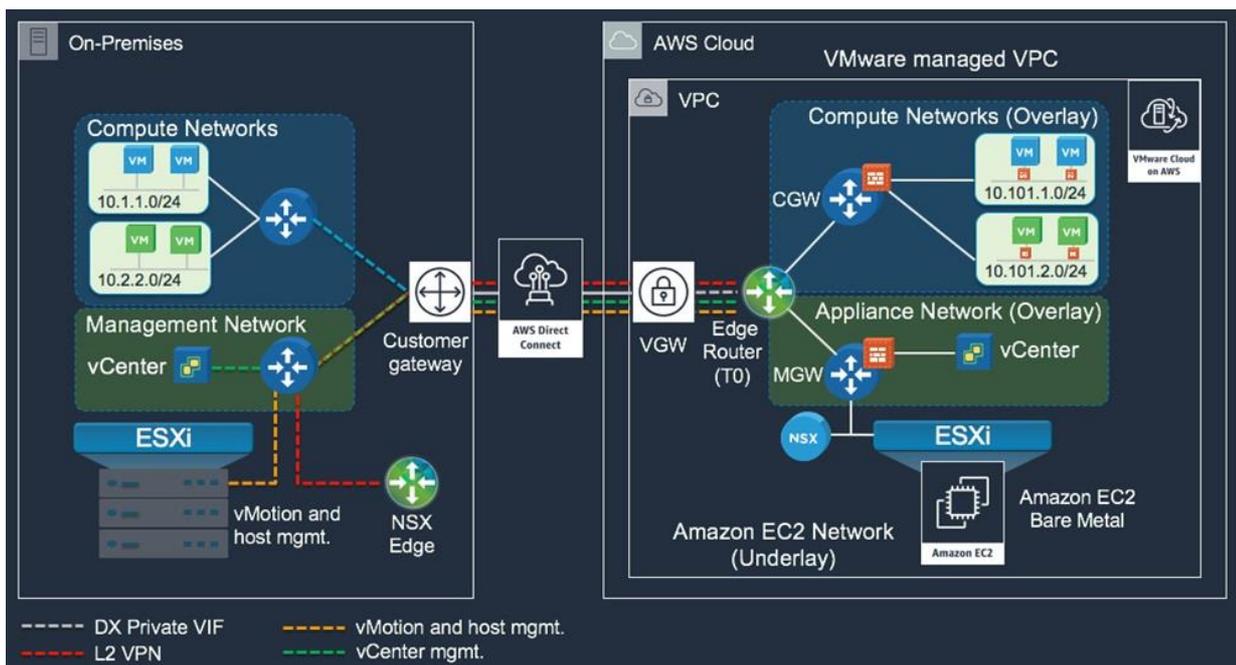


Figure 87 - Sample architecture showing some connectivity options

In this figure, we can see that the user in VMC is restricted to the Compute Networks in AWS (top right of the picture) which can only be connected to the CGW (a T1 Gateway). Given this

⁶ Starting with VMC on AWS's SDDC version 1.12 it is possible to have more than one Tier-0 Gateways using the so-called Multi-Edge SDDC topology but this is out of scope of this guide.

constraint, we will limit the proposed topologies to a modified Topology D which makes use of SNAT. We will also mention alternatives to avoid the use of SNAT.

Sample topology D for VMC on AWS – VMC configuration

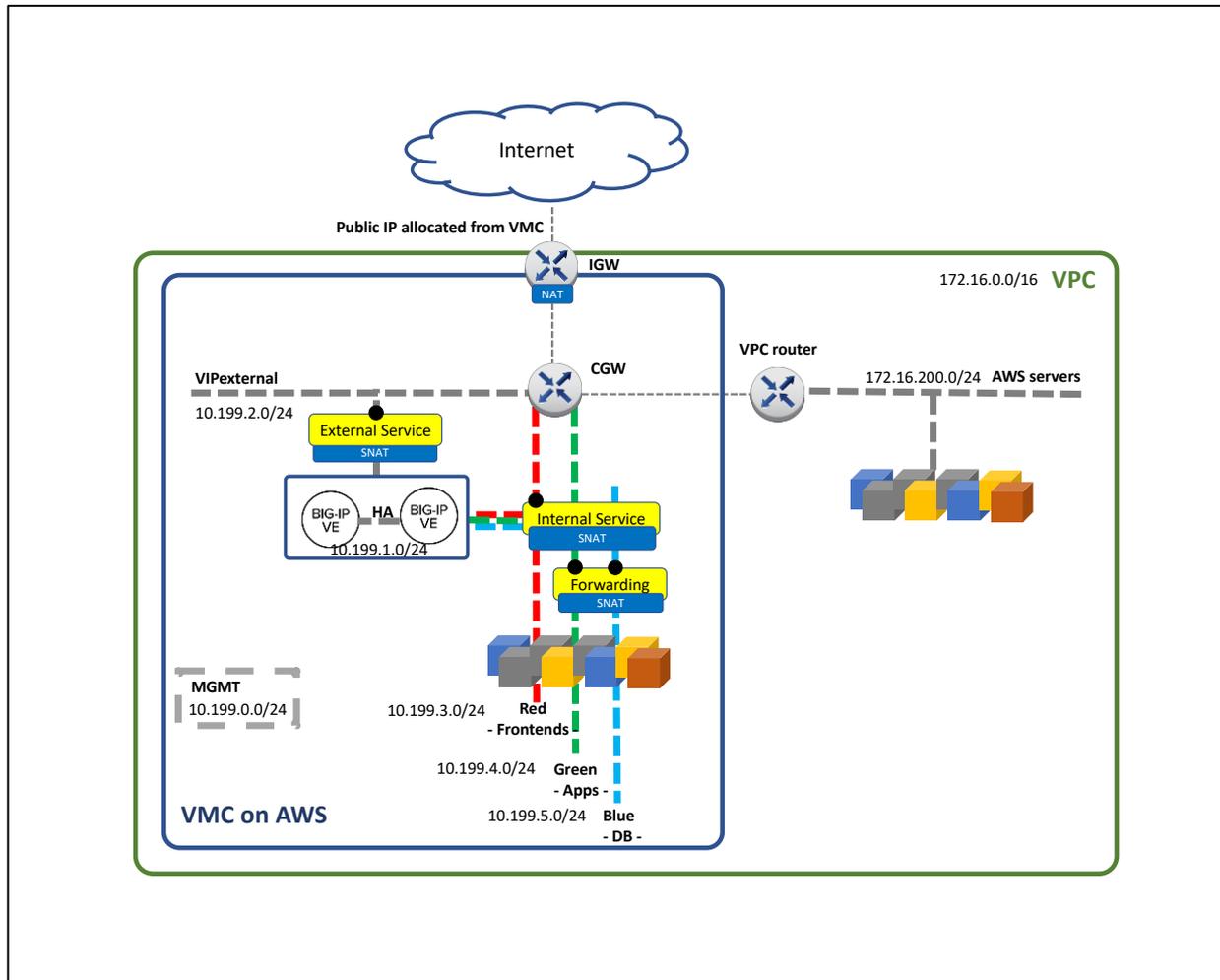


Figure 88 - Sample Topology D for VMC

In this sample topology, we create a typical 3-tier architecture with Frontend (External Service), Application (Internal Service) and Database tiers. Notice that the Database Tier is configured as “Disconnected” to provide an additional layer of secure by means of controlling the access through a VIP in the BIG-IP. The created segments can be seen in the next figure.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

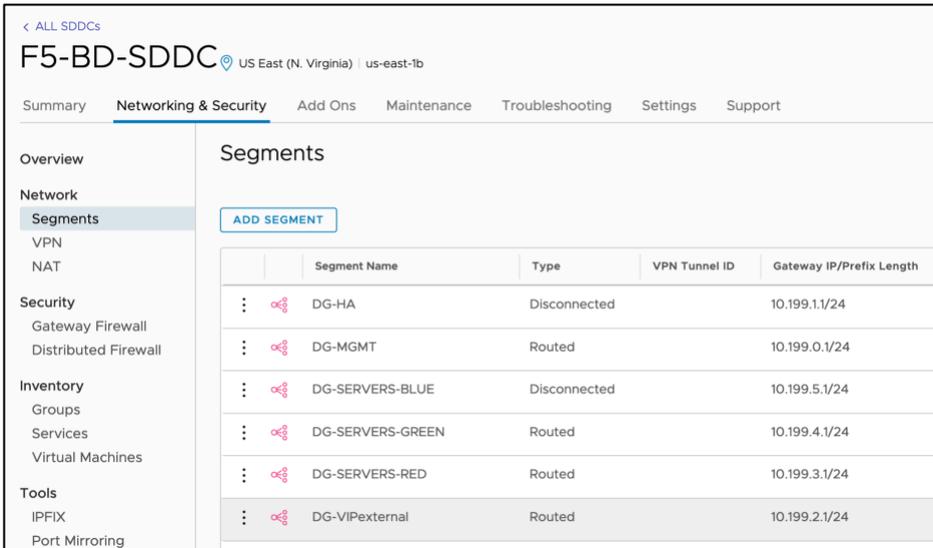


Figure 89 - Segments configuration in VMC

It is worth noting that VMC does not allow creating custom segment profiles, which inhibits the use of MAC Masquerading mechanism. See the subsection TODO-MAC Masquerading for more details.

The VPC in which the VMC deployment is hosted can be checked from the VMC console as shown in the next figure.

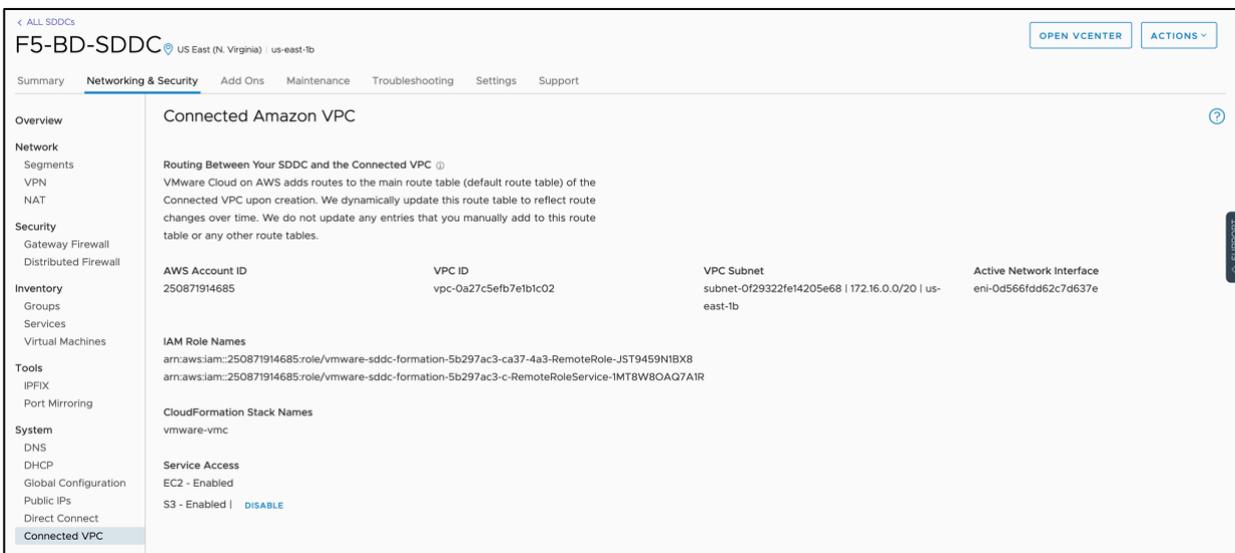


Figure 90 - Checking AWS VPC from the VMC console

If we want to check the routing table of the VPC, we need to use the AWS console. When we add new segments in VMC, routes will be automatically populated in the VPC router to provide connectivity from the non-VMC environment towards the VMC environment. We can see the configuration of this example in the next figure:

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

The screenshot shows the AWS Management Console interface for a VPC. At the top, there's a search bar with the VPC ID 'vpc-0a27c5efb7e1b1c02'. Below that, a table lists VPC route tables. The selected route table is 'rtb-0e55ba0579ab82544'. The 'Routes' tab is active, showing a list of routes. The 'View' dropdown is set to 'All routes'.

Destination	Target	Status
172.16.0.0/16	local	active
10.2.0.0/16	eni-0d566fdd62c7d637e	active
10.195.195.0/24	eni-0d566fdd62c7d637e	active
10.199.0.0/24	eni-0d566fdd62c7d637e	active
10.199.2.0/24	eni-0d566fdd62c7d637e	active
10.199.3.0/24	eni-0d566fdd62c7d637e	active
10.199.4.0/24	eni-0d566fdd62c7d637e	active

Figure 91 - Automatically created routing table of the VPC

Please note that this routing table is independent of the routing table within VMC. We can see this because the only VPC owned route/non-VMC owned route is marked as *local* in the **Target** column.

Lastly, we will configure a public address for the VMC deployment. This public address can be used as egress and ingress point for the non VMC deployment.

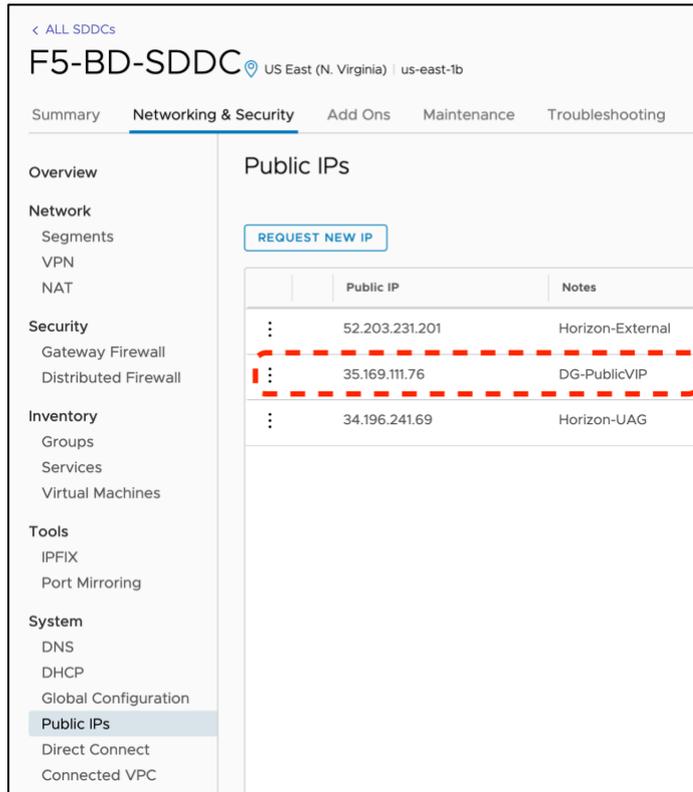


Figure 92 - Allocating an IP address for the VIP in the BIG-IPs

This public IP needs to be mapped into the VIP of the BIG-IP that we will configure later on. This is done by a 1:1 NAT which happens in the IGW of the VMC SDDC and is configured in the VMC console as shown in the next figure, where 10.199.2.100 will be the VIP in the BIG-IPs.

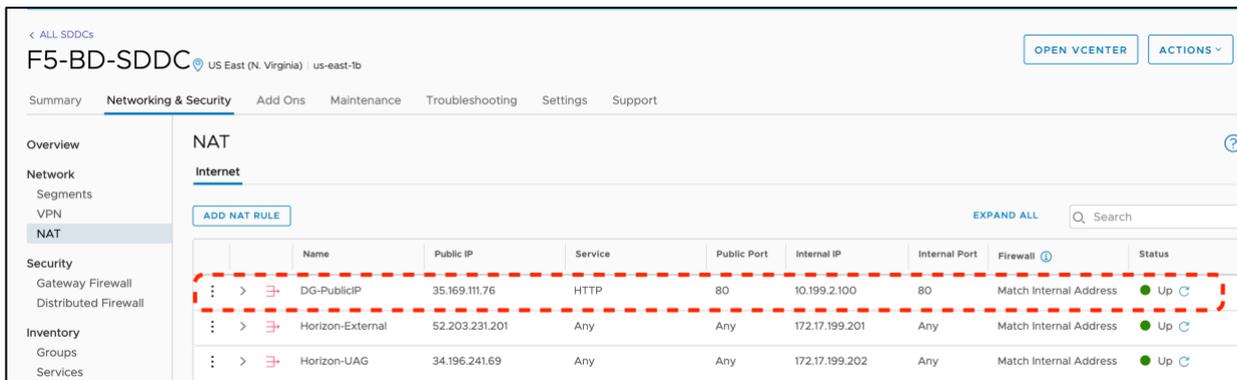


Figure 93 - Configuring the required 1:1 NAT for the BIG-IP VIP.

Sample topology D for VMC on AWS – BIG-IP configuration

The configuration in the BIG-IPs for this topology is a standard configuration, nothing differs from the Topology D shown in previous sections. Next it will be described the L3 configuration and then the Service configuration.

It has floating-IPs configured for all subnets with the exception of the HA segment but strictly speaking the floating-IP is only required for the blue segment used for the Database Tier which is disconnected from the CGW (NSX Tier-1 Gateway) and we use the BIG-IP as the default gateway, for an additional layer of security. The Frontend-Tier and the App-Tier use the CGW

as their default gateway. For the non-floating Self-IPs we use .11 for the BIG-IP unit #1 and .12 for the BIG-IP unit #2.

Name	Application	IP Address	Netmask	VLAN / Tunnel	Traffic Group	Partition / Path
<input type="checkbox"/> vip-floating		10.199.2.10	255.255.255.0	vip	traffic-group-1	Common
<input type="checkbox"/> vip		10.199.2.11	255.255.255.0	vip	traffic-group-local-only	Common
<input type="checkbox"/> red-floating		10.199.3.10	255.255.255.0	red	traffic-group-1	Common
<input type="checkbox"/> red		10.199.3.11	255.255.255.0	red	traffic-group-local-only	Common
<input type="checkbox"/> ha		10.199.1.11	255.255.255.0	ha	traffic-group-local-only	Common
<input type="checkbox"/> green-floating		10.199.4.10	255.255.255.0	green	traffic-group-1	Common
<input type="checkbox"/> green		10.199.4.11	255.255.255.0	green	traffic-group-local-only	Common
<input type="checkbox"/> blue-floating		10.199.5.1	255.255.255.0	blue	traffic-group-1	Common
<input type="checkbox"/> blue		10.199.5.11	255.255.255.0	blue	traffic-group-local-only	Common

Figure 94 - Directly connected segments. Self-IP configuration.

The connectivity to the non-directly connected segments, including the AWS workload segments in the VPC, is done by a single default route as shown next.

Name	Application	Destination	Netmask	Route Domain	Resource Type	Resource	Partition / Path
<input type="checkbox"/> default		Default IPv4		Partition Default Route Domain	Gateway	10.199.2.1	Common

Figure 95 - Routing required for non-directly connected segments, including AWS workload segments in the VPC.

For the service configuration the following setup is required:

- A VS for the Frontend (named Frontend) for which we previously configured the public IP and the 1:1 NAT.
- A VS for the App using the VMC compute (named App).
- A VS for an additional App using the AWS compute in the VPC (named AppAWS).
- A VS for forwarding between the App Tier and the DB Tier (named Forwarding).

All these VS with the exception of the forwarding VS are enabled only in the segment where the address belongs.

In the case of the Forwarding VS, it is enabled in the App Tier and the DB tier to allow traffic initiated from either of the two segments. The BIG-IP can be configured with additional controls to enhance the security between these two segments.

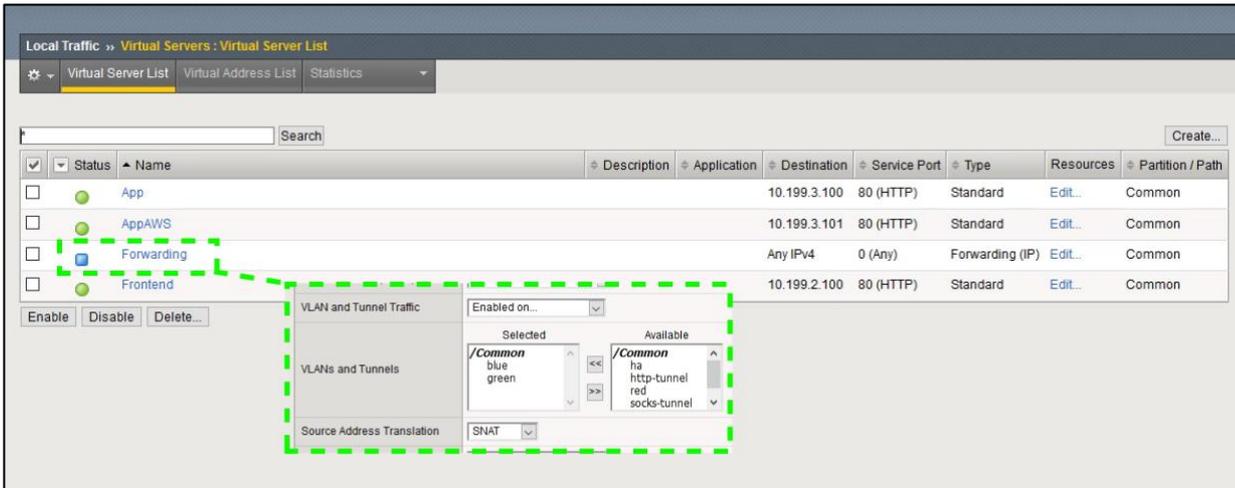


Figure 96 - Overview of the service configuration with detailing the additional segments where the Forwarding VIP is enabled.

Alternative topologies for BIG-IP in VMC on AWS

It is possible to configure services in the BIG-IP without SNAT but this requires that the servers are configured with the BIG-IP as their default gateway. In this scenario, the non-service traffic, just plain routed, is more complex because the traffic will be asymmetric (egress traffic will go through the BIG-IP and ingress traffic directly to the segment. Such an asymmetric forwarding Virtual Server can be configured in the BIG-IP if necessary.

Once VMC supports either modifying the routing table of the CGW or allows overlapping addresses with disconnected segments there are ways to do not require SNAT. When either of these features are available in VMC this guide will be updated with a non-SNAT topology.

Hybrid and Multi-cloud Design Considerations

Introduction and Use Cases

Multi-cloud allows for several use cases:

- High Availability by means of DC redundancy and Disaster Recovery.
- Load distribution and operational flexibility for continuous delivery.
- Traffic optimization bringing content closer to the customer.
- Regulatory compliance for data retention.
- Cloud Bursting.

As a consequence, many designs are possible. Ultimately the design will be highly dependent on the applications and on the databases, which most of the times require replication across sites. From the point of BIG-IP there are very few restrictions. The topic is so wide that this guide will give overall guidance and will consider three scenarios:

- A hybrid design using VMC on AWS with local VPC workloads.
- A generic design that can be applied to any public cloud or private data centers.
- A specific design focused in local data retention with cloud bursting.

Overall approach

There are several approaches to multi-cloud. IP Anycast is a transparent mechanism with high reliability and fast recovery times that relies in highly coordinated IP routing which is not possible across cloud vendors. IP anycast routing strategies are also possible but, in many cases, routes cannot be migrated across Autonomous Systems swiftly. IP addressing based strategies inherently do not allow a high degree of control on service publishing. F5 recommends Global Server Load Balancing (GSLB) because it has the following benefits:

- **Cross-cloud vendor.** It can be used in any public cloud or private data center and supports any IP service (not necessarily served by BIG-IP).
- **High degree of control.** Rules can be setup based on service name instead of IP address. Traffic is directed to specific data center based on operational decisions such as service load and also allowing canary, blue/green, and A/B deployments across data centers.
- **Stickiness.** Regardless the topology changes in the network, clients will be consistently directed to the same data center.
- **IP Intelligence.** Clients can be redirected to the desired data center based on client's location and gather stats for analytics.

GSLB is available by F5 in two form factors:

- **Software as a Service (SaaS)** with F5 Cloud Services' DNS LB service.
- **Self-managed** with F5 BIG-IP's DNS module. This offers automatic service discovery of Virtual Servers in BIG-IP. It can be deployed in Internet eXchanges, private data centers, or public clouds.

At time of this writing, we recommend F5 BIG-IP's DNS module for GSLB because its more sophisticated health probing and its automatic service discovery feature.

SaaS Security and multi-cloud

Several security functions such as anti-DDoS and WAF are available in BIG-IP. BIG-IP Scale-N and Two-Tier BIG-IP setups allow for great scalability of these functionalities. Nowadays, It is a common practice to use security services delivered as SaaS because they provide ultimate scalability to handle DDoS and are managed services. F5 Cloud Services provides them:

- [Silverline DDoS Protection >](#)
- [Silverline Shape Defense >](#)
- [Silverline Web Application Firewall >](#)
- [Silverline Threat Intelligence >](#)

These are cross-cloud vendor offerings not tied to BIG-IP but have an exceptional integration with BIG-IP. Both F5 BIG-IP and F5 Cloud Services provide Pay as You Go pricing options.

Please check the Silverline links for more detail on this SaaS Security topic.

Generic Public Cloud and VMC on AWS connectivity options

Currently, public clouds provide a wide range of inter-site connectivity options as a service. We can differentiate these in two main types:

- **Dedicated circuits** with low latency and high throughput where traffic is only IP routed. This is the case of local VPC connectivity from VMC through an ENI interface and Direct Connect which allows inter-site connectivity.
- **Shared circuits** with non-guaranteed latency and limited throughput where traffic is encapsulated (often encrypted too) via gateways. This is the case with VPNs.

An overview of these connectivity options can be seen in the next figure. In it we discourage VPN connectivity for BIG-IP data plane traffic. This is because BIG-IPs typically deal with application and frontend tiers where low latency and throughput cannot be constrained. These are critical for application performance. Lower performance connectivity such as VPNs should typically be limited for services such as management and databases which can handle the traffic asynchronously for database replication.

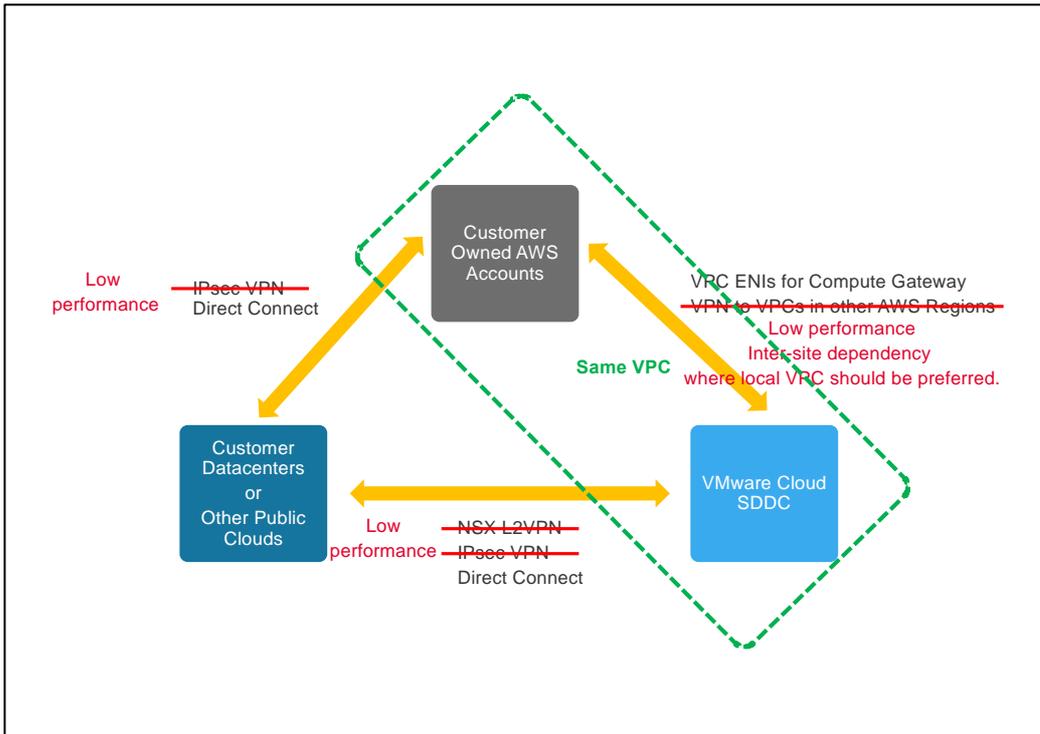


Figure 97 - Distilled connectivity options between the different types of clouds (squares). The less suitable connectivity options are stricken through and with annotations in red indicating the reason why they are less suitable.

Direct Connect, or even better VMC to local VPC connectivity can be used for stretching a cluster of servers across different infrastructures. Please note that this might create differently performant servers if pool members are spread amongst these infrastructures. Note as well that this also lowers reliability because there are more components and thus more points of failure involved. Whenever possible we will avoid these connectivity options too. In the design guidelines within this section we will indicate when these are suitable from BIG-IP data plane point of view.

VMware HCX – Hybrid Cloud Extension

A mention needs to be done on VMware’s HCX. HCX’s use cases⁷ are:

- Application migration.
- Change platforms or upgrade vSphere versions.
- Workload rebalancing.

⁷ <https://docs.vmware.com/en/VMware-HCX/services/user-guide/GUID-A7E39202-11FA-476A-A795-AB70BA821BD3.html>

- Business continuity and protection.

All these use cases make use of VM migration facilities provided by HCX. For the specific case of Workload rebalancing F5 recommends the use of GSLB instead.

In general, HCX doesn't mandate how the services are exposed externally therefore GSLB is always a valid option.

The VMware HCX Network Extension permits keeping the same IP and MAC addresses during a VM migration. This minimizes service disruption and is transparent to all devices including BIG-IP.

Design Guidelines – VMC on AWS with local VPC workloads

When using VMC on AWS direct connectivity to the VPC is available straight away. Moreover, reachability of the VMs is the same either from VMC to VPC or vice versa. The same applies to the Internet access. This opens the following dilemmas:

- Where to place the BIG-IPs?
- Where to place the Internet Gateway?

There is no definitive answer. We can choose whether we want each functionality in the AWS VPC or in the VMC side. This is shown in the next figure.

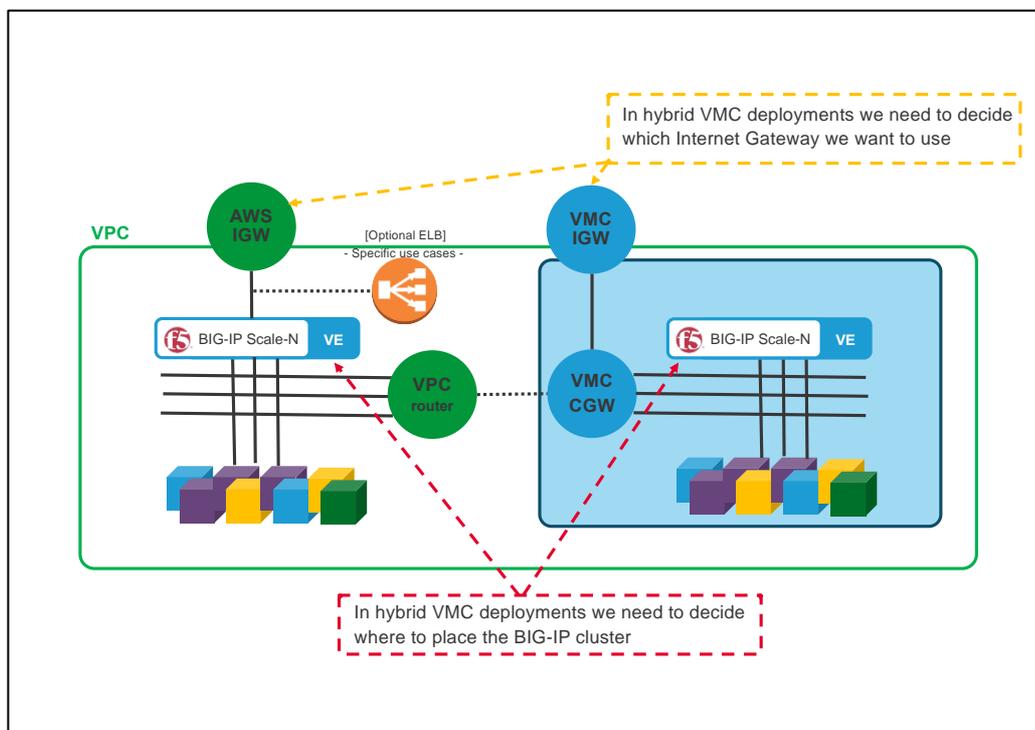


Figure 98 - Topology of VMC with local VPC workloads

The decision should consider the following aspects:

- At time of this writing, using an AWS IGW instead of an IGW via VMC has the possibility of using ELBs which provides Advanced Shield capabilities.
- The cost will depend where we have more traffic and where we have more compute resources.

Design Guidelines – multi-cloud

Designs depend on the applications and on databases. Inter-site dependencies play a crucial role. This guide recommends following the next design principles to minimize cost and maximize reliability while keeping simplicity in mind:

- Typically, ADCs like BIG-IP deal with Frontend-tier and App-tier servers which should not have to talk with peers in other sites. These tiers have the most throughput and latency demands so inter-site communication should be avoided. Otherwise, this could incur in uneven performance and increased and unnecessary costs.
- Identify strictly necessary inter-site dependencies. The typical case is DB replication which has much less throughput demands. Also, latency is a lesser issue because replication often happens asynchronously.
- There are other very relevant sources of inter-site traffic such as Automation, VM migration and data-store replication (for example a repository of images). VMware's HCX traffic fits in this category.

The first two items in this list deal with traffic that is generated upon client requests (blue arrows in the figure below). On the other hand, the third item is a new category of traffic (orange arrows) that is not expected to have dependencies when handling an ongoing customer request. Another characteristic of this traffic is that its traffic demands will greatly depend on frequency of updates in the applications.

- Simpler sites are easier to manage, scale, and replicate. GSLB allows for distribution of workloads based on a site's or a service's load and capacity so it is perfectly fine to have differently sized data centers. The most important attribute is to have them architecturally equal. Automations that are cross-cloud vendor capable are advised.

Using BIG-IP DNS and following the above guidelines we can create a cross-cloud vendor solution using GSLB. This is shown in the next figure.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

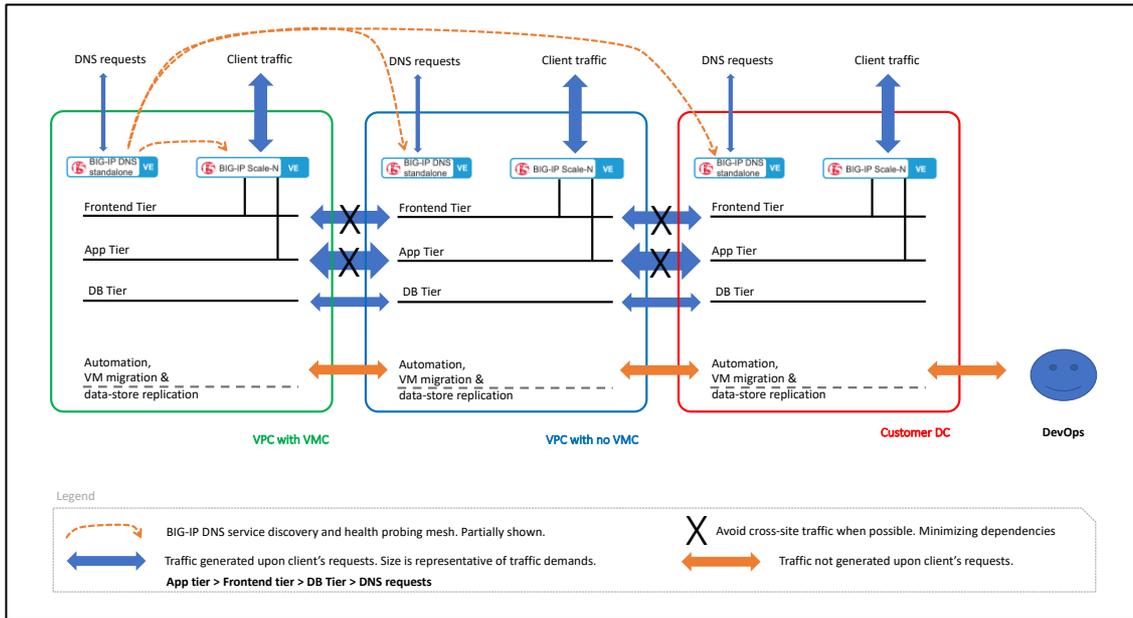


Figure 99 - Typical multi-cloud arrangement with most relevant traffic patterns.

Probably the most remarkable aspect of the diagram are the network dependencies and demands which drive the design. In this diagram Inter-site dependency is reduced to the minimum, typically DB replication only.

We can also see that there is additional inter-site traffic like the BIG-IP DNS iQuery (used for service discovery and health probing) but this traffic is different in nature because it is failure tolerant.

In the design above, the DNS functionality is implemented in a standalone BIG-IPs because redundancy is accomplished by having an independent BIG-IP DNS at each site. Having this BIG-IP DNS separated from the BIG-IP Scale-N cluster that handles client traffic gives clarity in the diagram and more relevantly sets a clear demarcation of functions. If desired, the BIG-IP DNS functionality can be consolidated in the BIG-IP Scale-N cluster at each site but a preferable approach is locate BIG-IP DNS outside of the data centers.

Ideally, BIG-IP DNS should be placed in Internet exchanges. This allows:

- To be closer to the clients. This only slightly improves DNS performance since client's local DNS resolvers usually reply from their DNS cache.
- To have a closer view to client's network performance and reachability to the clouds. This is very relevant.

A design with this approach can be seen in the next figure.

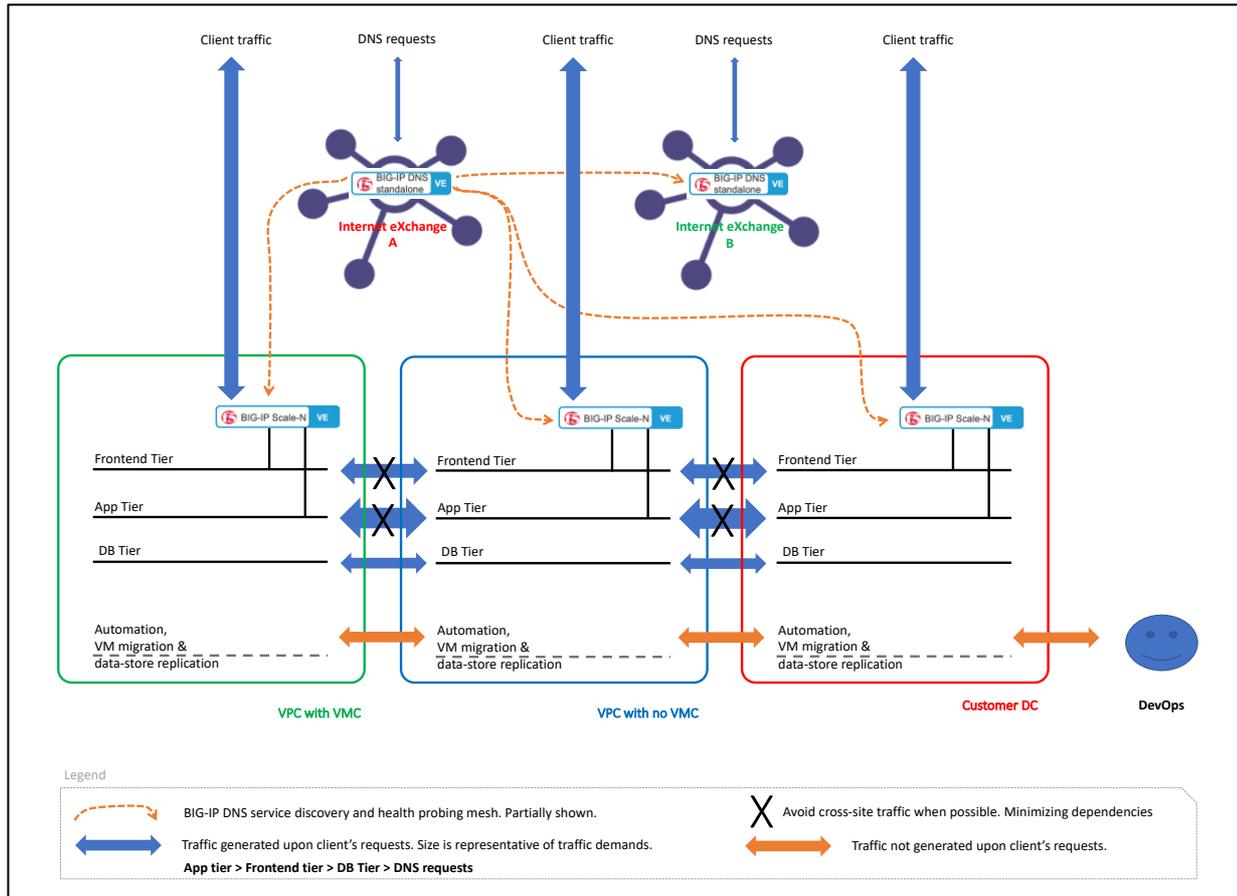


Figure 100 - Preferred multi-cloud arrangement by using Internet exchanges for BIG-IP DNS.

Cloud Bursting with multi-cloud

It is worth noting that the architecture being described in this section can be used for cloud bursting as well. Cloud Bursting refers to the use case when the main site has limited scalability and it is required to have increased capacity in peak periods. This cloud bursting capability is usually accomplished by spawning needed resources in Software Defined Data Centers/Public Clouds.

The approach described above in this section is preferred over adding compute from a Public Cloud by means of a Direct Connect circuit. This is because a GSLB multi-site approach has the following advantages:

- It automatically increases Internet traffic capacity. Each site has its own Internet access.
- It can reduce costs. Using a replica site uses almost the same compute resources and eliminates the need for a high performance Direct Connect.
- It provides increased reliability because of less inter-site dependency.
- Its automation is simpler because sites are architecturally similar.
- It is not necessary to deal with the bandwidth allocation management that the Direct Connect circuit will need over the time.
- An independent multi-site architecture can be easily replicated to additional sites when needed.
- It allows the use of more distributed regions, optimizing customer experience.
- The cloud bursting site can have alternative uses such as allowing migrations or new application roll outs.

An alternative Cloud burst architecture, specific to some use cases is described next.

Design Guidelines – single site with cloud bursting.

The topology to be described next is suitable for smaller deployments or when data must be stored on-premises, usually because of data retention policies or regulations. This can be observed in the next figure where the DB Tier is not stretched to the Public Cloud.

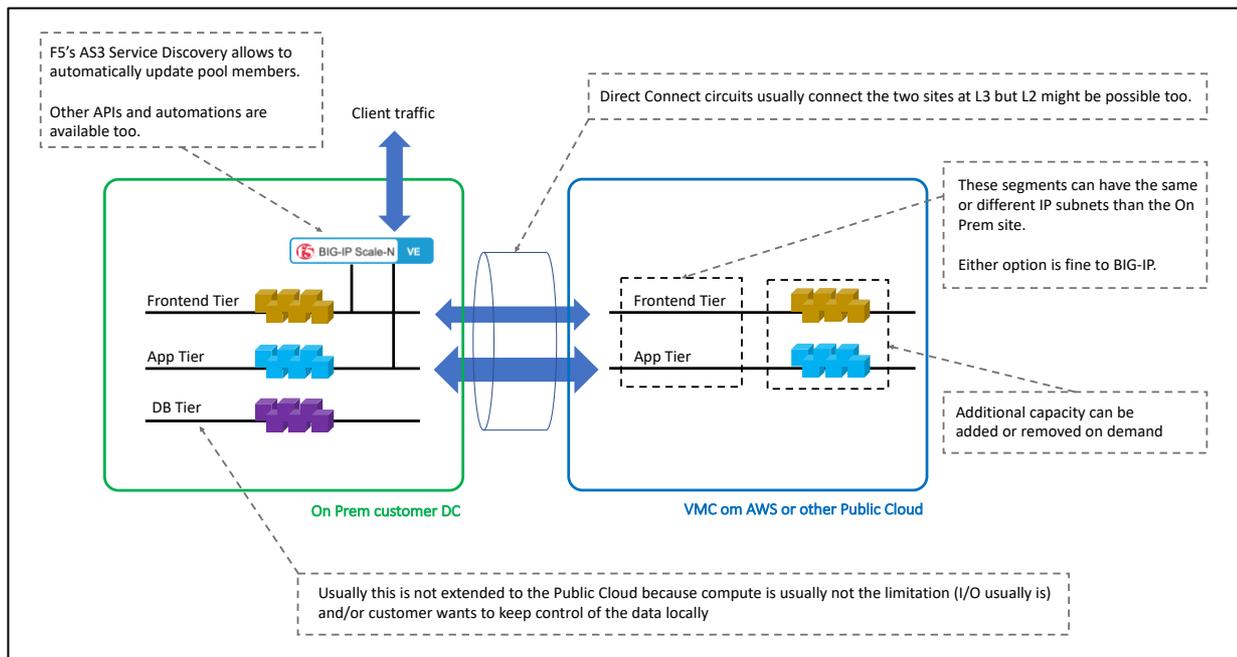


Figure 101 - Overall design of a single site with Cloud Bursting capability.

In this architecture the On-premises data center is stretched to a public cloud when load conditions require increasing the compute needs. In this scenario Internet access is kept in the On-premises data center. It requires the use of a high performance Direct Connect link with low latency. This is usually within the metropolitan area of the On-premises facility. This Direct Connect circuit needs to be established once and its capacity increased ahead of the peak periods. Some housing vendors allow to change circuit's capacity programmatically.

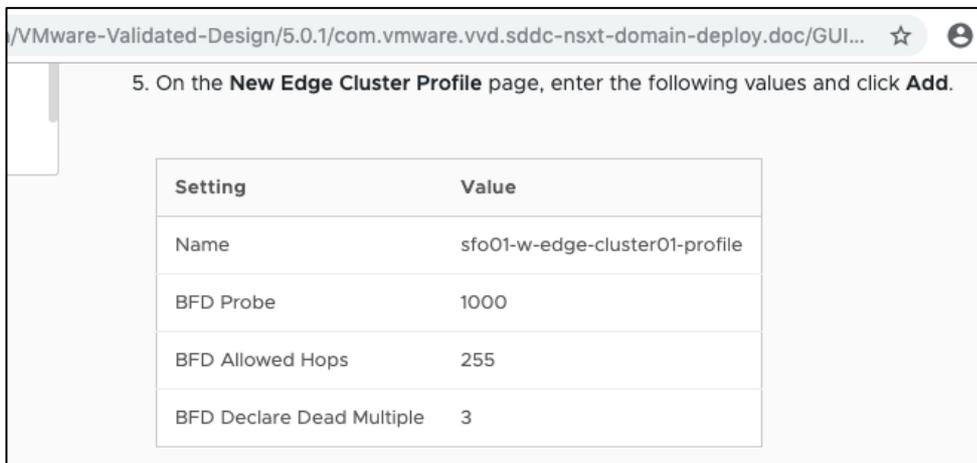
When compute changes dynamically, it is a perfect fit for F5's Service Discovery feature of AS3, automatically populating the pools with the added or removed computing instances. Please check the clouddocs.f5.com site for this and other automation options.

GENERAL NOTES

BGP configuration details

The following configuration settings follow VMware Validated Design 5.0.1⁸, see this guide for further details on these setting decisions:

- **NSXT-VI-SDN-033** – Use Bidirectional Forwarding Detection (BFD). VMware’s baseline recommendation is shown in the next figure (1000ms). When using NSX-T Edge in Bare metal servers the Probe timer can be reduced to 300ms to achieve sub-second convergence (see VMworld CNET1072BU session). These parameters are also appropriate when the F5 BIG-IPs are virtual machines (1000ms) or hardware (300ms) respectively.



The screenshot shows a web browser window with the URL `/VMware-Validated-Design/5.0.1/com.vmware.vvd.sddc-nsxt-domain-deploy.doc/GUI...`. The page content includes the instruction: "5. On the **New Edge Cluster Profile** page, enter the following values and click **Add**." Below this instruction is a table with the following data:

Setting	Value
Name	sfo01-w-edge-cluster01-profile
BFD Probe	1000
BFD Allowed Hops	255
BFD Declare Dead Multiple	3

Figure 102 - VMware's baseline settings for BFD

Virtualization is a potential source of latency and by using longer timers it is reduced the chance of false positives of link failures.

- **NSXT-VI-SDN-037** – Configure BGP Keep Alive Timer to 4 and Hold Down Timer to 12 seconds.
- **NSXT-VI-SDN-038** – Do not enable Graceful Restart between BGP neighbors.

⁸ <https://docs.vmware.com/en/VMware-Validated-Design/index.html>

Best practices for BIG-IP in VMware NSX-T

- **Management plane switch connectivity**
 - Following VMware general recommendations, the management interface (of either BIG-IP or BIG-IQ) should not be in an overlay network or use N-VDS at all. Typically, the management interface will be connected to a VDS switch, therefore isolating the management plane from the NSX-T networking.

- **Configure CPU reservation**
 - When deploying the BIG-IP OVA file using defaults, a specific amount of memory is reserved for the BIG-IP VE virtual machine. By default, CPU is not specifically reserved, but should be manually configured with an appropriate CPU reservation in order to prevent instability on heavily loaded hosts. This is done in vCenter.
 - The CPU must support a one-to-one, thread-to-defined virtual CPU ratio, or on single-threading architectures, support at least one core per defined virtual CPU. In VMware ESXi 5.5 and later, do not set the number of virtual sockets to more than 2.

- **VM placement in vCenter (on premises)**
 - BIG-IQs should be placed alongside other management functionalities of VMware (ie: NSX-T manager and vCenter). In large deployments, these management functions are in their own Management Cluster.
 - BIG-IPs used for North-South traffic should be placed in the same cluster as NSX-T Edge nodes in order keep traffic affinity. This might be a dedicated “Centralized Services” cluster, a shared “Management & Edge” cluster or in an all-shared “Collapsed” cluster depending the size of the deployment.
 - BIG-IPs used for East-West traffic should be distributed across the Compute Clusters to distribute their workload as much as possible. In the case that each tenant has their own nodes, the BIG-IPs should be run just as another tenant VM maximizing affinity of the traffic flows.
 - Very importantly, the previous recommendations should be complemented by making sure that the VMs of a given BIG-IP cluster should reside in different ESXi hosts. This is typically referred to as anti-affinity.

The above VM placement best practices can be achieved with the Dynamic Resource Scheduler (DRS). In the next picture, the creation of anti-affinity rules is shown to avoid two BIG-IPs of the same cluster running on the same hypervisor. Note: the anti-affinity rules should be “must” rather than “should” to guarantee anti-affinity and therefore high availability.

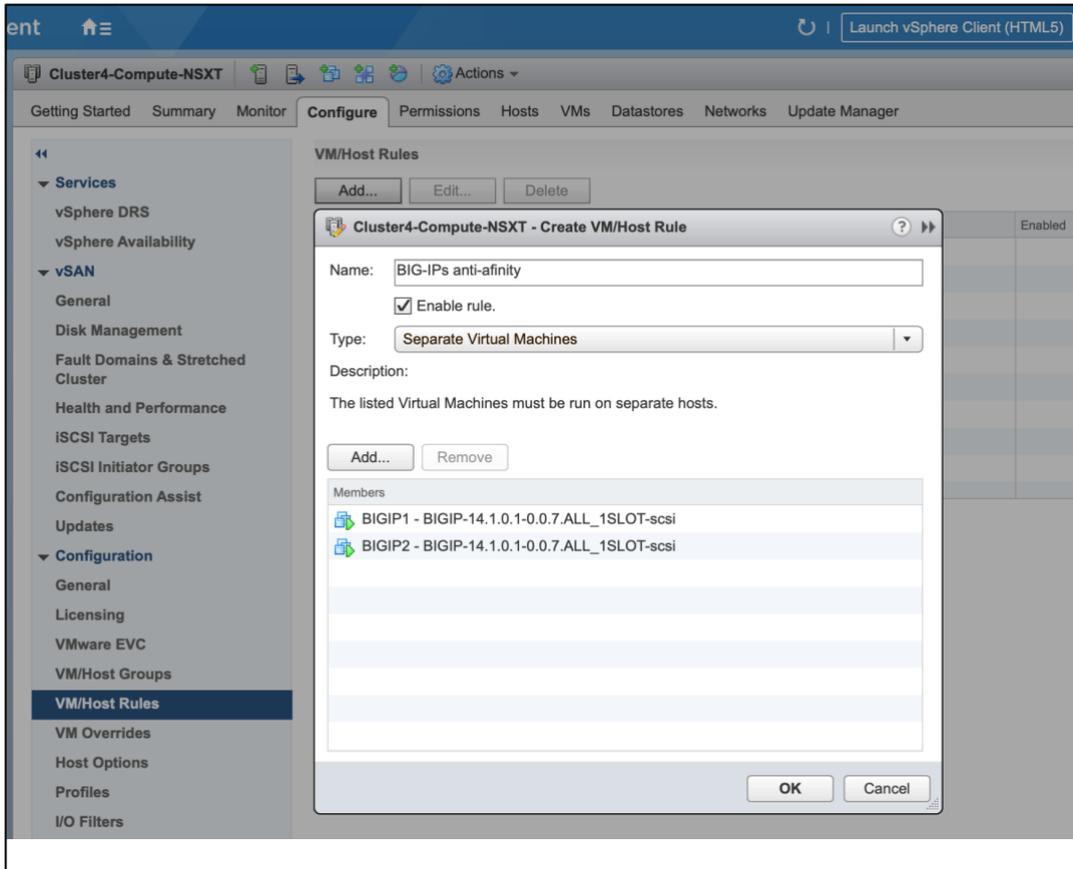


Figure 103 - Setting anti-affinity rules with VMware's Dynamic Resource Scheduler.

- VM placement in VMC for AWS

- High Availability of VMs in VMC requires using the stretched cluster deployment type. When deploying a VM you can choose an ESXi host in the desired Availability Zone (AZ). In case of failure, the VM will stay in its original AZ if possible. Each site in a stretched cluster resides in a separate fault domain. See the VMC FAQ⁹ and this community article¹⁰ for more details. A screenshot of this configuration is shown next.

⁹ <https://cloud.vmware.com/vmc-aws/faq#stretched-clusters-for-vmware-cloud-on-aws>

¹⁰ <https://cloud.vmware.com/community/2018/05/15/stretched-clusters-vmware-cloud-aws-overview/>

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

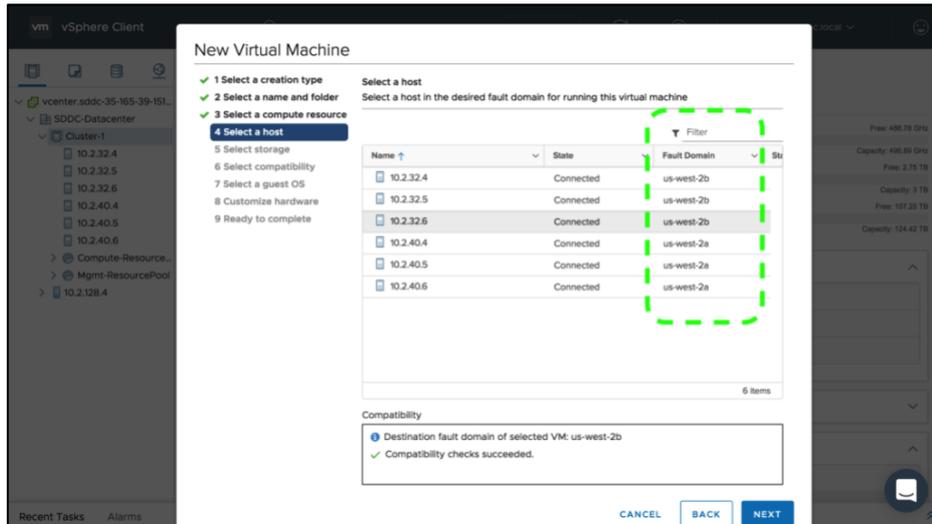


Figure 104 - Configuring High Availability of VMs in VMC stretched clusters.

MAC Masquerade in NSX-T

MAC Masquerade is a mechanism in BIG-IP that eliminates the need of neighbor L3 devices updating ARP entries of the BIG-IPs when a traffic-group is shifted from one BIG-IP to another. Only the L2 devices (switches) need to update their L2 forwarding databases. Traffic-group shifts happen when workloads are redistributed within the Scale-N BIG-IP cluster or when there is a failover event.

Please note that this feature is an optimization to slightly reduce the time for the traffic to be sent to the appropriate BIG-IP when a traffic-group shift occurs. This optimization, although it is a slight reduction in time might be critical for some applications. Usually this feature is not needed and is not noticeable when configured because the GARP mechanism used by default is fast enough for the vast majority of applications.

MAC Masquerade is achieved by having a single MAC addresses for each traffic-group which is shared by the BIG-IPs of the Scale-N cluster (by default each BIG-IP has a different MAC address for each traffic-group). This BIG-IP feature is further described in [K13502: Configuring MAC masquerade \(11.x - 15.x\)](#)¹¹.

NSX-T has a very security tight L2 configuration and requires adjustment. More precisely, a new MAC Discovery Profile needs to be created with the following settings changed from their default:

- MAC Learning: Enabled.
- Unknown Unicast Flooding: Enabled.

These settings can be seen in the following figure. This profile has to be applied to all the segments of the traffic group where MAC masquerading is going to be used.

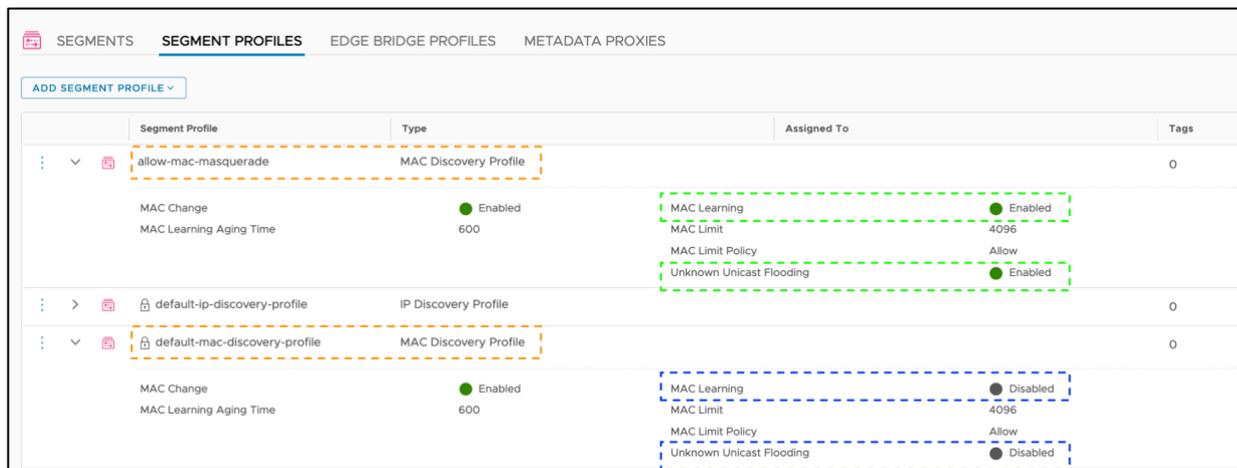


Figure 105 - Creating a new MAC Discovery Profile for MAC Masquerade.

¹¹ <https://support.f5.com/csp/article/K13502>

VMC on AWS

At time of this writing VMC on AWS doesn't allow this customization hence MAC Masquerade cannot be used.

Considerations for container platforms

General guidelines

This section takes into account Red Hat OpenShift and Kubernetes in general. At present, handling Pivotal PKS differently than any other Kubernetes flavor is not required, and as long as Pivotal PKS aligns to the Kubernetes API, this will be supported by F5 Networks like any other Kubernetes flavor. Red Hat OpenShift and Pivotal PKS are able to use NSX-T's load balancer natively. In this release of the guide, the focus is in replacing the LBs for workloads and not for the management and control plane of these platforms.

As described in previous sections, for any of these container platforms the POD's IP addresses should be routable from the BIG-IP. In other words, there cannot be any NAT between the BIG-IP and the PODs. Moreover, there are two ways which POD workers can be exposed with a resource of kind Service: via NodePort or via ClusterIP. Although both are supported it is highly recommended to use ClusterIP¹². This is because when using NodePort mode the BIG-IP (or any other external host) cannot send traffic directly to the PODs which means for the BIG-IP that:

- There is an additional layer of load balancing (at node level) which adds latency and complexity, which makes troubleshooting and observability more difficult.
- Some functionalities like L7 persistence would not behave as expected.
- The BIG-IP has limited visibility of PODs actual health.

Exposing container services

Once the PODs that compose the workers of a given Service are defined, the BIG-IP must be automatically configured and updated when the PODs of the service are created, updated or deleted. This is performed by the F5 Container Ingress Services (CIS)¹³ which is installed as a Kubernetes POD that monitors configuration changes in Kubernetes. F5 CIS automatically updates the BIG-IP configuration by translating orchestration commands into F5 SDK/iControl REST calls. The overall architecture is shown in the next picture.

¹² <https://clouddocs.f5.com/containers/v2/kubernetes/kctr-modes.html#kctr-modes>

¹³ <https://clouddocs.f5.com/containers/v2/>

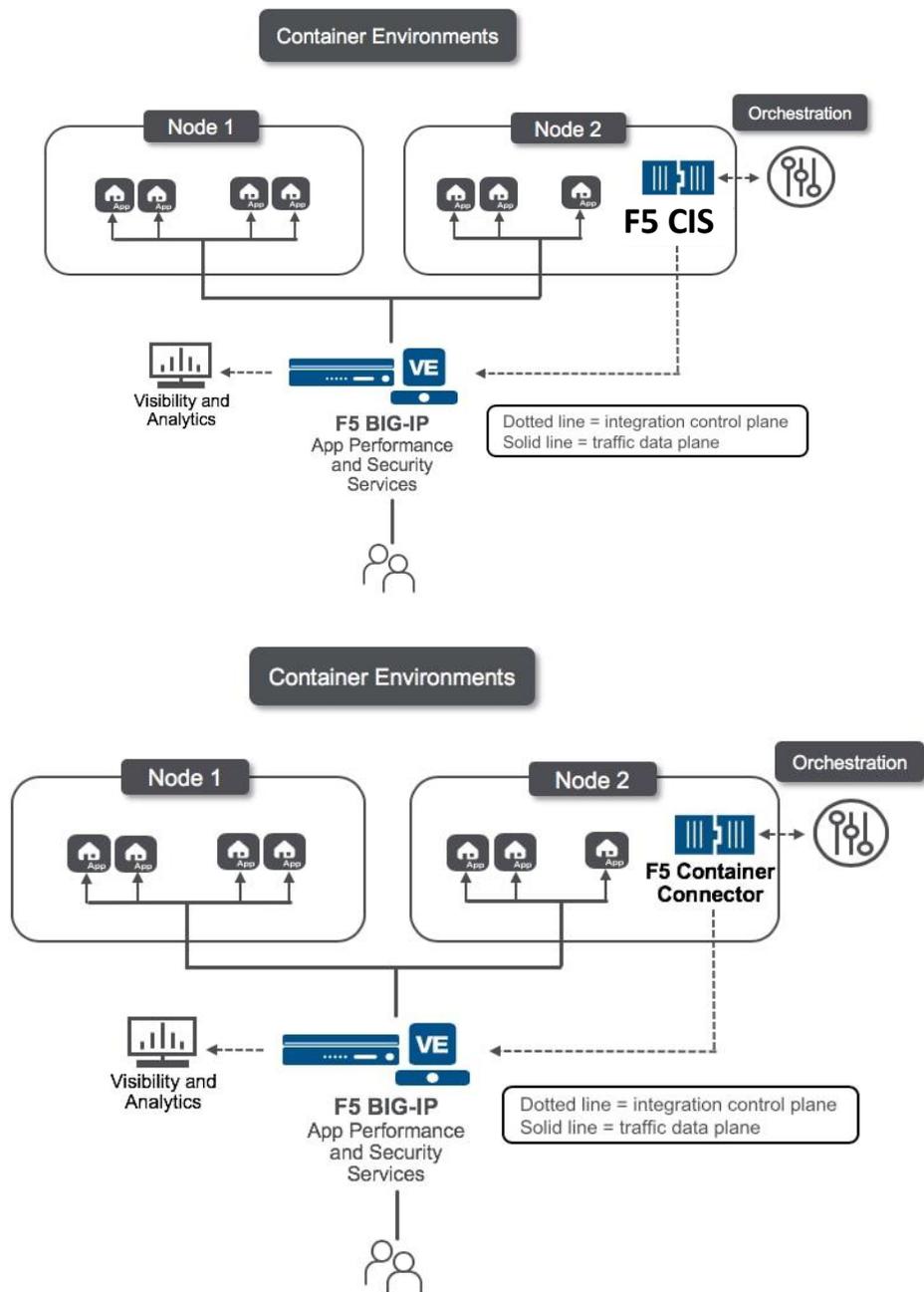


Figure 106 - F5 BIG-IP integration with container platforms with F5 Container Ingress Services (CIS)

Although in the diagram above only one CIS instance is shown, it is possible that a single instance of F5 BIG-IP can be managed by several CIS instances associating different container namespaces or projects to different partitions in the F5 BIG-IP.

Kubernetes services can be exposed in F5 BIG-IP using several resource types, these are shown in the next table:

	Red Hat OpenShift	Pivotal PKS	Vanilla Kubernetes
OpenShift routes	Yes	NA	NA
Ingress	Yes	Yes	Yes
AS3 declaration	Yes	Yes	Yes

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

F5 BIG-IP ConfigMap	Yes	Yes	Yes
----------------------------	-----	-----	-----

These options can be combined in the same deployment. Note that in the above table, the LoadBalancer Service type is not mentioned. This is out of scope because it is meant to be implemented by a cloud provider's load balancer. Also note that the LoadBalancer Service type is also not efficient in the use of IP address ranges because it requires an IP address for each instance.

Relevant configuration settings when using Red Hat OpenShift

NSX Container Plug-in (NCP) provides integration between NSX-T Data Center and OpenShift (also to other PaaS/CaaS). In this section, the settings of ncp.ini are described (or the related YAML ConfigMap file at installation time) that should be taken into account:

To make use of BIG-IP instead of NSX-T's Load Balancer it is needed to set

```
use_native_loadbalancer = False
```

In order to have PODs that do not require SNAT, it is necessary to indicate either the desired CIDR address blocks or the UUIDs of previously defined address blocks in the next variable:

```
no_snat_ip_blocks = <comma separated list of UUIDs or CIDRs>
```

When creating projects/namespaces these will need to be created with the ncp/no_snat=true annotation. This way the subnets will be taken from these IP blocks and there will be no SNAT for them. These IP blocks are expected to be routable. An example namespace is shown next:

```
apiVersion: v1
kind: Namespace
metadata:
  name: no-nat-namespace
  annotations:
    ncp/no_snat: "true"
```

External IP Pools will not be used because any SNAT or Ingress/LoadBalancer resource will be handled by the BIG-IP. Further details can be found in the following documents:

- VMware's "NSX Container Plug-in for OpenShift - Installation and Administration Guide".
- Red Hat's "Deploying and Managing OpenShift on a VMware Software-Defined Data Center".

Relevant configuration settings when using Pivotal PKS

Like any other container platform, NAT must be disabled within the container environment. This is to allow the BIG-IP to have direct visibility to the container's IP address.

In the case of Pivotal PKS this is indicated with the PKS Ops Manager UI while performing PKS installation. Following the regular PKS configuration, it is needed to unset the NAT option in the Networking Tab as shown in the next screenshot.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

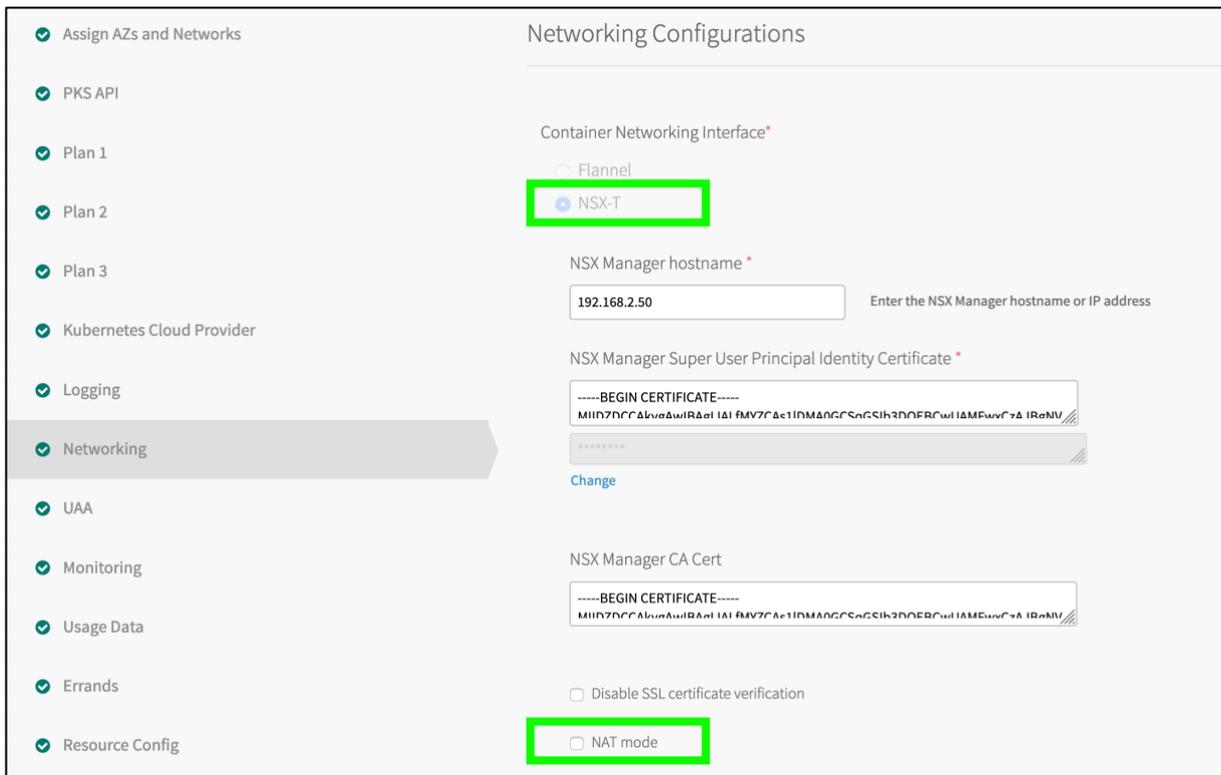


Figure 107 - Indicating PKS networking options at installation time. The NAT option must be unset.

Verifying the deployment

Basic testing

The first test to perform is ping connectivity from the F5 BIG-IPs to the adjacent next-hops.

	Adjacent next-nops
Topology A Impl. static routing	Northbound – 10.105.217.1 Southbound – 10.106.53.1
Topology A Impl. dynamic routing	Northbound – 10.105.217.1 Southbound – 10.106.53.{1,2}
Topology A Impl. dynamic routing +ECMP	Northbound – 10.105.217.1 Southbound Uplink Red – 10.106.53.{1,2} Southbound Uplink Blue – 10.106.54.{1,2}
Topology B	Northbound – 10.106.49.1 Southbound – 10.106.{51,52}.10 (Servers)
Topology C	Northbound – 10.10.216.1 Southbound – 10.106.48.1
Topology D	NorthBound – 10.106.32.1 (default route's next-hop) Southbound – 10.106.{32,33,34},100 (Servers)

The next step will be creating a test VM that will be attached to the tenant networks where the workload servers will reside.

	Segment / IP address
Topology A	10.106.32.10
Topology B	10.106.{51,52}.10
Topology C	10.106.51.10
Topology D	10.106.{32,33,34},100

Configuring the VM's network interface should allow pinging the NSX-T Tier-1 Gateway's router port (or the BIG-IP in the case of Topology B) as shown in the next figure. The next test will be to ping BIG-IP's closest IP.

The IP addresses to be used in these two tests are shown in the next table.

	Test VM's NSX-T next-hop	BIG-IP's closest IP to test VM
Topology A Impl. static routing	10.106.32.1	10.106.53.10
Topology A Impl. dynamic routing	10.106.32.1	10.106.53.10
Topology A Impl. dynamic routing +ECMP	10.106.32.1	Southbound Uplink Red – 10.106.53.10 Southbound Uplink Blue – 10.106.54.10
Topology B	10.106.{51,52}.1	10.106.{51,52}.100
Topology C	10.106.51.1	10.106.48.100
Topology D	10.106.{32,33,34}.1	10.106.{32,33,34}.10

If testing BIG-IP's closest IP doesn't succeed it is recommended to 1) ping from the BIG-IP end instead and check the port-lock down in the Self IPs, 2) ping the floating Self IP address from the BIG-IPs themselves and 3) ping the non-floating IPs as well.

Dynamic routing testing

First, verify that the BFD is established properly. This is a prerequisite for the dynamic routing to work properly and BFD will also show us that connectivity at IP level for the NSX-T Uplinks is operational.

Login in the imish cli and run the following command in both BIG-IP units and verify that the `Session State` is `Up` for all BFD sessions (one per BGP peering configured):

```
bigipla.nsxt.bd.f5.com[0]#show bfd session
Sess-Idx  Remote-Disc  Lower-Layer  Sess-Type  Sess-State  UP-Time  Remote-Addr
3         458133421   IPv4         Single-Hop  Up          2d19h49m 10.106.53.1/32
4         211353312   IPv4         Single-Hop  Up          2d19h49m 10.106.54.1/32
Number of Sessions: 2
```

Figure 108 - Verification of the NSX-T uplinks by checking the BFD sessions.

Next, verify that the BGP peerings are in Established state by running the following command:

```
bigipla.nsxt.bd.f5.com[0]#show bgp neighbors | include BGP state
BGP state = Established, up for 2d19h50m
BGP state = Established, up for 2d19h50m
```

Figure 109 - Verifying that the BGP sessions are Up.

As you can see in Figure 109, it is expected to see two lines with Established state (one line per BGP peering). This command must be run in both BIG-IPs as well. If the output shown is not the same as above, verify that: BGP's TCP port 179 is open, the peering IP addresses for each BIG-IP are correct and the BGP password is correct.

The next step is to verify that the routes are exchanged through BGP as expected. You should expect two next-hops for the NSX-T routes (in blue) and one for the default route (in green).

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

```
bigipla.nsxt.bd.f5.com[0]#show ip bgp
BGP table version is 9, local router ID is 192.174.70.111
Status codes: s suppressed, d damped, h history, * valid, > best, i - internal, l
- labeled
                S Stale
Origin codes: i - IGP, e - EGP, ? - incomplete

   Network        Next Hop          Metric      LocPrf     Weight Path
*> 0.0.0.0/0      10.105.217.1      0           0          32768  ?
* 10.106.32.0/24  10.106.54.1       0           0          65001  ?
*> 10.106.32.0/24  10.106.53.1       0           0          65001  ?
* 10.106.33.0/24  10.106.54.1       0           0          65001  ?
*> 10.106.33.0/24  10.106.53.1       0           0          65001  ?
* 10.106.34.0/24  10.106.54.1       0           0          65001  ?
*> 10.106.34.0/24  10.106.53.1       0           0          65001  ?

Total number of prefixes 4
```

Figure 110 - Verifying BGP route exchange.

Finally, if using an NSX-T Edge Active-Active setup, verify that the NSX-T routes are ECMP routes by checking in the BIG-IP tmsh cli with the following command (again in both BIG-IP units).

```
root@(bigipla) (cfg-sync In Sync) (Active) (/Common) (tmsh)# show net route | grep ecmp
10.106.32.0/24    10.106.32.0/24    gw          10.106.53.1      dynamic ecmp
10.106.32.0/24    10.106.32.0/24    gw          10.106.54.1      dynamic ecmp
10.106.33.0/24    10.106.33.0/24    gw          10.106.53.1      dynamic ecmp
10.106.33.0/24    10.106.33.0/24    gw          10.106.54.1      dynamic ecmp
10.106.34.0/24    10.106.34.0/24    gw          10.106.53.1      dynamic ecmp
10.106.34.0/24    10.106.34.0/24    gw          10.106.54.1      dynamic ecmp
```

Figure 111 - Verifying NSX-T ECMP routes learned via dynamic routing (BGP).

End to End testing: test egress forwarding connectivity through the BIG-IP.

Note that this end-to-end testing doesn't apply to Topologies C and D because in these the BIG-IPs are not inline.

Create a forwarding type virtual server in the F5. This virtual server will service outbound traffic flows from the NSX-T environment. The configuration of this virtual server is shown in the following Figure 112, where the parameters are in red are mandatory.

The screenshot shows the configuration for a Virtual Server named 'egress_forwarding'. The configuration is divided into two main sections: General Properties and Configuration.

General Properties:

- Name: egress_forwarding
- Partition / Path: Common
- Description: (empty)
- Type: Forwarding (IP)
- Source Address: 0.0.0.0/0 (highlighted in red and green)
- Destination Address/Mask: 0.0.0.0/0 (highlighted in red)
- Service Port: 0, * All Ports (highlighted in red)
- Notify Status to Virtual Address:
- Availability: Unknown (Enabled) - The children pool member(s) either don't have service checking enabled, or service check results are not available yet
- Synccookie Status: Inactive
- State: Enabled

Configuration: Basic

- Protocol: * All Protocols (highlighted in red)
- Protocol Profile (Client): fastL4
- VLAN and Tunnel Traffic: Enabled on... (highlighted in red)
- VLANs and Tunnels: Selected: /Common Services; Available: /Common, EastWestVIPs, HA, Transit, http-tunnel
- Source Address Translation: None (highlighted in green)

Optional fields are indicated by green boxes and the word 'Optional' in green text.

Figure 112 - Creating a Forwarding Virtual Server for testing egress traffic.

Note that in the case of the Topology A with the Active-Active setup the two VLANs used for the NSX-T uplinks must be specified.

The optional parameter **Source Address** can be used to restrict from which source addresses the VIP is limited. This could be changed to NSX-T's address range (10.106.0.0/16) to tighten security.

The optional **Source Address Translation** parameter can be used in the case you want to hide the NSX-T's address range and NAT these addresses when going north of the F5 BIG-IPs.

After applying this configuration, you can reach the spine router's IP address which is the default gateway of the F5 BIG-IPs. If the spine routers provide Internet connectivity at this stage, it should be possible to ping an Internet address as shown in the next figure.

```

TestPoint guide
Enforce US Keyboard Layout View Fullscreen Send Ctrl+Alt+Delete

root@testpoint-guide:~# ping -c 1 10.105.197.1
PING 10.105.197.1 (10.105.197.1) 56(84) bytes of data:
64 bytes from 10.105.197.1: icmp_seq=1 ttl=61 time=12.5 ms

--- 10.105.197.1 ping statistics ---
1 packets transmitted, 1 received, 0% packet loss, time 0ms
rtt min/avg/max/mdev = 12.578/12.578/12.578/0.000 ms
root@testpoint-guide:~# ping -c 1 8.8.8.8
PING 8.8.8.8 (8.8.8.8) 56(84) bytes of data:
64 bytes from 8.8.8.8: icmp_seq=1 ttl=46 time=9.80 ms

--- 8.8.8.8 ping statistics ---
1 packets transmitted, 1 received, 0% packet loss, time 0ms
rtt min/avg/max/mdev = 9.807/9.807/9.807/0.000 ms
root@testpoint-guide:~# _
    
```

Figure 113 - Ping test using spine router's IP address and the well-known Internet address 8.8.8.8 for checking egress connectivity.

	Closest's spine router IP address
Topology A Impl. static routing	10.105.217.1
Topology A Impl. dynamic routing	10.105.217.1
Topology A Impl. dynamic routing +ECMP	10.105.217.1
Topology B	10.105.216.1

In all the example topologies, the same spine routers are used so the IP address to use for this testing is the same. If this test doesn't succeed it is recommended to 1) In the case of using Topology A, check the advertised networks in the NSX-T Tier-1 Gateway, 2) verify the routing

table in the NSX-T Tier-0 Gateway, 3) verify the routing table in the BIG-IPs and 4) run a `tcpdump -ne1 -i 0.0` in the Active BIG-IP to see what is actually happening.

End to End testing: test egress forwarding connectivity without the BIG-IP.

This testing applies only when using the BIG-IP in parallel path configuration where the egress forwarding traffic doesn't go through the BIG-IPs. In this case it will be tested that the NSX-T networking works as expected, and that NSX-T is properly connected to its upstream next-hops.

	Closest spine router's IP address
Topology C	10.105.216.1
Topology D	10.105.217.1

If these tests doesn't succeed it is recommended to 1) Check the advertised networks in the NSX-T Tier-1 Gateway, 2) verify the routing table in the NSX-T Tier-0 Gateway, 2) verify the routing table in the BIG-IPs and 3) use NSX-T tracing & packet capture tools.

End to End testing: test Ingress connectivity through the BIG-IP.

For this test, a Standard type virtual server is used listening in BIG-IP's external facing network. A pool with a web servers will be configured. The overall process is the same for all topologies and a table with the settings that are specific to each topology is shown next. How to install a web server is not described here.

	IP address for the webserver virtual server	SNAT	Pool member address (actual web server)
Topology A	10.105.217.100	No/Optional	10.106.{32,33,34}.10
Topology B	10.105.216.100	No/Optional	10.56.{51,52}.10
Topology C	10.106.49.100	Yes	10.56.{51,52}.10
Topology D	10.106.32.100	Yes	10.106.{32,33,34}.10

The overall configuration of this webserver virtual server is shown next following Topology B. The values for all topologies are shown at the end of the graphical example.

The screenshot shows the configuration page for a virtual server named 'webservice'. The breadcrumb trail is 'Local Traffic >> Virtual Servers : Virtual Server List >> webservice'. The 'Properties' tab is selected. Under 'General Properties', the following fields are visible:

- Name: webservice
- Partition / Path: Common
- Description: (empty text box)
- Type: Standard (dropdown menu)
- Source Address: Host (selected), Address List (radio buttons); 0.0.0.0/0 (text box)
- Destination Address/Mask: Host (selected), Address List (radio buttons); 10.106.64.10 (text box)
- Service Port: Port (selected), Port List (radio buttons); 80 (text box), HTTP (dropdown menu)

The screenshot shows the 'VLANs and Tunnels' configuration section. It features two columns: 'Selected' and 'Available'. The 'Selected' column contains '/Common Transit'. The 'Available' column contains '/Common', 'EastWestVIPs', 'HA', 'Services', and 'http-tunnel'. There are '<<' and '>>' buttons between the columns. Below this, the 'Source Address Translation' is set to 'None' (dropdown menu).

Figure 114 - Creating a Standard Virtual Server for testing Ingress services' connectivity.

Before clicking the Finished button for creating the virtual server it is needed to attach a pool with the test VM as member. This is done by clicking the '+' button shown next:

The screenshot shows the 'Default Pool' configuration section. It includes a '+' button and a dropdown menu currently set to 'None'.

Figure 115 – Creating a new pool that will be used for the connectivity test with the Ingress Virtual Server.

Then specifying the pool as shown in the next picture. Please note that the default HTTP health monitor is used.

Local Traffic » Pools : Pool List » New Pool...

Configuration: Basic

Name: pool_webserver

Description:

Health Monitors:

Active: /Common http

Available: /Common gateway_icmp, http_head_f5, https, https_443

Resources:

Load Balancing Method: Round Robin

Priority Group Activation: Disabled

New Members:

New Node
 New FQDN Node
 Node List

Node Name: (Optional)

Address: 10.106.66.100

Service Port: 80 HTTP

Add

Node Name	Address/FQDN	Service Port	Auto Populate	Priority
10.106.66.100	10.106.66.100	80		0

Edit Delete

Figure 116 - Specifying pool member details for the test Ingress Virtual Server.

This pool health monitor already tests the connectivity from the BIG-IP to the web server when it is shown as green as in the next figure at virtual server level.

If the pool health monitor doesn't succeed it is recommended to 1) perform a ping test from the BIG-IP to the pool member, 2) verify that the web server is up and the socket listening in the expected address and 3) there is no distributed firewall rule that inhibits the connectivity between the Self IP of the BIG-IPs used for sending the probes and the pool member.

Local Traffic » Virtual Servers : Virtual Server List

Virtual Server List Virtual Address List Statistics

Search Create...

✓	Status	Name	Description	Application	Destination	Service Port	Type	Resources	Partition / Path
<input type="checkbox"/>		egress_forwarding			Any IPv4	0 (Any)	Forwarding (IP)	Edit...	Common
<input type="checkbox"/>		webserver			10.106.64.10	80 (HTTP)	Standard	Edit...	Common

Figure 117 - virtual server status after creating the webserver VS for Ingress traffic.

DESIGN GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP

This 'green' status doesn't validate end to end traffic path for this it is needed send an HTTP request from a host upstream of the spine-router.

If this doesn't succeed it is recommended to 1) perform the HTTP request locally using the pool member's address (not `127.0.0.1`), 2) perform a ping test to the BIG-IP's virtual server address and 3) verify that the virtual server is enabled in the expected VLANs, these are the VLANs where the connection to the virtual server are established and not the VLANs towards the pool members. Also, if there is a routing problem many times enabling SNAT might solve these and would reveal that there is a routing miss-configuration.