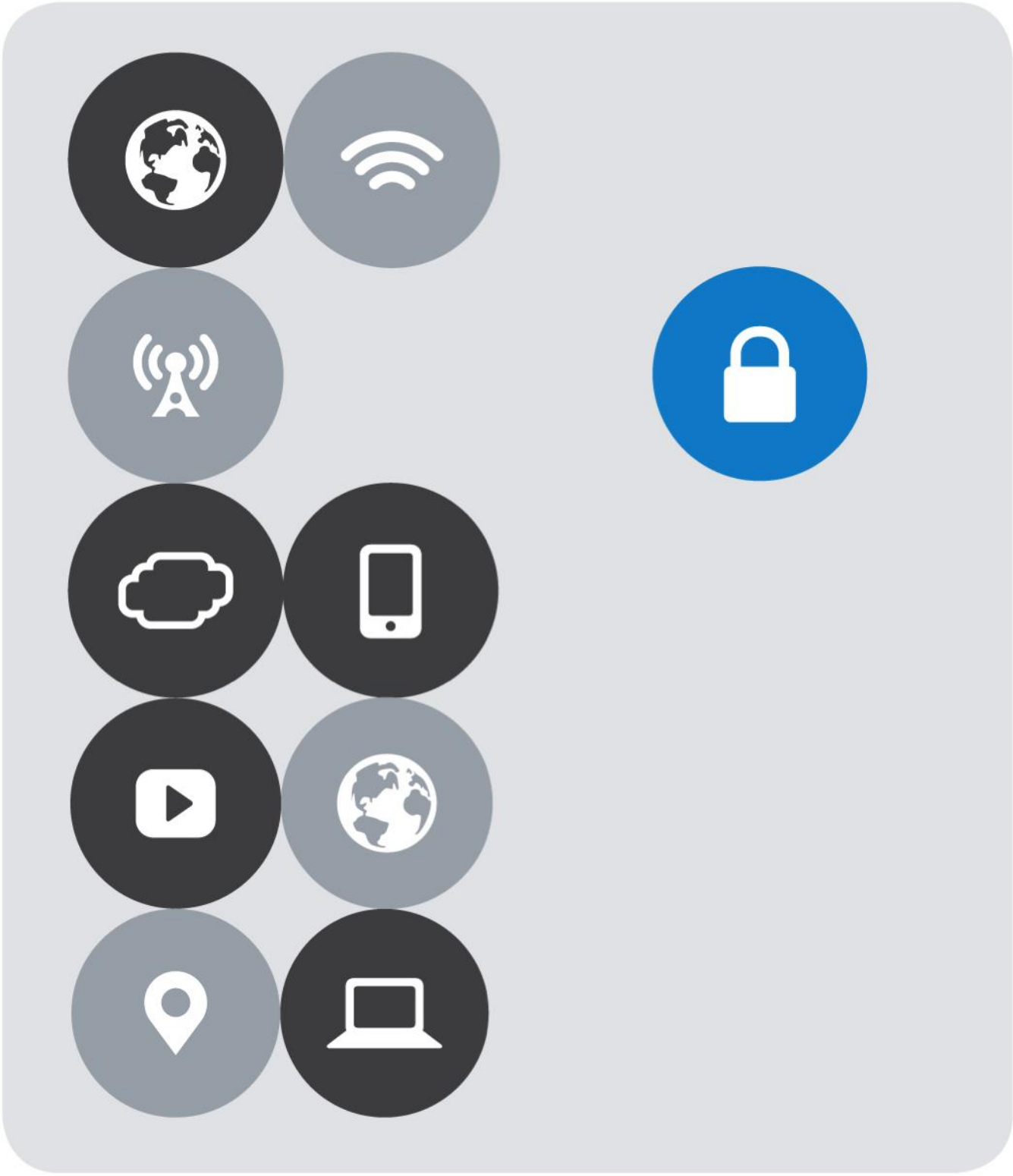
DEPLOYMENT GUIDE AND BEST PRACTICES

VMware NSX-T and F5 BIG-IP



Version History

Date	Version	Author	Description
December 2019	1.0	Ulises Alonso Camaró,	Initial public version.
June 2020	2.0	Ulises Alonso Camaró	Validated with NSX-T 3.0. Updated all screenshots and configuration flows to match NSX-T 3.0. Changed network addressing to allow a lab with all topologies simultaneously. Changed Logical Router/LR nomenclature for Gateway following new NSX-T convention. 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. Extended information on BIG-IP version compatibility with a dedicated section.
September 2020	2.1	Ulises Alonso Camaró	Extended topology suitability matrix based on flow's direction with the inter-tenant E-W flows case. Added MAC masquerading information. Added VMC on AWS section. Added section on Hybrid and Multi-Cloud design considerations. Renamed from "Integration guide" to "deployment guide"
June 2021	2.2	Matt Mabis Paul Pindell Ulises Alonso Camaró	Proofreading review and minor changes
December 2022	2.3	Ulises Alonso Camaró Proofreading review by Paul Pindell	Added section "Considerations for Tanzu Kubernetes Grid (TKG)". Deleted previous container section. New best practices: - HA connectivity. - Failover and HA Groups. - BIG-IP VE hypervisor optimization. Added "Topology B extended". Added section "BIG-IP multi-tenant considerations". Added multi-tenant subsection to each topology. Fixed metadata and headings.

June 2023	2.4	Ulises Alonso Camaró Proofreading review by Paul Pindell	Modified “Topology B extended” so the T1 is placed below the VE. New or updated best practices: - Hypervisor optimization for BIG-IP VE. - Monitoring VM hypervisor performance. - Distributed firewall rules (micro-segmentation): add F5 VEs to the exclusion list.

Introduction

The Software-Defined Data Center (SDDC) is characterized by server virtualization, storage virtualization, and network virtualization. The first pillar, server virtualization has already proved the value in reducing costs and complexity of the compute infrastructure. In VMware environments this is provided by VMware vSphere. The second pillar, storage virtualization is provided by VMware with the vSAN product. VMware NSX network virtualization provides the third 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 the 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 must be connected is small and is not expected to change often.

- For East-West traffic the lack of a publicly documented API inhibits the possibility of using F5 BIG-IP hardware. 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 vNICs[1] 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) or 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 hypervisor hosts 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/WAAP, SSL-VPN, anti-DDoS protection, Secure Web Gateway with Identity Management, and many other solutions with unified management & visibility via F5 BIG-IQ.

[1] For checking vSphere's limits consult the link https://configmax.vmware.com/guest?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 2.4-3.2 but given that the F5 BIG-IP integration is transparent from NSX-T point of view[2] 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 is mainly structured around the 5 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.
    - Topology B extended: BIG-IPs inline-connected to 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.

Topologies where the BIG-IPs are placed below a Tier-1 Gateway are usually meant per tenant deployments, following NSX-T's multi-tenancy model. For these topologies, usually the best approach is to have a dedicated BIG-IP HA pair for scalability and isolation purposes.

The F5 BIG-IPs shown in the figures should be considered logical F5 BIG-IPs that might be shared amongst tenants when appropriate. Please check the Multi-tenant considerations section for details.

[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.

There is a section with implementation details for each topology, and for Topology A there are three implementation options. At the end of the guide there is a Verifying the deployment summary for these.

These implementations sections don't detail with the cases where F5 BIG-IPs are used for multiple tenants. For these cases a section named Multi-tenant considerations has been introduced.

For a successful implementation, it is encouraged to check the following topics in the GENERAL NOTES section:

- General best practices for BIG-IP in VMware NSX-T.
- Performance best practices in VMware NSX-T.

This section also contains the following deployment specific topics:

- MAC Masquerade in NSX-T.
- Considerations for Tanzu Kubernetes Grid (TKG).
- BGP configuration details with NSX-T Edge nodes.

If you find there is any topic that should be covered or enhanced, please contact your sales representative.

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.

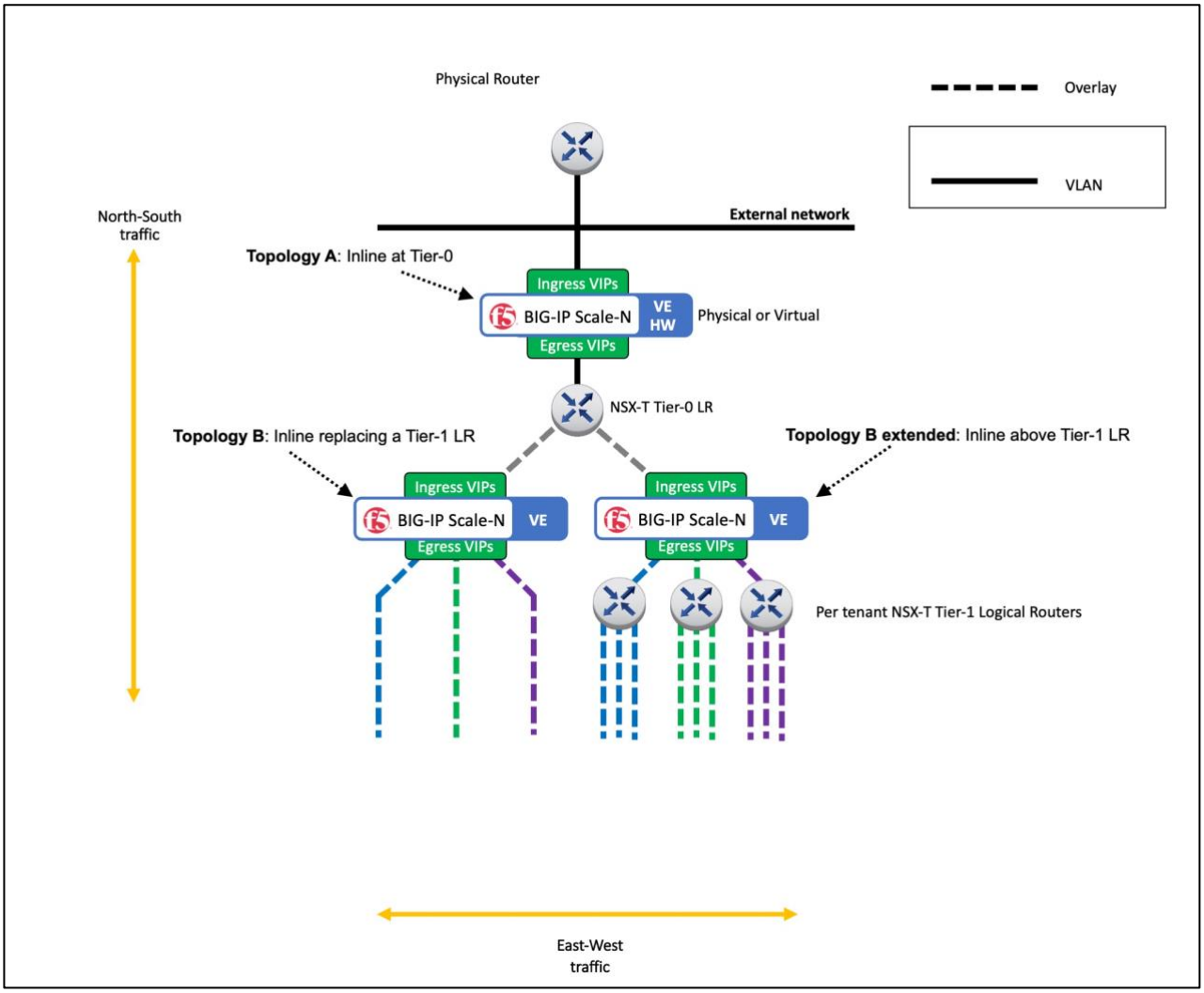


Figure 1 - BIG-IPs in inline-connected topologies A, B and B extended.

- **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 advanced WAF/WAAP, Firewall and anti-DDoS) and NAT if needed.

  For this topology three possible configurations are described:

  - NSX-T Edge cluster in Active-Standby HA mode using a static routing.

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

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

  In this topology it is proposed to eliminate NSX-T's Tier-1 Gateways to keep a 2-tier routing model while keeping BIG-IPs inline to the traffic path. This topology is like Topology A but allows per-tenant BIG-IP clusters, therefore allowing hard isolation between tenants with VEs.

  This topology only uses BIG-IP Virtual Editions and is only recommended when each tenant is going to have its own BIG-IP cluster. This topology is not recommended for multi-tenant in a single cluster, due to the limitations of the ESXi hypervisor, providing only 10 vNICs.

- **Topology B extended – BIG-IPs inline-connected to an NSX-T's Tier-1 Gateway.**

  This is the recommended per-tenant inline topology due to its greater flexibility. This topology allows the use of NSX services in the tenant's segments and allows a greater number of segments for the tenants. Because of this latter, it is also more suitable when a single BIG-IP cluster is going to be used for multiple tenants.

  Examples of NSX-T's services that this topology allows at Tier-1 are the Gateway Firewall or DHCP. When only doing plain packet forwarding the additional Tier-1 Gateway doesn't incur noticeable latency or throughput impact. On the other hand, when the Tier-1 Gateway makes use of the underlying Service Router (implemented in the NSX-T Edge nodes), it might have an impact. 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.

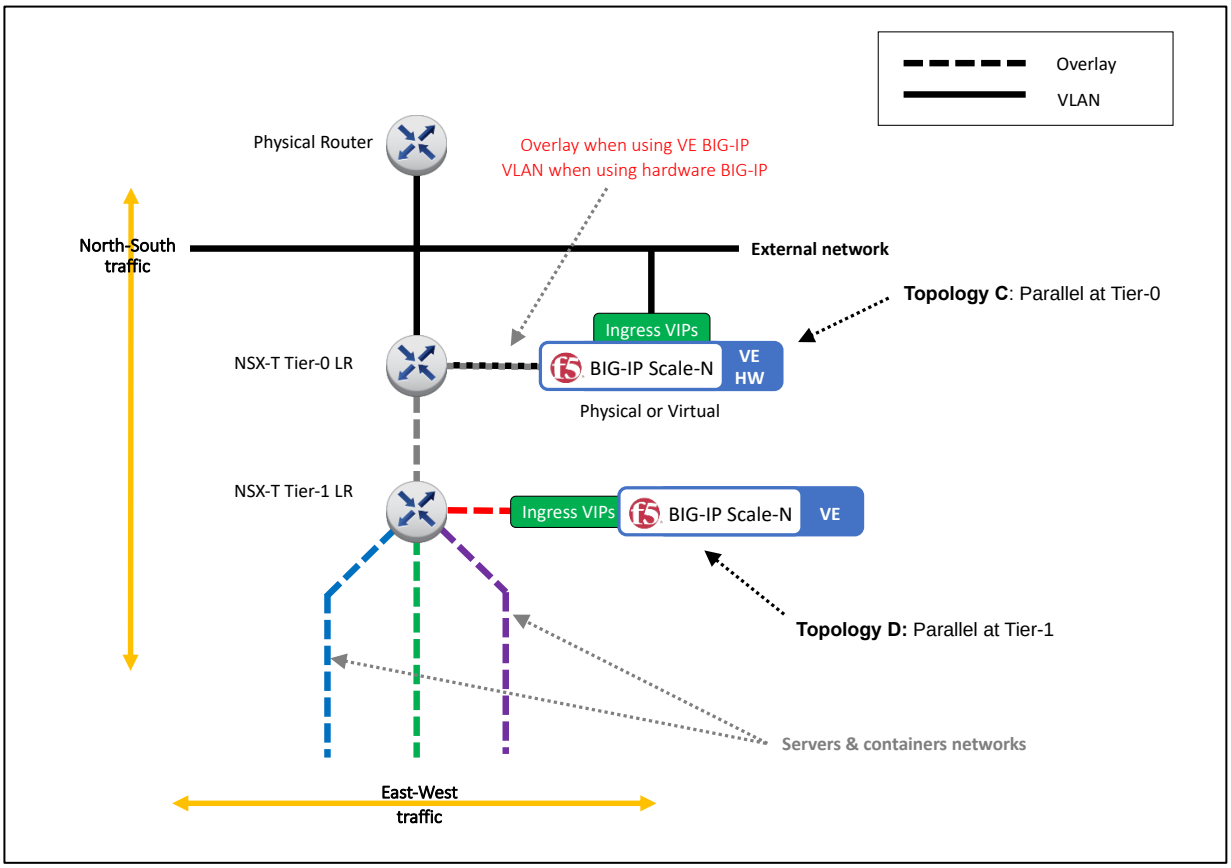


Figure 2 - BIG-IPs in parallel-connected topologies C and D.

- **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 HA 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 advanced WAF/WAAP, 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)	✓
B extended	Inline	Tier-1	✓		✓		✓ (per tenant)[3]	✓
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 can be an enforcement point for all traffic, and it is guaranteed no traffic will by-pass BIG-IP's topologies.

Tier: Whether 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 client address: The source IP address of the clients doesn´t need to be changed in the server side to guarantee symmetric return traffic through the BIG-IP. This avoids the need for 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

[3] It additionally allows the use of NSX-T Gateway firewall in a per tenant basis.

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 and NSX-T FW are enforcement points)	✓ (BIG-IP and NSX-T FW are enforcement points)
B extended	✓	NA (NSX-T FW is the enforcement point)	✓ (BIG-IP and NSX-T FW are the enforcement points)
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/extended or D should be chosen depending on if it is required that the BIG-IP is a tenant enforcement point. Note that Topology D doesn't allow the BIG-IP to be an enforcement point yet it allows distributed L3 forwarding by using only Tier-1 gateways for these flows.

In the case of topology B extended, each tenant is directly connected to its Tier-1 Gateway and the BIG-IP only participates in this traffic when it is not only routed within the Tier-1 Gateway, ie: the address of the virtual server is owned by the BIG-IPs.

Suitable for inter-tenant East-West: traffic that flows between Tier-1 Gateways.

For most topologies this means traffic that goes through the Tier-0 Gateway. In this case, the actual data path is typically different. When forwarding these flows, NSX-T typically takes advantage of distributed processing and traffic goes directly from VM to VM. From isolation point of view, the most natural place to setup FW rules is in the Tier-1 Gateways. Inline topologies such as Topology B and Topology B extended can provide additional advanced WAF/WAAP L7 security.

In the case of topology B extended, connected to the BIG-IPs there is a fan-out of Tier-1 Gateways and therefore traffic between tenants in different Tier-1 Gateways can happen without Tier-0 Gateway intervention.

Multi-tenant considerations

Unless otherwise stated in this guide, the regular NSX-T 2-Tier nomenclature is used. Following this convention, a Tier-0 Gateway is a provider construct and a T1 Gateway is tenant construct. A tenant will have all its objects connected to a single Tier-1 Gateway and the Tier-0 Gateway will provide isolation between them.

Sometimes, further isolation is required, and tenants have their own Tier-0. A use case of this is to avoid tenants sharing CPU resources in the Edge Nodes. This guide focuses on topologies with a single Tier-0. Multi-Tier-0 designs can be extrapolated from these.

In the topologies discussed in this guide, the F5 BIG-IPs shown in the designs should be contemplated as logical entities. A single F5 BIG-IP can be partitioned in a way that several tenants can be managed by it independently:

- Using F5 BIG-IP hardware with the vCMP feature allows several instances of F5 BIG-IP to run with **hard isolation**.

  Using this hard isolation between the tenants there are no shared resources and per-tenant performance is guaranteed. vCMP hardware also provides the highest security isolation.

- When using F5 Virtual Edition (VE), the general recommendation is to have dedicated VMs for each tenant. This achieves the same **hard isolation** as vCMP.

  Sometimes, the latter is not found appropriate for some deployments. A single F5 BIG-IP VE can also have several tenants by means of using logical partitions and route domains. Please note this is considered **soft isolation** and therefore hardware resources are not isolated or guaranteed between tenants. Single F5 BIG-IP VE multi-tenancy is limited by the ESXi hypervisor which limits the number of vNICs to 10 for each VM.

The PROs and CONs of each option are outlined in the next figure.

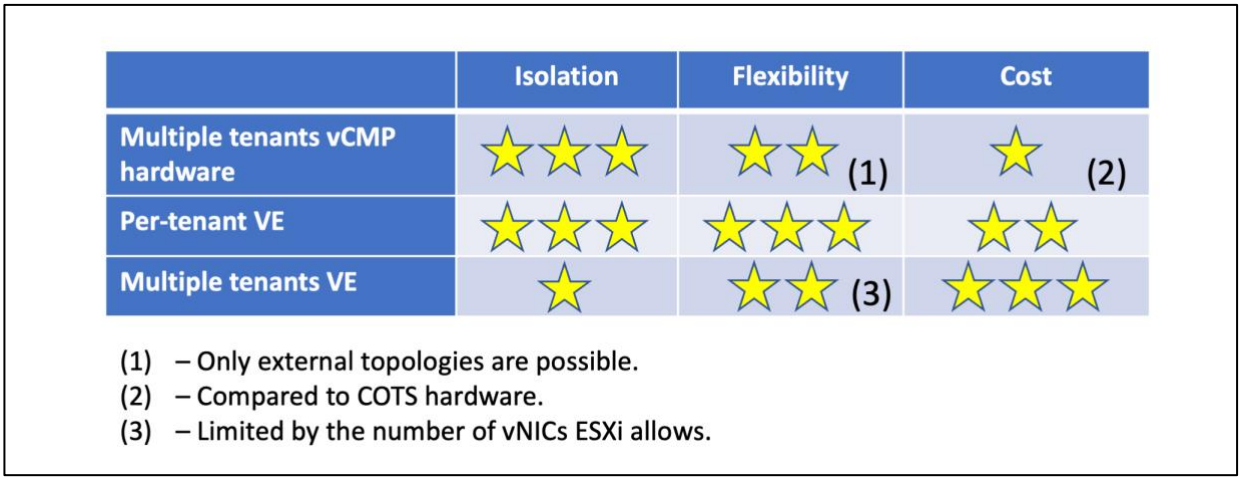


Figure 3 - Comparing different multi-tenant options.

At the end of each topology's section, there is a multi-tenancy subsection outlining how to implement multi-tenancy in each case.

For general information on multi-tenancy features, please check the following guides:

- vCMP – feature that allows to provision and manage multiple, hosted instances of the BIG-IP software on a single hardware platform.
- Administrative partitions – a logical container that you create, containing a defined set of BIG-IP system objects.
- Route domains – feature that isolates network traffic.

NSX-T network & general 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 LACP[4] 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 HA mode with ECMP without restrictions - NAT in Tier-0 can only run in Active-Active when using Reflexive (stateless) mode[5].

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

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

[4] LACP - Link Aggregation Control Protocol is an IEEE standard.
[5] Reflexive NAT - https://docs.vmware.com/en/VMware-NSX-T-Data-Center/2.5/administration/GUID-46900DFB-58EE-4E84-9873-357D91EFC854.html

- When using NSX-T Edge in Active-Active HA 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, and 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.

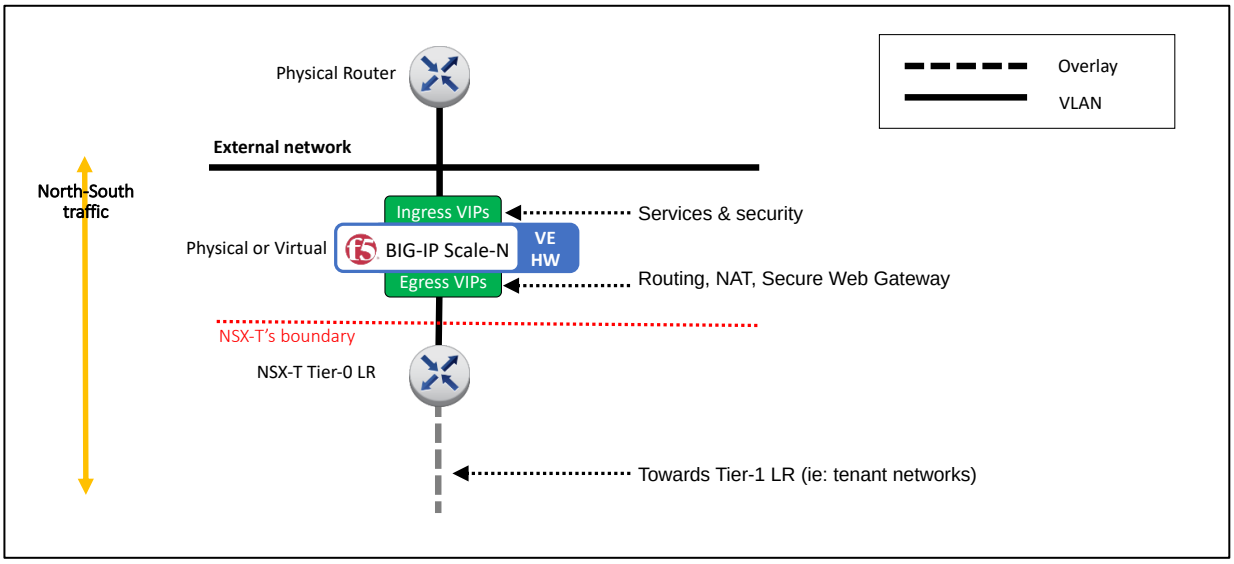


Figure 4 – Topology A overview (simplified view without HA components).

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 BIG-IP clusters can be either deployed as hardware or as 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 HA mode with ECMP, it is not able to run stateful services (ie: edge firewall, load balancing, or NAT except for 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).

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

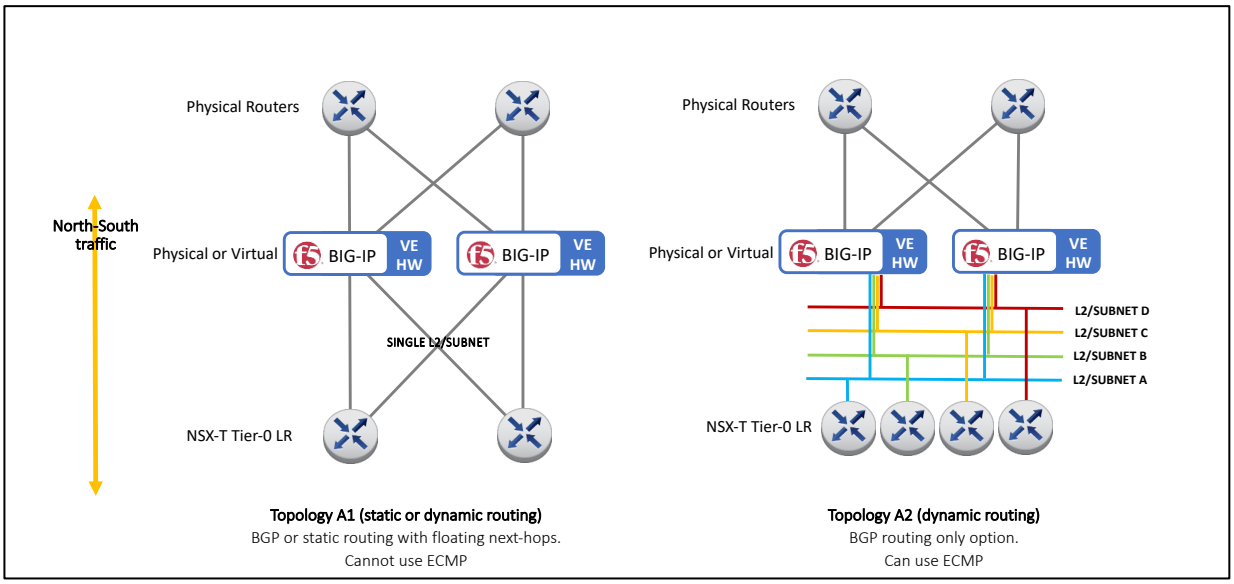


Figure 5 - L1/L2 options for Topology A.

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

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

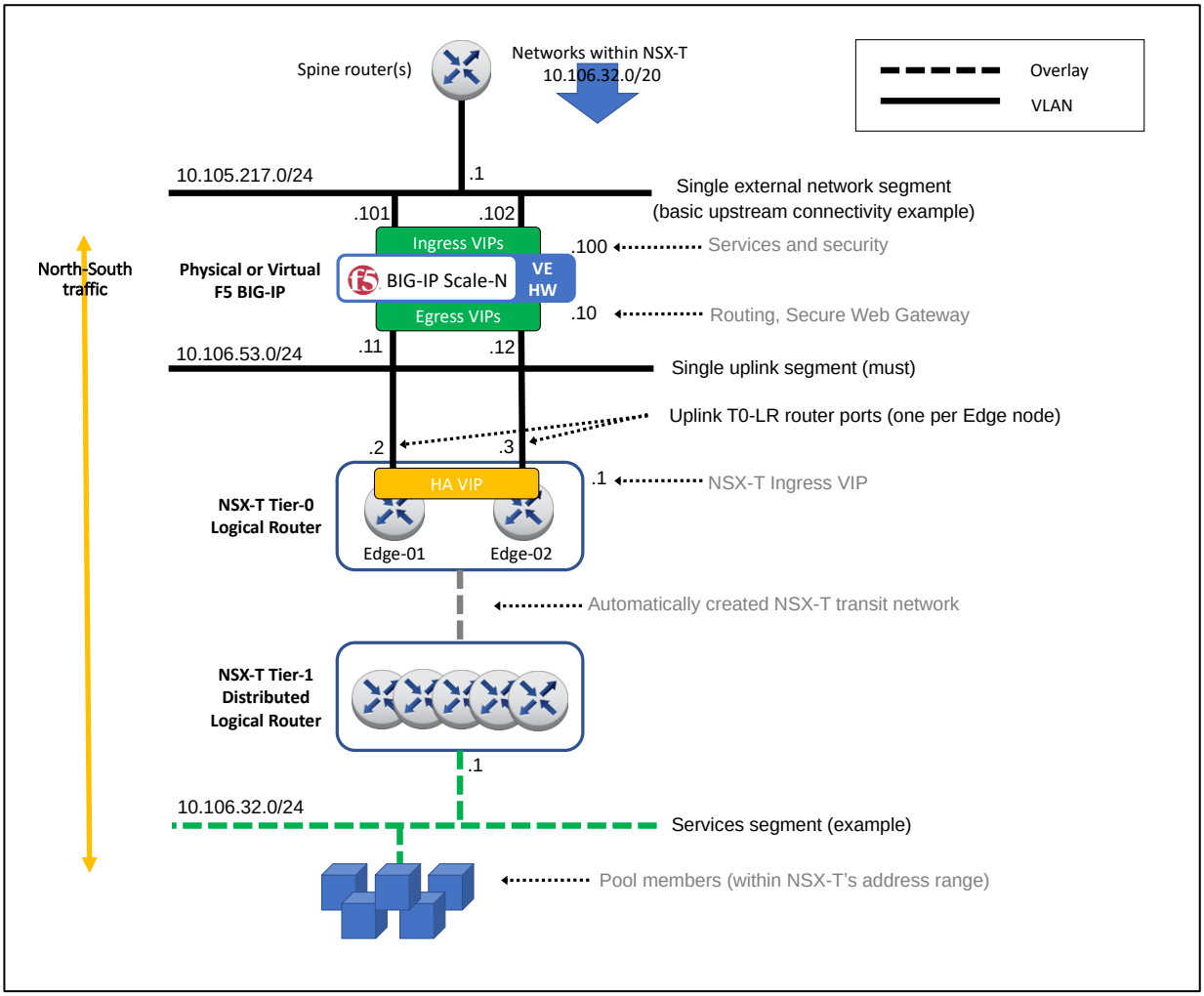


Figure 6 – Example of topology A using static routing used through this section.

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 on 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 HA mode.

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

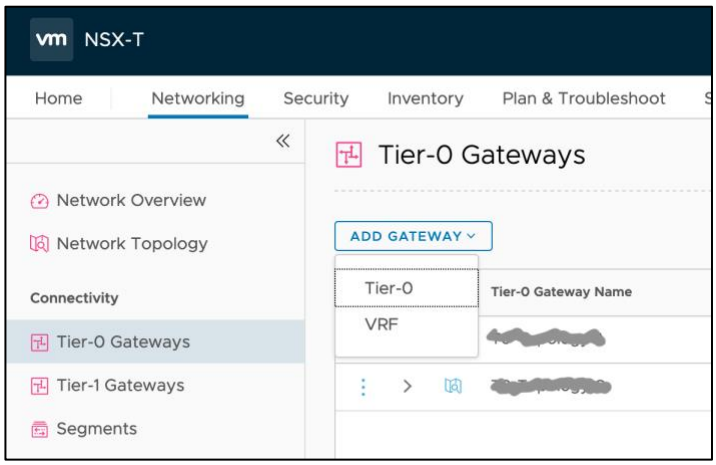


Figure 7 - Adding a Tier-0 Gateway/Gateway.

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).

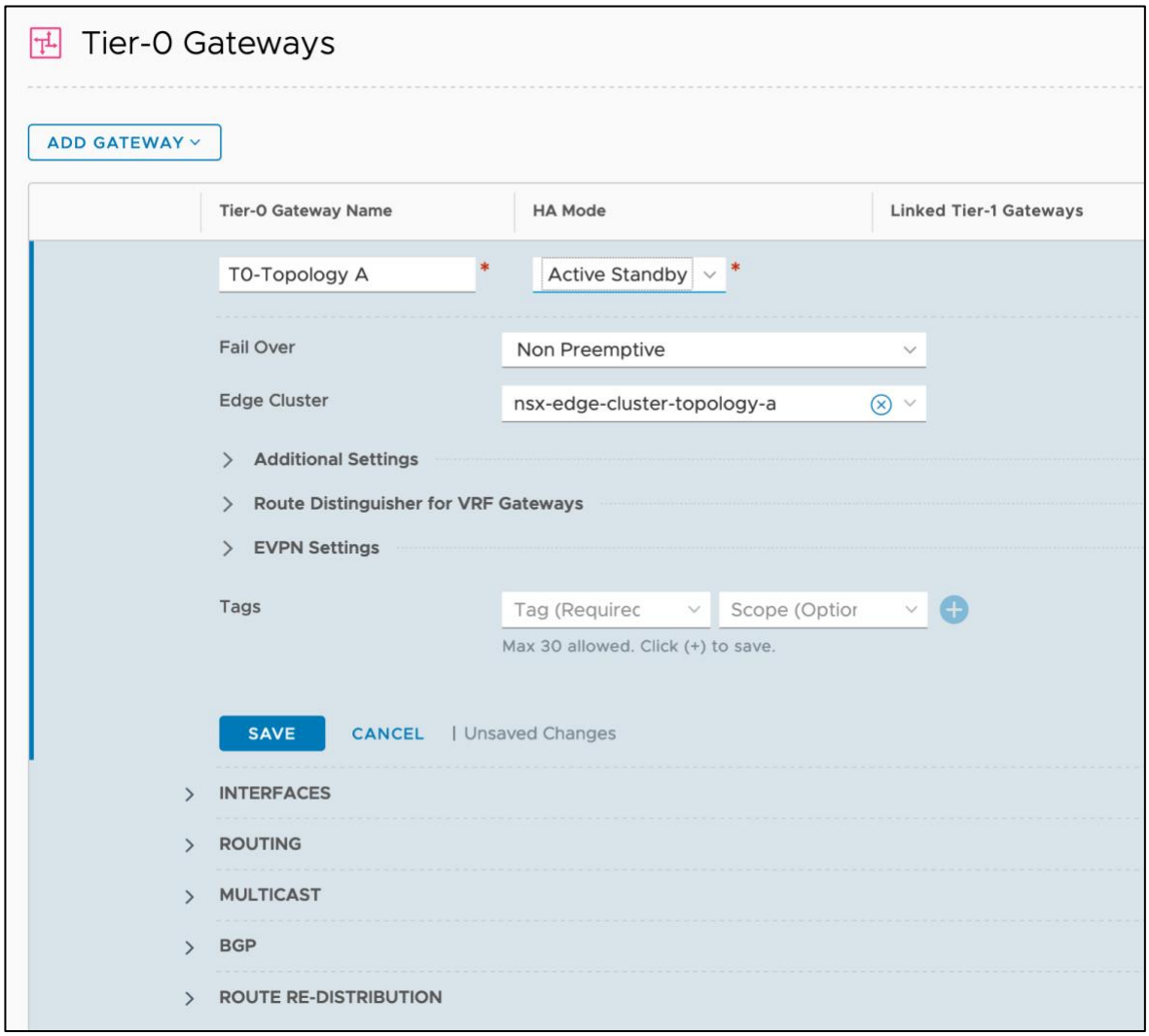


Figure 8 - Filling the details of a Tier-0 Gateway/Gateway.

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

`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 pre-requisite 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.

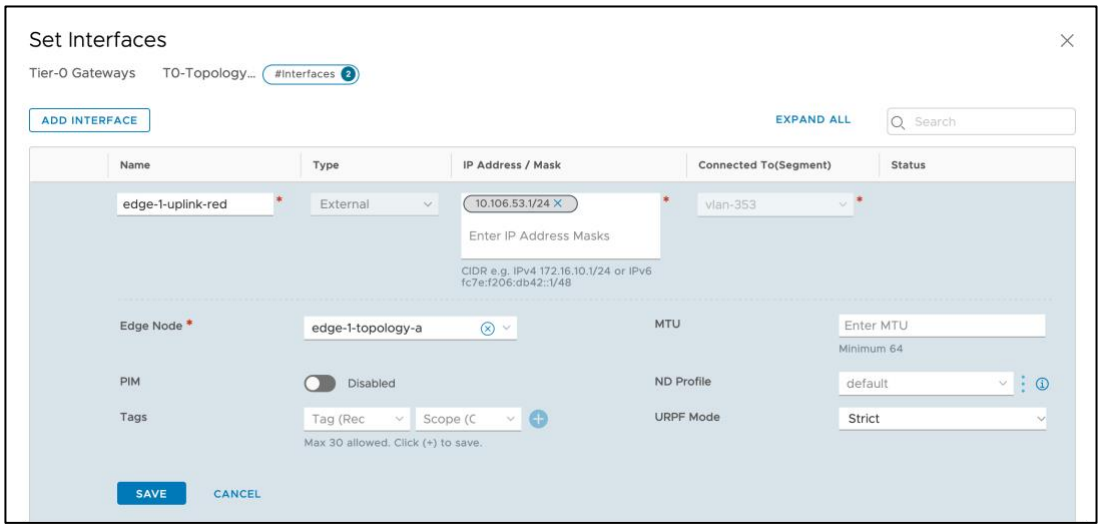


Figure 9 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.

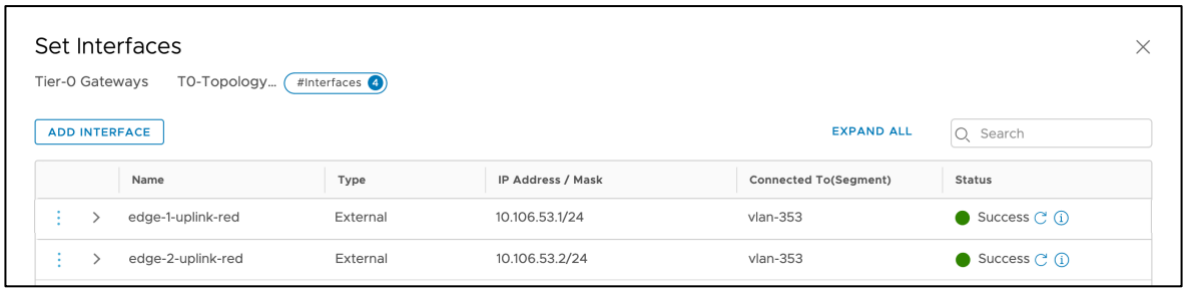


Figure 10 – Final Gateway Port configuration of the Tier-0 Gateway.

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:

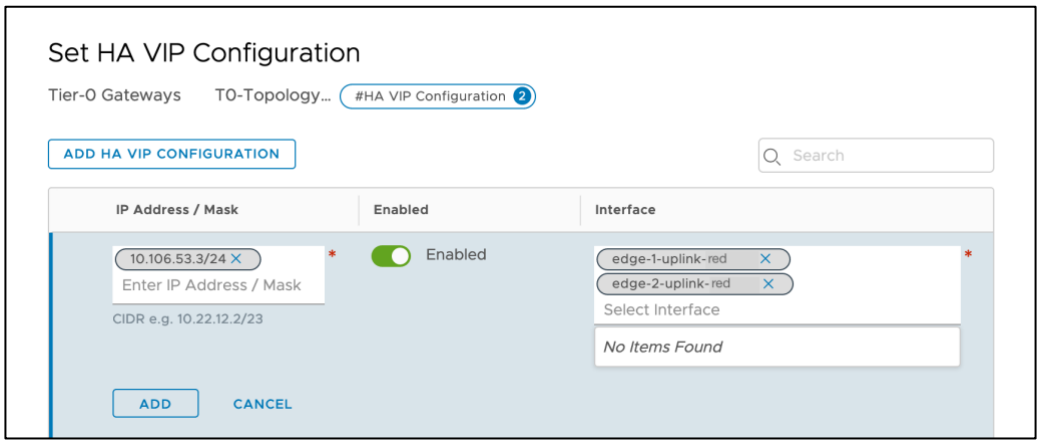


Figure 11 - Adding an HA VIP to NSX-T's T0 Gateway.

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`:

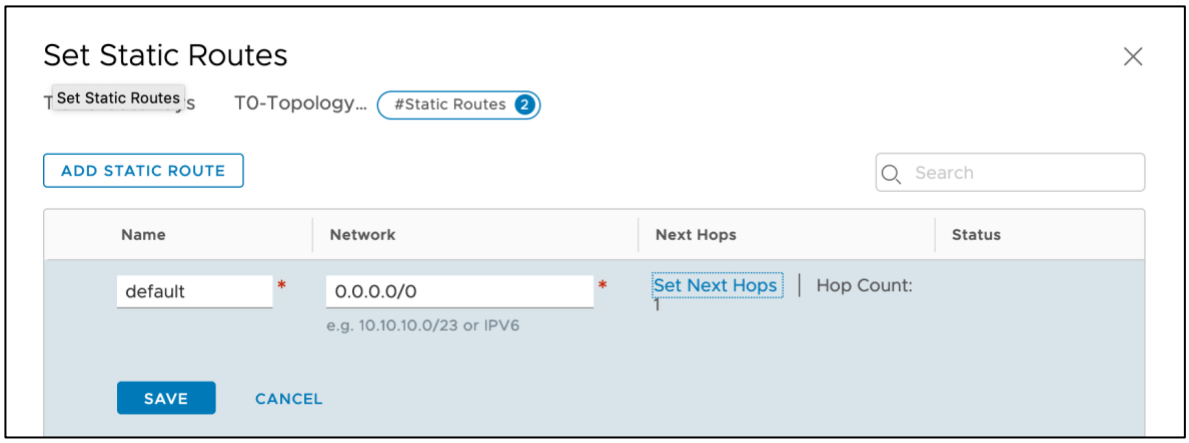


Figure 12 – Adding Tier-0 Gateway's default route.

2. Create a Tier-1 Gateway.

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

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`.

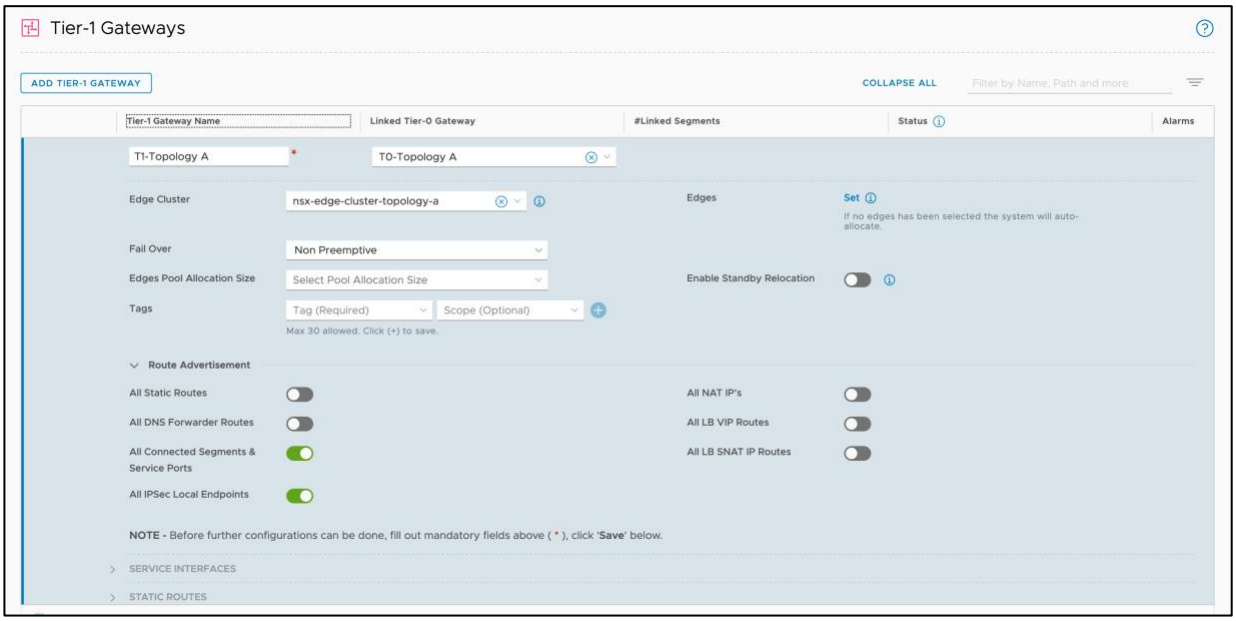


Figure 13 – Filling the properties when creating a Tier-1 Gateway.

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:

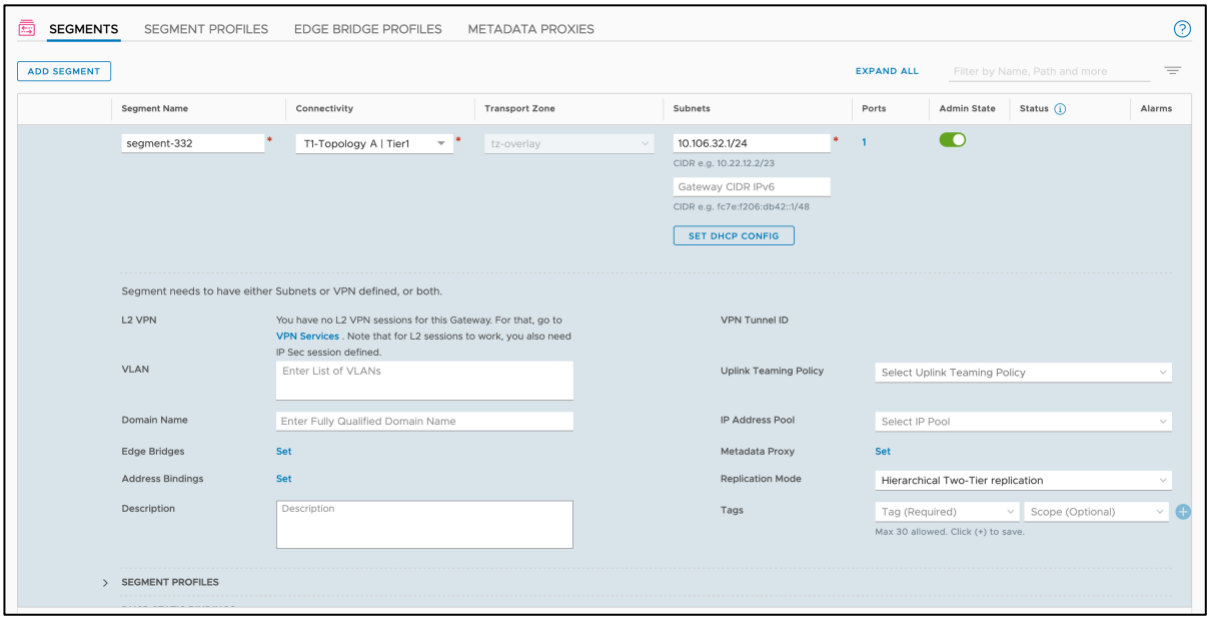


Figure 14 - Adding a segment to the T1 Gateway.

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.

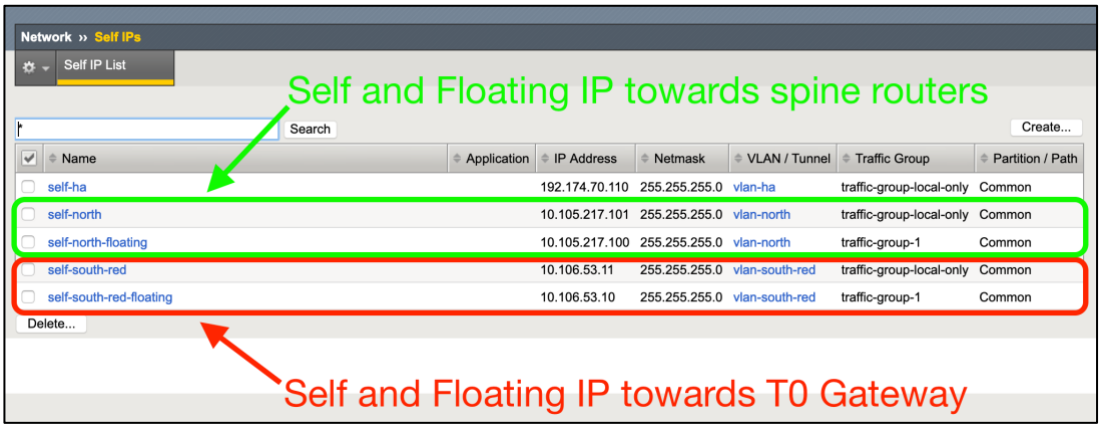


Figure 15 – Self IPs and floating Self IPs required (shown in BIG-IP unit 1).

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).

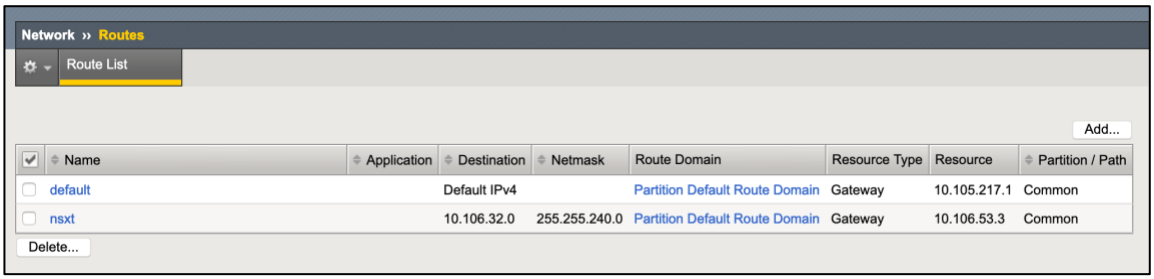


Figure 16 – Static routes required in the BIG-IP units.

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.

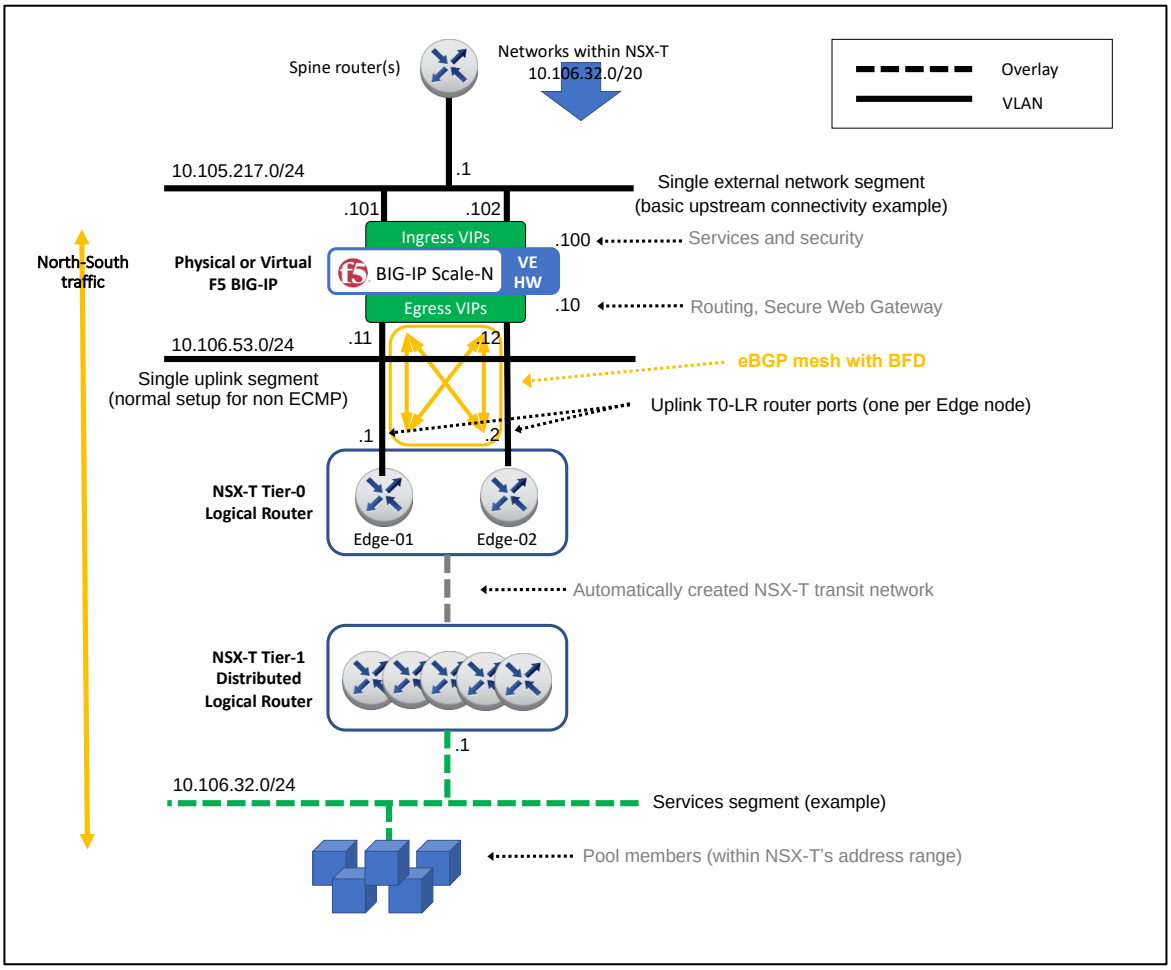


Figure 17 – Example of topology A using BGP routing used through this section.

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 on 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 HA mode.

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

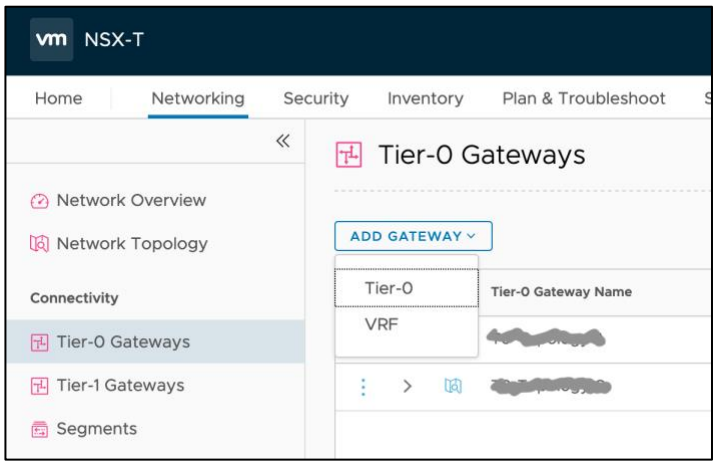


Figure 18 - Adding a Tier-0 Gateway.

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).

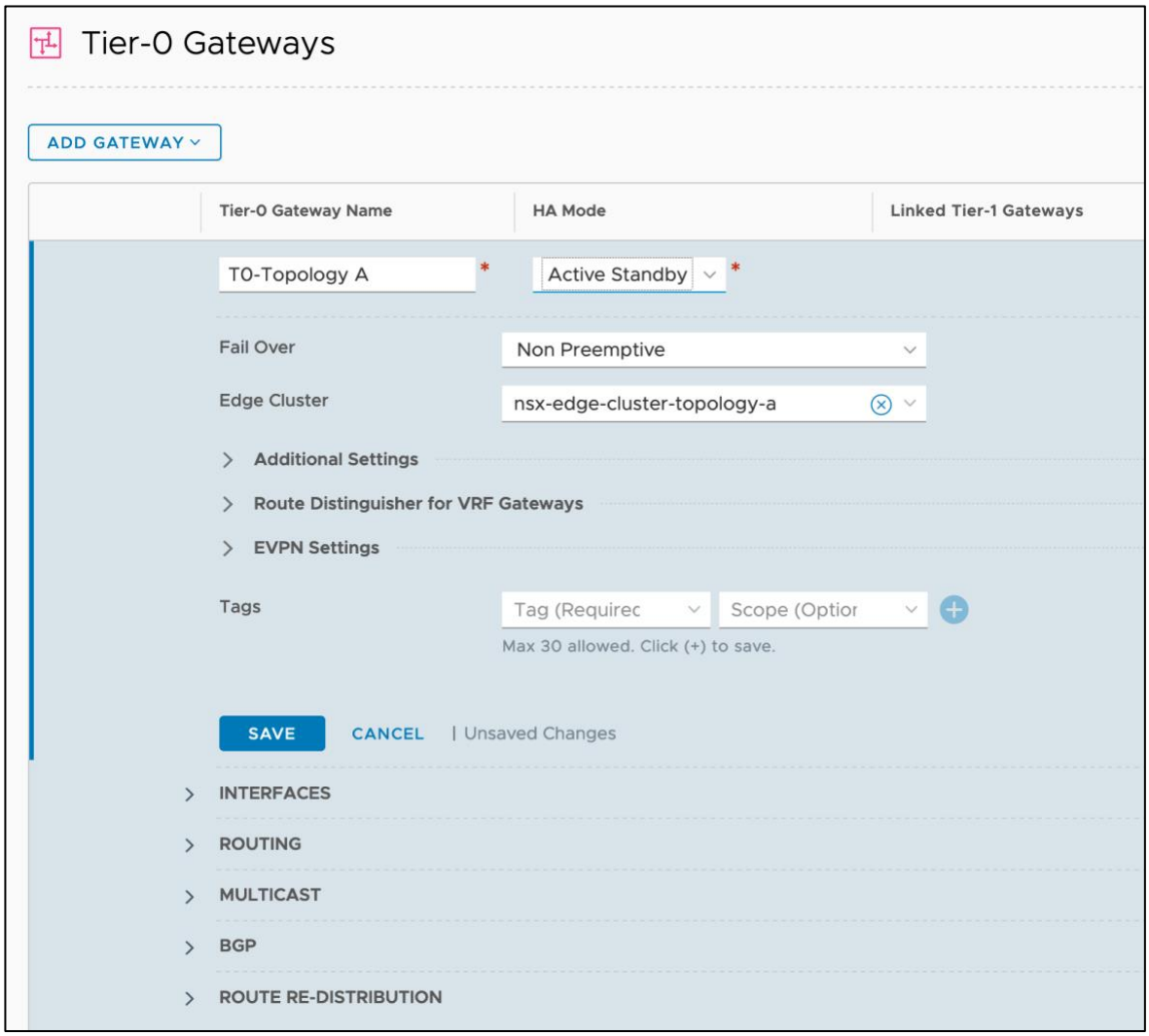


Figure 19 - Filling the details of a Tier-0 Gateway/Gateway.

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

`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 pre-requisite 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.

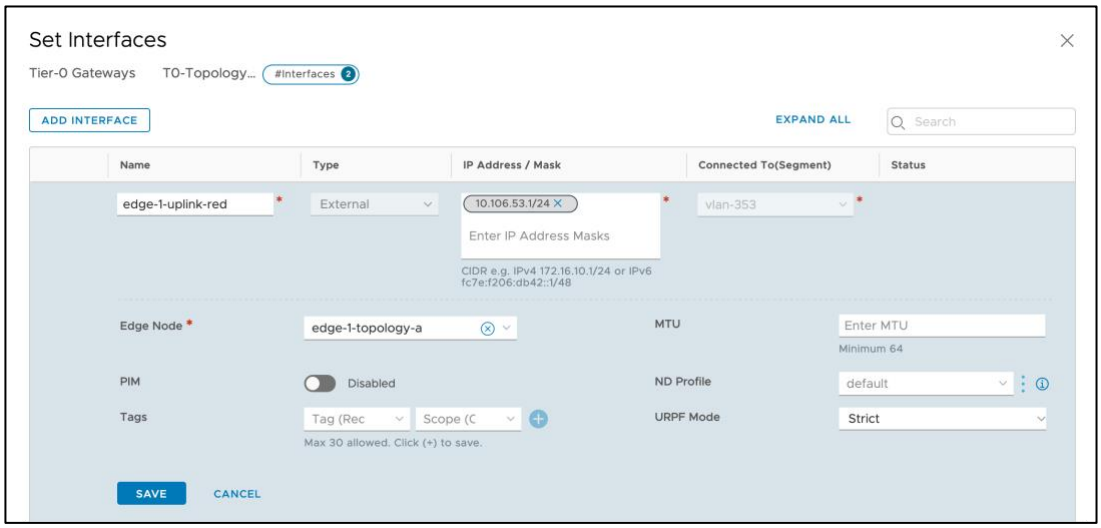


Figure 20 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.

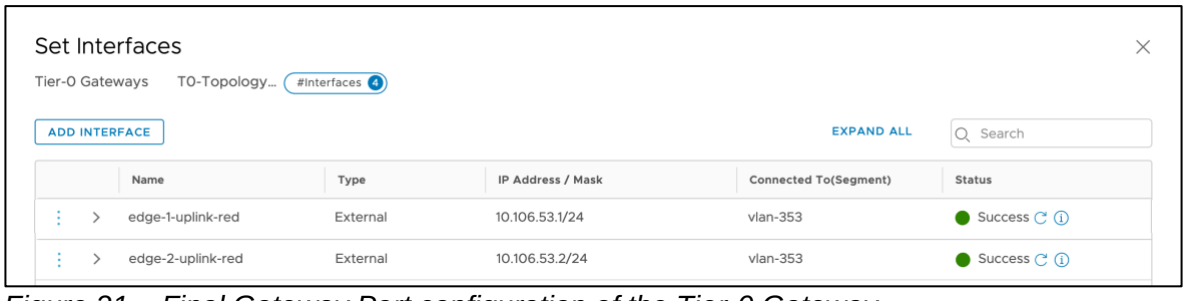


Figure 21 – Final Gateway Port configuration of the Tier-0 Gateway.

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.

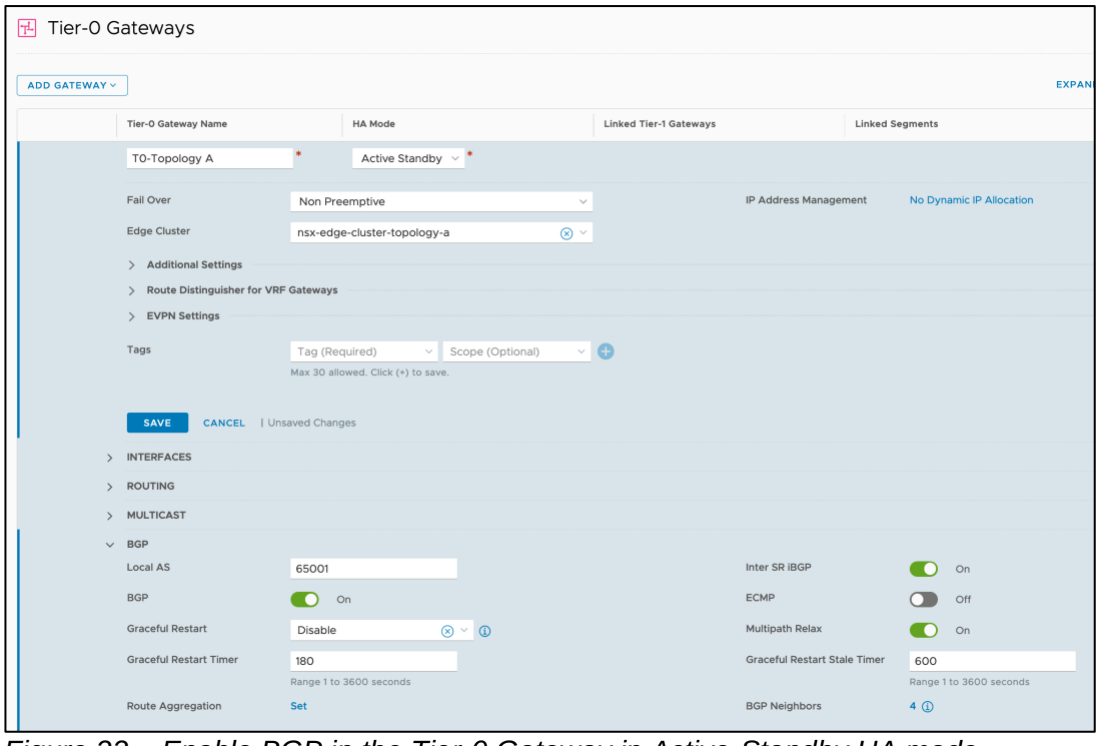


Figure 22 – Enable BGP in the Tier-0 Gateway in Active-Standby HA mode.

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.

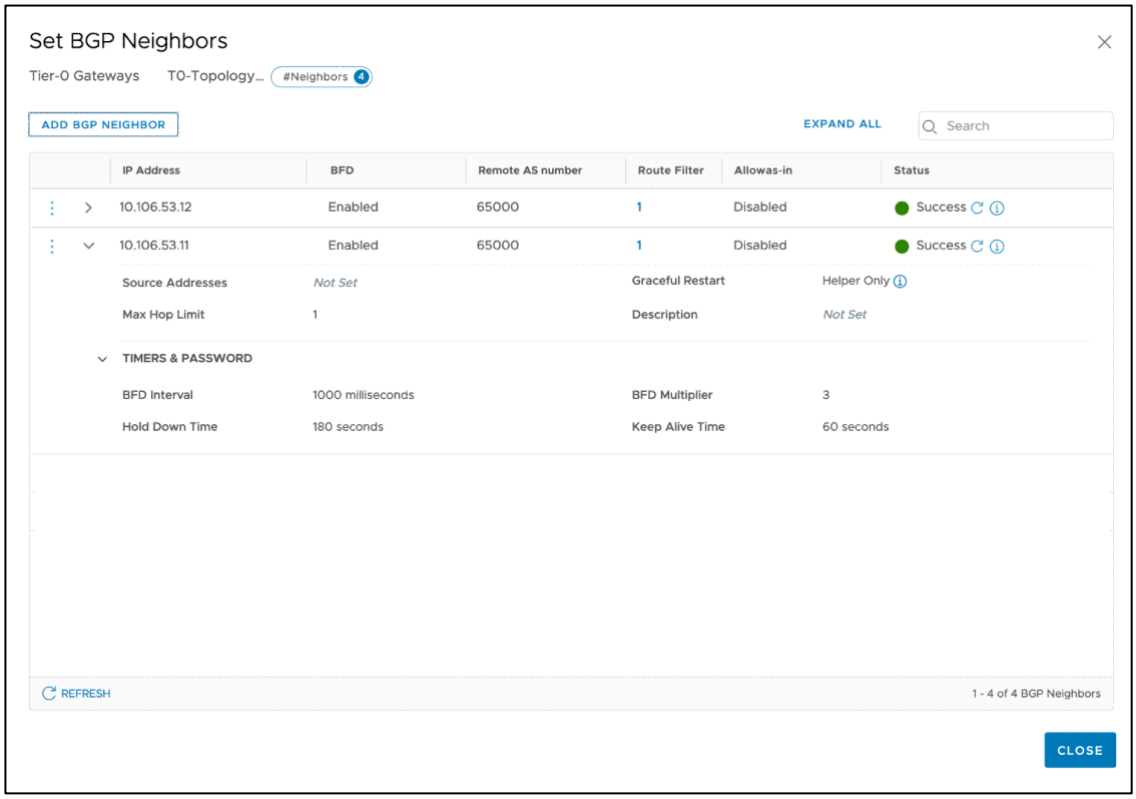


Figure 23 – Adding a BGP neighbor.

In this figure, the default values are used except for 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 24 - Enabling Route redistribution at T0 Gateway

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

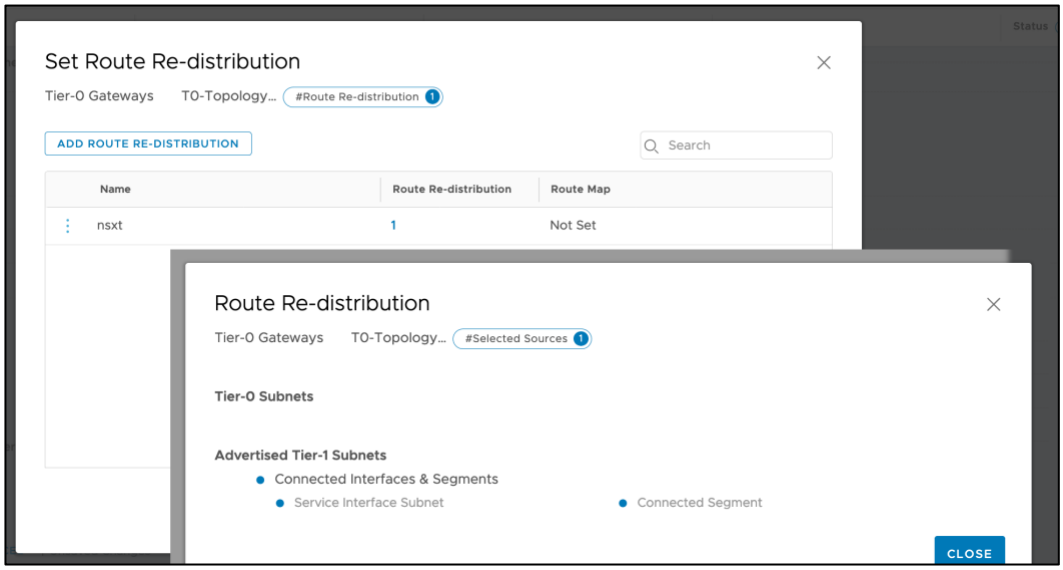


Figure 25 - Route redistribution settings at T0 Gateway

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.

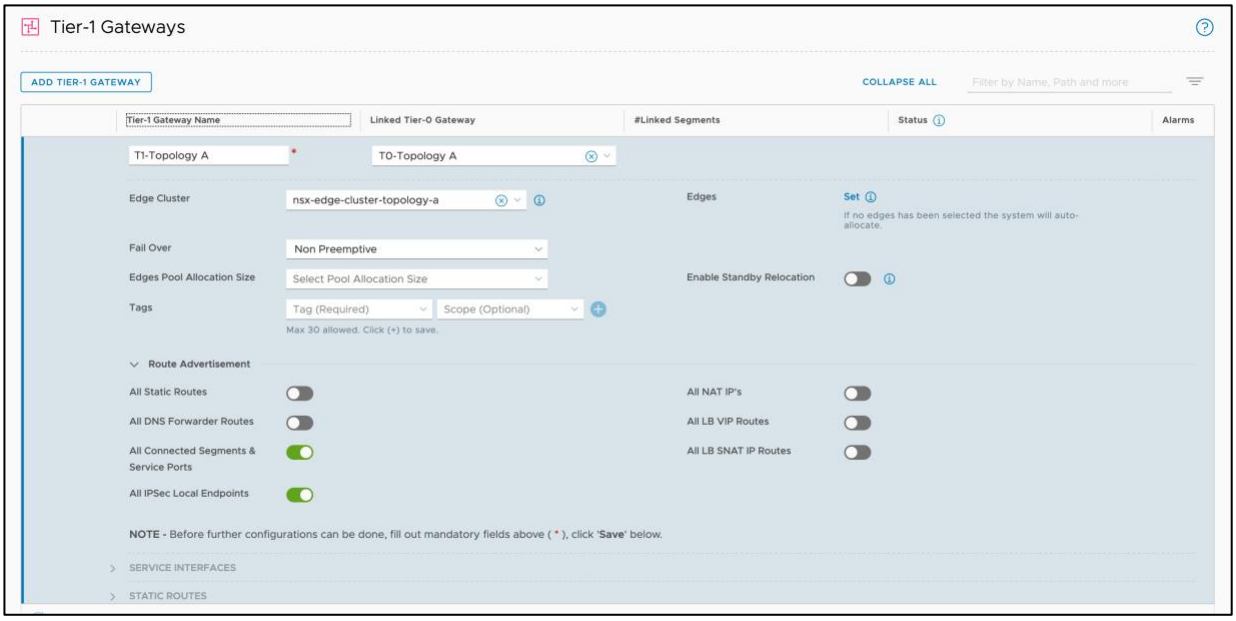


Figure 26 – Filling the properties when creating a Tier-1 Gateway.

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:

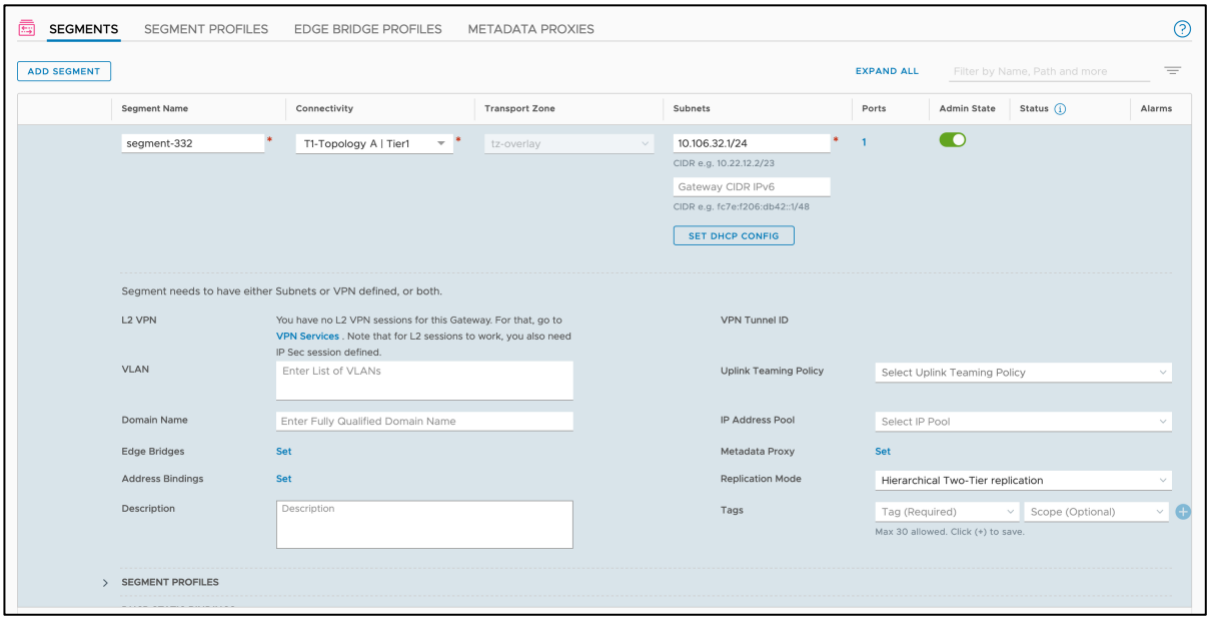


Figure 27 - Adding a segment to the T1 Gateway.

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.

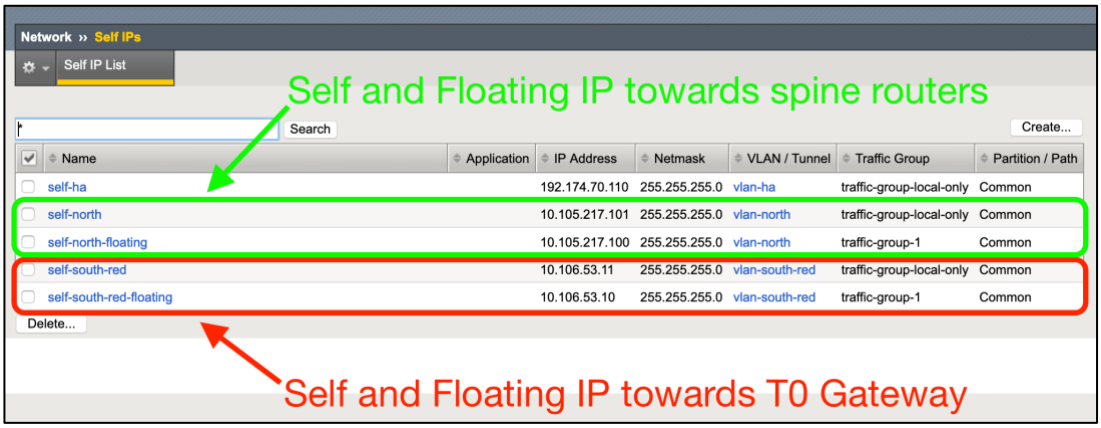


Figure 28 – Self IPs and floating Self IPs required (shown in BIG-IP unit 1).

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.

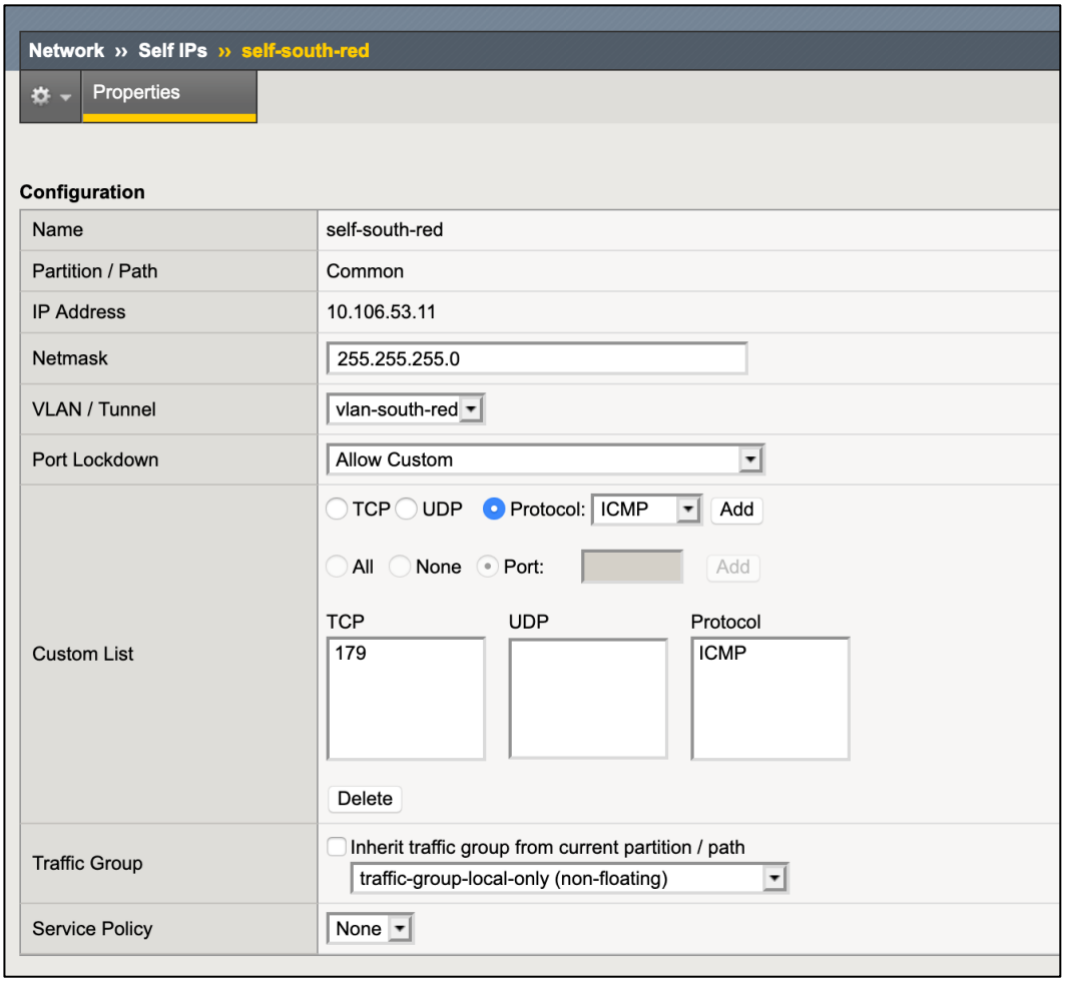


Figure 29 – Allowing BGP in the non-floating Self IPs. ICMP is added for allowing basic connectivity tests.

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.

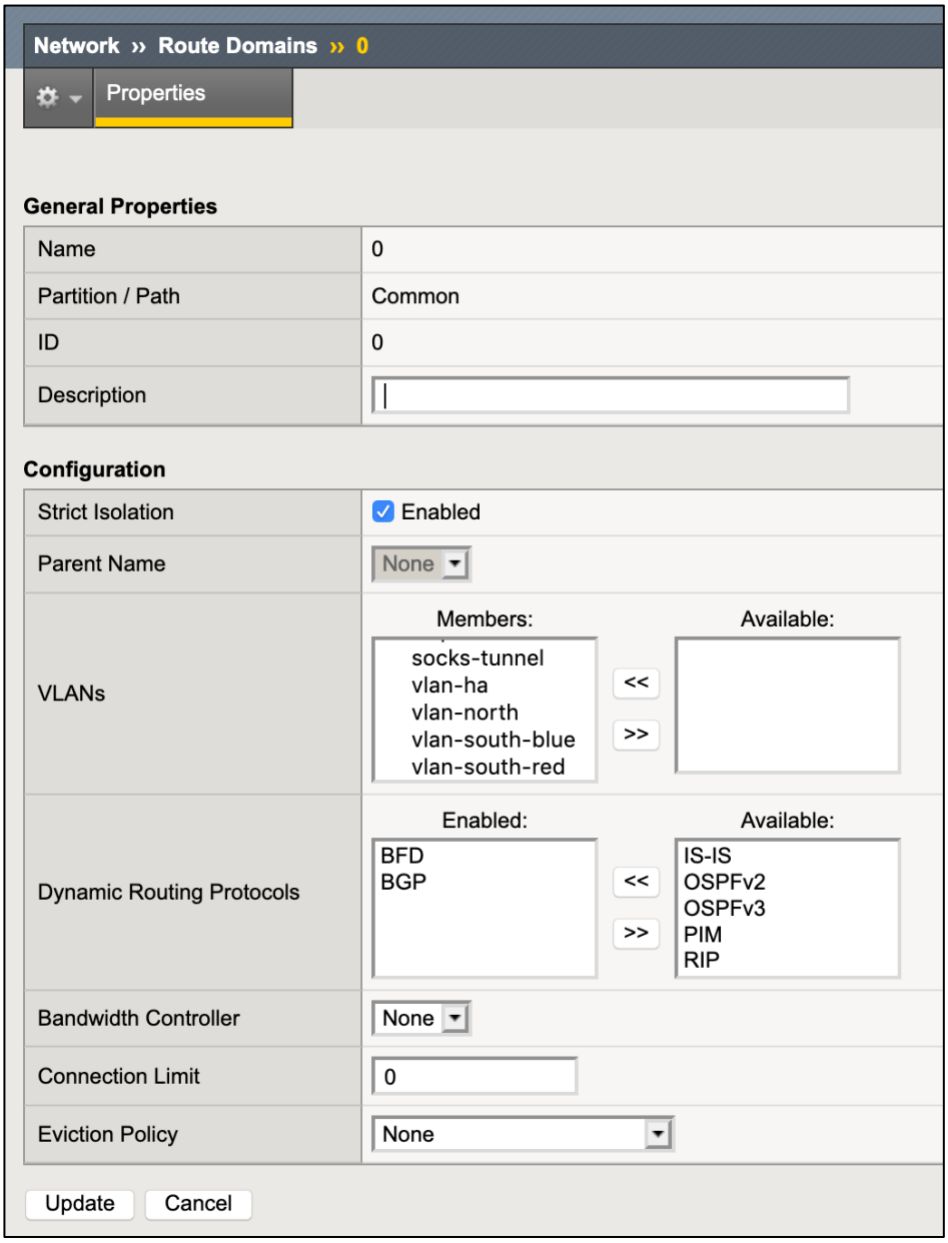


Figure 30 – Enabling BFD and BGP in F5 BIG-IP. This must be performed in all units.

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.

```
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 gtsm 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 31 – 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 HA mode. 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 ECMP[6] 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.

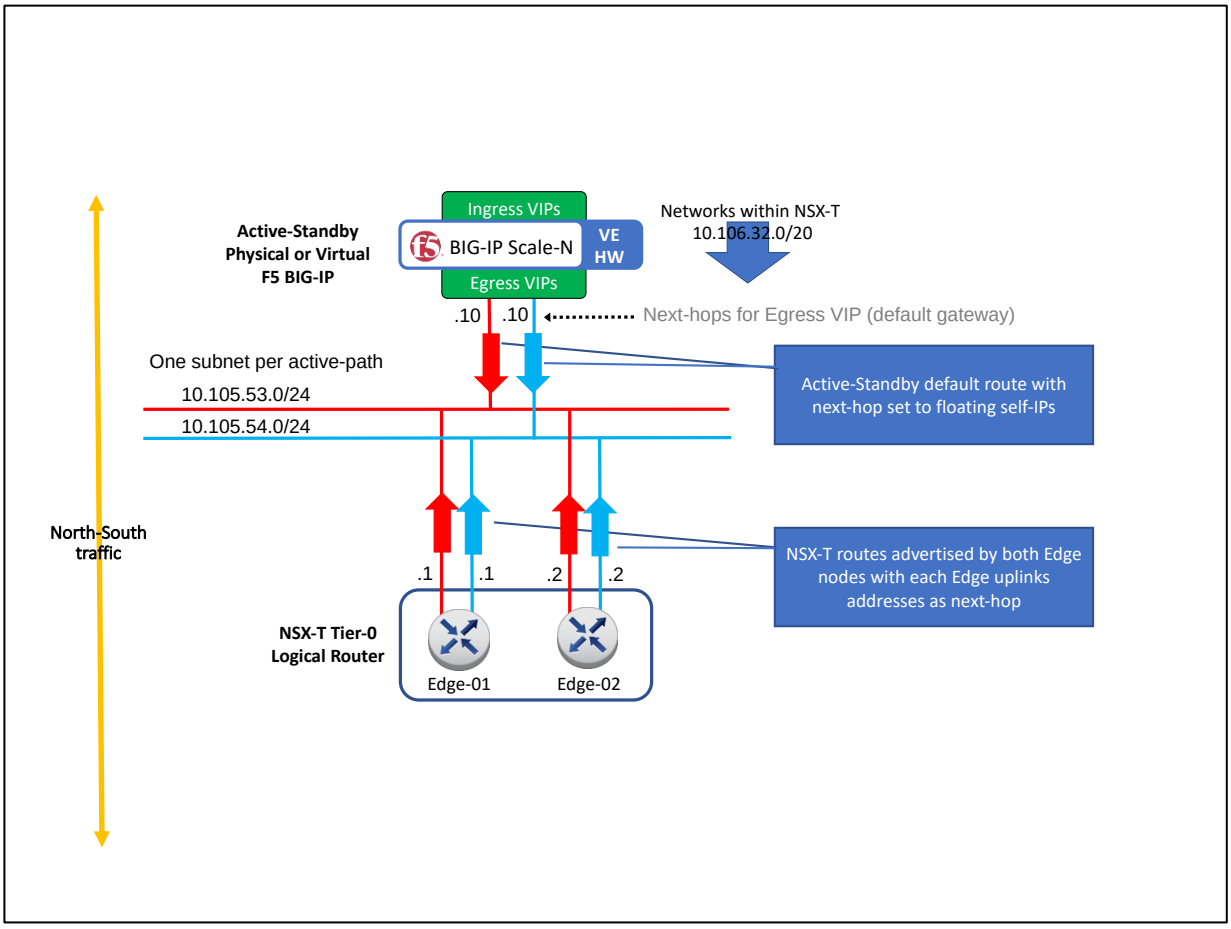


Figure 32 – Active-Active NSX-T Edge with two ECMP uplinks and BIG-IP in Active-Standby.

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

[6] 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 33 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.

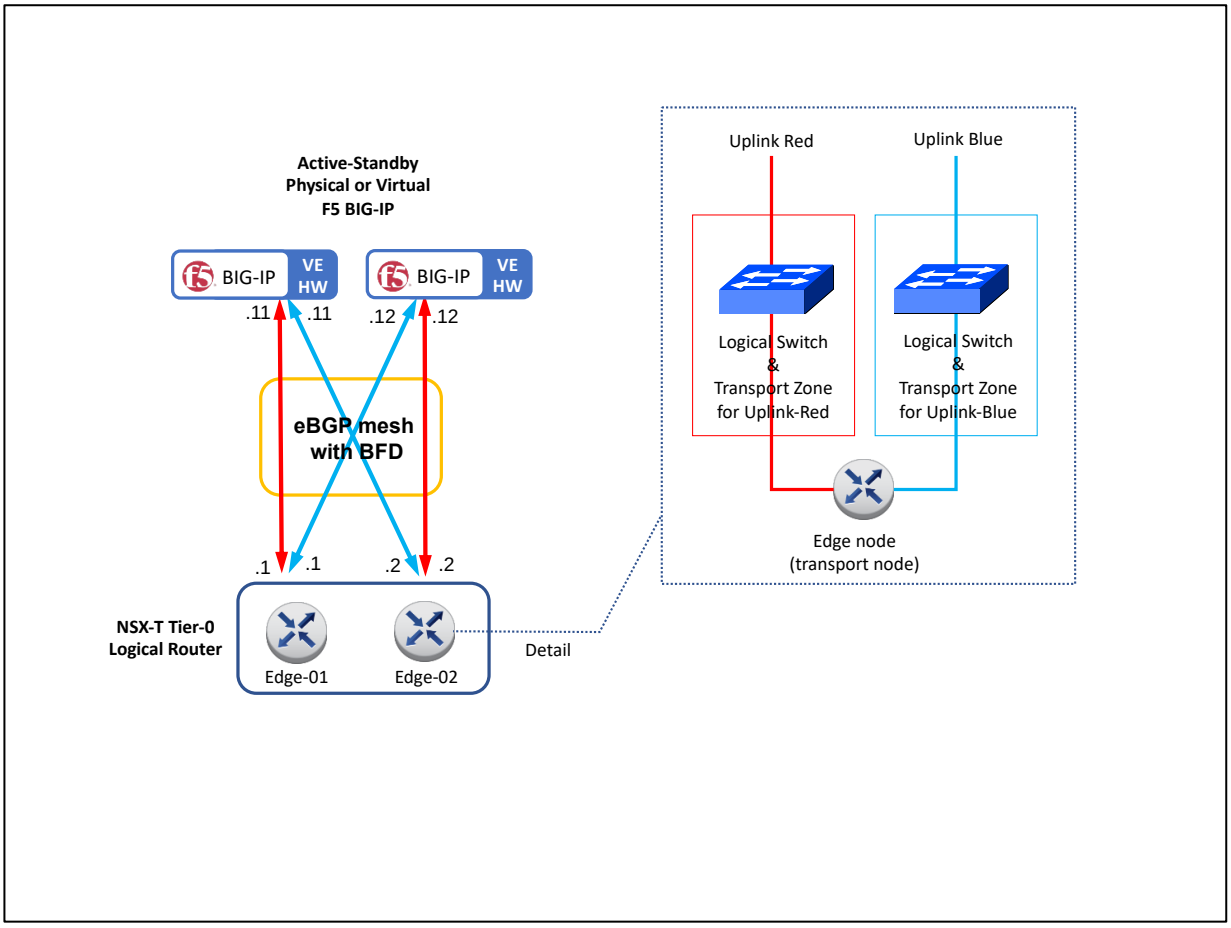


Figure 33 – BGP peering detail with two uplink Layer 3 paths & transport zones for ECMP.

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 on 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

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.

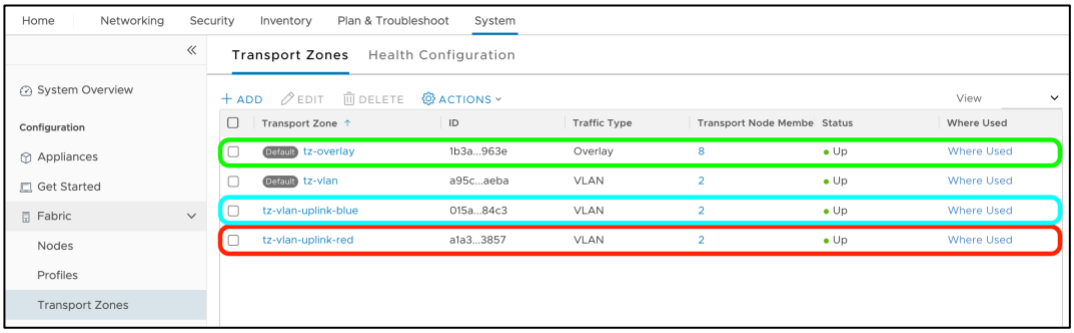


Figure 34 – Overall configuration of transport zones. The used ones by this topology are highlighted (red and blue for the uplinks).

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.

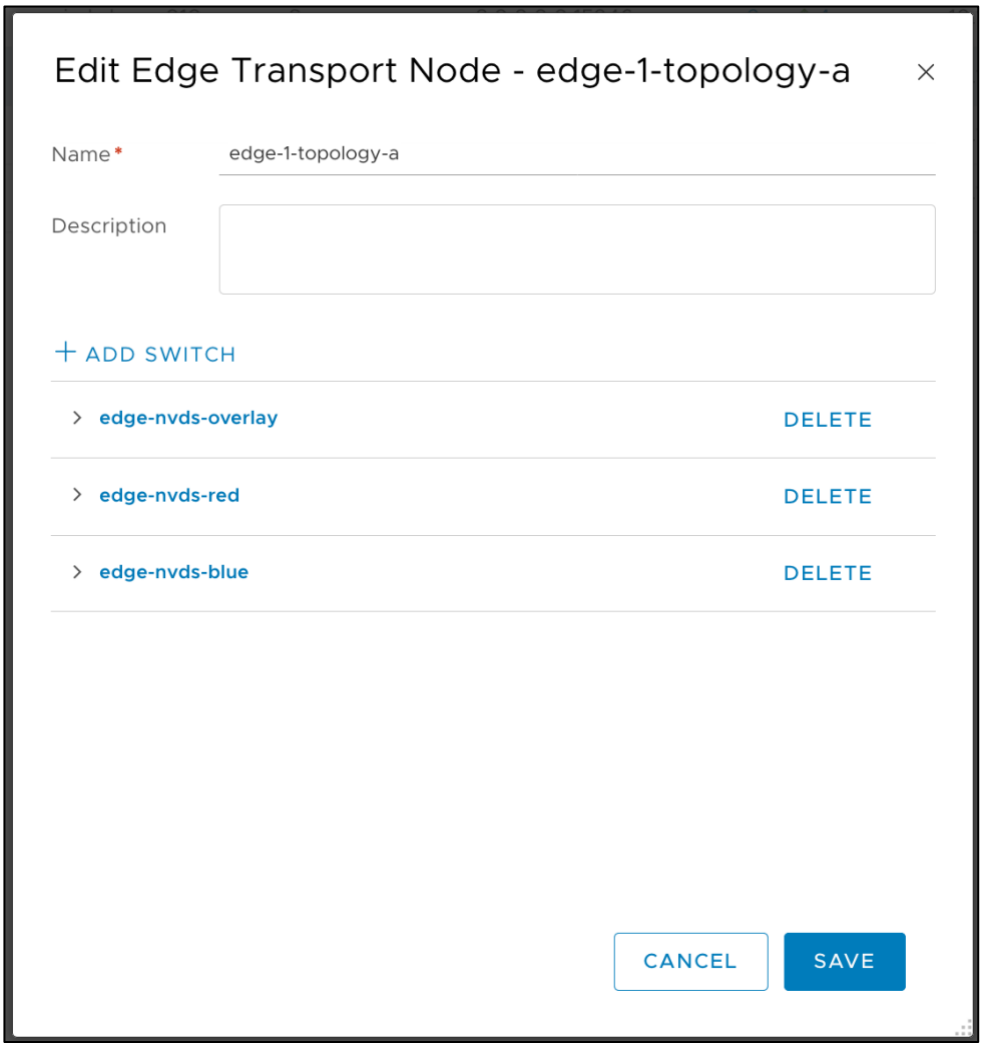


Figure 35 – Adding the switches for each Uplink transport zone in each Edge transport nodes.

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.

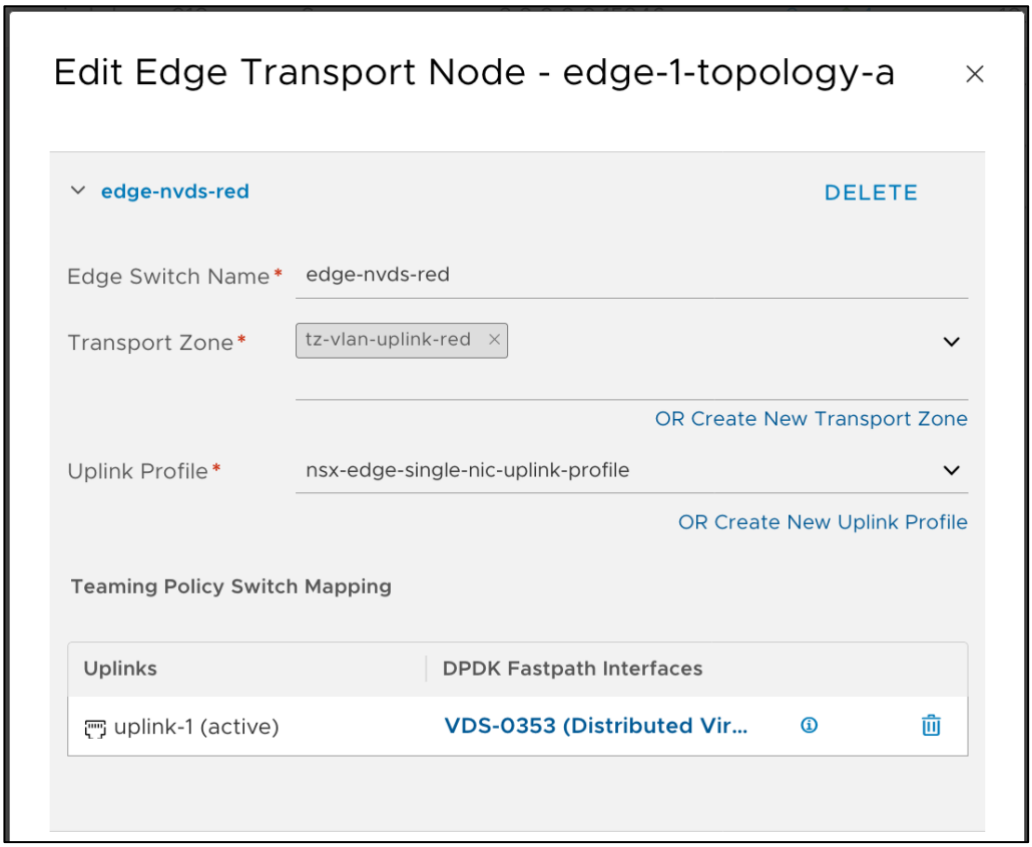


Figure 36 – N-VDS switch configuration for a sample Uplink transport zone.

3. Create a Tier-0 configuration.

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

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

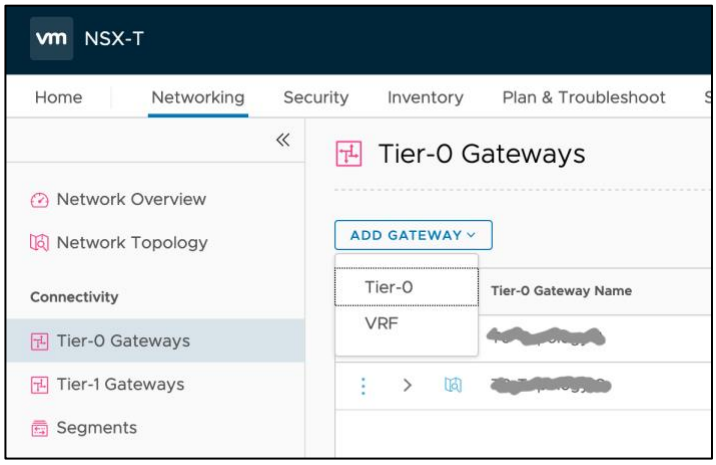


Figure 37 - Adding a Tier-0 Gateway.

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`.

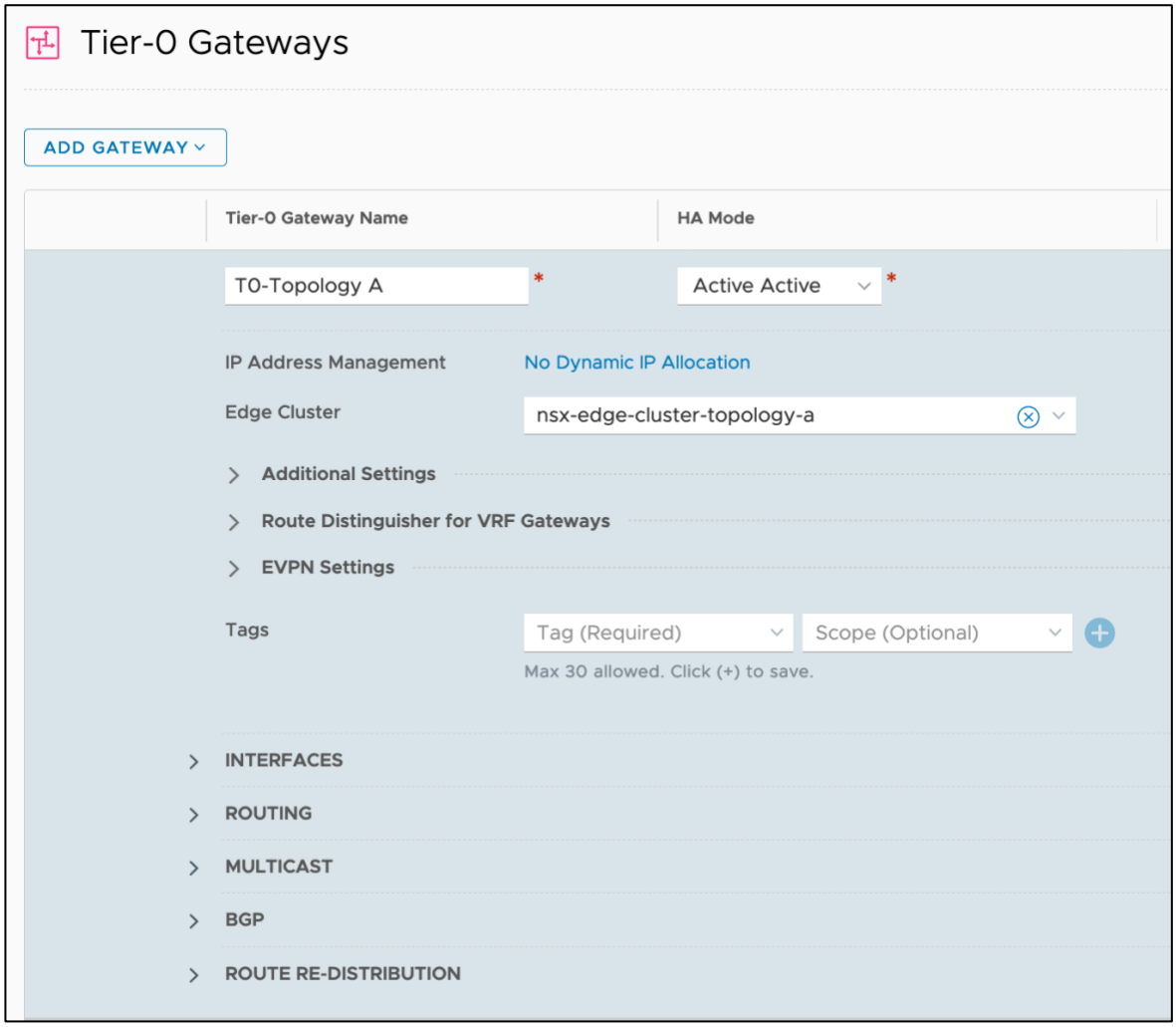


Figure 38 - Filling the details of a Tier-0 Gateway in Active-Active HA mode.

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.

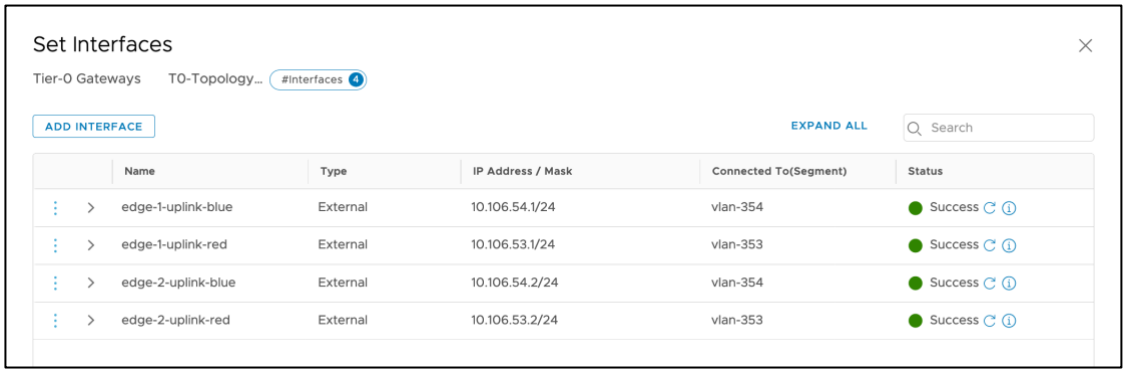


Figure 39 – Adding the Gateway's interfaces for the Uplink paths.

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

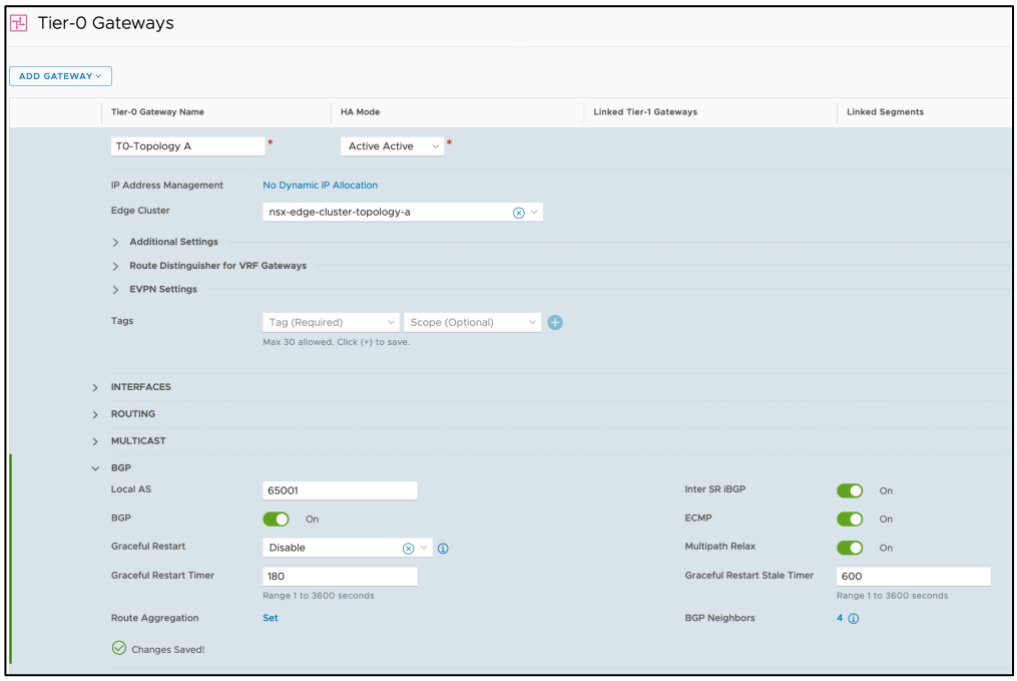


Figure 40 - Enable BGP with ECMP in the Tier-0 Gateway in Active-Active HA mode.

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:

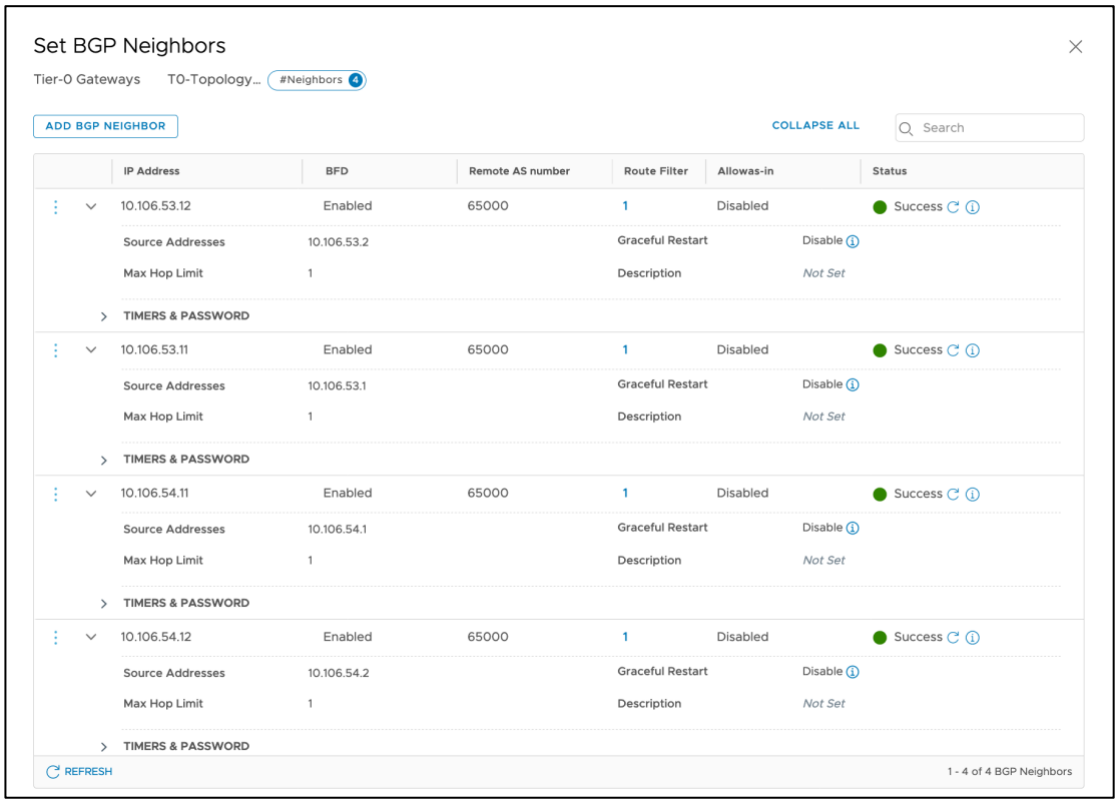


Figure 41 – BGP peerings for ECMP.

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 42 - Enabling Route redistribution at T0 Gateway

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

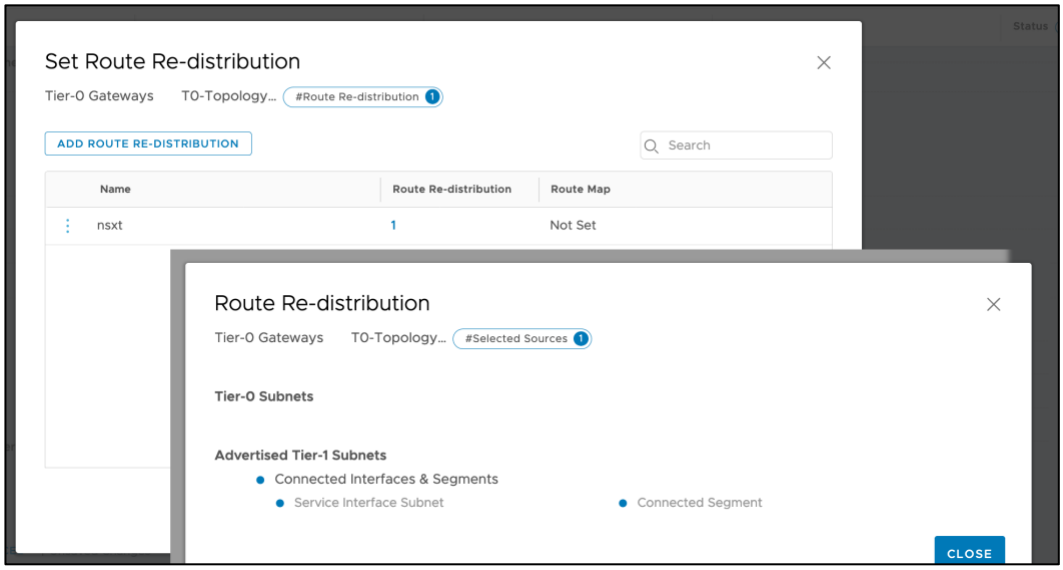


Figure 43 - Route redistribution settings at T0 Gateway

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.

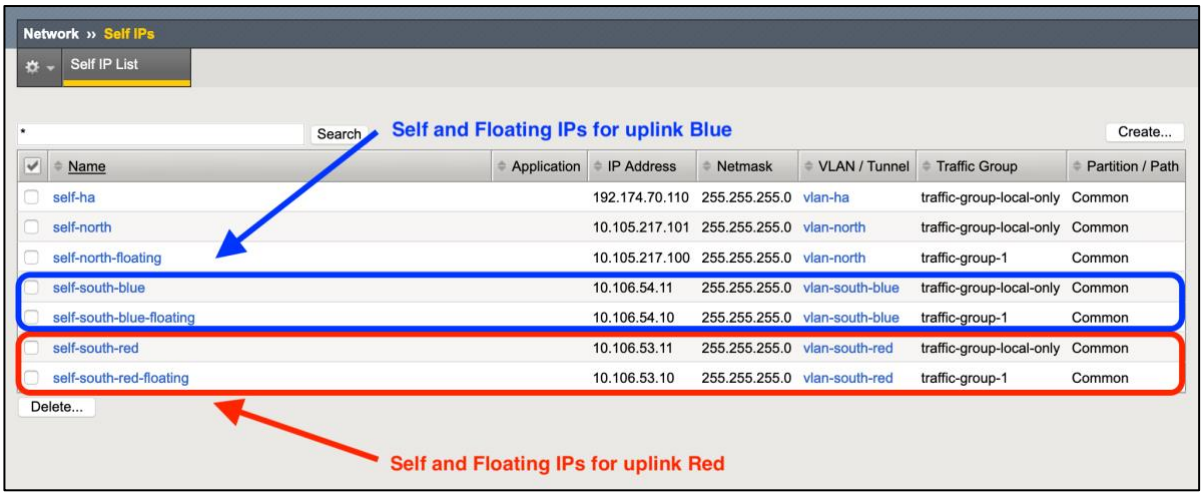


Figure 44 – Self IP in BIG-IP unit #1 for the NSX-T uplinks.

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

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
 match ip address prefix-list default-route
 set ip next-hop 10.106.53.10 primary
!
route-map default-route-uplink-blue permit 5
 match ip address prefix-list default-route
 set ip next-hop 10.106.54.10 primary
!
```

← route-map to set the next-hop to the floating-IP, one per load sharing path.

← route-map to set the next-hop to the floating-IP, one per load sharing path.

Figure 45 – 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.

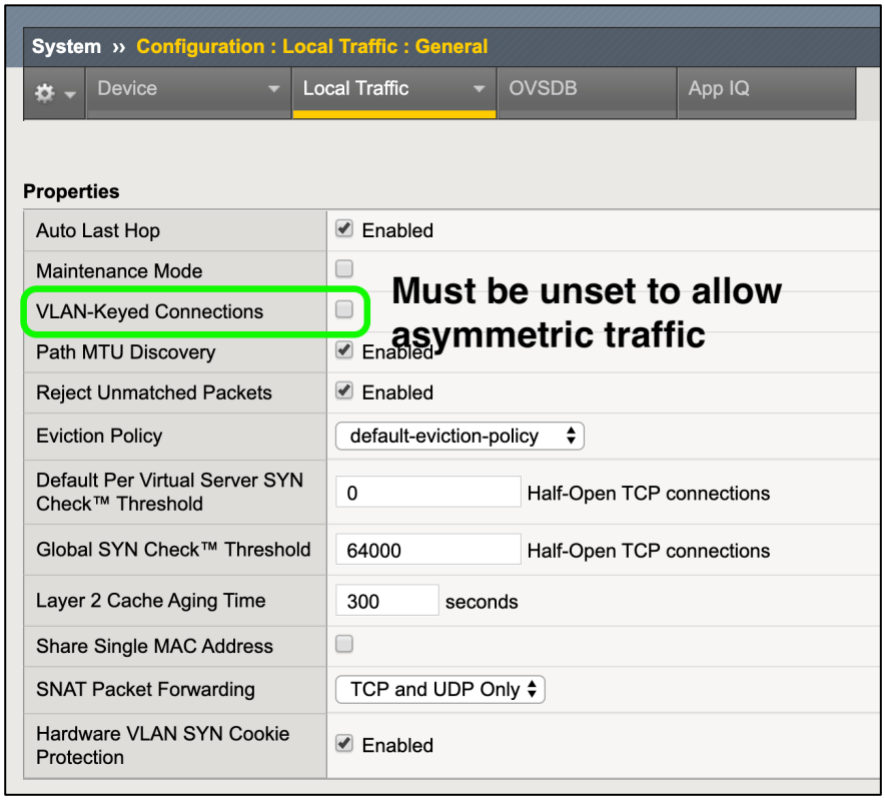


Figure 46 – Configuration required for ECMP which might generate asymmetric traffic.

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

Multi-tenant considerations for Topology A

Topology A supports using both vCMP hardware and software versions of F5 BIG-IP. Next, it is outlined the deployment using vCMP hardware and multiple tenants per VE. In the case of using Virtual Edition, using a per-tenant VE doesn't require any special configuration.

In an NSX environment, it is usual to isolate tenants using dedicated T0 Gateways. This model can be replicated using vCMP hardware or per-tenant VEs. It is up to the local security considerations if the External Networks or the NSX-T boundary networks (see next figure) can be shared between the tenants. F5 BIG-IP can accommodate both models.

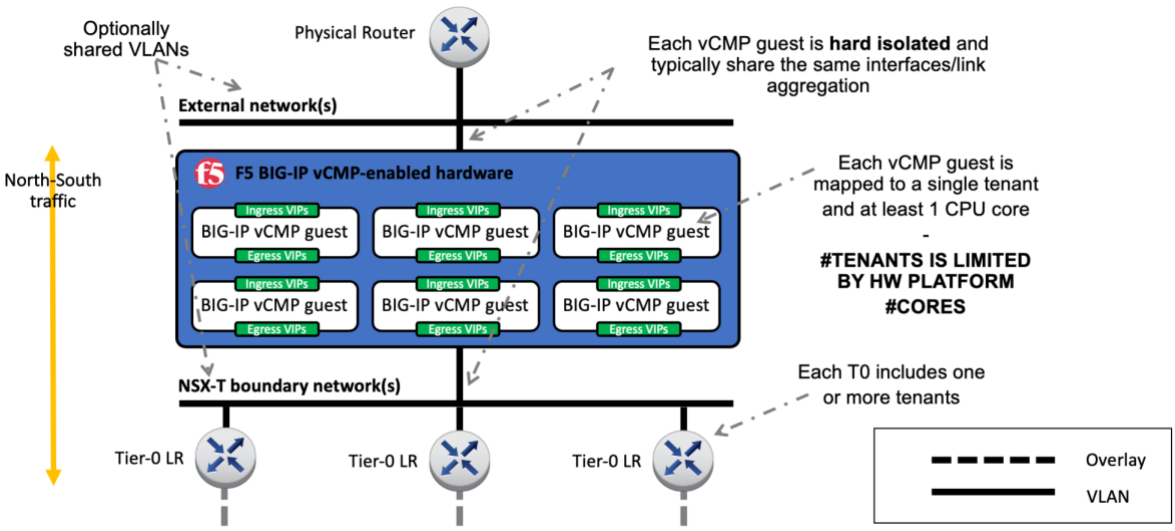


Figure 47 - Topology A with multiple tenants using vCMP hardware.

Although it is possible to use a shared VE model by using partitions and routing domains, it is discouraged. This is because it is expected these deployments handle high throughputs. In the case of test environments, it is discouraged as well because the configuration (partitions, route domains) will not match exactly the production deployment using per-tenant VE or using vCMP hardware.

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.

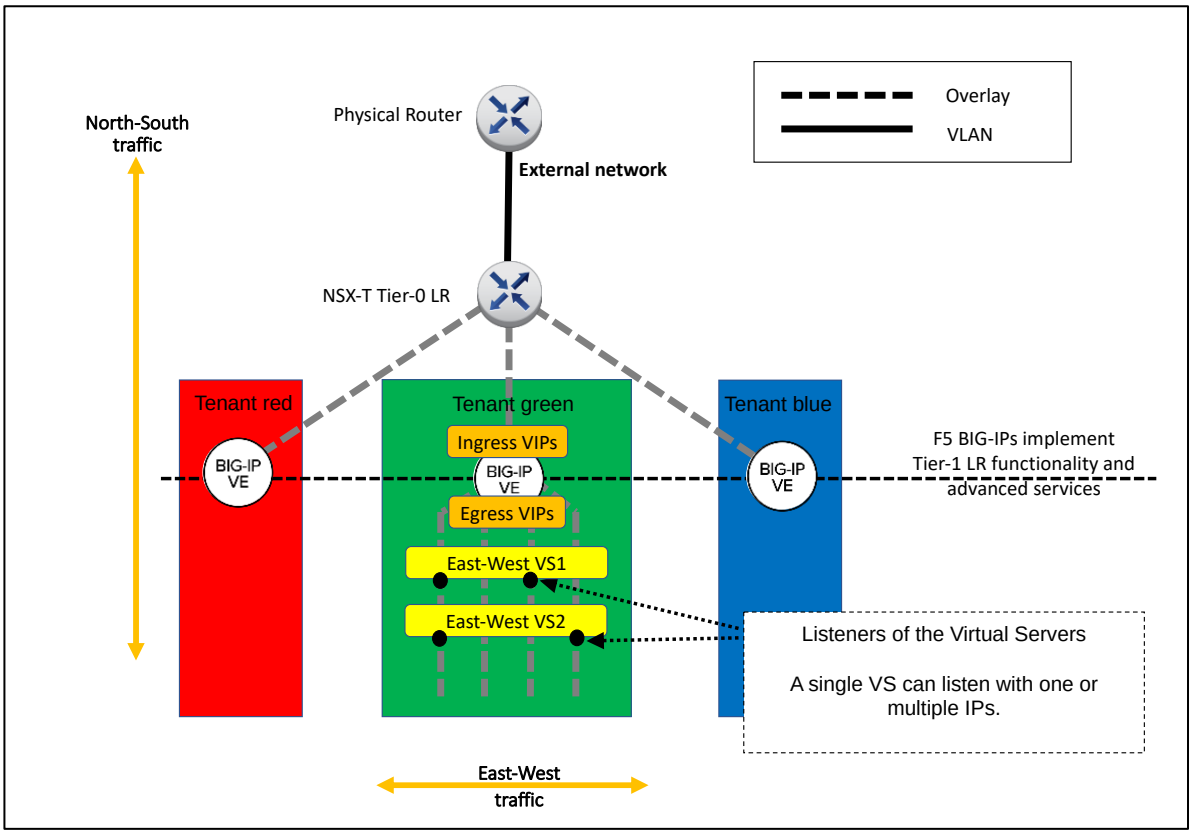


Figure 48 – Overview of BIG-IPs inline-connected like an NSX-T's Tier-1 Gateway.

The main characteristic of this topology is that NSX-T's Tier-1 Gateways are replaced by BIG-IPs. Eliminating the NSX-T's Tier-1 Gateway keeps a simpler 2-tier routing but it has several limitations which are overcome by Topology B extended described in the next section.

Topology B is recommended when simplicity is preferred over flexibility and when each tenant is going to have a dedicated BIG-IP cluster. The benefits of using BIG-IQ for centralized management and visibility are more relevant in this this topology. Additionally, when having several BIG-IP clusters, these can be distributed across any compute node unlike NSX-T's LBs which are limited to running in NSX-T Edge's nodes only[7].

[7] This applies to native NSX-T's LBs not based in AVI LB.

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

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

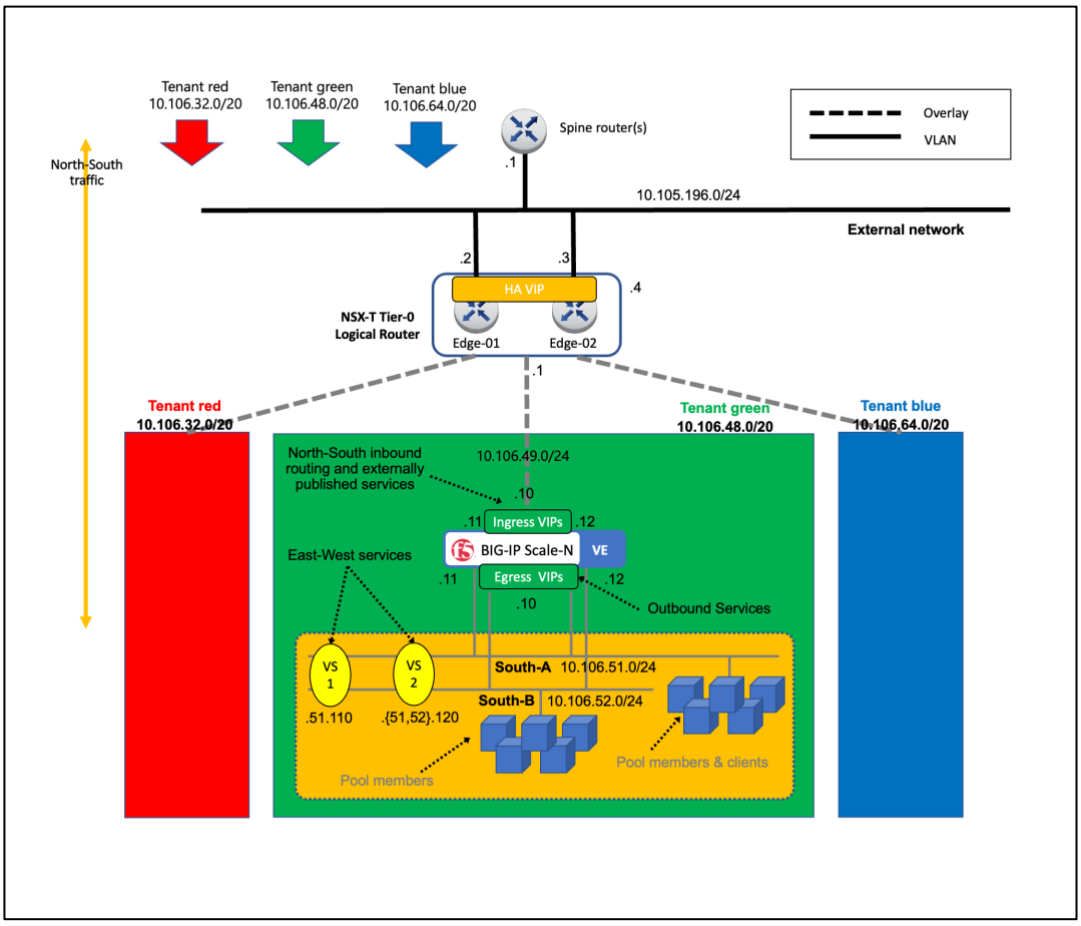


Figure 49 – Topology B example implementation.

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, this topology is only supported by BIG-IP VE. The configuration will be detailed next. As with all other topologies, this guide focuses on 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 HA mode.

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

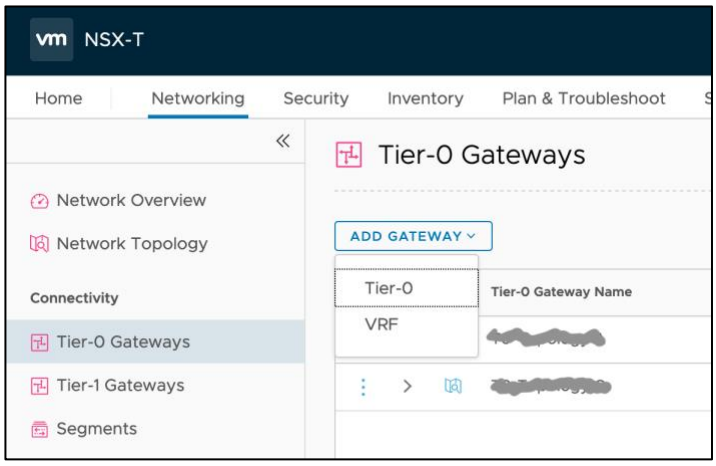


Figure 50 - Adding a Tier-0 Gateway.

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).

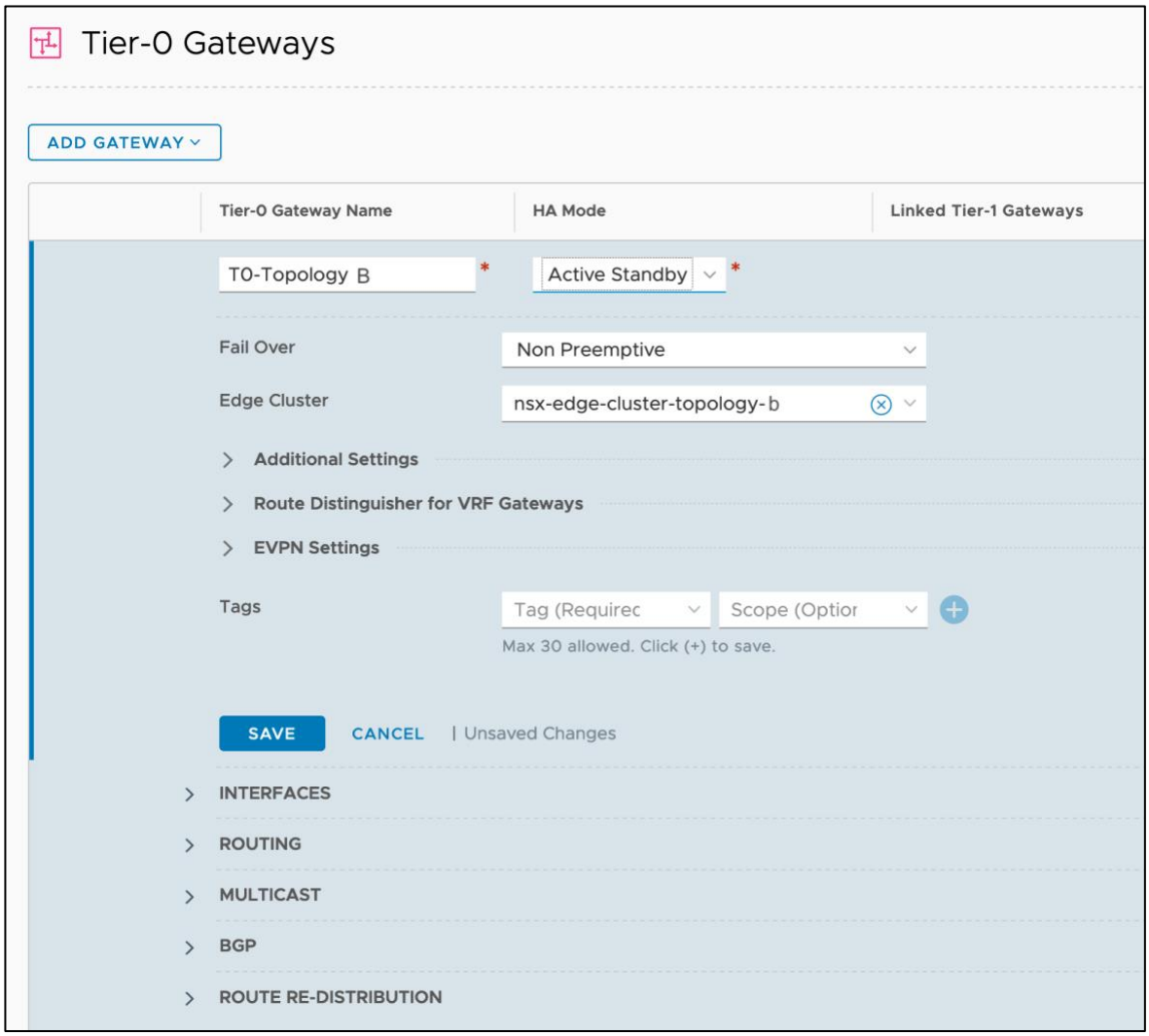


Figure 51 – Filling the details of a Tier-0 Gateway.

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

`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 pre-requisite 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.

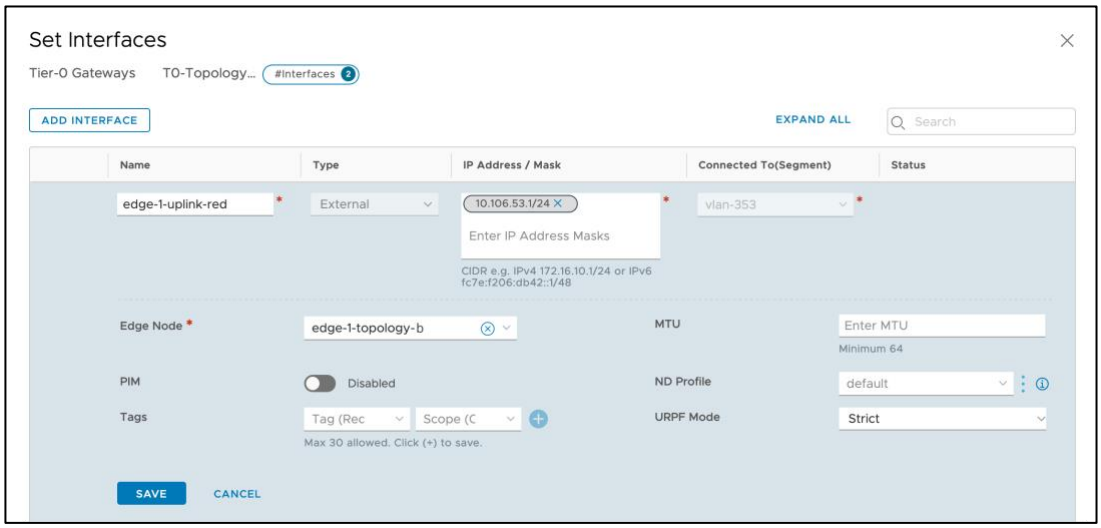


Figure 52 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.

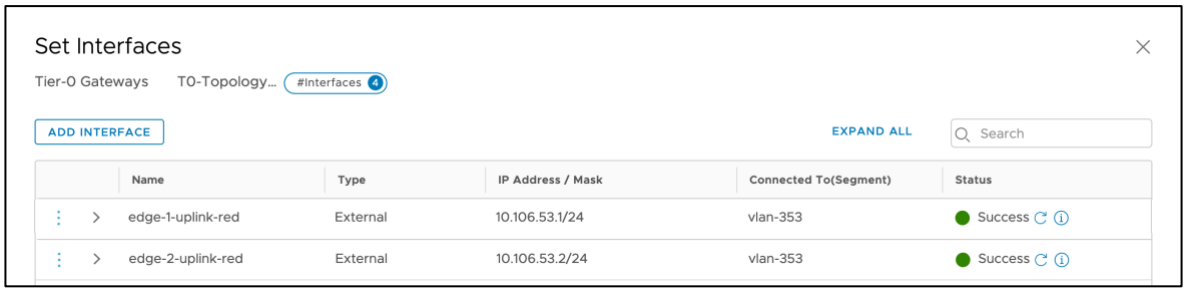


Figure 53 – Final Uplink interface configuration of the Tier-0 Gateway.

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:

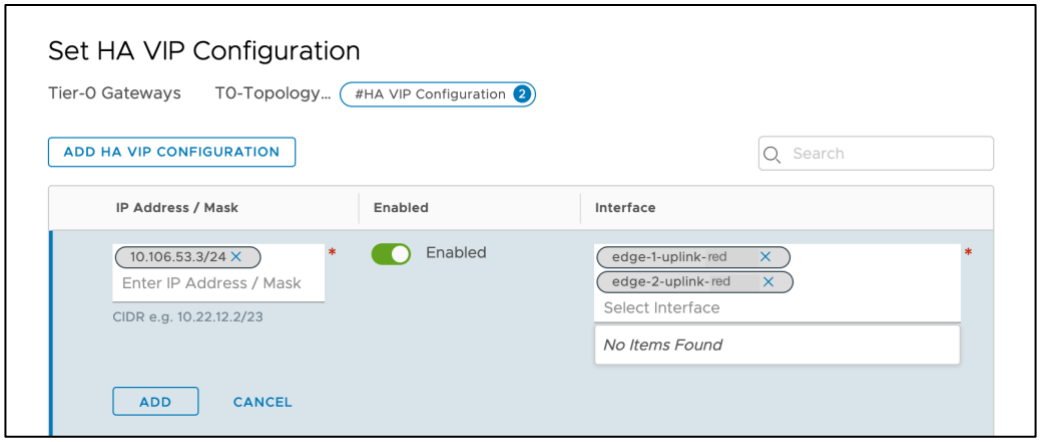


Figure 54 - Adding an HA VIP to NSX-T's T0 Gateway.

Selecting the two external interfaces just created.

Add a default route in the Tier-0 Gateway towards the upstream router.

In our example, the BIG-IP cluster floating address to use as the next hop is 10.105.196.1. 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) `10.105.196.1`.

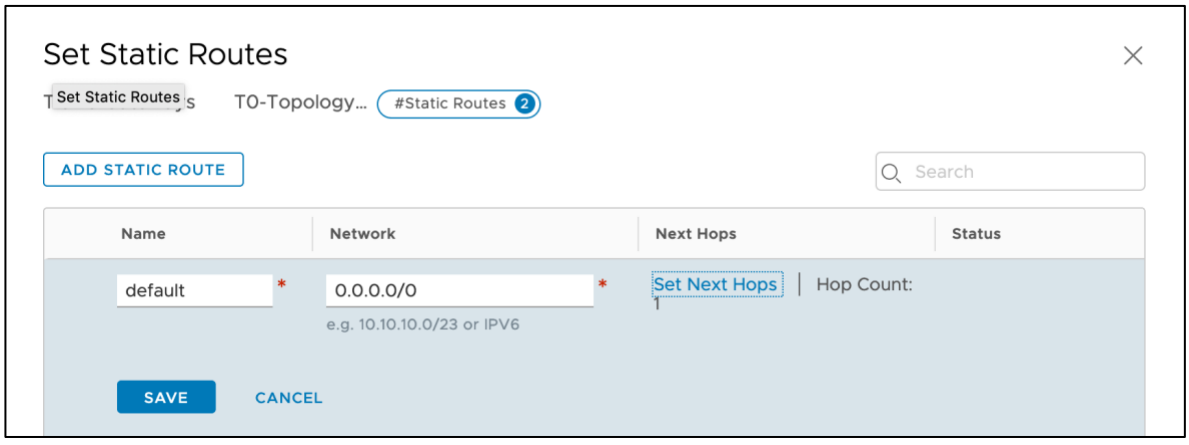


Figure 55 – Adding Tier-0 Gateway's default route.

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

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:

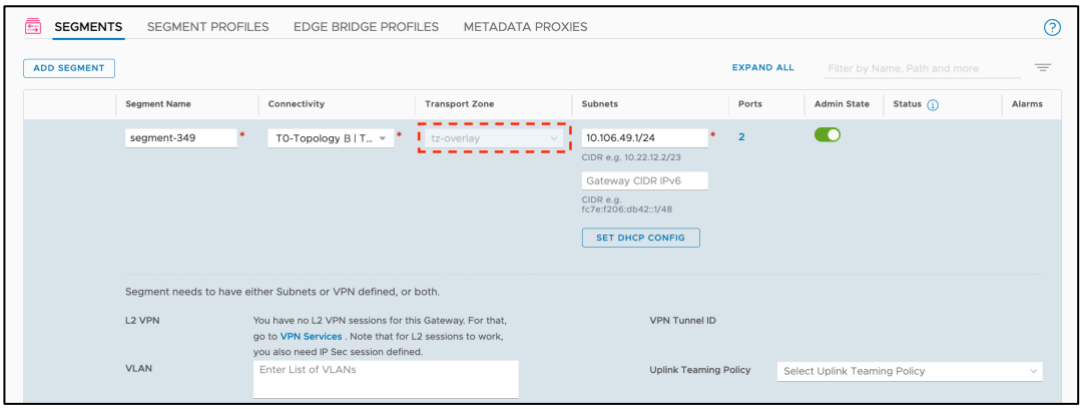


Figure 56 – Creating an overlay segment for the transit network between the Tier-0 Gateway and the BIG-IPs.

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.

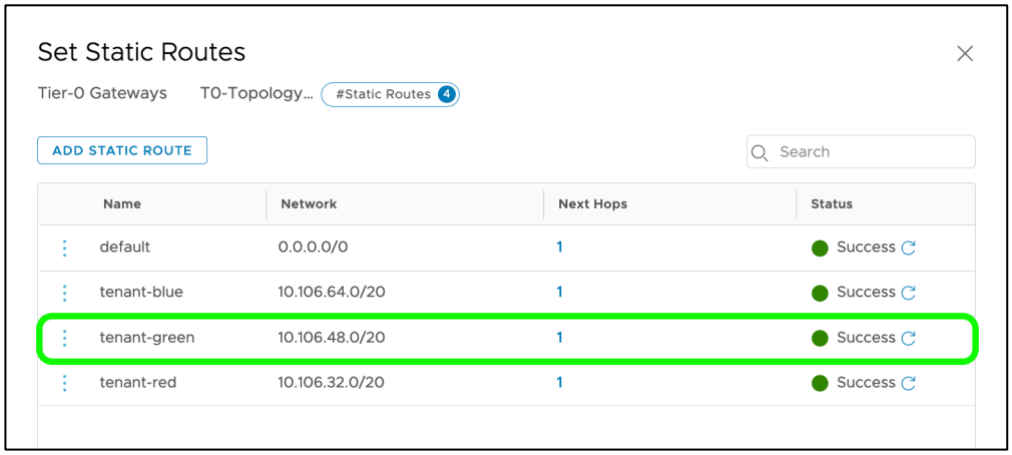


Figure 57 – Adding tenant's routing entries. Highlighted is the routing entry for tenant green for which BIG-IPs are configured in this section.

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

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

- segment 351/tenant network - 10.106.51.10/24
- segment 352/tenant network - 10.106.52.10/24

Where .10 are BIG-IPs' floating IPs in these networks.

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:

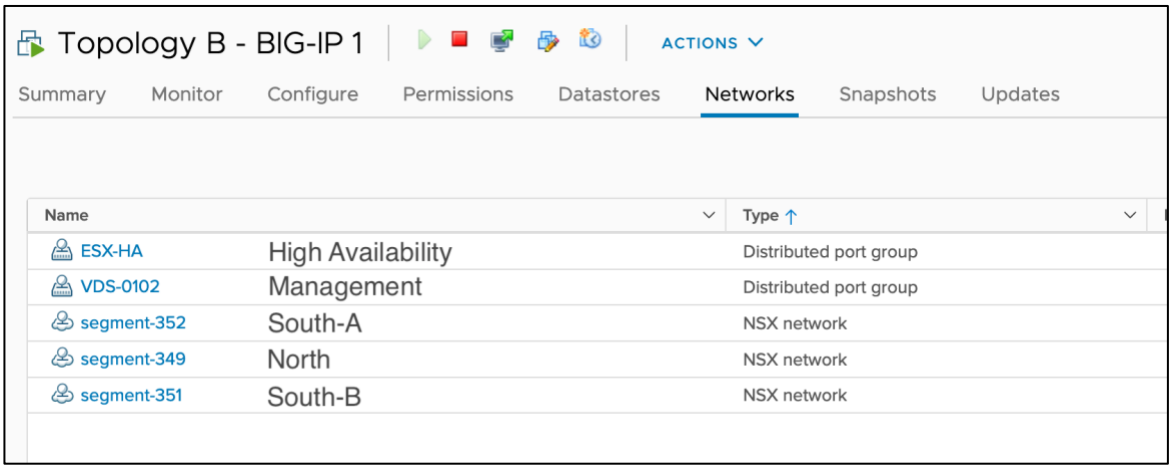


Figure 58 – Attaching the BIG-IP to NSX-T segments.

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:

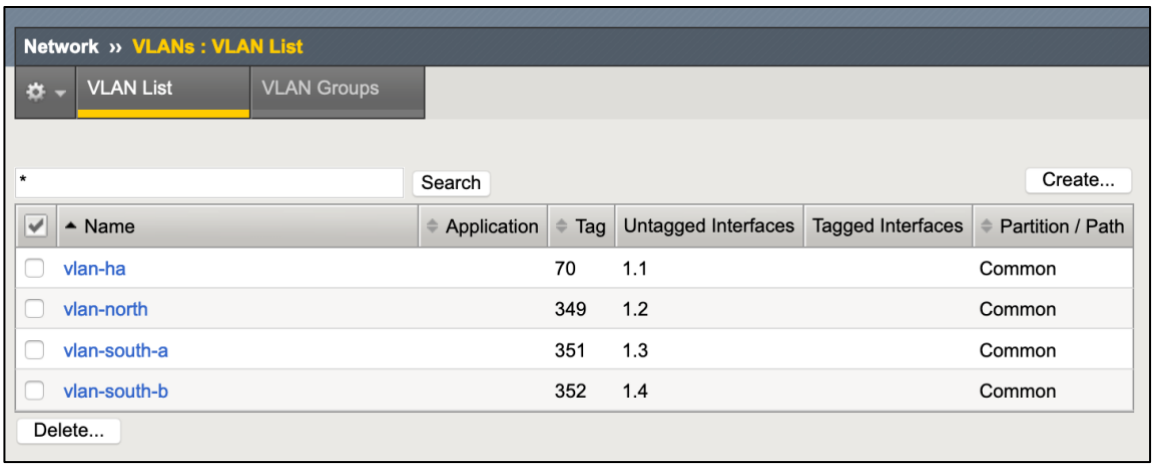


Figure 59 – Adding the NSX-T segment to the BIG-IP is just like adding a regular untagged VLAN.

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.

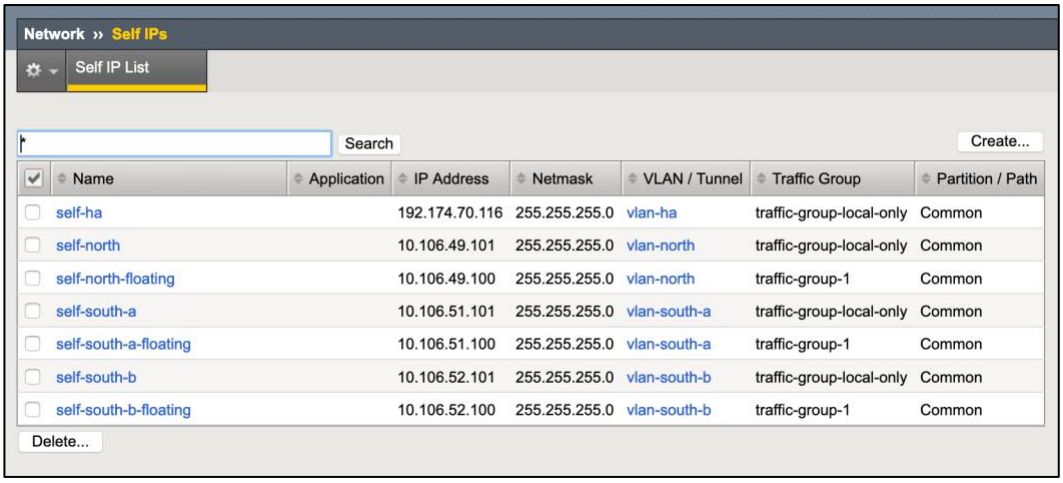


Figure 60 – Self IPs and floating Self IPs required (shown in BIG-IP unit 1).

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).

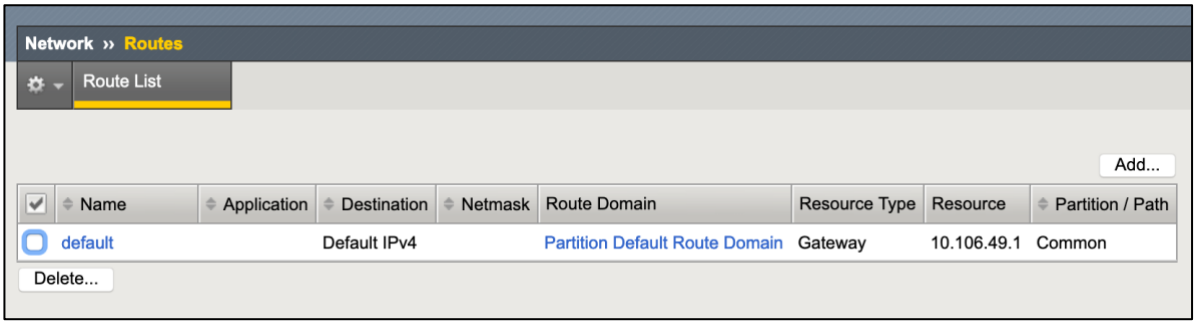


Figure 61 – Static route required in the BIG-IP units.

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 there are

shown two Virtual Servers:

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

These would be implemented as follows

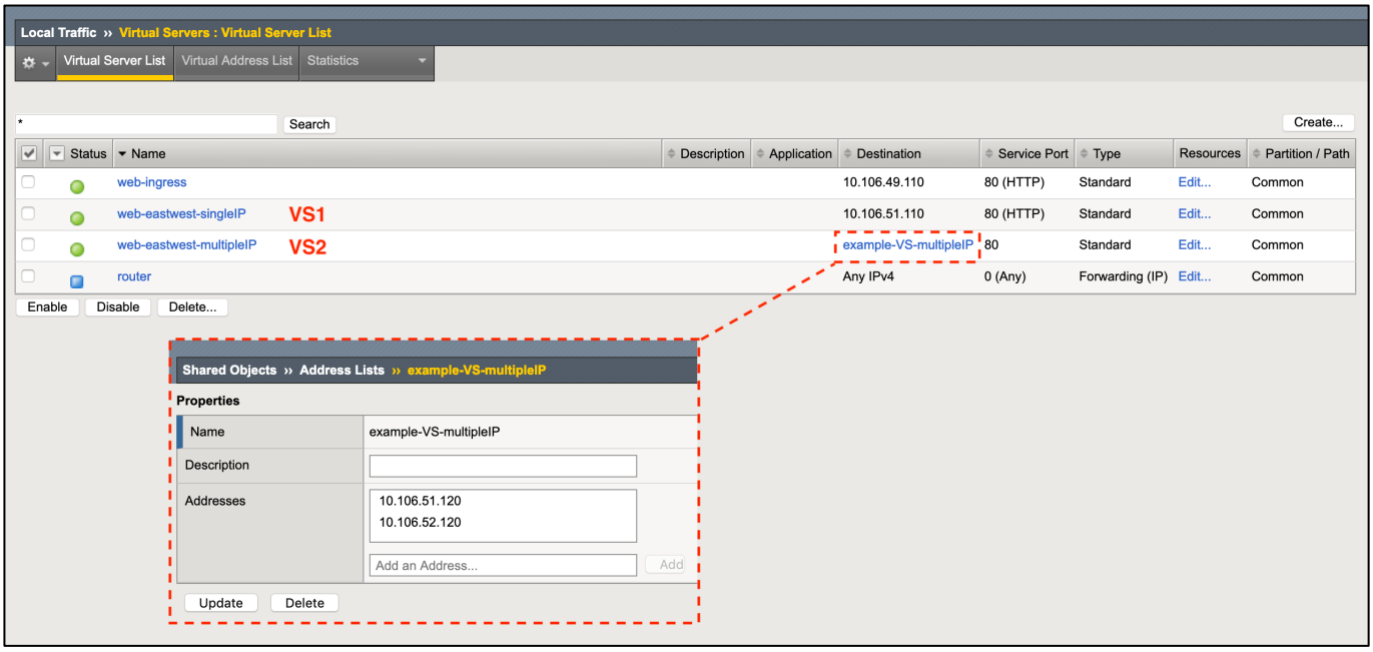


Figure 62 – Example of East-West Virtual Severs with multiple IP addresses.

It is important to differentiate the following Virtual Server Settings:

- The destination addresses 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.

Multi-tenant considerations for Topology B

It is recommended to use Topology B extended instead of plain Topology B because the former doesn´t limit the number of segments each tenant can have.

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

Note: Topology B extended has been modified from version 2.3 of this document by moving the location of the Tier-1 Gateway.

As the name indicates, this topology is an extension of Topology B, where an additional Tier-1 Gateway is placed south of the F5 BIG-IPs. Therefore, the workload VMs are no longer directly connected to the F5 BIG-IPs (in version 2.3 the BIG-IP was below the T1 Gateway). This topology has the following benefits compared to the plain Topology B:

- Tenants are not limited by the number of segments.
- The additional Tier-1 Gateway provides additional firewall layering.
- Workload VMs can make use of T1's DHCP.
- Additional services can be implemented in the T1 Gateway. Note that these services will likely be implemented in the Service Routers located in the Edge nodes.

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

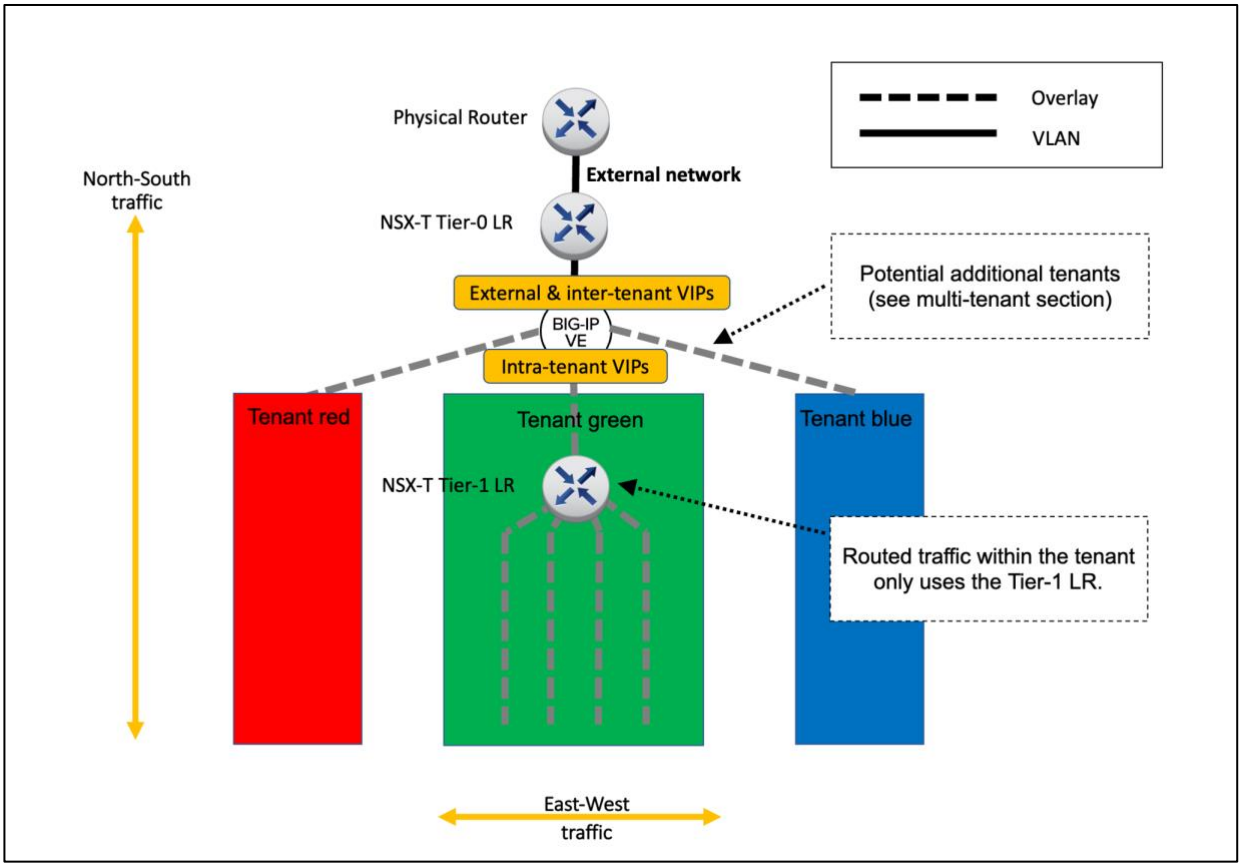


Figure 63 – Overview of BIG-IPs inline-connected to an NSX-T's Tier-1 Gateway.

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

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

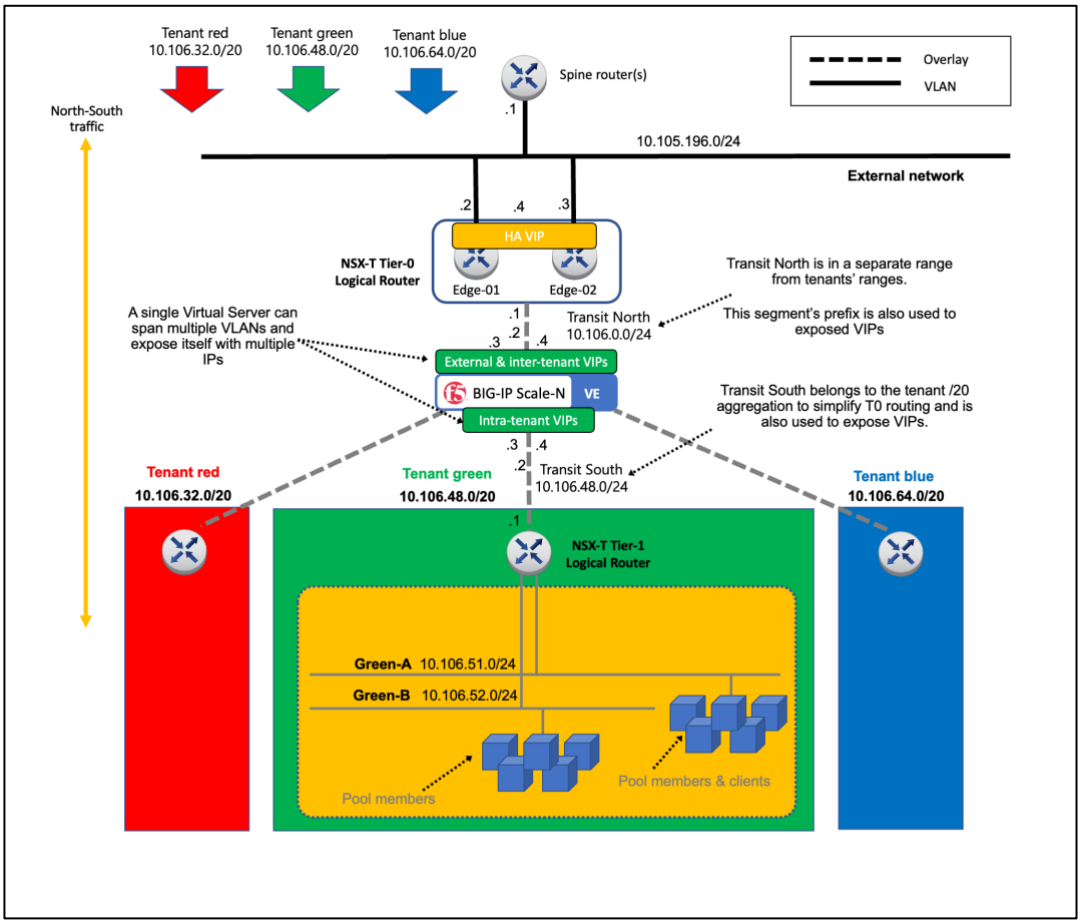


Figure 64 – Topology B extended, example implementation.

Likewise in the case of Topology B, to simplify the routing in the Tier-0, contiguous networks are used for each tenant. In this example, /20 prefixes are used.

The segment Transit North, between the Tier-0 Gateway and the BIG-IP uses a /24. Using a /24 prefix is larger than strictly necessary for an HA-pair (only 3 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. This also applies to the segment Transit South, between the BIG-IP and the Tier-1 Gateway. This latter segment belongs to the tenant's /20 prefix.

It is recommended to use addresses from the Transit North prefix for exposing services to multiple tenants. This allows service mobility regardless which tenant is implementing the service. On the other hand, addresses for intra-tenant services should be taken from the Transit South prefix. Please note that if required, a single Virtual Server can listen to multiple addresses and VLANs and could discriminate between them using LTM policies or iRules.

From the figure above, this topology is only supported by BIG-IP VE. The configuration will be detailed next. As with all other topologies, this guide focuses on 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 HA mode.

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

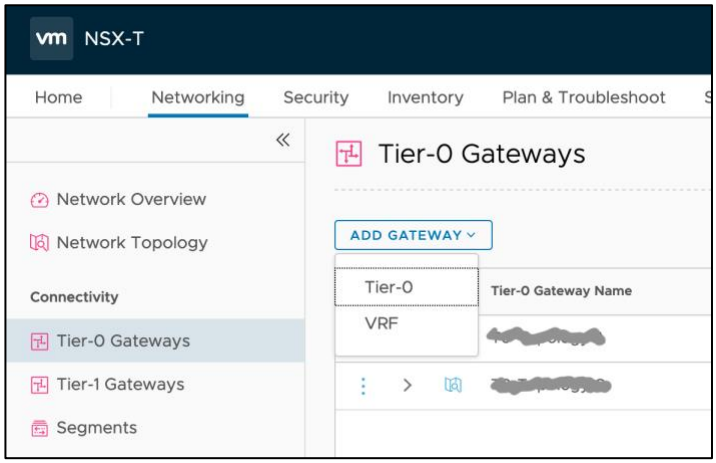


Figure 65 - Adding a Tier-0 Gateway.

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

- Name: `T0-topology B-extended` 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).

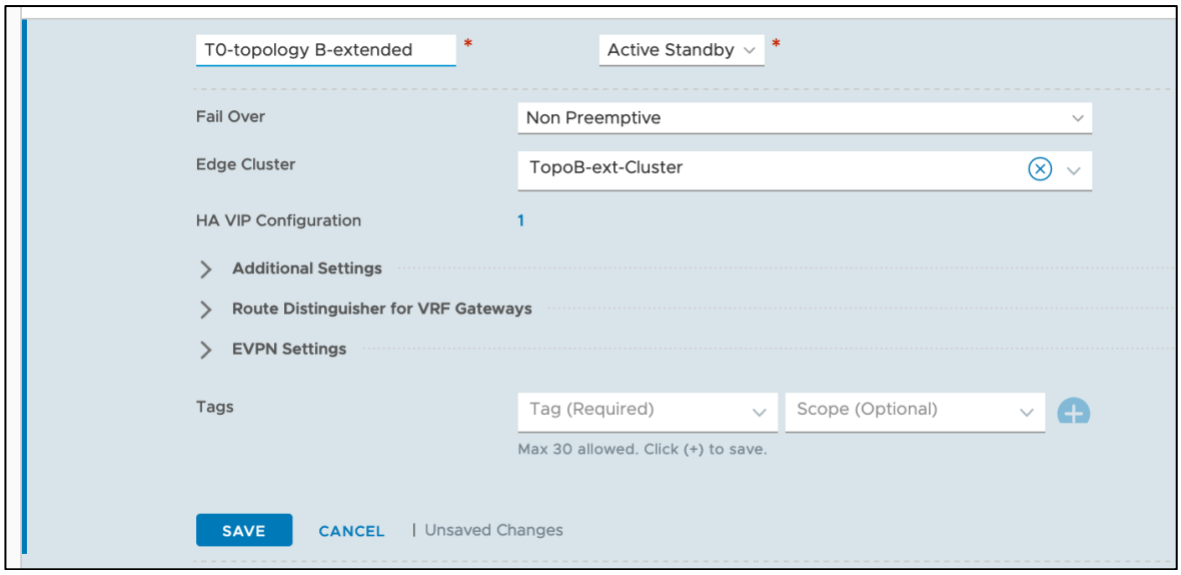


Figure 66 – Filling the details of a Tier-0 Gateway.

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

Select the router created (`T0-topology B-extended` 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 `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 pre-requisite 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.

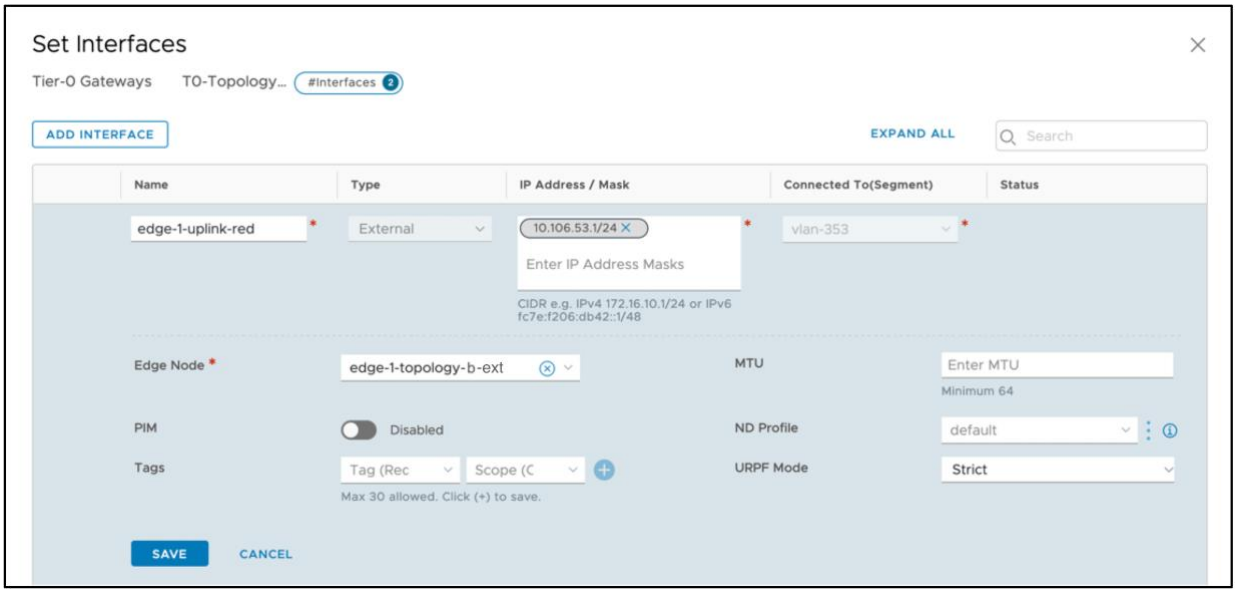


Figure 67 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.

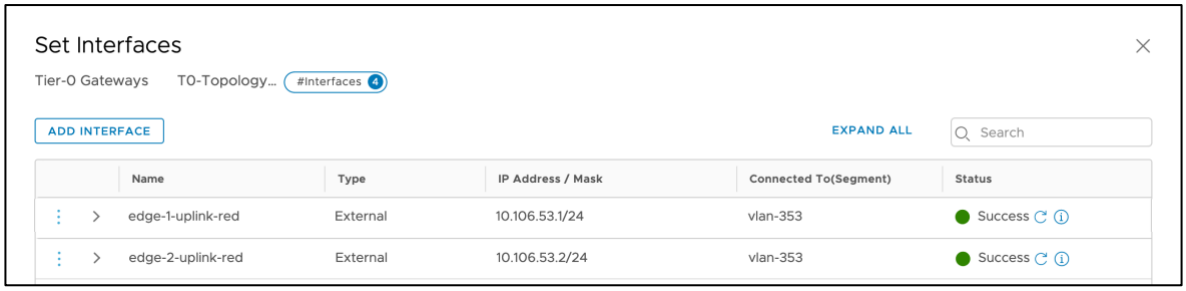


Figure 68 – Final Uplink interface configuration of the Tier-0 Gateway.

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:

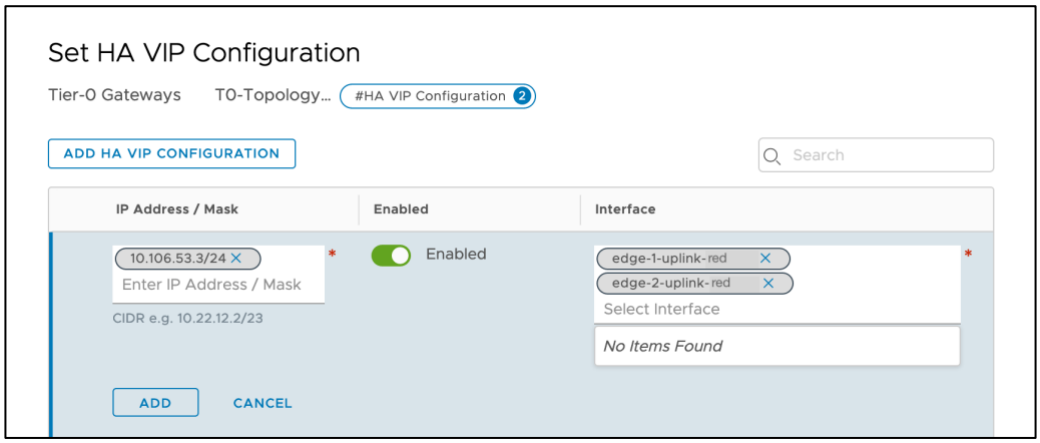


Figure 69 - Adding an HA VIP to NSX-T's T0 Gateway.

Selecting the two external interfaces just created.

Add a default route in the Tier-0 Gateway towards the upstream router.

In our example, the BIG-IP cluster floating address to use as the next hop is 10.105.196.1. 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 `10.105.196.1` (not shown in the figure).

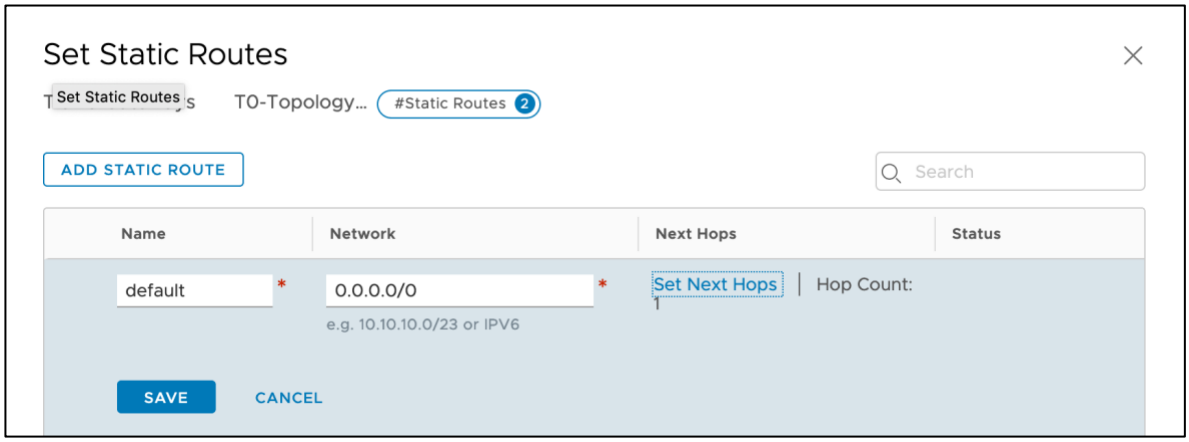


Figure 70 – Adding Tier-0 Gateway's default route.

Likewise the plain Topology B, in this case please note that we will not create any static routes in the Tier-0 Gateway. These will be created in the Tier-1 Gateway instead.

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

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:

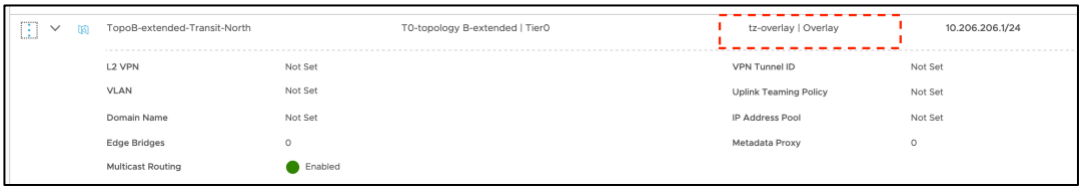


Figure 71 – Creating the overlay segment for the transit network between the Tier-0 Gateway and the BIG-IPs.

Where .1 is the Tier-0 Gateway and .2 will be the floating IP address of the BIG-IP.

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

By using a contiguous prefix per tenant, one only needs to add a single route to the existing routing table. The next hop is the floating IP in the BIG-IP cluster. Ultimately the routing table will look like as shown next.

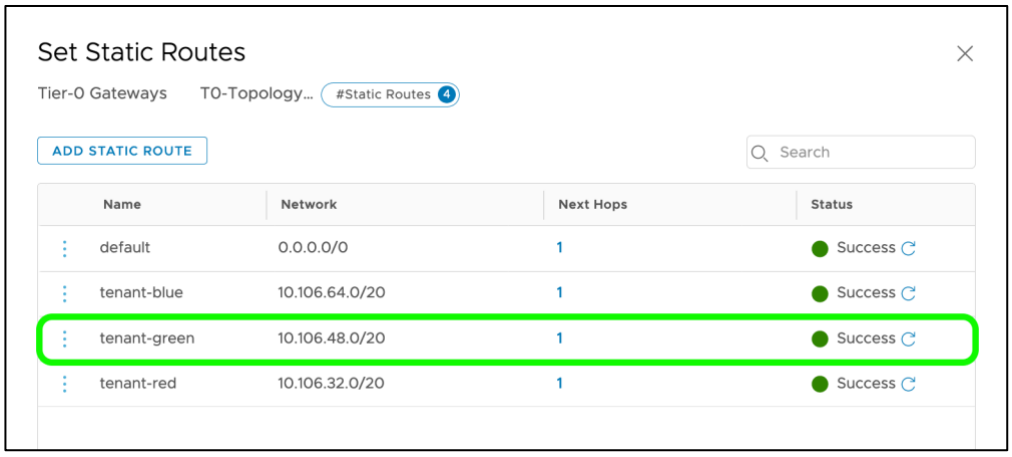


Figure 72 – Adding tenant's routing entries. Highlighted is the routing entry for tenant green for which BIG-IPs are configured in this section.

3. Create tenant's Tier-1 Gateway and its networks

Go to `Networking > Tier-1 Gateway > ADD TIER-1 GATEWAY` and create a Gateway attached to the `T0-topology B-extended` Tier-0 Gateway as follows:

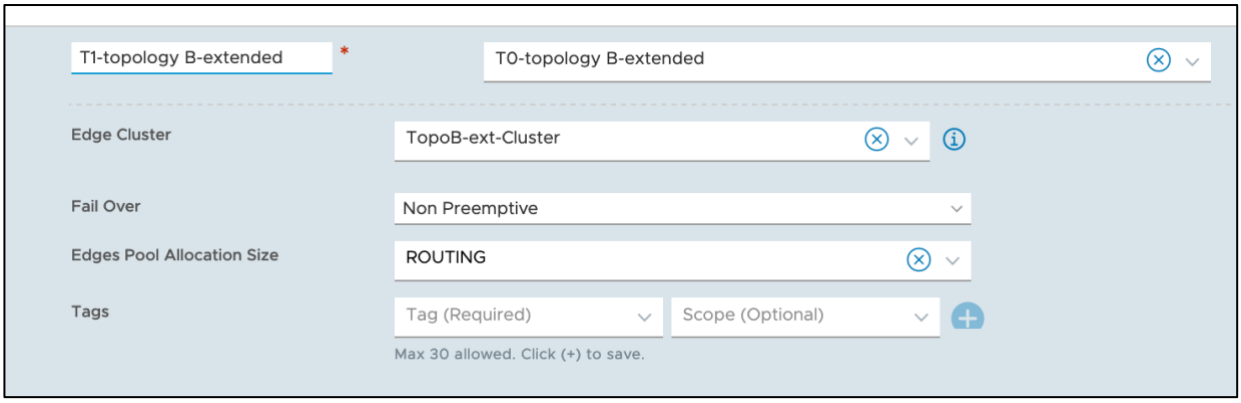


Figure 73 - Creating the Tier-1 Gateway.

This Tier-1 Gateway has a default route pointing to the F5 BIG-IPs, all other networks are directly attached.

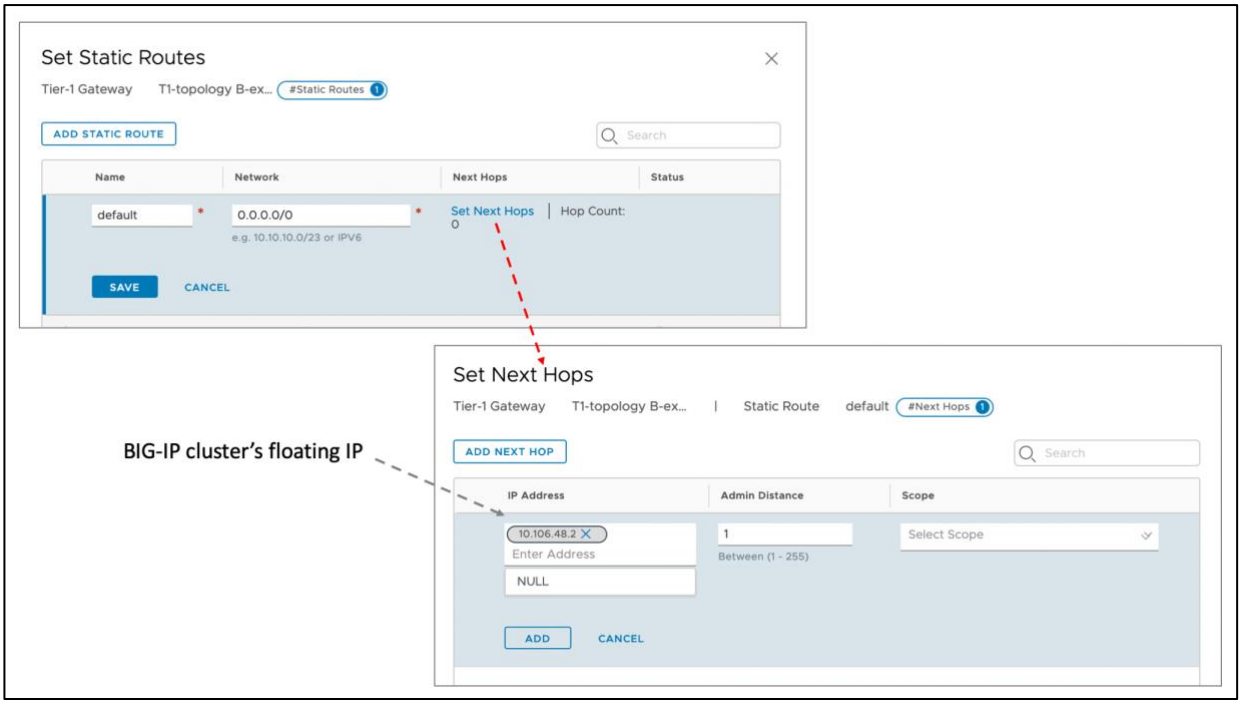


Figure 74 - Adding tenant's default route.

3.1. Create the transit segment between the BIG-IP and the Tier-1 Gateway

Go to `Networking > Segments > ADD SEGMENT` and create a Logical Switch within the overlay Transport Zone and attaching it to the just created Tier-1 Gateway as follows:

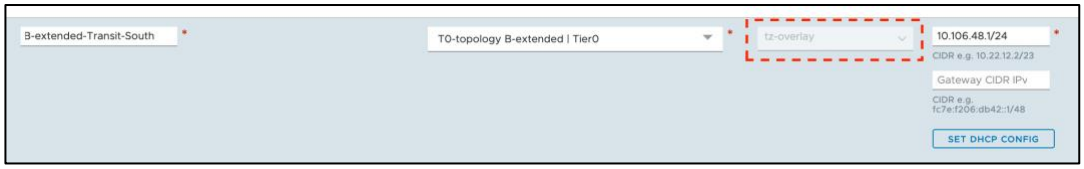


Figure 75 – Creating the overlay segment for the transit network between the BIG-IPs and the Tier-1 Gateway.

Where .1 is the Tier-1 Gateway and .2 is the floating IP address of the BIG-IP.

3.2. Create tenant's workload segments.

Use the same steps as for the previous transit segment and create as many workload segments as needed where .1 is always the Tier-1 Gateway.

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

After creating the segments in the NSX manager, the BIG-IP VE can be attached to these normally. In this case, besides the Management and HA segments, the BIG-IP will be attached to two transit segments:

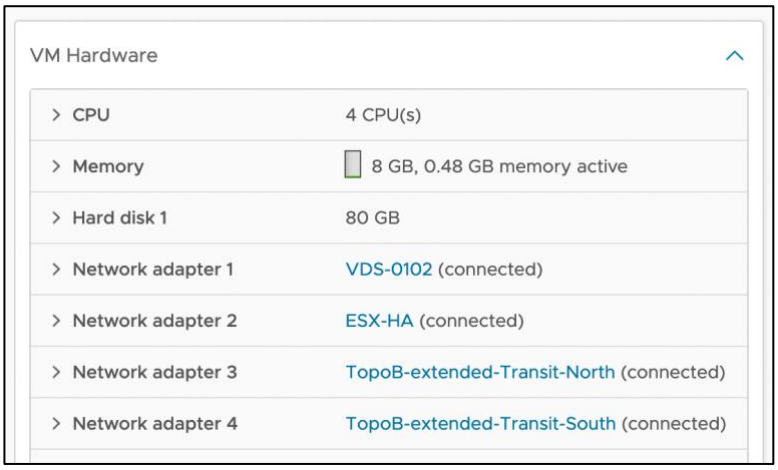


Figure 76 – Attaching the BIG-IP to NSX-T segments.

Although we are using NSX and regular/non-NSX segment types these are used likewise in the BIG-IPs and are configured just like any regular untagged VLAN as shown in the next figure:

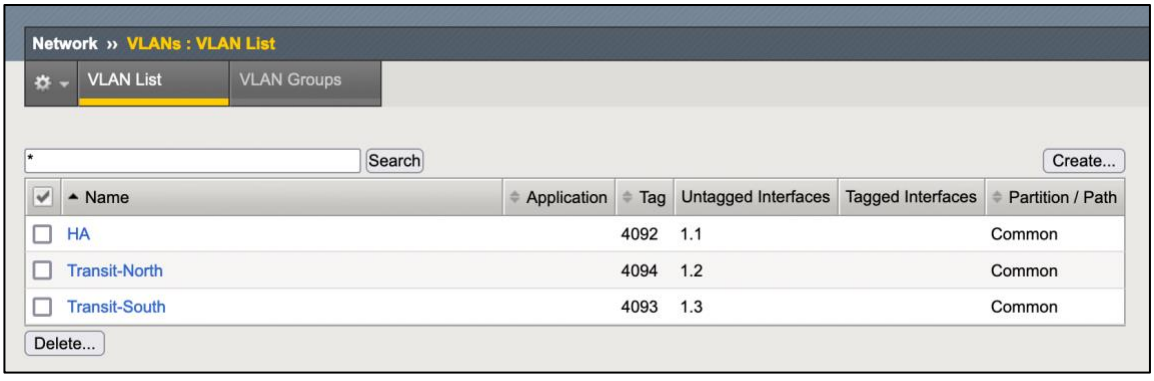


Figure 77 – Adding the NSX-T segment to the BIG-IP is just like a regular untagged VLAN.

Where the VLAN tags are not relevant, yet they should match across all BIG-IP units in the cluster.

Next, create the Self IPs and floating Self IPs towards the Tier-0 Gateway (north-bound) and for the Tier-1 Gateway (south-bound). None of these require any special configuration. An example of the first BIG-IP unit is shown next.

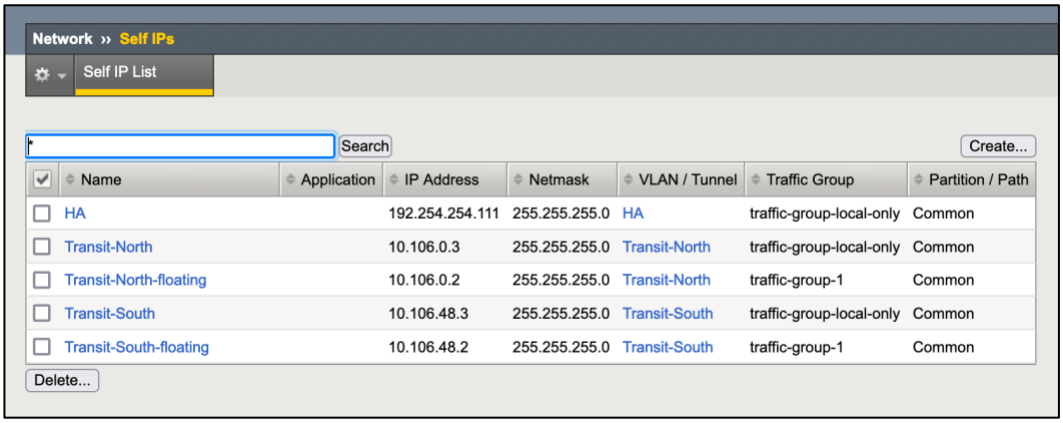


Figure 78 – Self IPs and floating Self IPs required (shown in BIG-IP unit 1).

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. Two static routes are needed:

- A default route via the Tier-0 Gateway.

- A route towards the tenant prefix via the Tier-1 Gateway. This is the /20 prefixes in the example. Note that the connecting network between the BIG-IP and this Tier-1 Gateway (Transit-South) is already within this tenant /20 prefix.

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

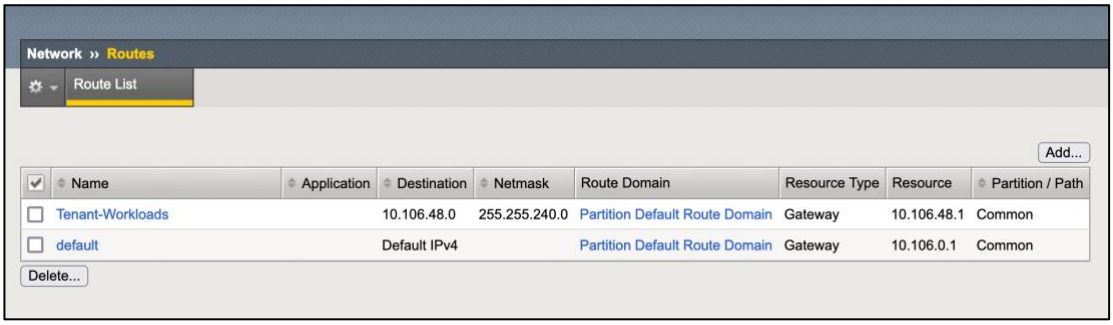


Figure 79 – Static route required in the BIG-IP units.

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

Multi-tenant considerations for Topology B extended

Topology B extended only supports VE because it makes use of NSX-T overlays. The general recommendation is to use per-tenant VE when hard isolation is required and when higher performance scalability might be required. Per-tenant VE doesn't require any special configuration.

On the other hand, a shared VE model (multiple tenants in the same VE) can be interesting to reduce resource utilization or expenditures depending on the licensing used. The number of tenants is limited to 7 per VE cluster because the ESXi hypervisor limits the number of vNICs to 10 per VM (the remaining vNICs are used for management, HA, and the upstream link).

If desired to perform some isolation between the tenants, the mechanisms firewall rules, Virtual Server's source address prefixes, and VLAN listener filters are considered the most appropriate instead of using route domains which are less granular and apply to a wider scope.

It is possible to have either shared or dedicated virtual servers for the tenants. As a good practice, each tenant should be in its own partition. When the same configuration is required for several tenants, it is possible to span a Virtual Server across multiple tenants using the VLAN selector. In this latter case the configuration would reside in a shared partition.

On the other hand, It should be considered as well the option of having duplicate configurations for each tenant (instead of having shared VS). This could be handled with automation and each tenant could have exclusively all its configuration in its own partition. This latter approach provides fine-grained configuration and stats.

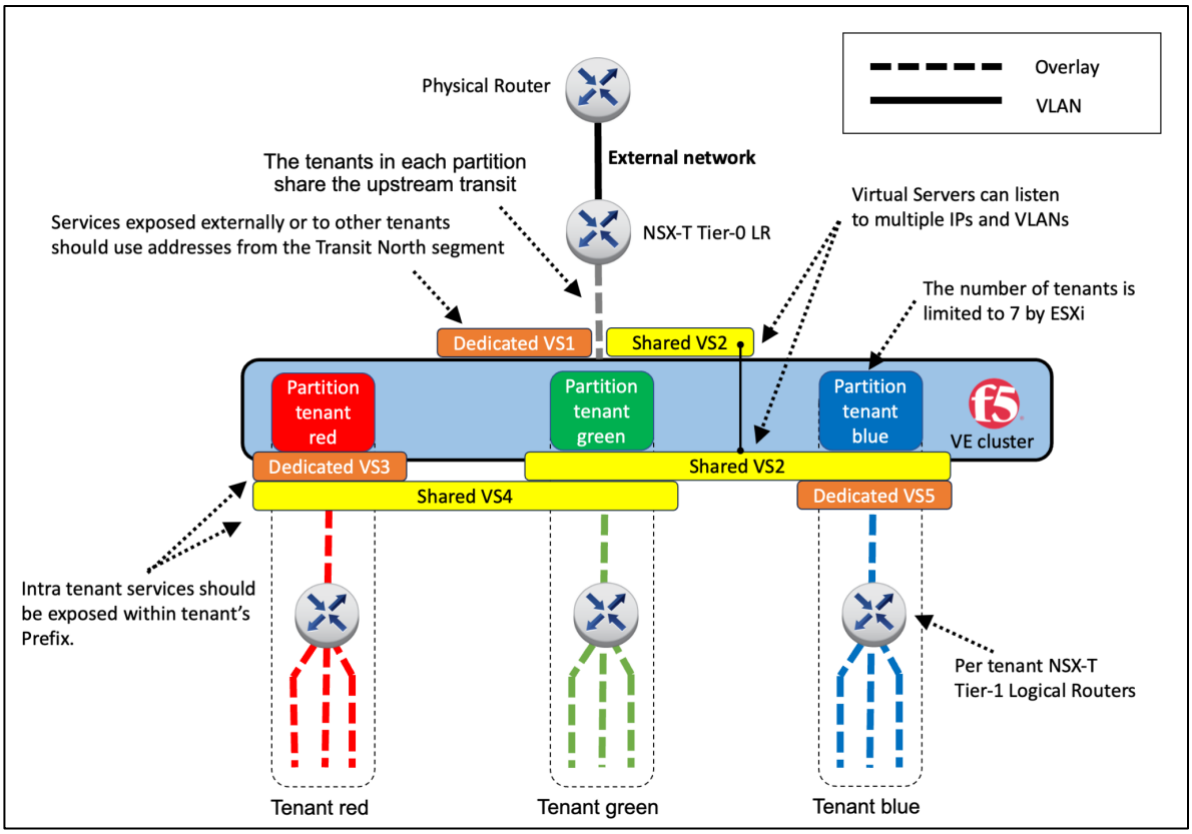


Figure 80 - Topology B extended with multiple tenants.

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

In the next figure, an overview of this topology with its traffic flows is shown.

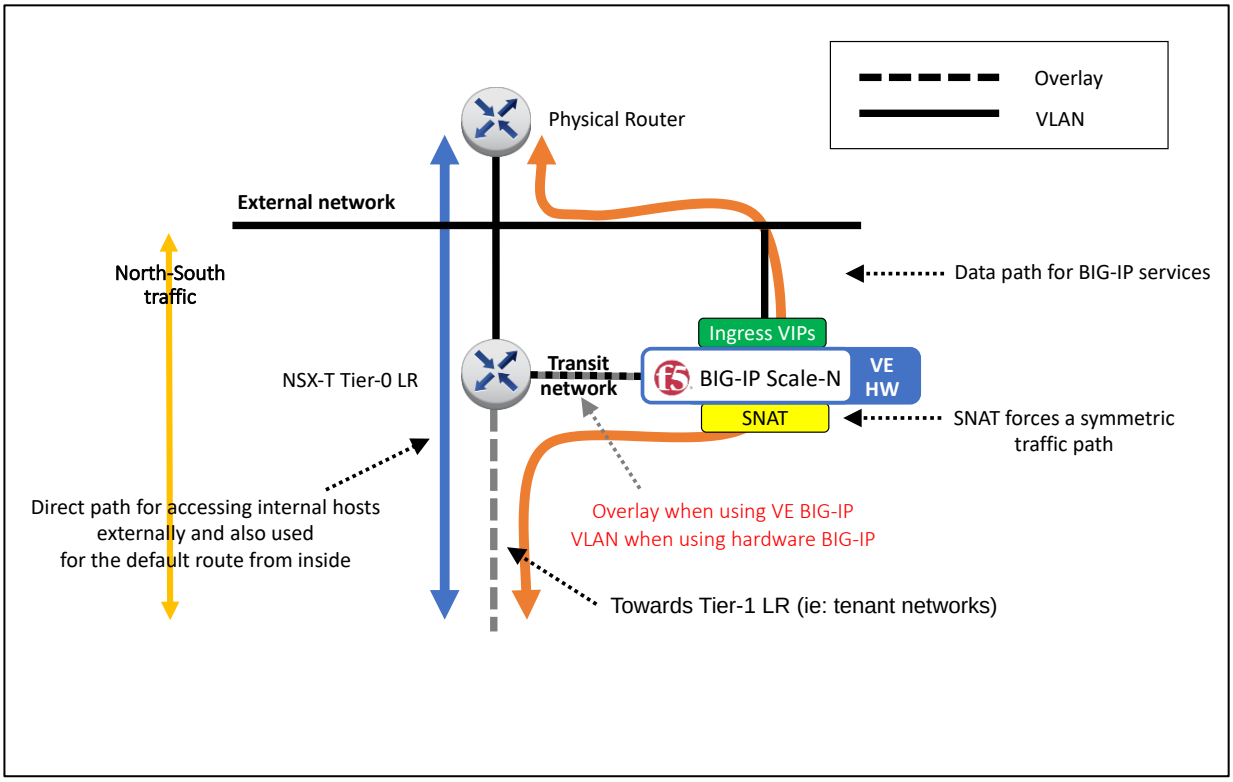


Figure 81 – Topology C overview.

Traffic-path wise, the main characteristic of this topology is that it allows direct access to the workloads without going through the BIG-IPs (BIG-IP bypass). Performance reasons should not drive the selection of this topology: the logical additional hop that the F5 BIG-IP represents incurs very little added latency with no throughput reduction. Moreover, when using F5 BIG-IP hardware the added latency is negligible compared to the latency impact that virtualization infrastructures imply.

In the previous figure, depending on the choice of a hardware or virtualized BIG-IP, the NSX-T boundary will differ. When using a hardware BIG-IP, the connectivity between the Tier-0 and the BIG-IPs will be done with an NSX-T Edge uplink. When using a virtualized BIG-IP, this connectivity can be done with a regular segment.

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.

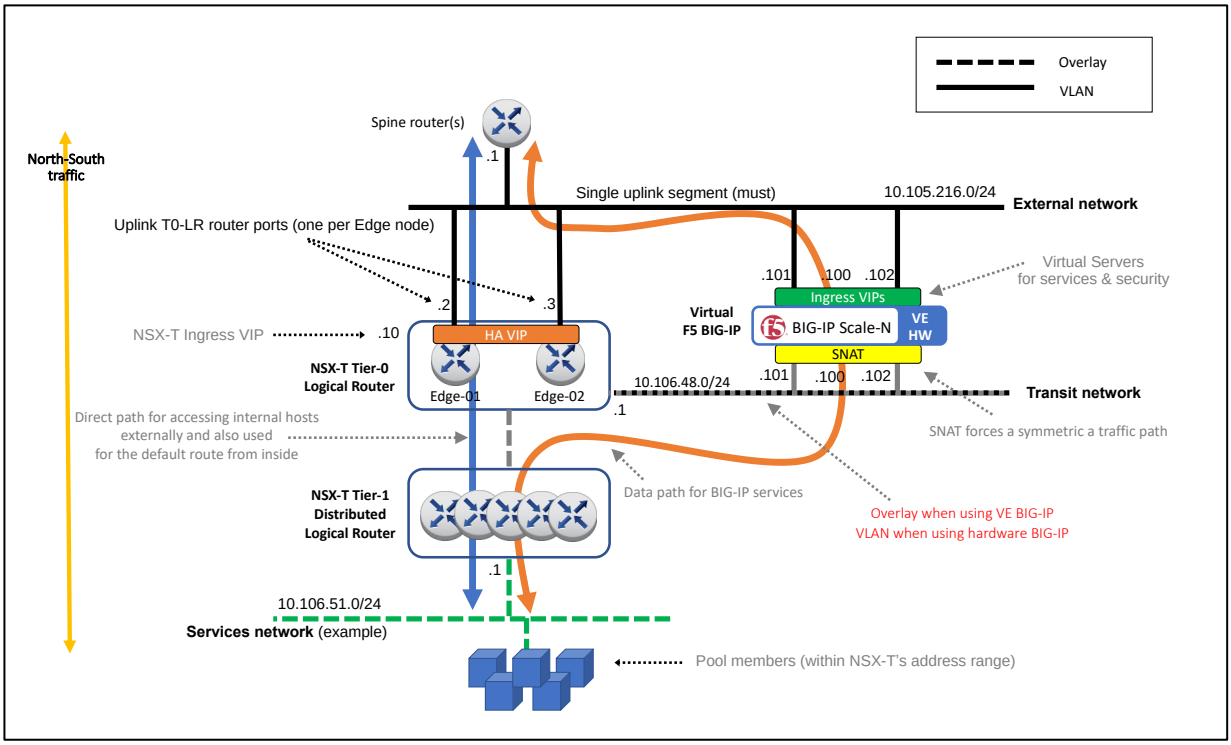


Figure 82 – Topology C example implementation.

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 on 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 HA mode.

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

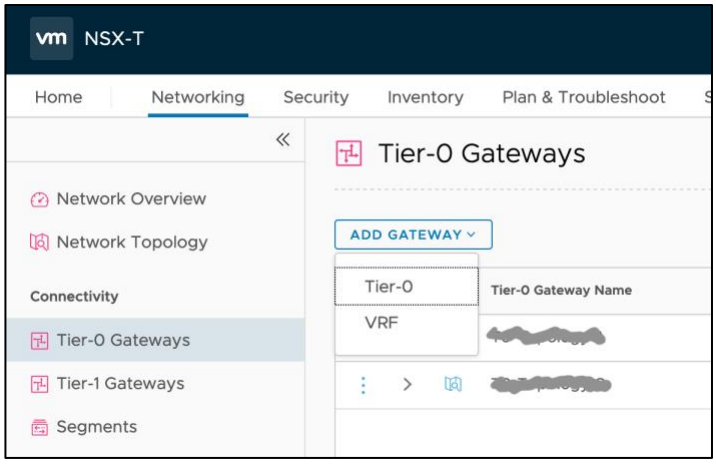


Figure 83 – Adding a Tier-0 Gateway.

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).

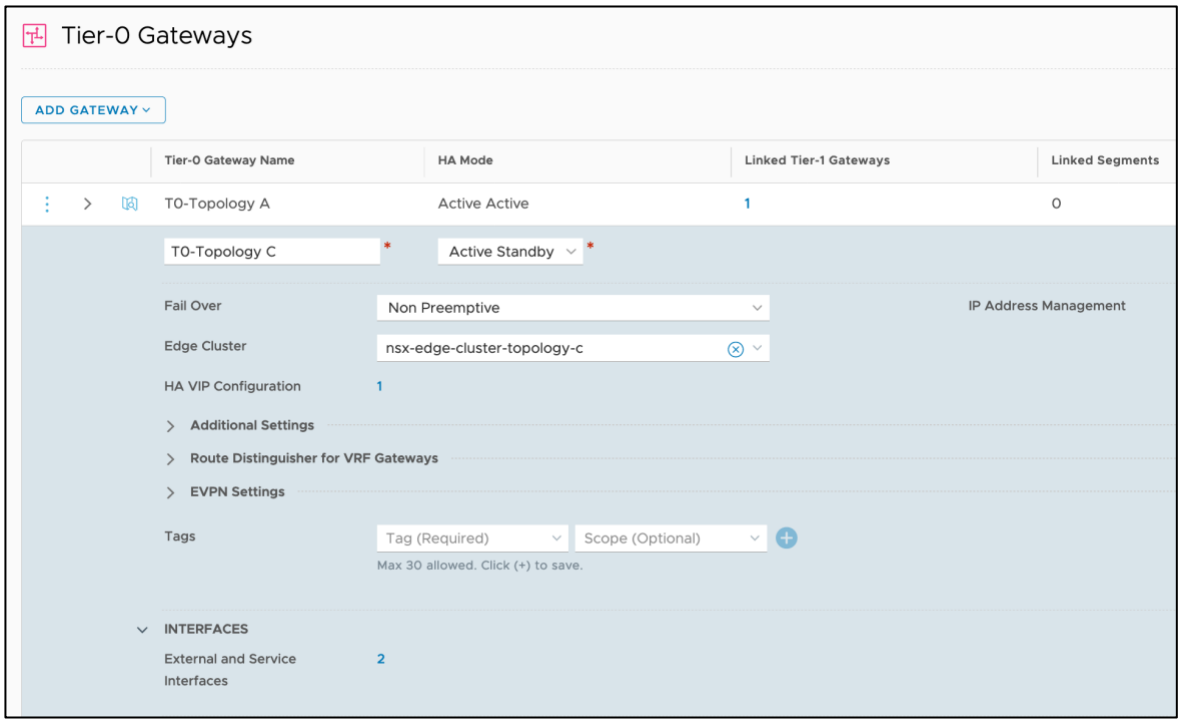


Figure 84 – Filling the details of a Tier-0 Gateway.

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

`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 pre-requisite 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.

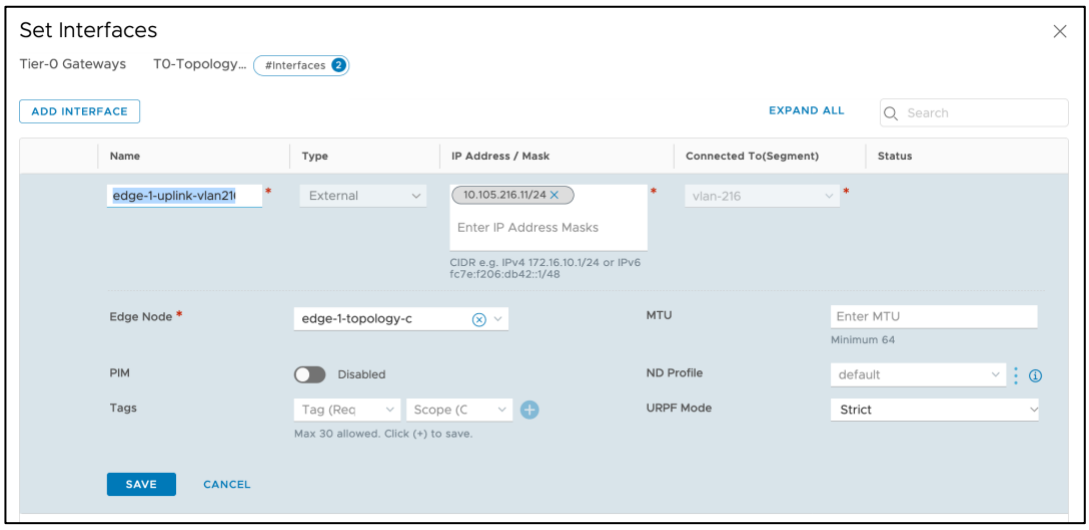


Figure 85 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.

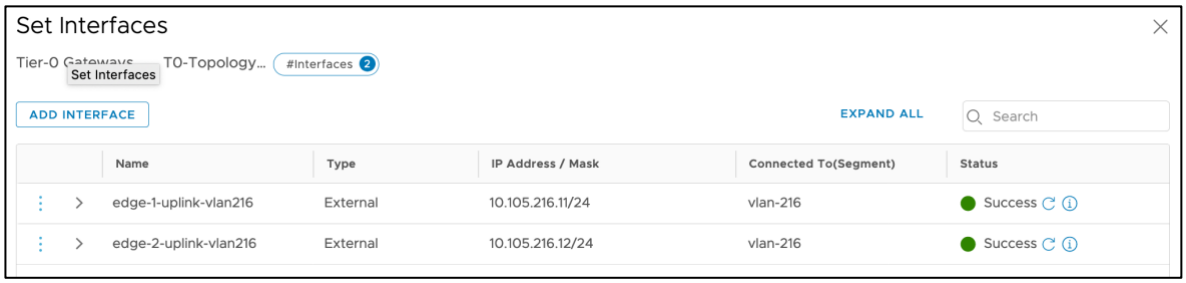


Figure 86 – Final Uplink interface configuration of the Tier-0 Gateway.

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:

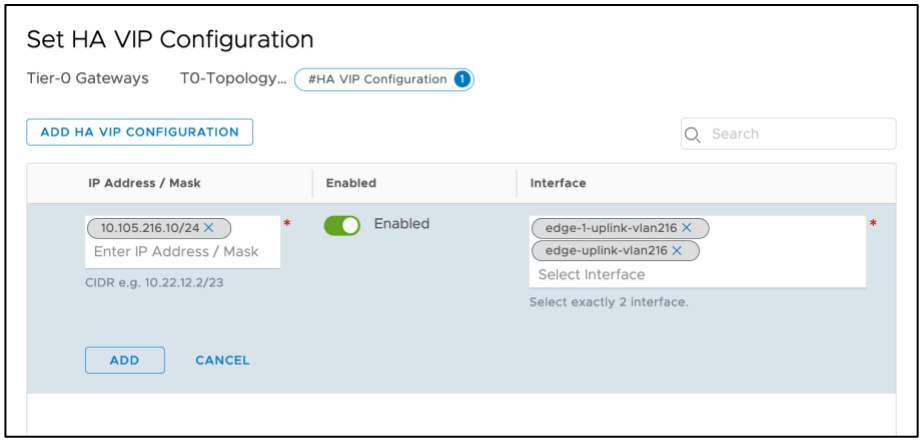


Figure 87 – Adding an HA VIP to NSX-T's T0 Gateway.

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`:

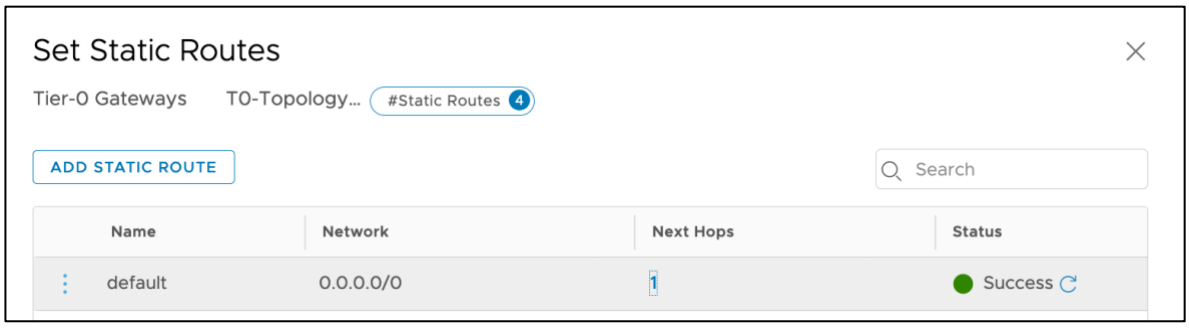


Figure 88 – Adding Tier-0 Gateway's default route.

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 on whether 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.

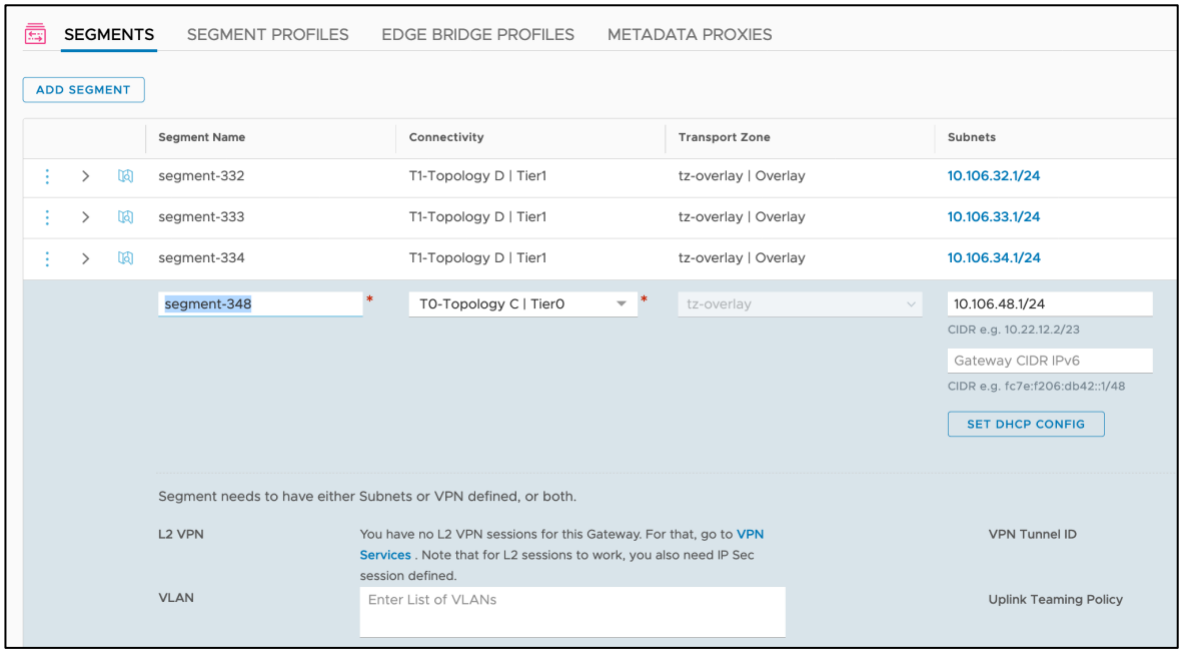


Figure 89 – Creating the Transit segment (segment-348) within the Overlay Transport Zone for a BIG-IP VE

2. Create a Tier-1 Gateway.

Although not part of this topology, this configuration 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 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.

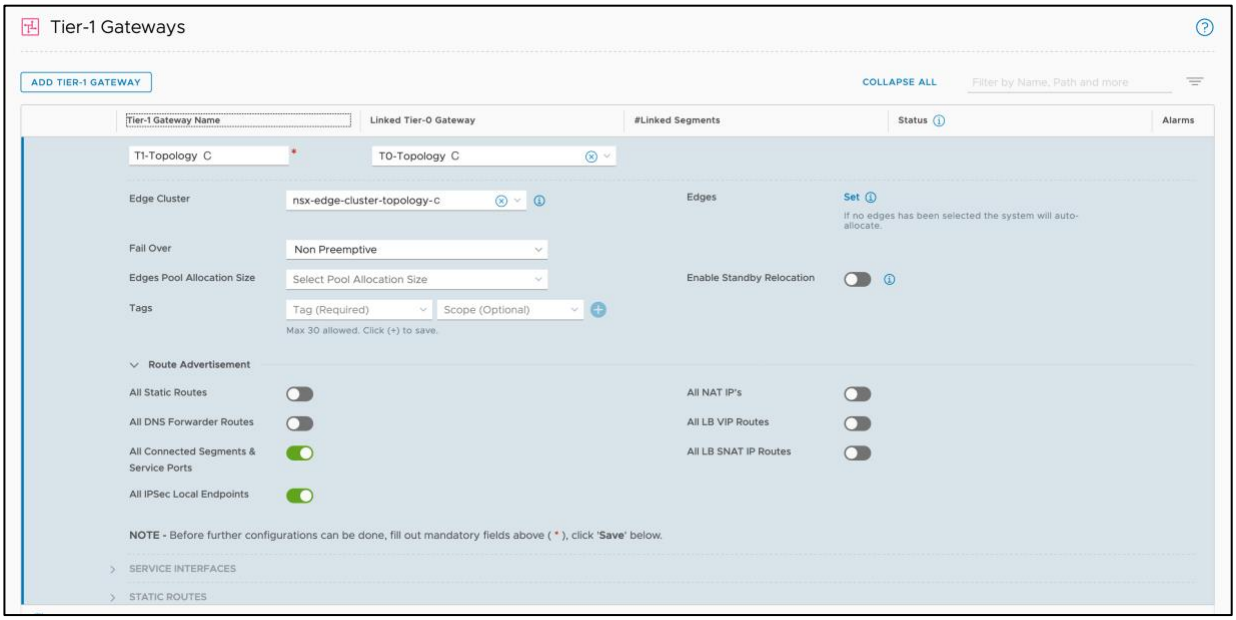


Figure 90 – Filling the properties when creating a Tier-1 Gateway.

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:

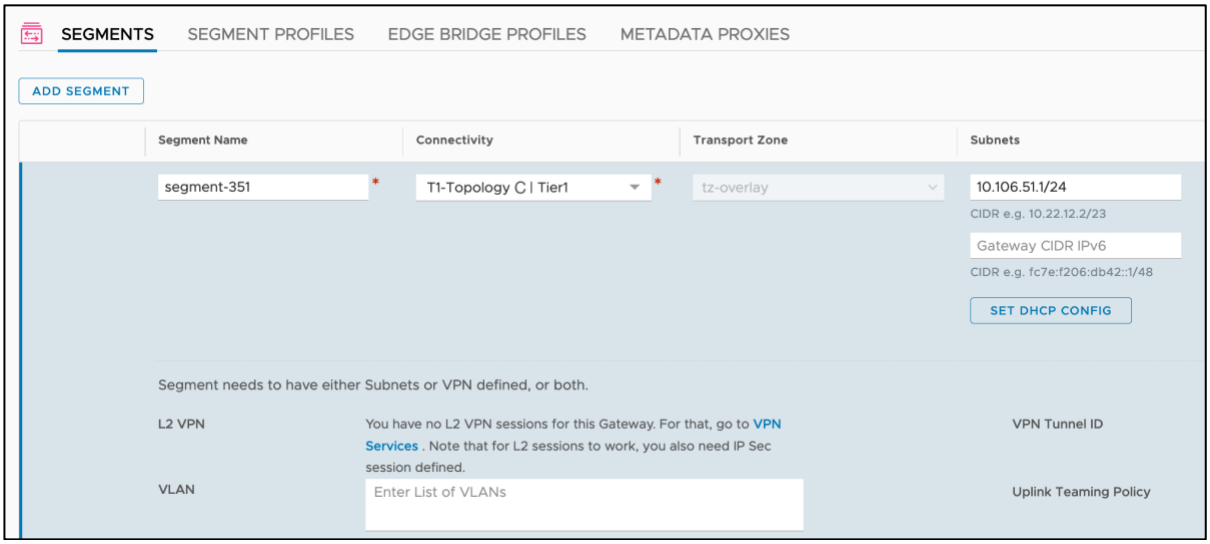


Figure 91 – Adding a segment to the T1 Gateway.

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:

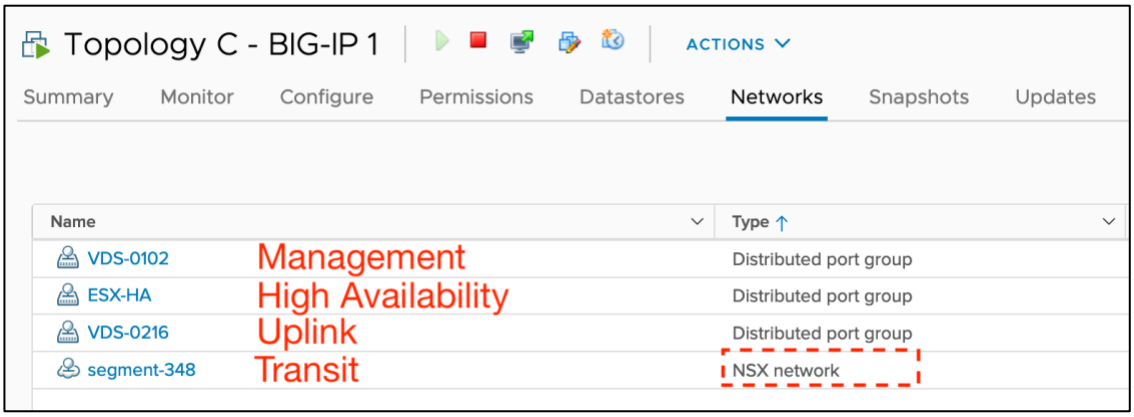


Figure 92 – Attaching the BIG-IP to an NSX-T overlay segment for the transit network.

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

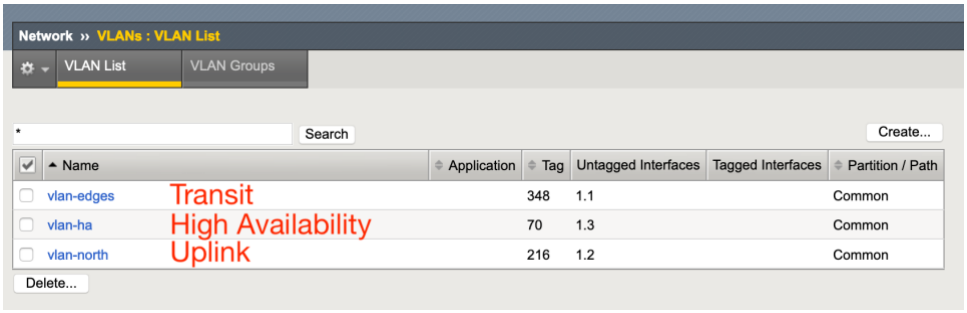


Figure 93 – Adding the Layer 2 networks to the BIG-IPs.

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.

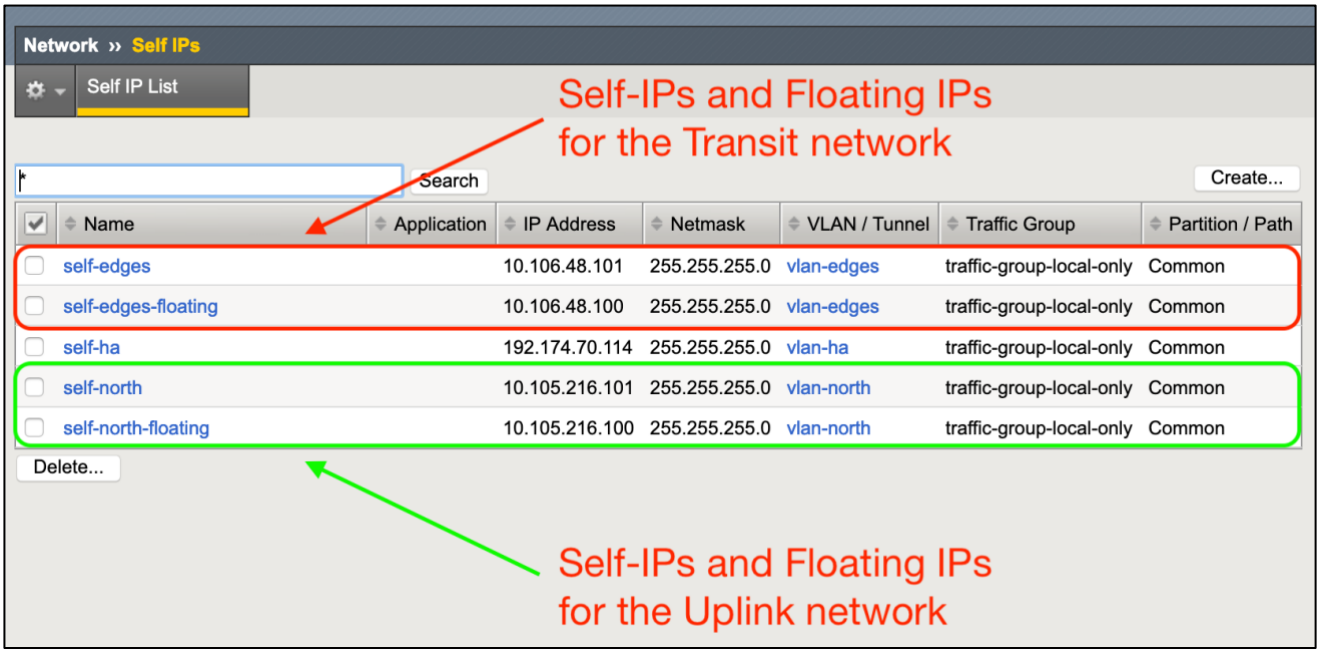


Figure 94 – Self IPs and floating Self IPs required (shown in BIG-IP unit 1).

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).

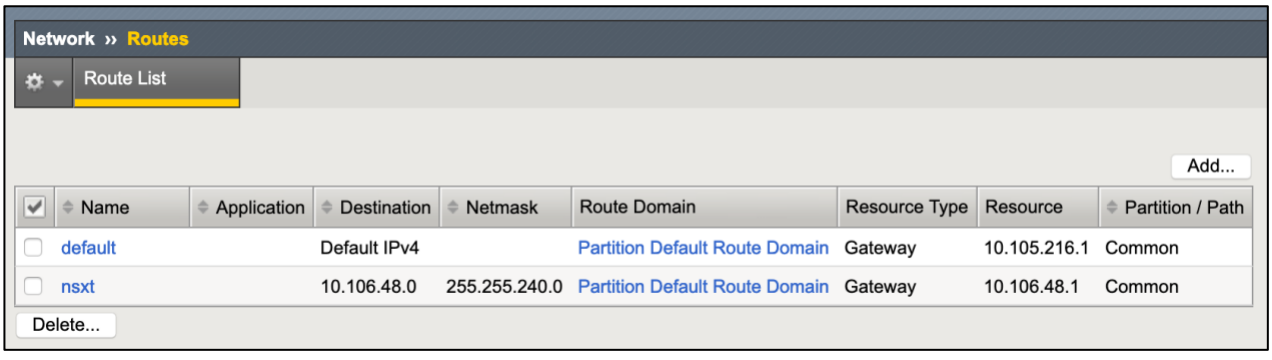


Figure 95 – Static routes required on the BIG-IP units.

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

Multi-tenant considerations for Topology C

One advantage of this topology is each tenant can have its own North-South BIG-IP VE, which can be managed independently, without interfering with the routed flows. A multi-tenant setup with full isolation is shown next.

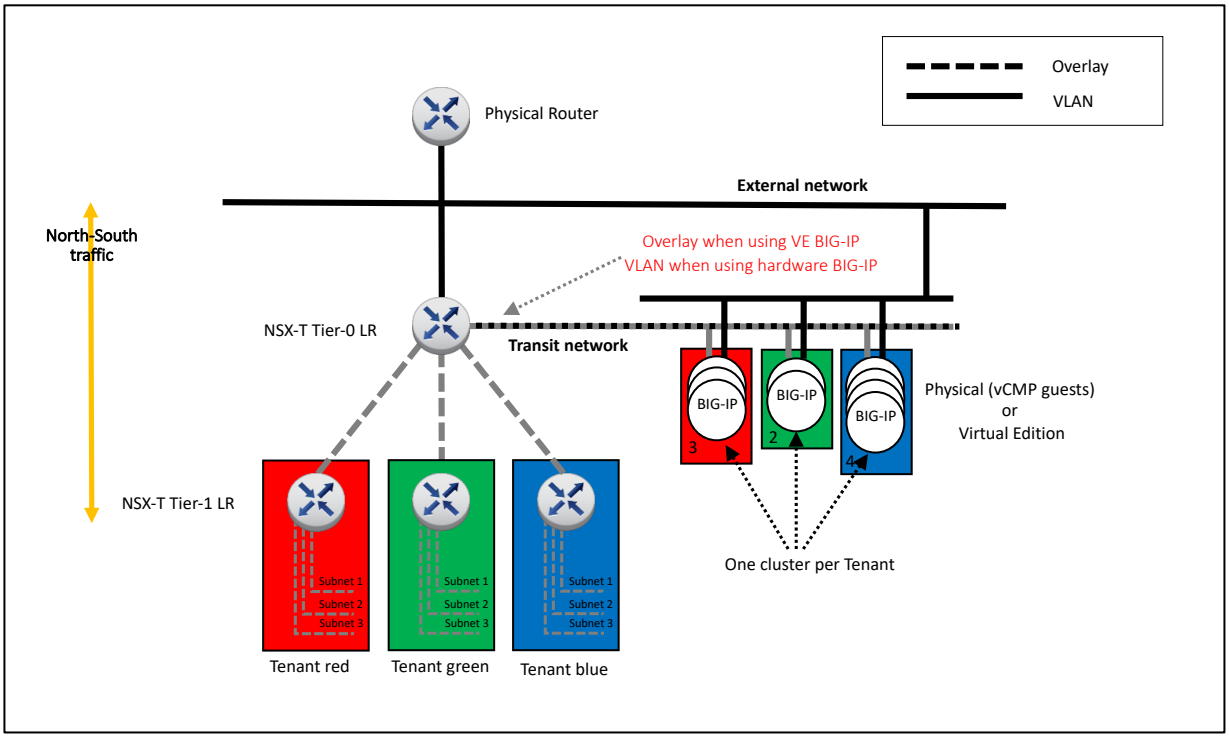


Figure 96 – Topology C with multiple tenants shown.

This topology has the following benefits:

- Allows a 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 BIG-IP VE.

On the other hand, it has the following drawbacks:

- It is a more complex topology, with two paths for the same endpoints.
- Requires SNAT, hiding client's IP addresses.
- This topology is suitable for ADC, advanced WAF/WAAP & Identity management use cases but requires that the direct path is tightly controlled in NSX-T's firewall otherwise security functionalities would be bypassed.

Similar to Topology A, the use of a single VE for multiple tenants with Topology C is discouraged. This is because it is expected these deployments will handle high throughputs. In the case of test environments, it is discouraged as well because the configuration (partitions,

route domains) will not match exactly the production deployment using per-tenant VE or using vCMP hardware.

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

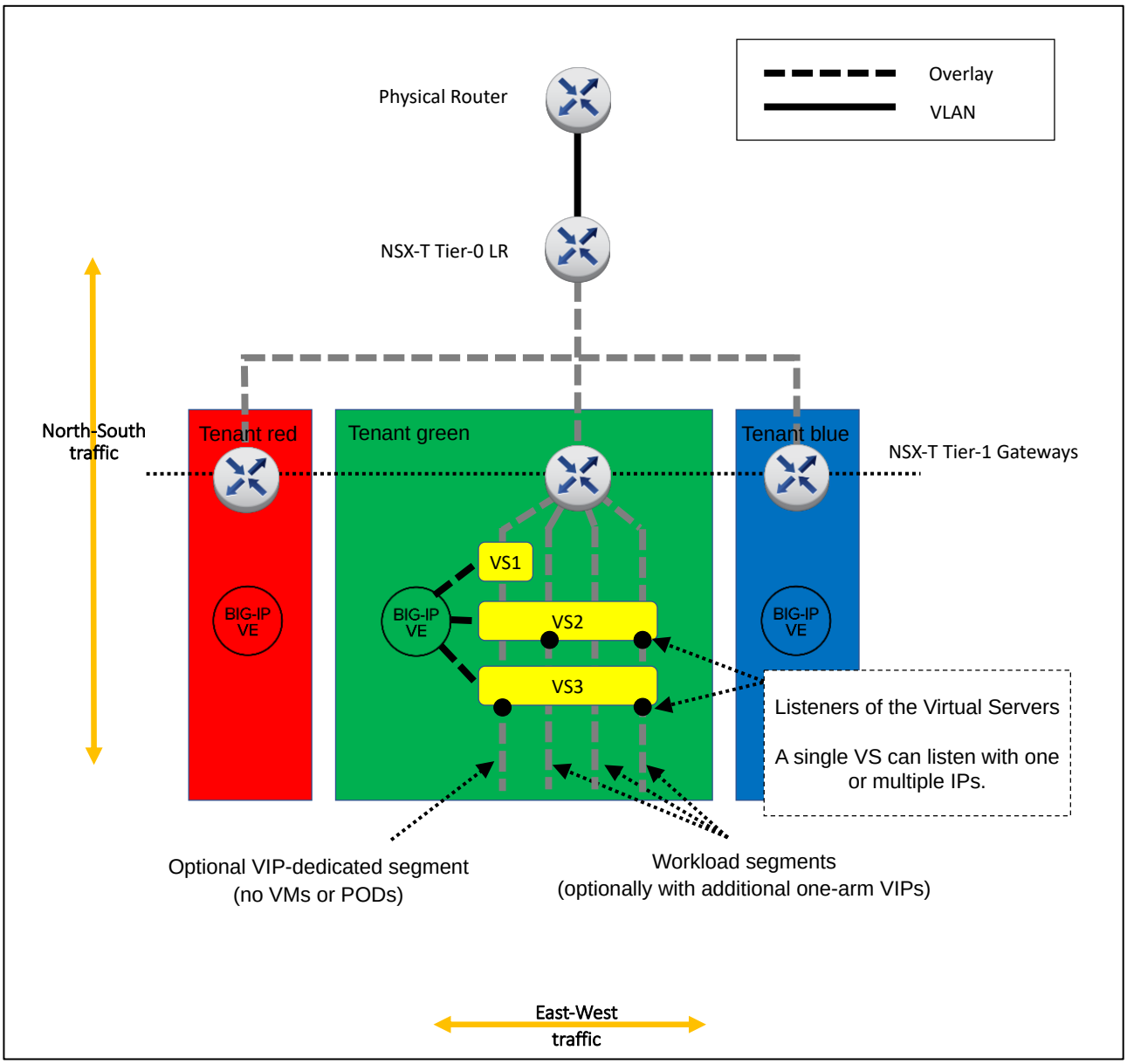


Figure 97 – Topology D overview (simplified view without HA components).

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-IQ 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:

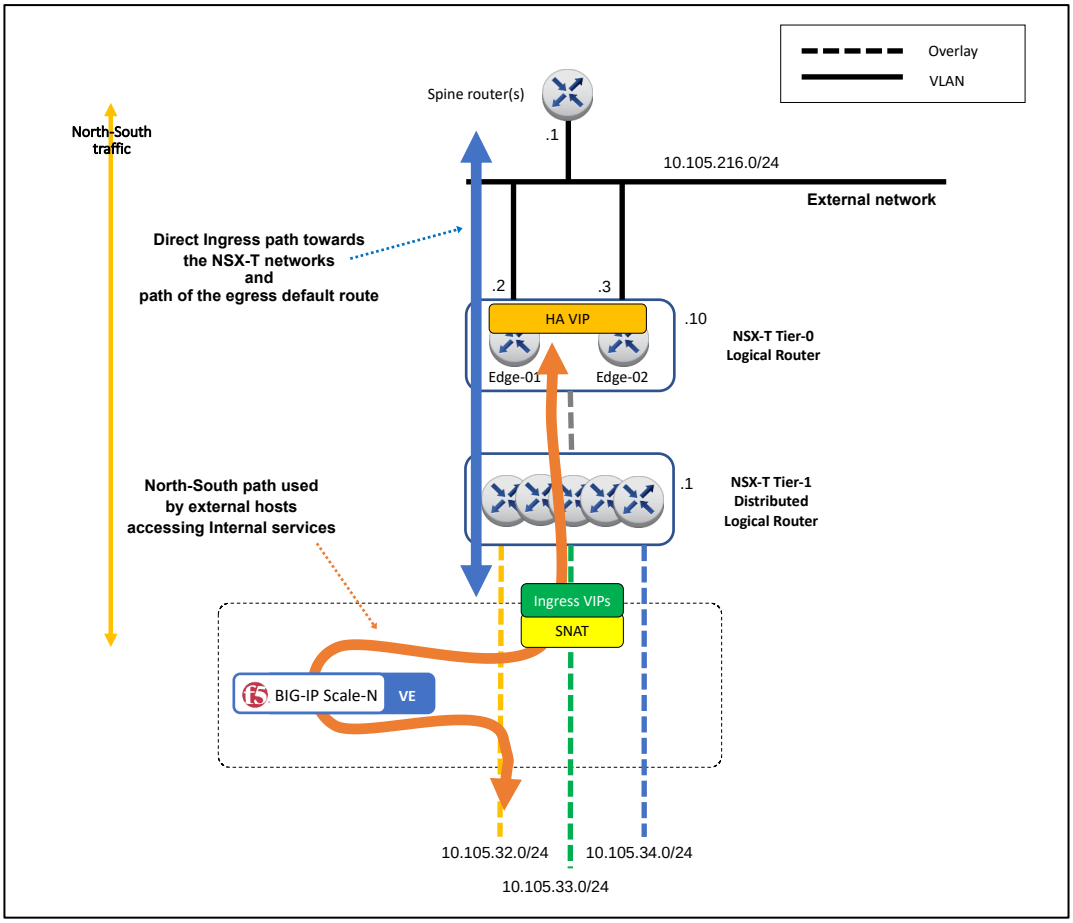


Figure 98 – Topology D example implementation – North/South traffic flows.

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:

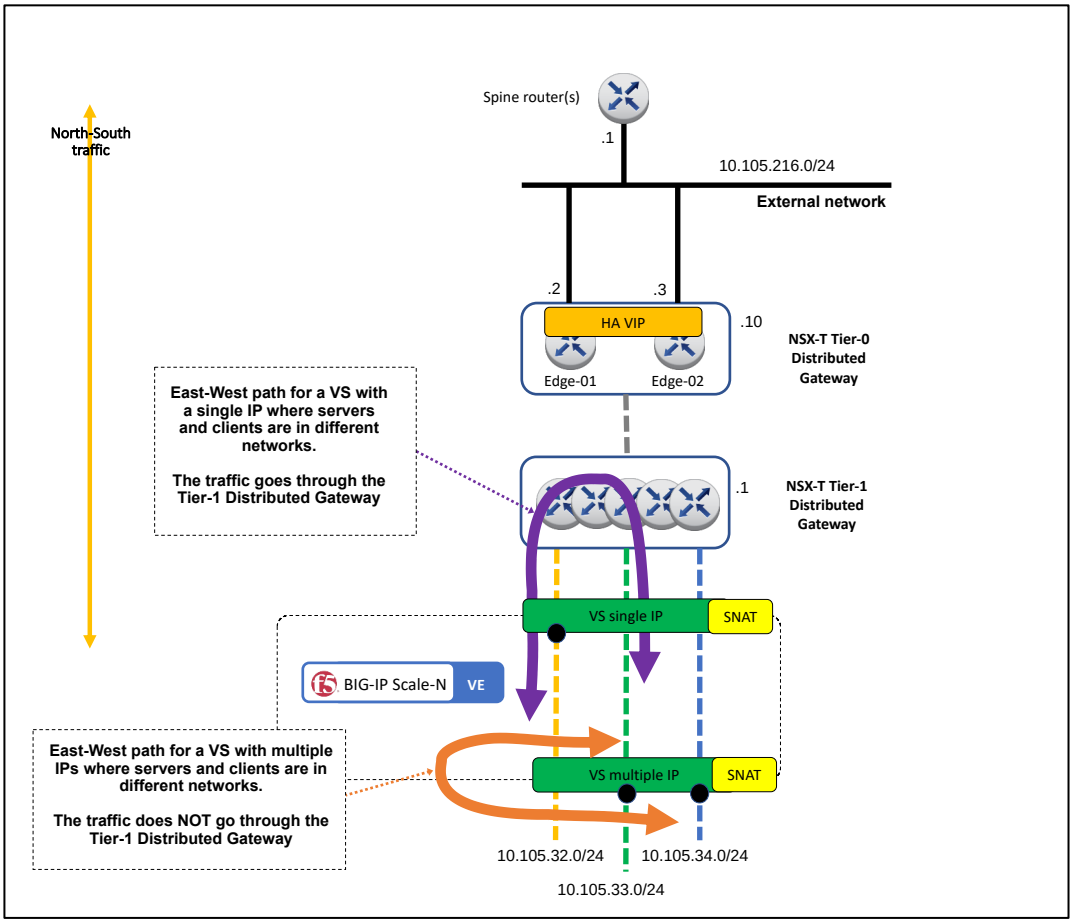


Figure 99 – Topology D example implementation – East/West traffic flows.

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. For the client to reach the VIP it has go to 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.

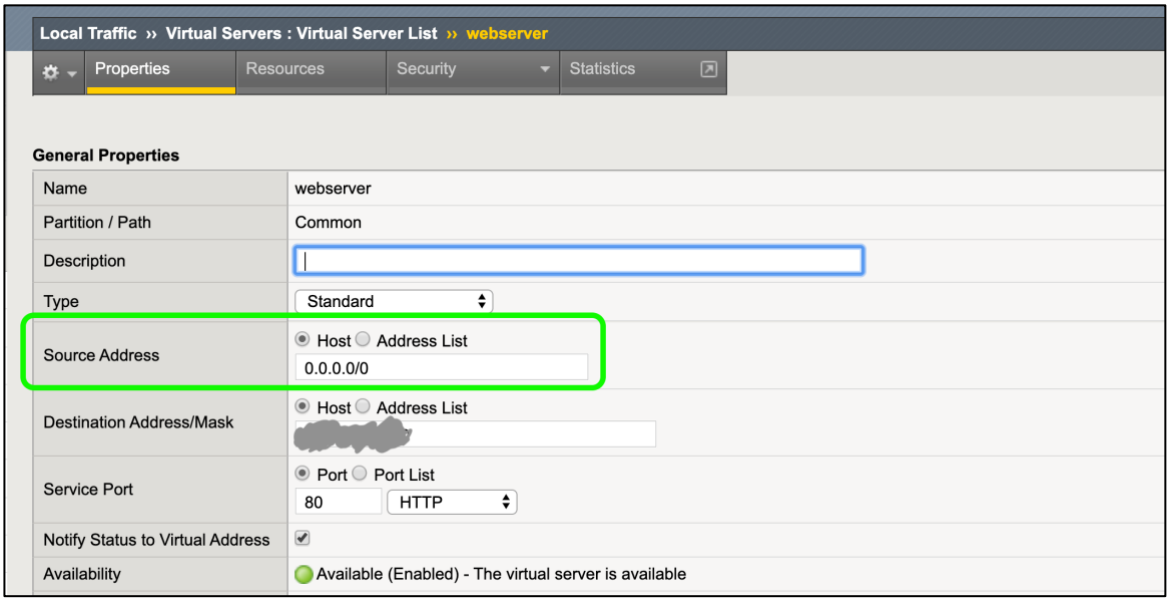


Figure 100 – Source Address setting to discriminate the prefixes to which the Virtual Server applies.

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

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

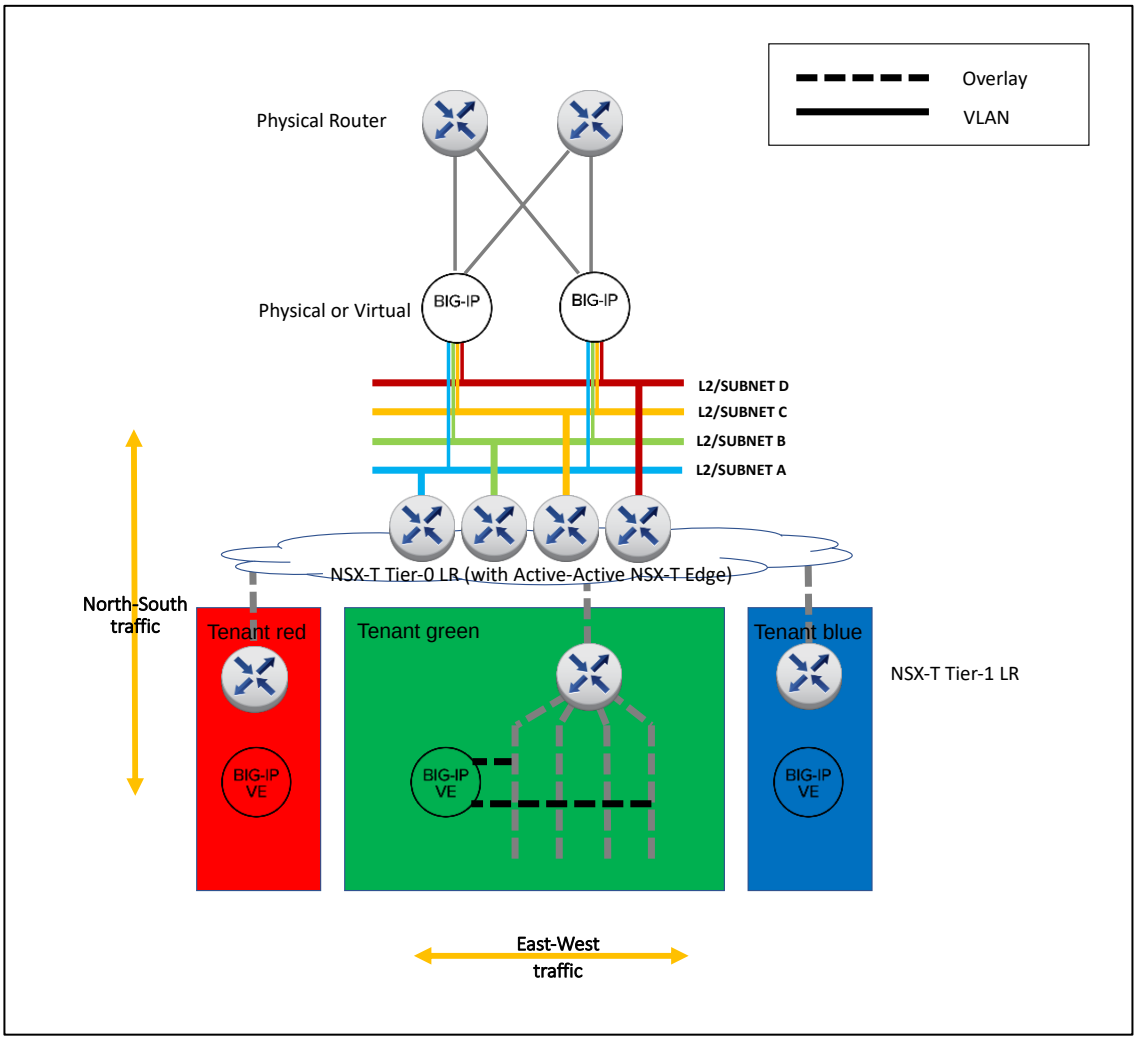


Figure 101 – Combined Topology A and D.

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

The figure below shows the configuration implemented in this section.

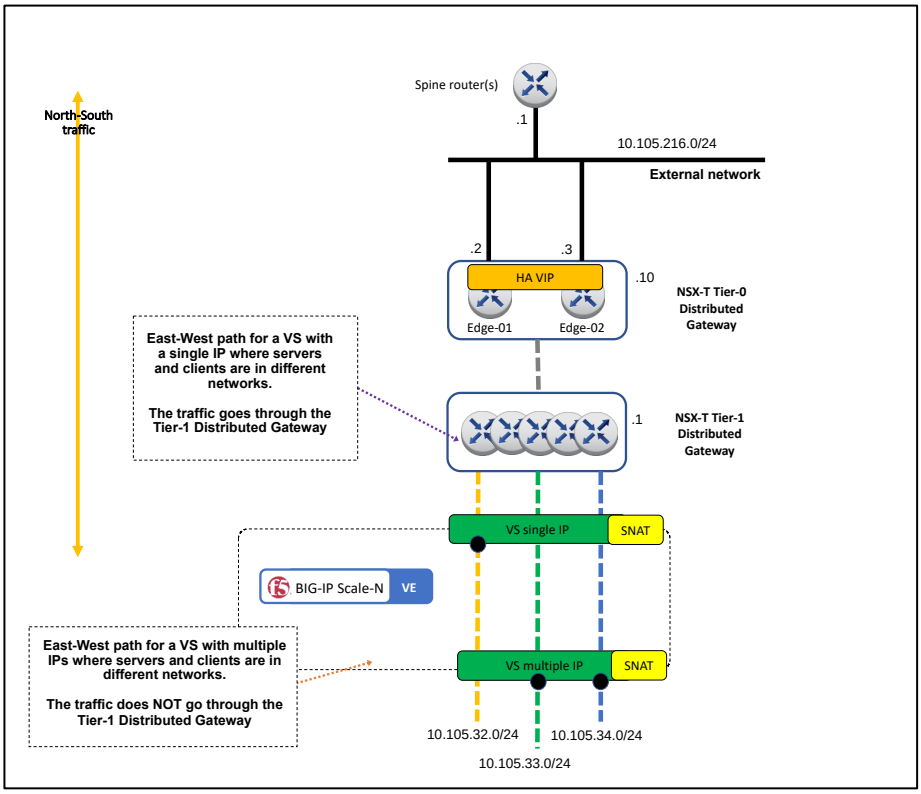


Figure 102 – Topology D implementation used through this section.

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).

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.

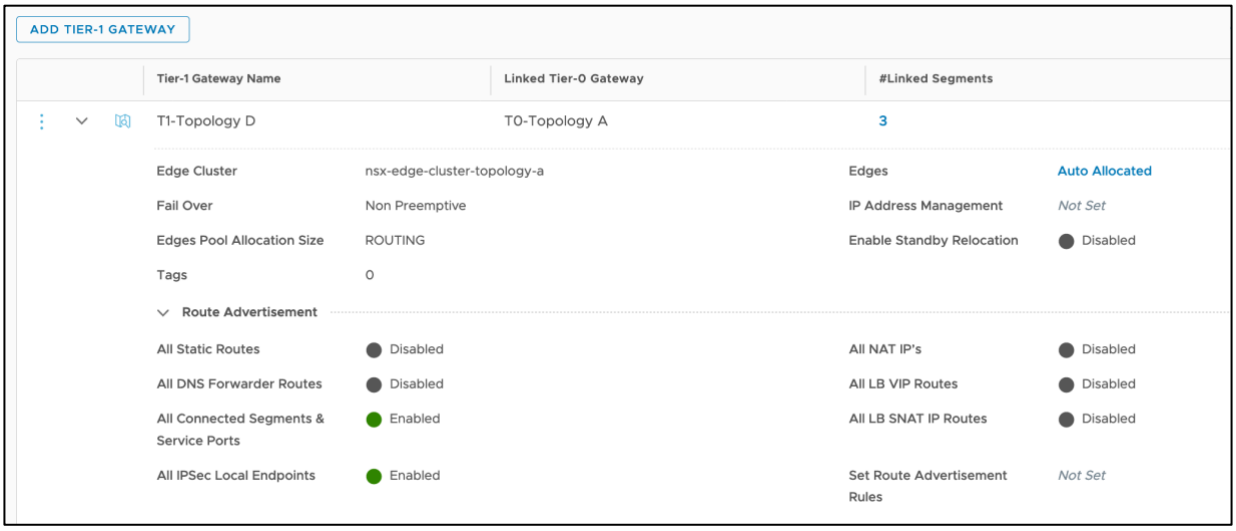


Figure 103 – Filling the properties when creating a Tier-1 Gateway.

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:

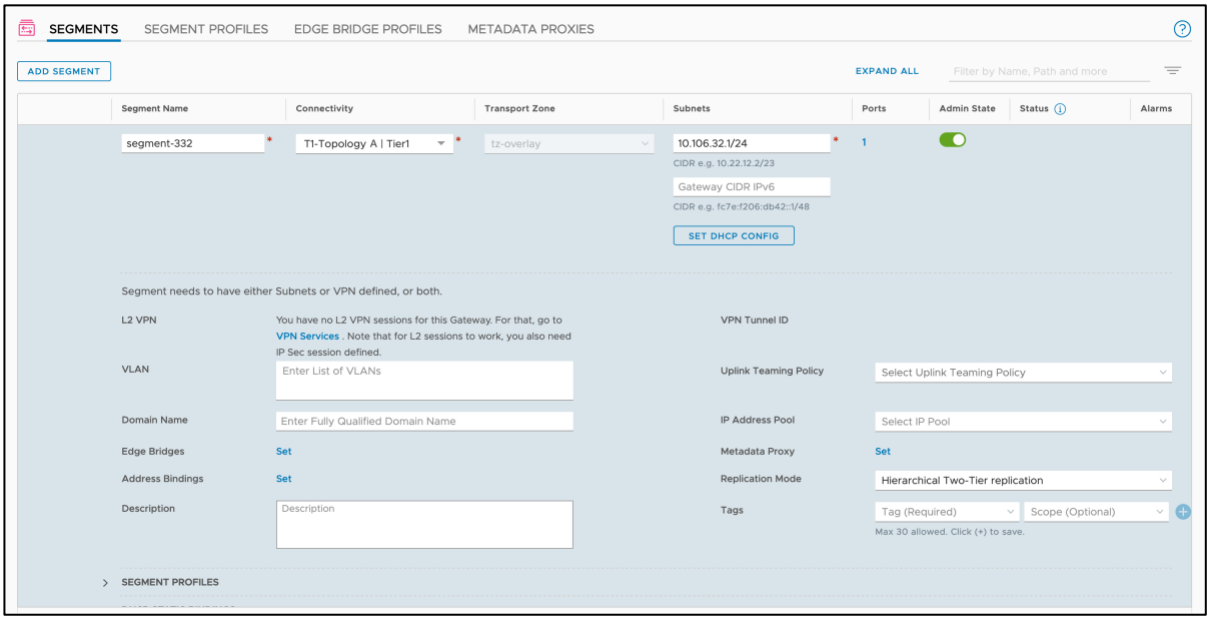


Figure 104 – Adding a segment to the T1 Gateway.

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 the figure below.

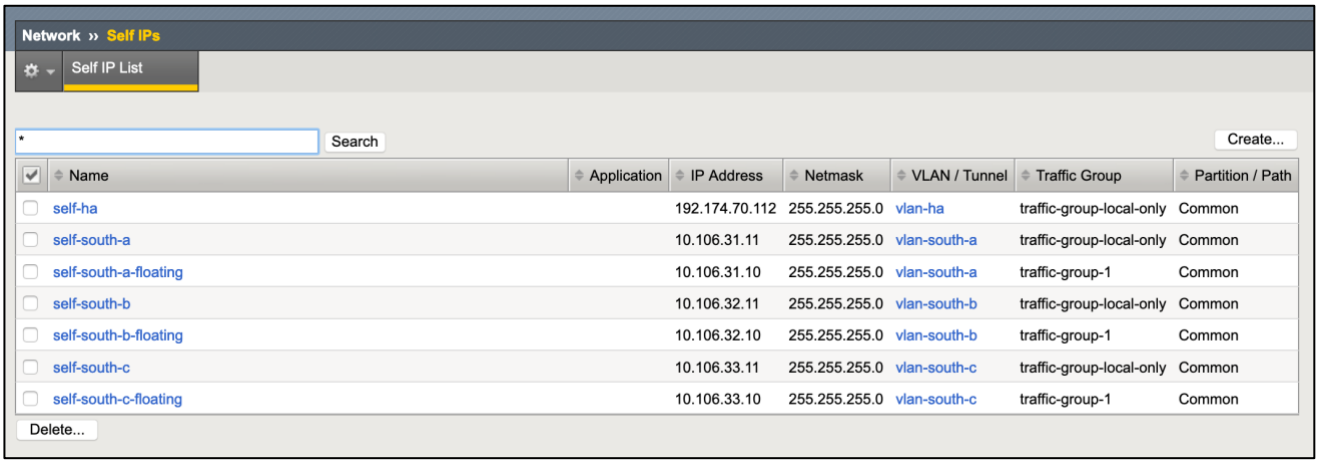


Figure 105 – Self IPs and floating Self IPs required (shown in BIG-IP unit 1).

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

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 the figure below:

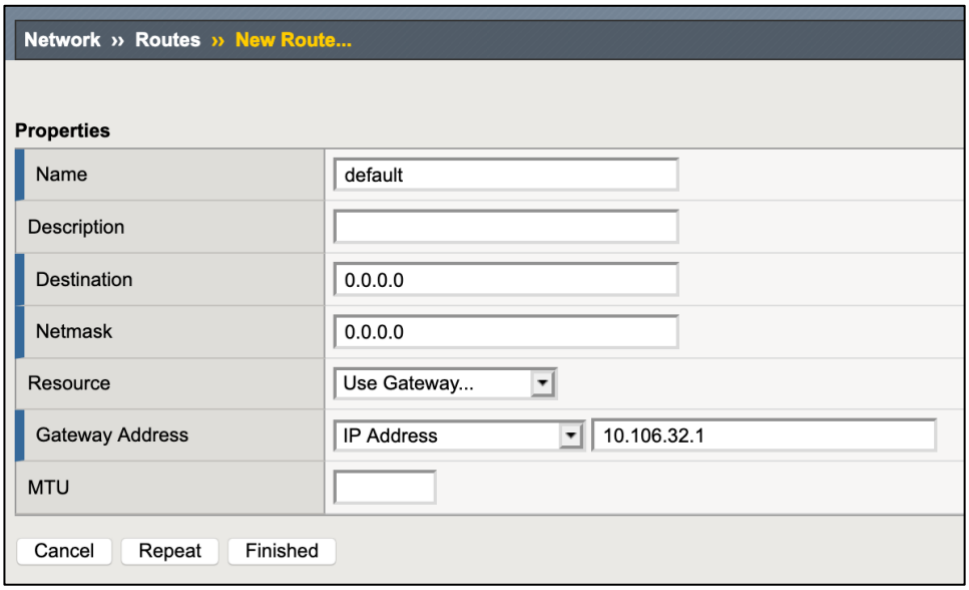


Figure 106 – Static routes required in the BIG-IP units.

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

Multi-tenant considerations for Topology D

Topology D only supports VE because it makes use of NSX-T overlays. The general recommendation is to use per-tenant VE because when it is required, it provides hard isolation and provides higher scalability. Per-tenant VE doesn't require any special configuration.

Alike topology B extended, a shared VE model can be interesting to reduce footprint or expenditures depending on the licensing used. The number of tenants is limited to 7 per VE cluster because the ESXi hypervisor limits the number of vNICs to 10 per VM (the remaining vNICs are used for management, HA and the upstream link).

When using multi-tenancy in VE it is always encouraged to use partitions. Given that this is not an inline topology, it is not expected to use forwarding virtual servers and hence it is not expected that route domains with strict isolation are required, yet the use of this latter is a good practice which avoids any unexpected flows due to mis-configurations.

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 Gateway[8]. 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.

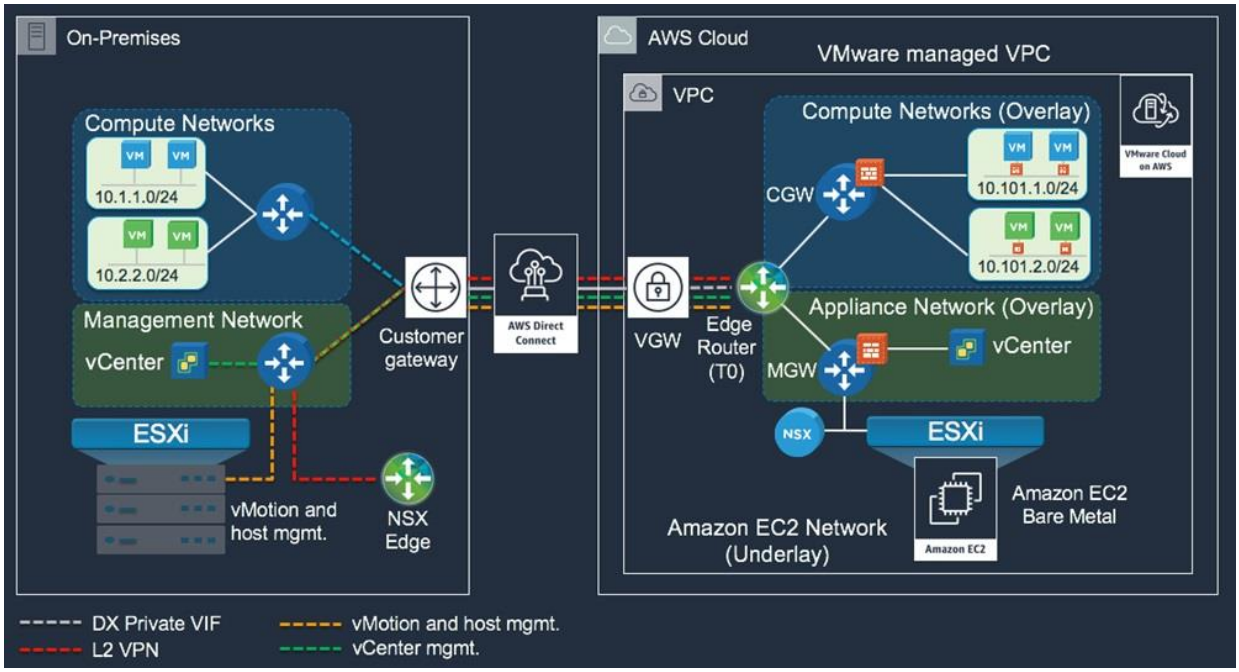


Figure 107 - Sample architecture showing some connectivity options

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

[8] 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 for 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

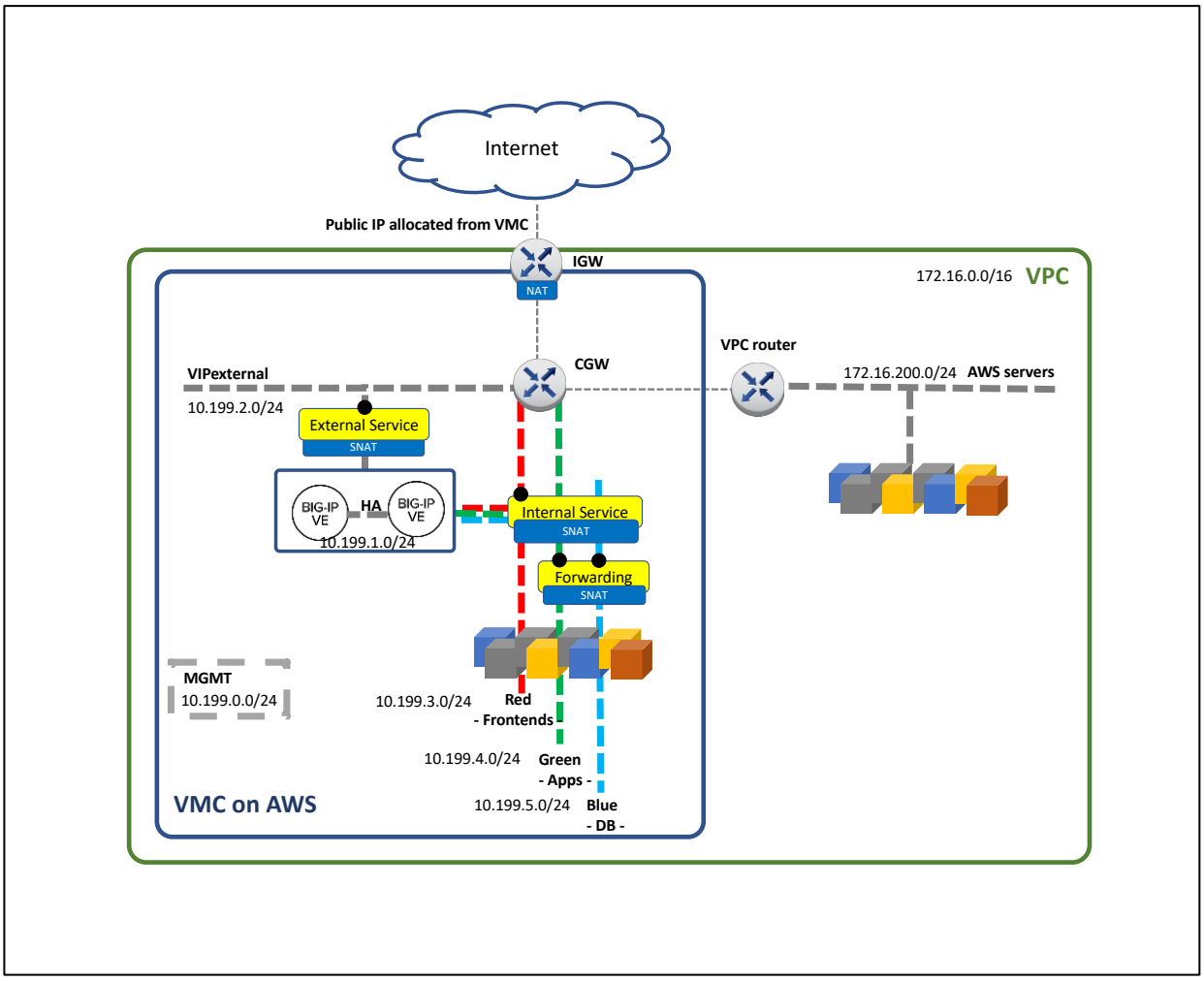


Figure 108 - Sample Topology D for VMC

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 security by means of controlling the access through a VIP on the BIG-IP. The created segments can be seen in the next figure.

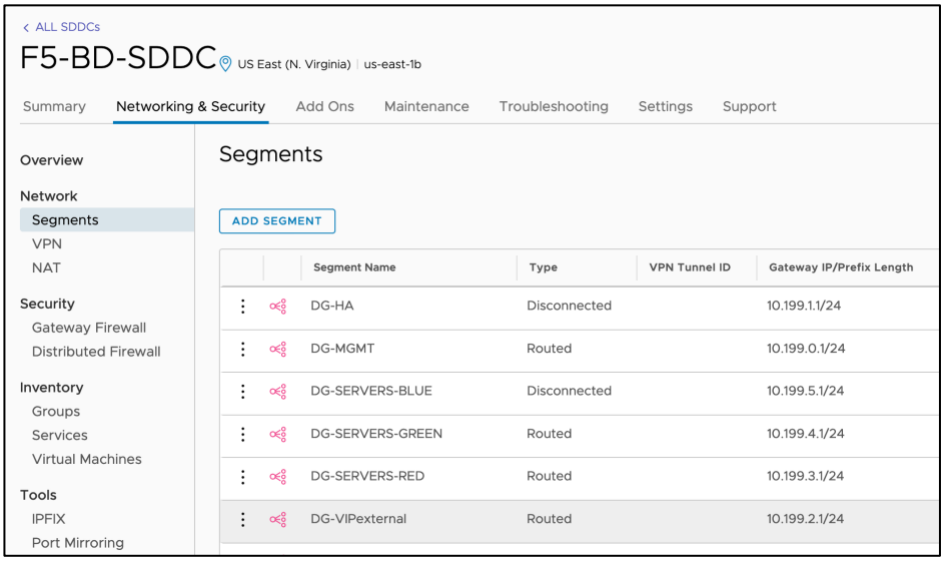


Figure 109 - Segments configuration in VMC

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

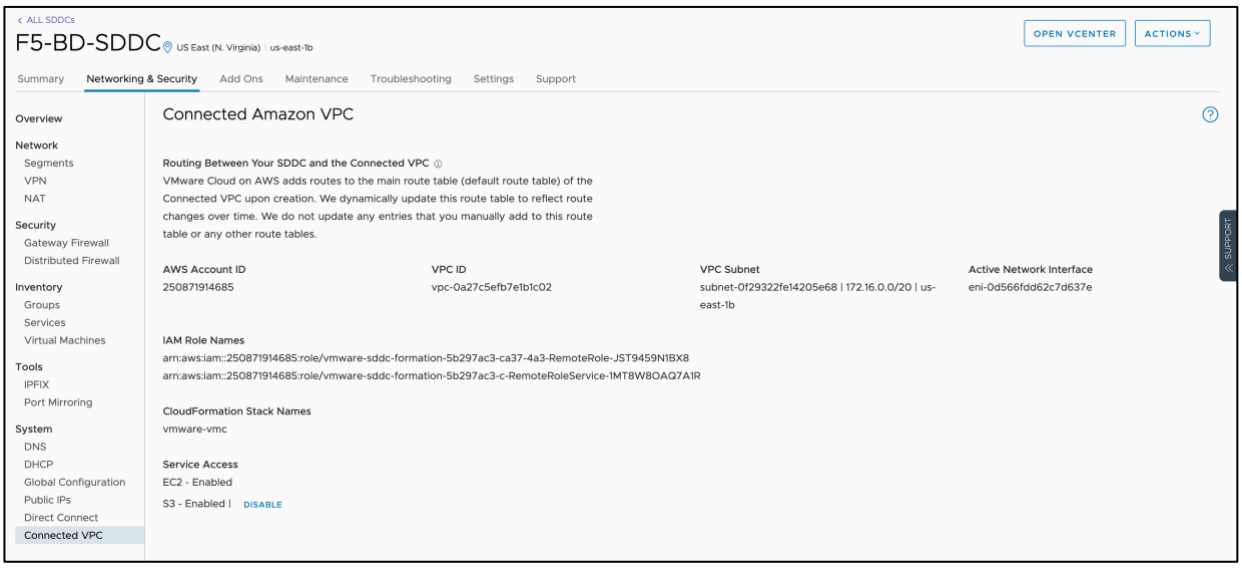


Figure 110 - Checking AWS VPC from the VMC console

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:

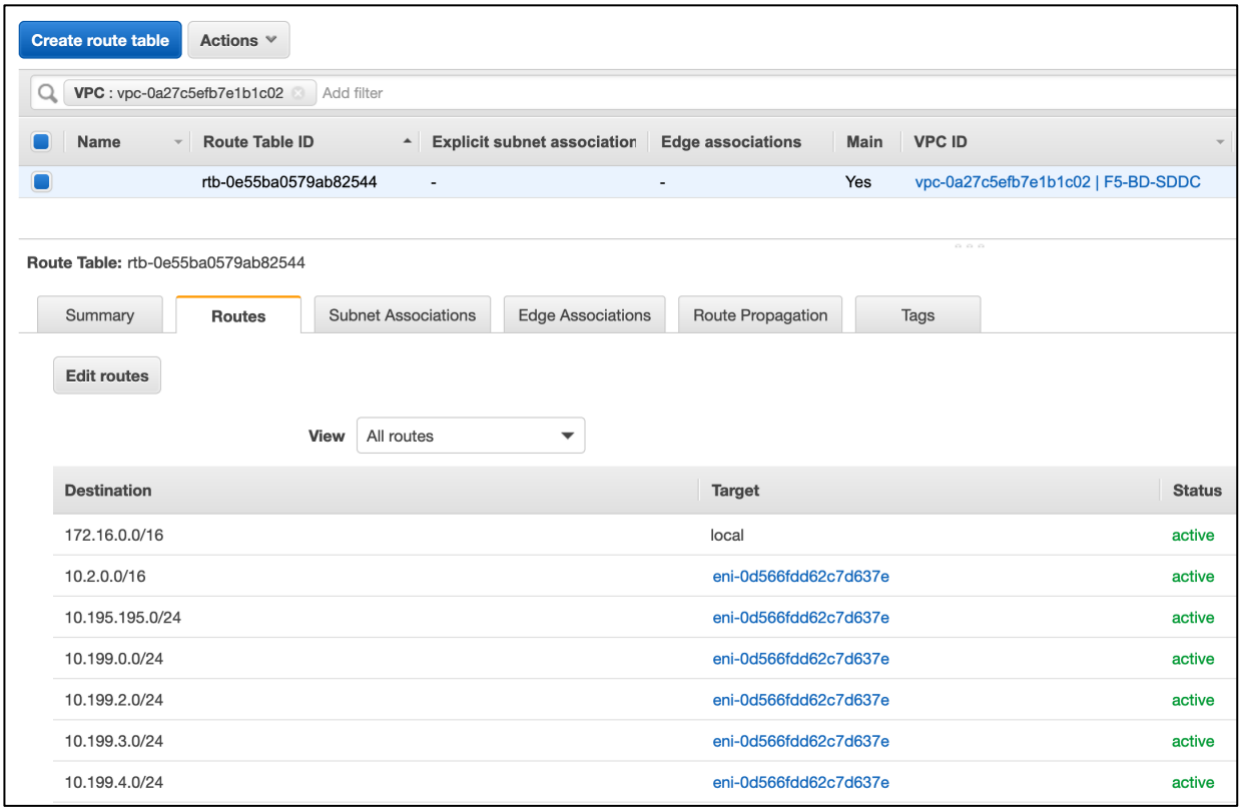


Figure 111 - Automatically created routing table of the VPC

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.

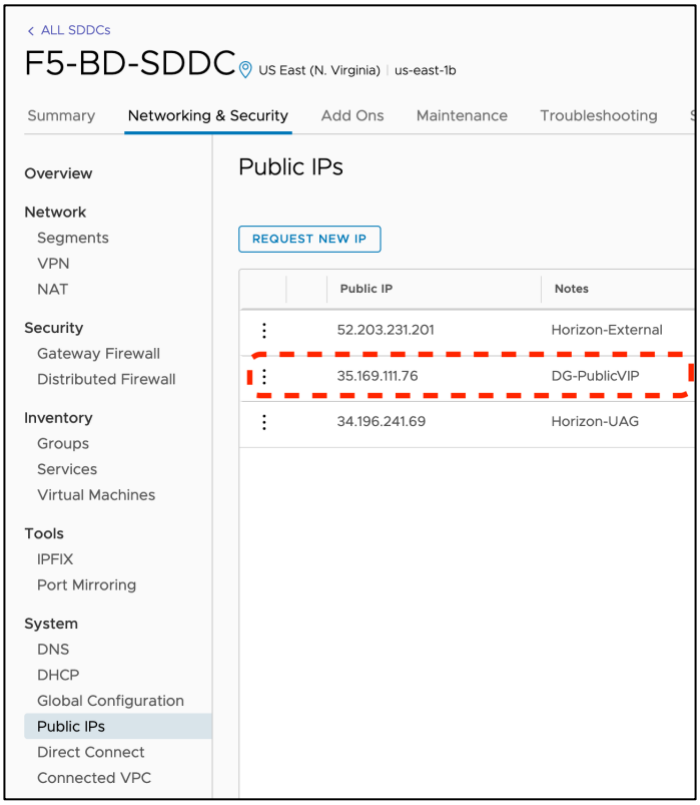


Figure 112 - Allocating an IP address for the VIP in the BIG-IPs

This public IP needs to be mapped into the VIP of the BIG-IP that we will configure later. 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.

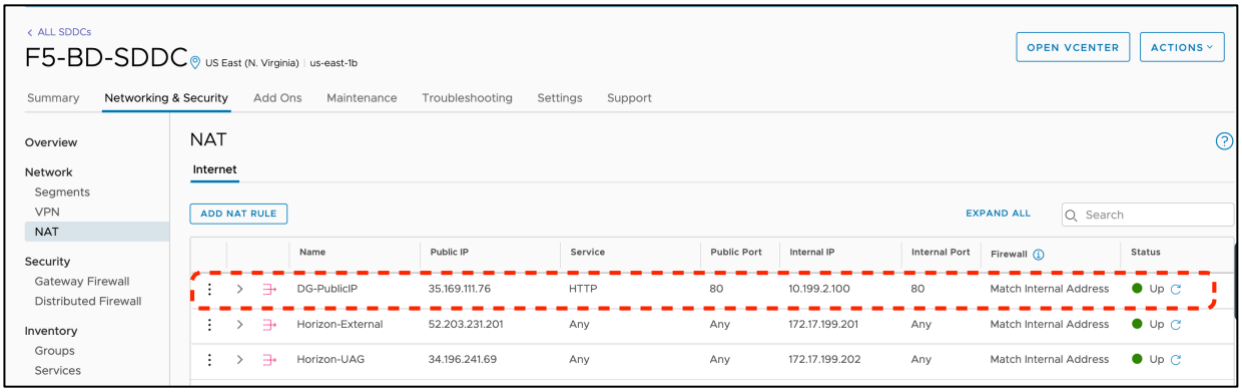


Figure 113 - Configuring the required 1:1 NAT for the BIG-IP VIP.

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 Self-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.

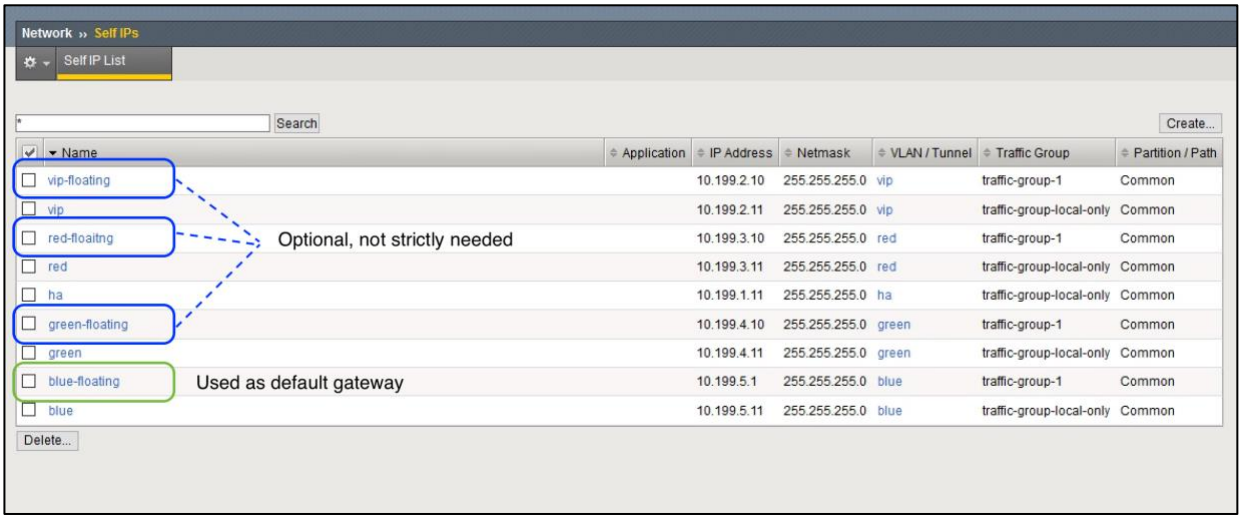


Figure 114 - Directly connected segments. Self-IP configuration.

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.

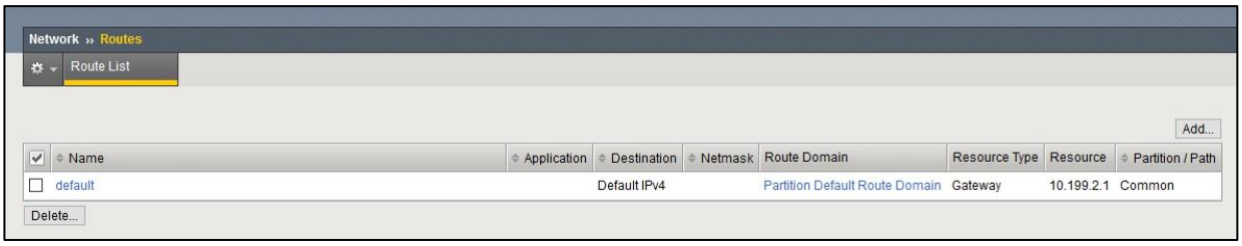


Figure 115 - Routing required for non-directly connected segments, including AWS workload segments in the VPC.

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.

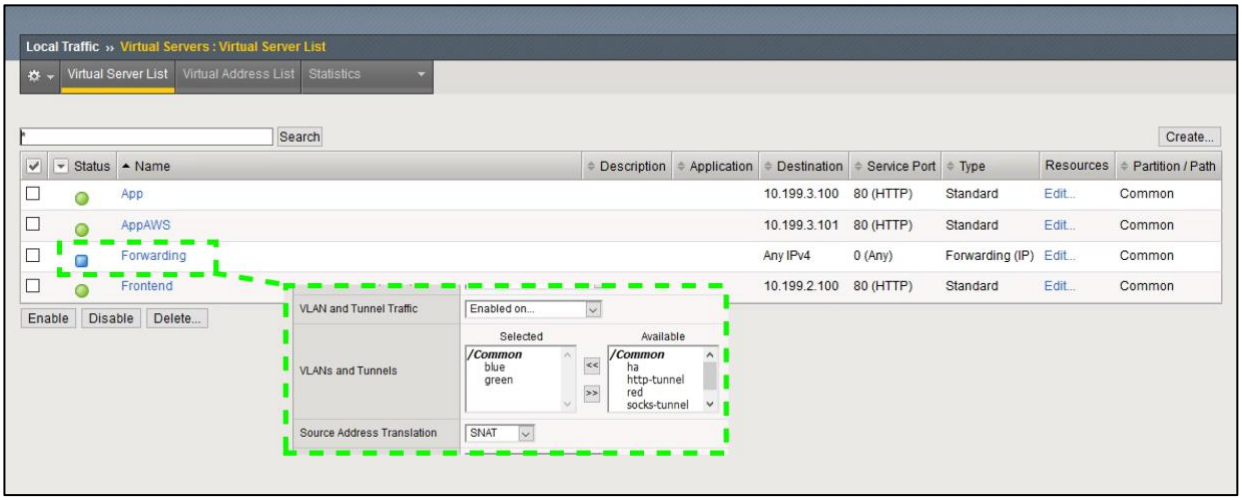


Figure 116 - Overview of the service configuration with detailing the additional segments where the Forwarding VIP is enabled.

Alternative topologies for BIG-IP in VMC on AWS

It is possible to configure services on 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 on 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 for 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 (IX), private data centers, or public clouds.

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

SaaS Security and multi-cloud

Several security functions such anti-DDoS and advanced WAF/WAAP 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.

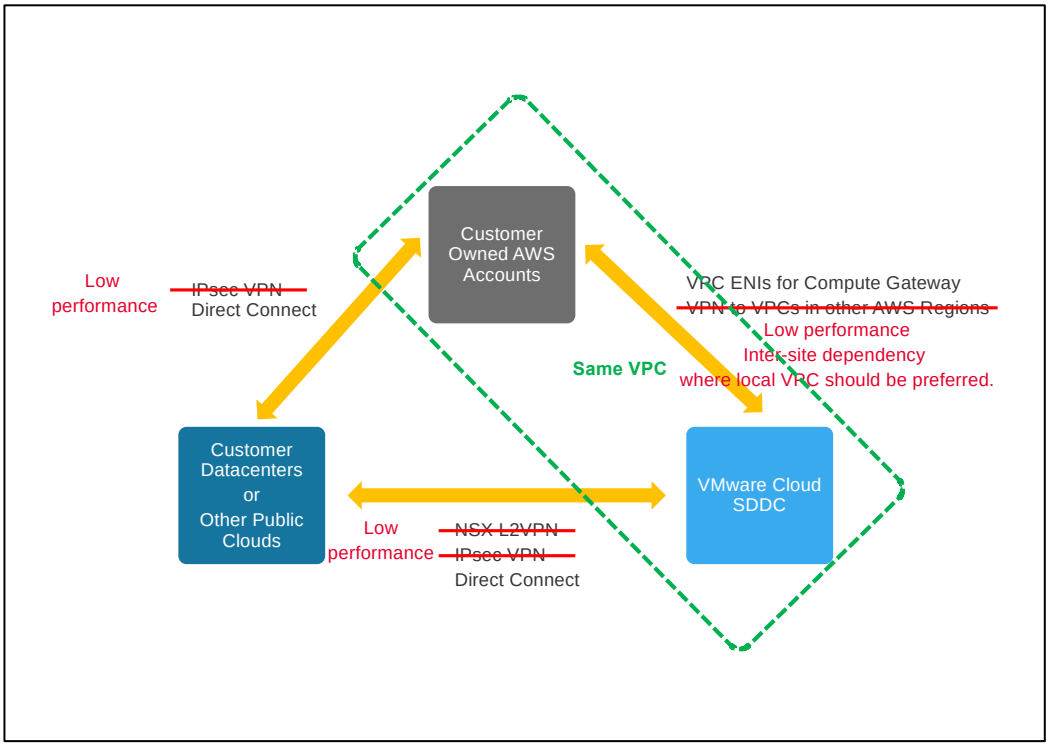


Figure 117 - Distilled connectivity options between the different types of clouds (squares). The less suitable connectivity options are stricken through and with annotations in red indicating the reason why they are less suitable.

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 cases[9] are:

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

[9] 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.

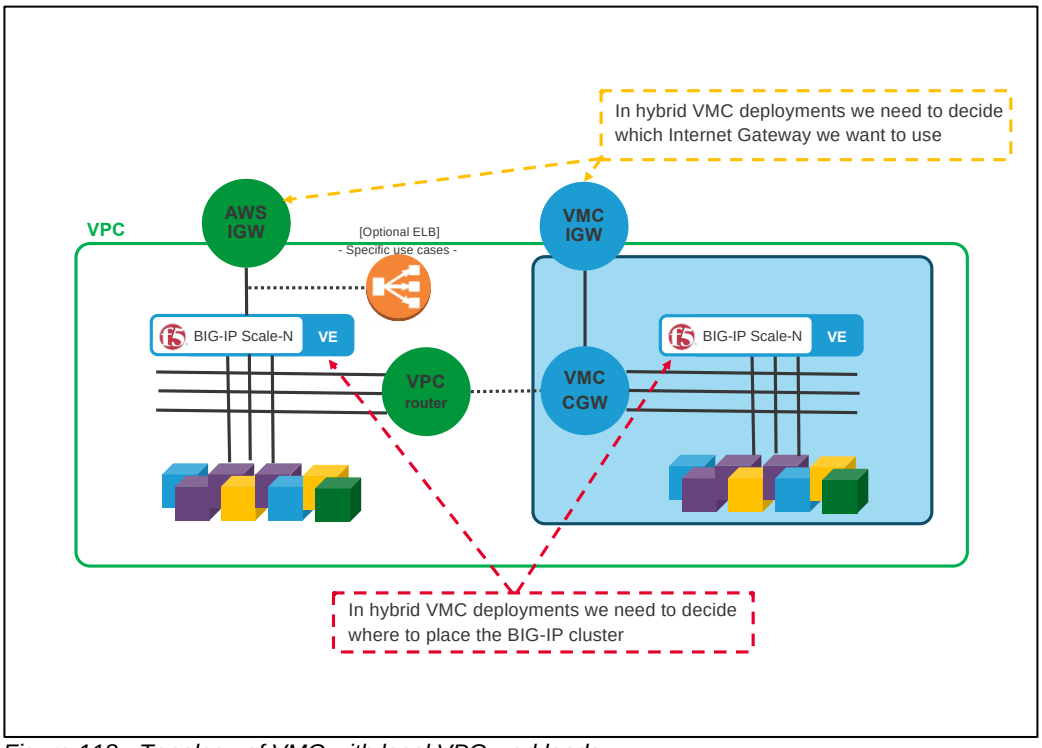


Figure 118 - Topology of VMC with local VPC workloads

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 on 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 guidelines 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 lower 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 guidelines deal with traffic that is generated upon client requests (blue arrows in the figure below). On the other hand, the third guideline 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 (orange arrows) 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 sites' or services' 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.

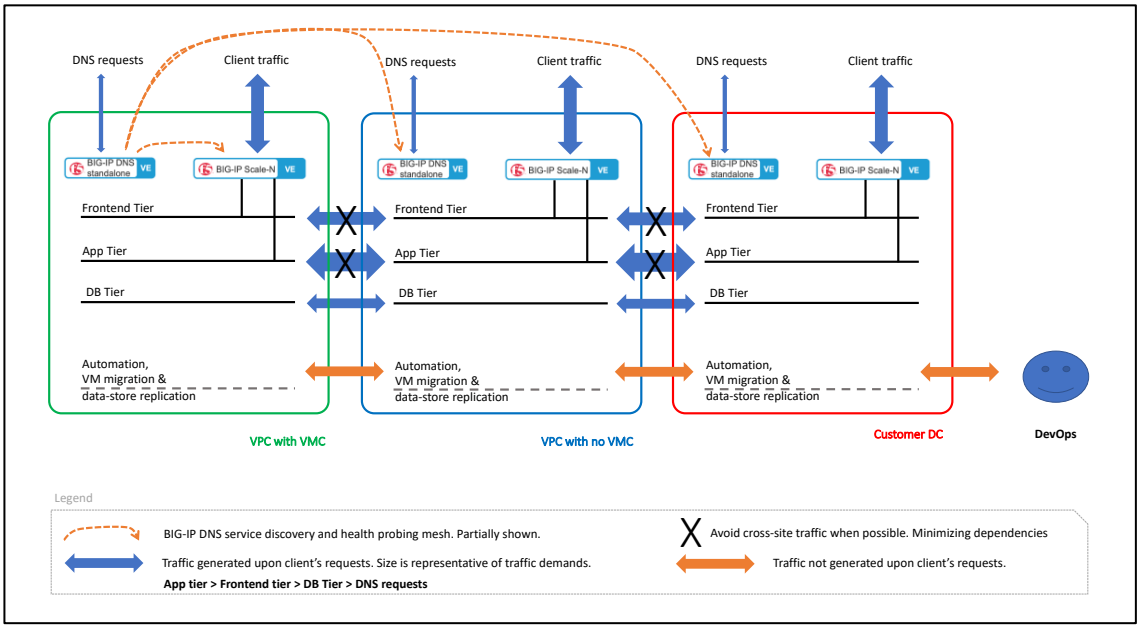


Figure 119 - Typical multi-cloud arrangement with most relevant traffic patterns.

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 to 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.

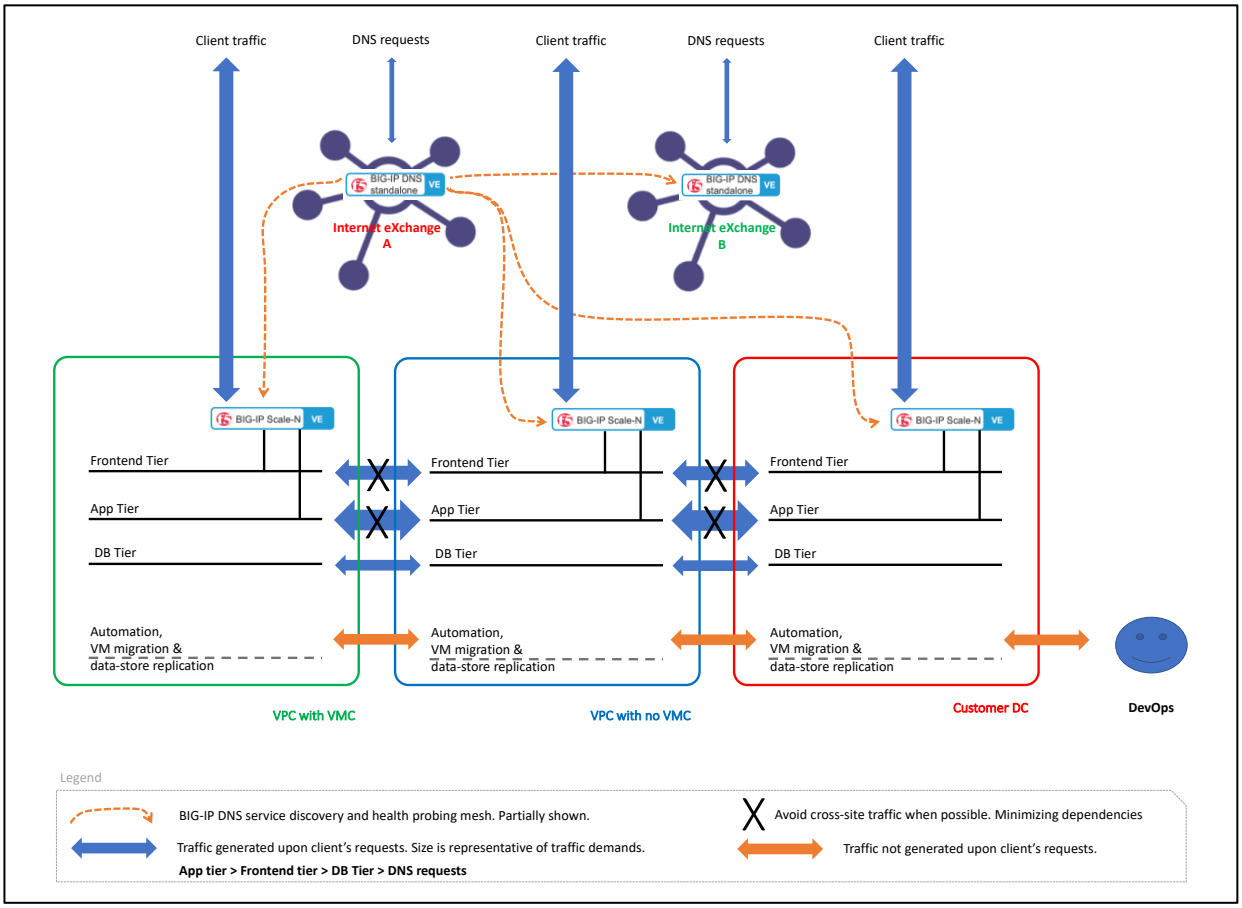


Figure 120 - Preferred multi-cloud arrangement by using Internet exchanges for BIG-IP DNS.

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.

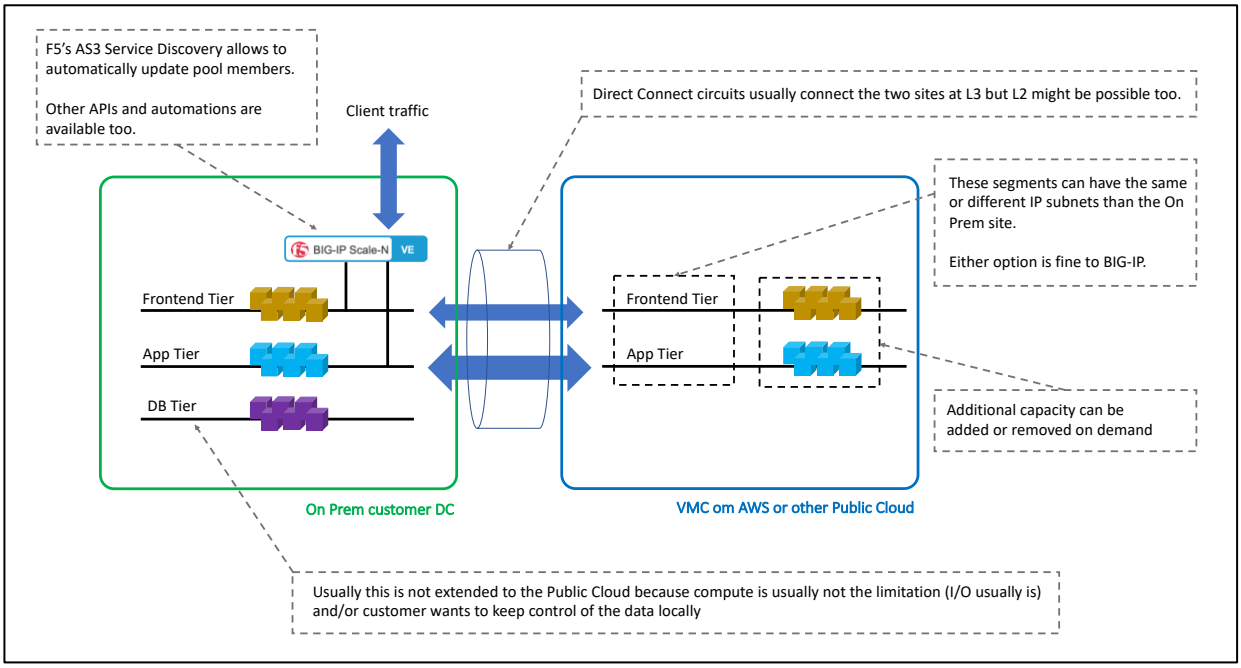


Figure 121 - Overall design of a single site with Cloud Bursting capability.

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

General best practices for BIG-IP in VMware NSX-T

- Management interface connectivity

  Following VMware's general recommendations, the management interface should not be in an overlay network. This minimizes the number of components in the NSX software stack required to access critical functions usually found in the management interfaces.

- HA connectivity

  A BIG-IP HA cluster uses two channels for HA communication, one through the management port and one through a dedicated data plane interface. Both are used for HA failover signaling. The data plane interface is also used for connection and persistence mirroring.

  In this guide there is no strong recommendation on whether this data plane interface should be in an overlay or a VLAN backed segment given that one of the failover channels (the management interface) already doesn't rely on an overlay. The ultimate decision might rely on the underlying infrastructure.

- VM placement in vCenter (on premises deployments)

  The following recommendations should be followed:

  - BIG-IQs should be placed alongside the 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 on 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 on 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.

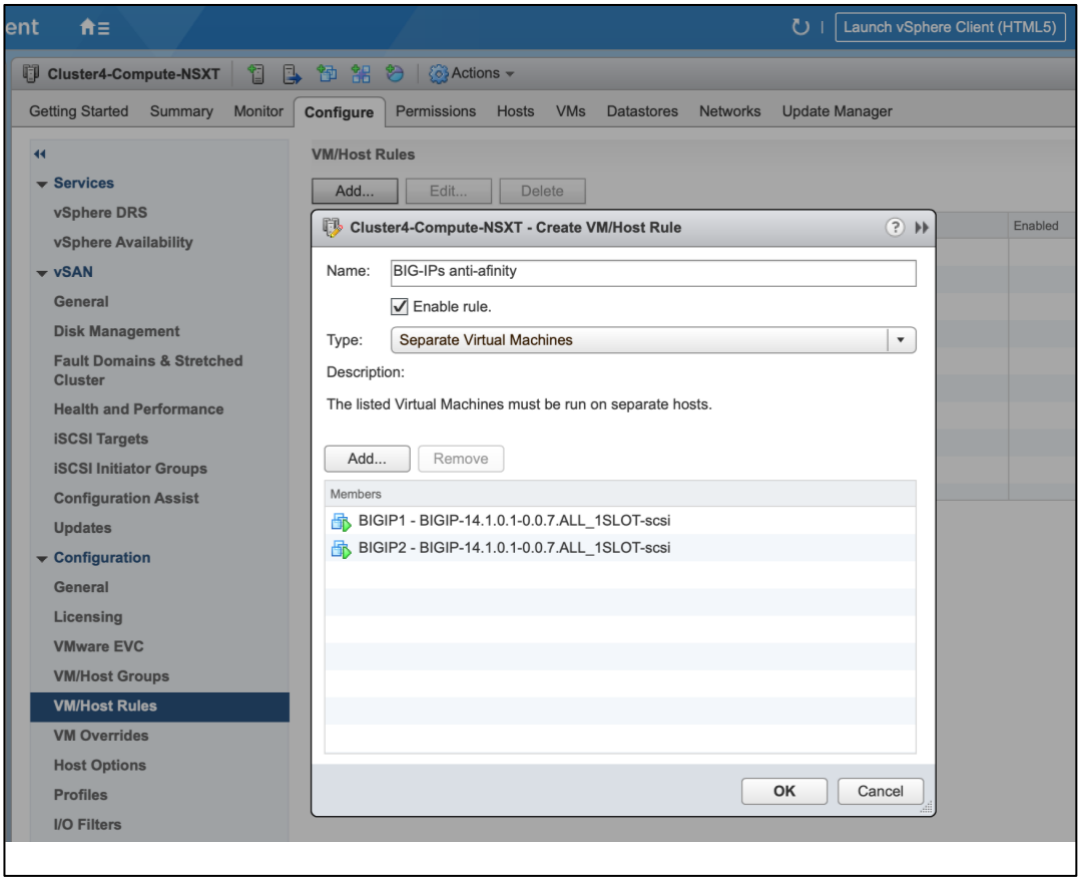


Figure 122 - Setting anti-affinity rules with VMware's Dynamic Resource Scheduler.

- 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 FAQ[10] and this community article[11] for more details. A screenshot of this configuration is shown next.

[10] https://cloud.vmware.com/vmc-aws/faq#stretched-clusters-for-vmware-cloud-on-aws
[11] https://blogs.vmware.com/cloud/2018/05/15/stretched-clusters-vmware-cloud-aws-overview/

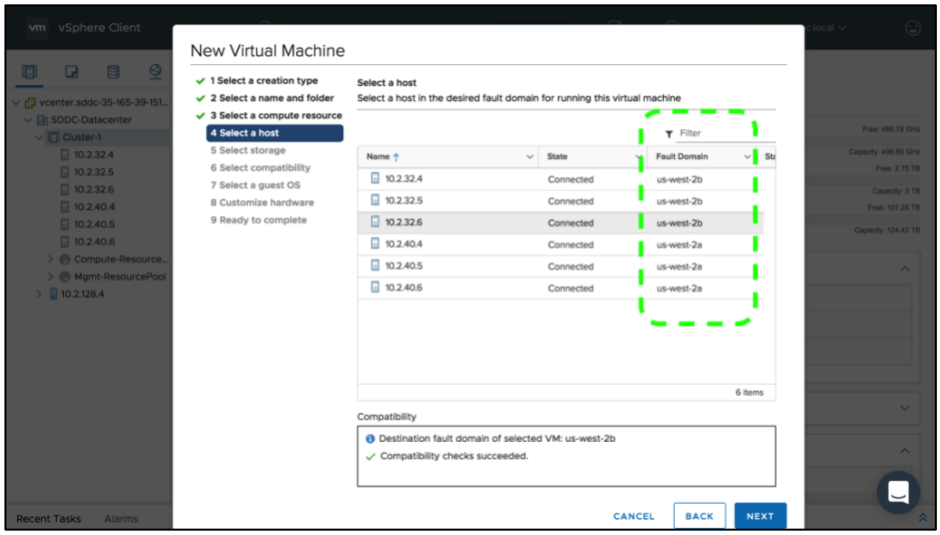


Figure 123 - Configuring High Availability of VMs in VMC stretched clusters.

- Failover and HA Groups

HA groups is a high availability feature that allows to specify a set of configuration objects such as pools that may be used to raise failover for redundant F5 BIG-IP systems.

It is recommended to use this feature to validate the reachability of the NSX-T Gateways. In a nutshell, a pool will be added to an HA Group where the pool members will be NSX-T Gateway's address(es). ICMP monitoring will be done on these pool members. This ICMP monitoring will induce a failover in the BIG-IPs if the Active unit is unable to reach the Gateways monitored (and the Standby is able to do it).

In this section, it will be discussed the key aspects to achieve the recommended configuration. For further details of HA Group feature, the following reading is recommended:

- Manual: BIG-IP System: Maintaining High Availability through Resource Monitoring
- K16947: F5 recommended practices for the HA group feature

The suggested monitoring tests the path up to the Edge nodes. Using ICMP probing to both T0 and T1, the ICMP packets will be processed by the associated T0 and T1 Service routers which are in the Edge nodes. Using NSX-T's traceflow it has been concluded the following guarantees these tests:

- When the F5 BIG-IPs are directly connected to a T0 Gateway, use as probing destination the T0's IP address facing the BIG-IP.

- When the F5 BIG-IPs are directly connected to a T1 Gateway, use as probing destination the T1's IP address facing the upstream T0 address. In this case it is recommended the use of a Gateway ICMP monitor in transparent mode as shown next.

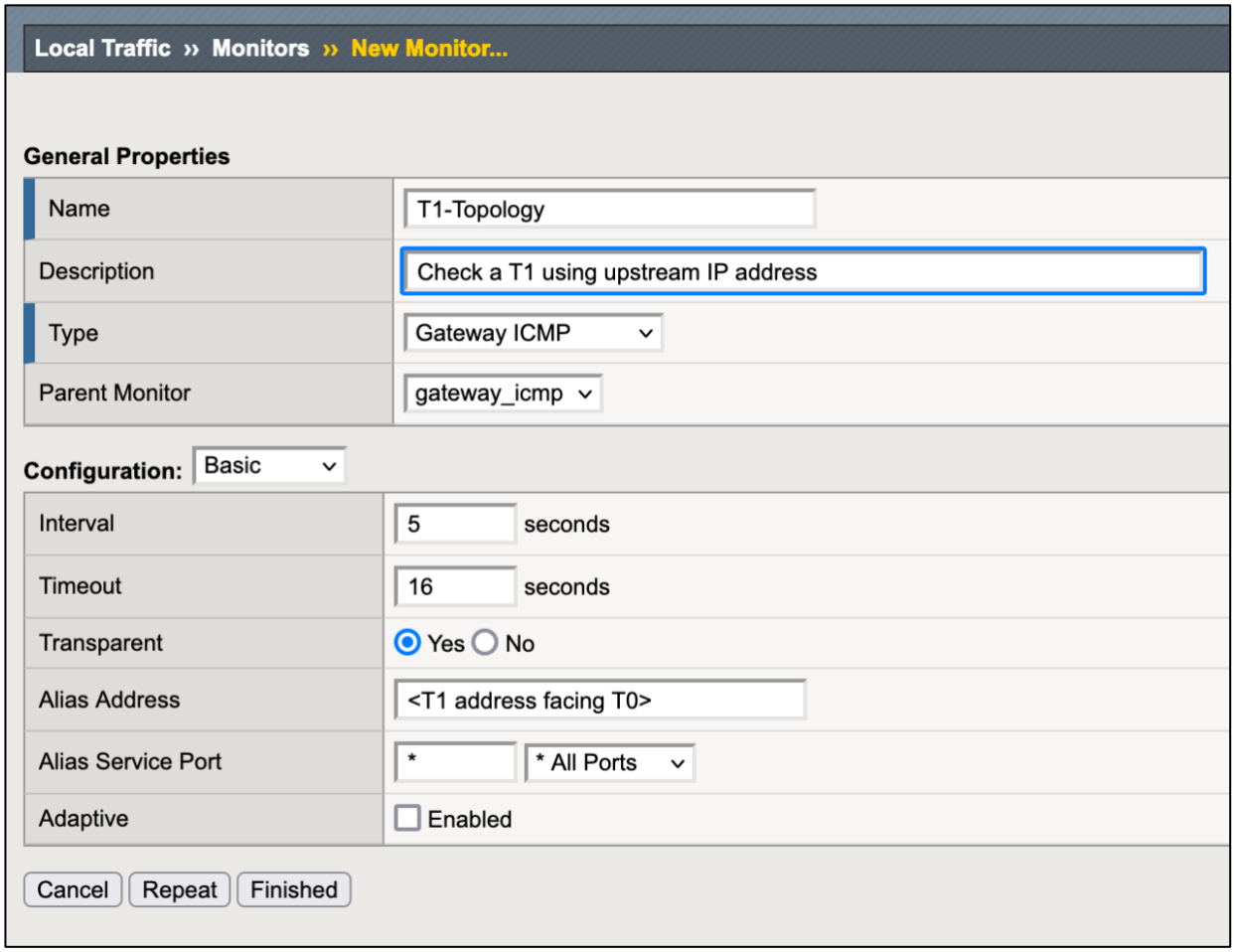


Figure 124 - pool monitor for the HA Group.

From the above figure please note:

- All settings are defaults except for the `Transparent` setting (set to enable), and the `Alias Address`.

- The `Alias Address` is the T1 IP address facing the T0.

For additional information in `Transparent` setting, please check Transparent and Reverse modes.

The next table summarizes the pool members to be configured for each topology:

Topology	HA Group pool members	Notes
A	Attached T0	These are the T0 external interfaces facing the F5 BIG-IP. This is a single VIP address or BGP peers IPs depending on if static or dynamic routing is being used respectively.
B	Attached T0	The T0's IP in the segment connected to the F5 BIG-IPs.
B extended	Attached T0	This is T0's IP address of the segment between the T0 and the F5 BIG-IP.
C	Attached T0	This is T0's IP address of the segment between the T0 and the F5 BIG-IP.

D	Attached T1	This is T1's IP address of the transit segment between the T0 and the T1. This makes use of transparent monitors to force the ICMP being routed through the T1 Gateway.

Please note that the pool and monitors are configured and synchronized across a cluster, but the HA configuration is performed individually on each BIG-IP.

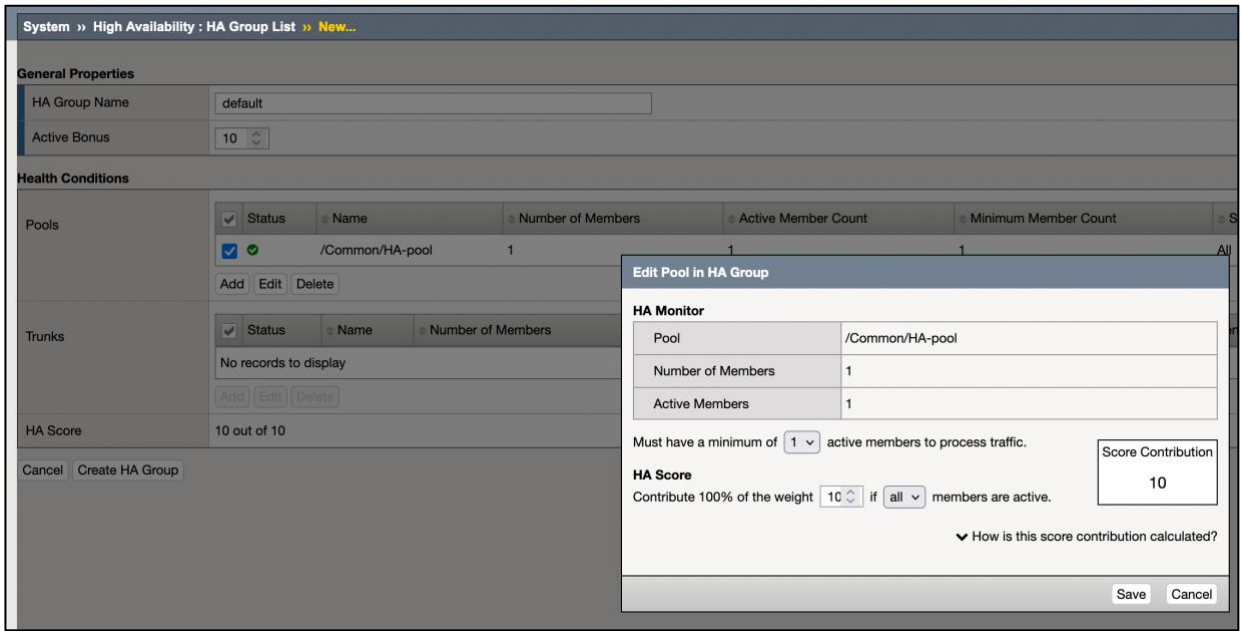


Figure 125 - Configuring HA Groups to monitor upstream connectivity up to the Edge nodes.

Performance best practices for BIG-IP in VMware NSX-T

A VM's performance depends on the hypervisor. For latency critical applications such as network functions, it is crucial that the VMs have the required resources allocated and the hypervisor doesn't pre-empt VM execution while running hypervisor related functions.

NSX-T's distributed functions are one example of hypervisor functionality that might pre-empt VM execution. These are executed in the VM's hypervisor running with higher priority than the VMs themselves, regardless of VM's configuration. This is the case of the Distributed Firewall which is worth special mention. A large number of Distributed Firewall Rules might impact VM´s performance without notice.

In this guide, it is recommended to:

- configure BIG-IP VE with the appropriate resources in the hypervisor.
- Monitor VM performance.
- If using the Distributed Firewall, place the BIG-IP VEs in the exception list.

These are detailed in the next sections.

- Hypervisor optimization for BIG-IP VE

  By default, VMs in VMware are configured to not have guaranteed resources, sharing resources so overall performance is achieved. This default policy doesn't apply well to network appliances such as BIG-IP VE because these require real-time performance. In these cases, it is recommended following VMware best practices extracted from

  - section "Running Network Latency Sensitive Workloads" in the guide "Performance Best Practices for VMware vSphere 7.0" and,

  - section "Summary of Recommendations" in the guide "Tuning vCloud NFV for Data Plane Intensive Workloads".

  In these guides BIOS, Hypervisor, and VM settings are suggested. In an IT environment It is usual that BIOS and Hypervisor general settings cannot be tunned because network appliance VMs usually run in the same Hypervisors as regular compute VMs. That is, in most of the cases only the VM-specific settings can be tuned.

  Resource allocation

  Next are the VM settings recommended for the BIG-IP VE from the mentioned VMware guides:

  - Set "Latency Sensitivity" to high, which will effectively disable hyperthreading for the VM. For effective use of this setting requires setting the next settings.
  - vCPU reservation: set the value to the Max value. This Max value calculated by VMware is the #vCPUs multiplied by processor base frequency.
  - Memory shares: set to high.
  - Memory reservation: set to the memory allocated to the VM.
  - Memory limit: set to unlimited.

  vSphere Network I/O Control

  In general, it is not recommended to use vSphere Network I/O Control as this might affect the overall BIG-IP VE performance.

  Storage I/O requirements

  Typically, F5 VEs have very low storage I/O requirements but care should be taken in the case of using the ASM module which has higher I/O demands. It is left to the infrastructure administrator that the BIG-IP VEs access the storage without contention.

- Monitor hypervisor's performance

When using the hypervisor optimizations recommended in the previous section, it is not expected that:

- BIG-IP VE CPUs will be pre-empted by other VMs.
- BIG-IP VE CPUs to share the physical core with other logical cores (HyperThreading).
- BIG-IP VE CPUs to run in power saving management modes other than C0.

Because of these, it is expected that % CPU Ready times are extremely low at all the times. Some VMware troubleshooting guides such as https://kb.vmware.com/s/article/2001003 suggest % CPU Ready times below 5% but this threshold is not acceptable to network

appliances. For the CPU cores directly used by the BIG-IP VE, values below 0% are expected and will guarantee stability.

Note that the aggregated % CPU Ready values of the VM's "Group Cpu" include all Hypervisor's threads associated to the VM which greatly increase the statistics. This aggregated statistic will vary greatly depending on the vNIC configuration[12] and overall hypervisor load.

If unexpectedly a BIG-IP VE ceases running, a number of problems might occur besides performance degradation. These are described in K22658205: Unexpected failover or other instability in BIG-IP VE due to CPU starvation by the hypervisor. Before these issues arise, it is common to see in the logs the error message `01010029:5: Clock advanced by <ticks> ticks` which is described in K10095: Error Message: Clock advanced by <number> ticks. It is recommended to monitor these messages. It has been observed that the Distributed Firewall might induce these problems. It is highly recommended to follow the best practice "Distributed Firewall: add BIG-IP VEs to the Exclusion List. described next.

- Distributed Firewall: add BIG-IP VEs to the Exclusion List.

Large number of NSX-T Distributed Firewall (DFW) rules can hurt VM performance and even stability[13]. For this reason, latency critical VMs such as NSX-T Edge nodes are by default configured in the NSX-T's DFW Exclusion List. In general, micro-segmentation is meant for end hosts and not for gateways such as NSX-T Edge or BIG-IP VE. Because of these, it is recommended that BIG-IP VEs are also placed in the Exclusion List.

To configure the BIG-IP VEs in the Exclusion List, first an NSX-T tag has to be created. This NSX-T tag is created in the Inventory > Tags section of the UI shown next.

[12] Specially the ethernetX.ctxPerDev setting.
[13] Please see the section "Monitor hypervisor's performance" for more information.

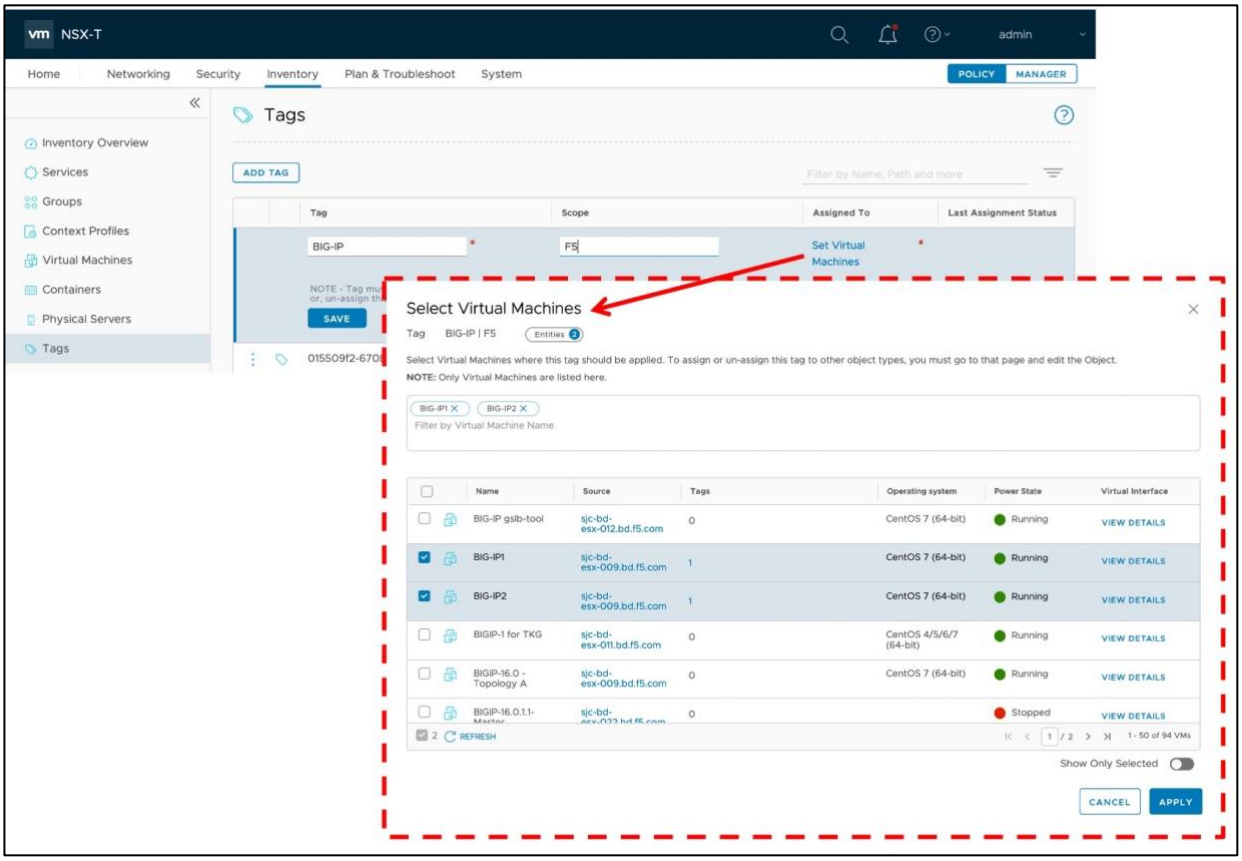


Figure 126 - Creating a NSX-T tag for F5 VEs.

In the example above, a tag with the name “BIG-IP” is created with the (optional) scope “F5” set. The scope might be useful to have additional tags, all within the F5 scope, and differentiate groups of BIG-IPs or BIG-IQs.

Once the NSX-T tag is created, this can be used it in the Exclusion List. Creating the Exclusion List is found in the Security > Distributed Firewall (East-West security) section of the UI. Once in the Distributed Firewall screen, access the Exclusion List from Actions menu as shown next.

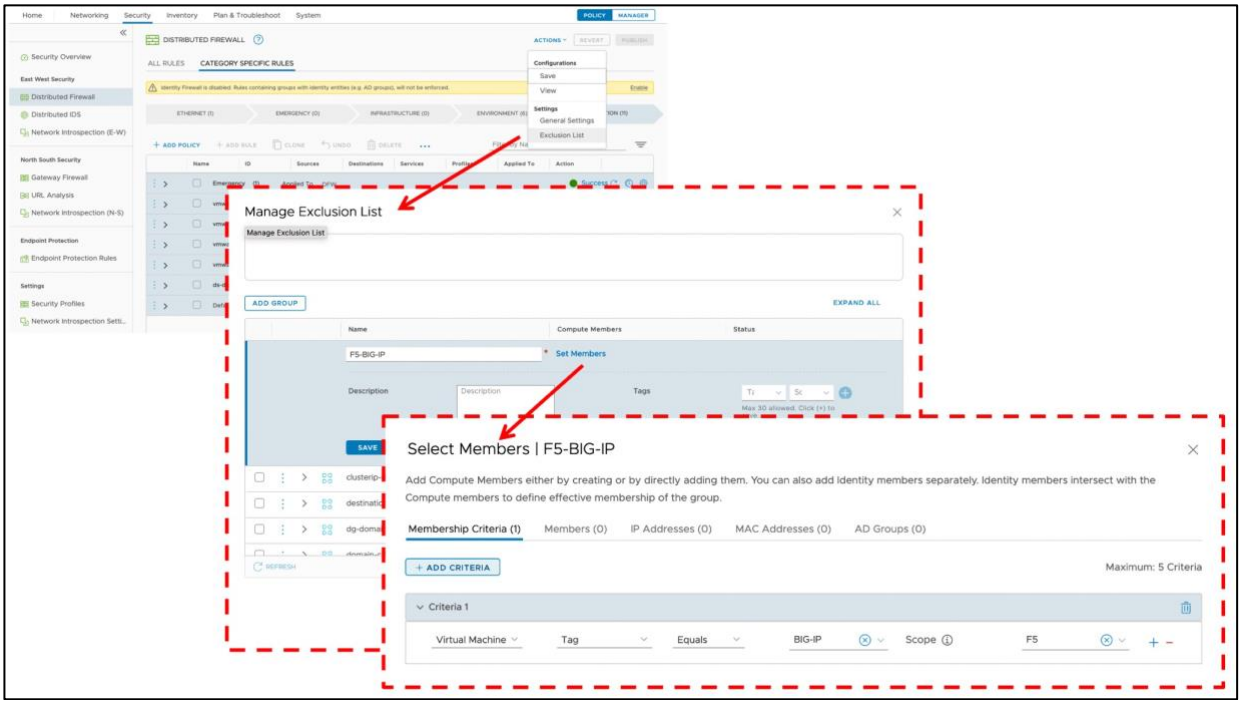


Figure 127 - Creating an entry for F5 VEs in the exclusion list.

In the Manage Exclusion List dialog, create a new group (in the example “F5 BIG-IP") and a membership of type Virtual Machine. The “BIG-IP” tag should be found in the tag drop-down, and the scope automatically filled. Once completed, click “Save”.

With this, the configuration of the BIG-IP VE in the Exclusion List is completed. This can be verified by using the “View Members” option in the “Manage Exclusion List dialog as shown next.

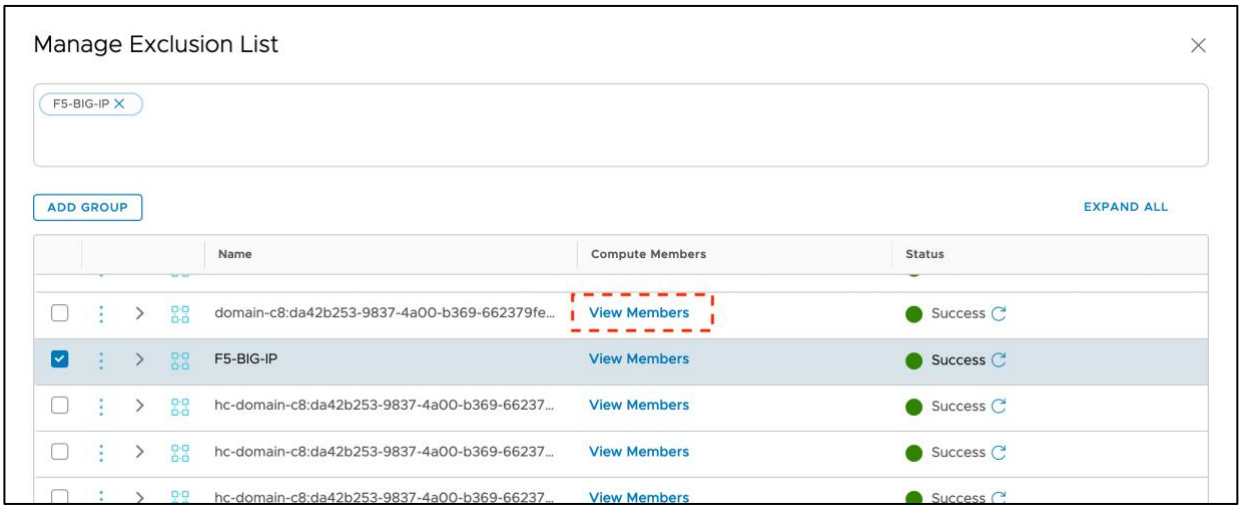


Figure 128 - Verifying the Exclusion List.

If specific security rules were previously applied to the F5 VE using the NSX-T Distributed Firewall, these can be replaced by BIG-IP security own mechanisms. Consider the following documents as a guideline:

Data plane related:

- K17333: Overview of port lockdown behaviour (12.x - 17.x)

- K42075438: Restricting access to a virtual server by IP subnet

Management plane related:

- K13092: Overview of securing access to the BIG-IP system

- K5380: Specify allowable IP ranges for SSH access

- K13309: Restricting access to the Configuration utility by source IP address

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)[14].

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 must be applied to all the segments of the traffic group where MAC masquerading is going to be used.

[14] https://support.f5.com/csp/article/K13502

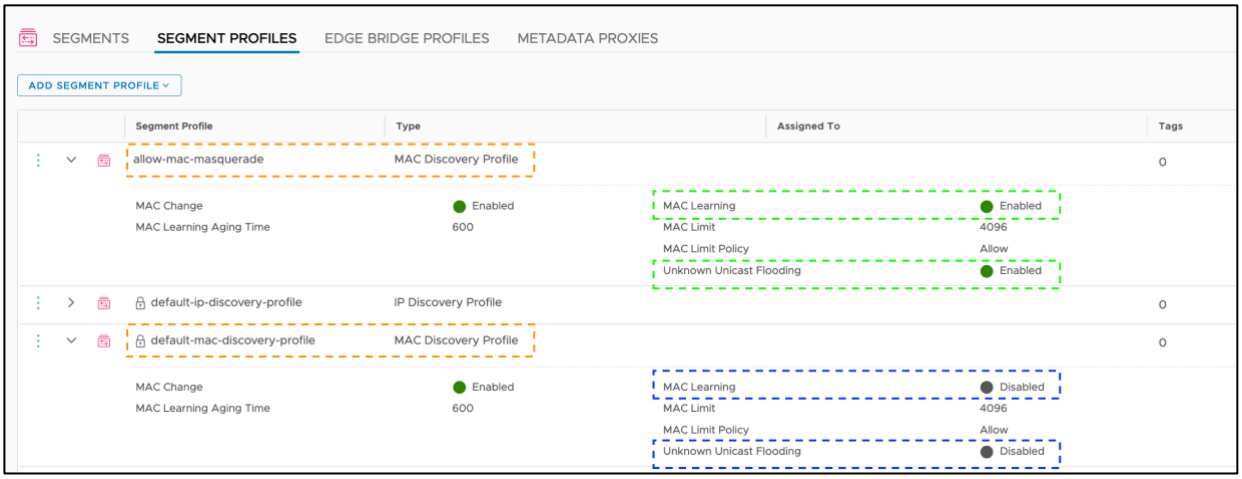


Figure 129 - Creating a new MAC Discovery Profile for MAC Masquerade.

VMC on AWS

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

Considerations for Tanzu Kubernetes Grid (TKG)

Overall architecture

When deploying a TKG workload cluster, a dedicated T1 Logical Router for the cluster is created by the installer (either TKGm or TKGs). Overall, the topology used for TKG clusters is Topology D (parallel). The general recommendation is to have a separate BIG-IP cluster for each TKG workload cluster. This allows for strict isolation between workloads and higher scalability. This 1:1 BIG-IP to TKG cluster architecture is shown next.

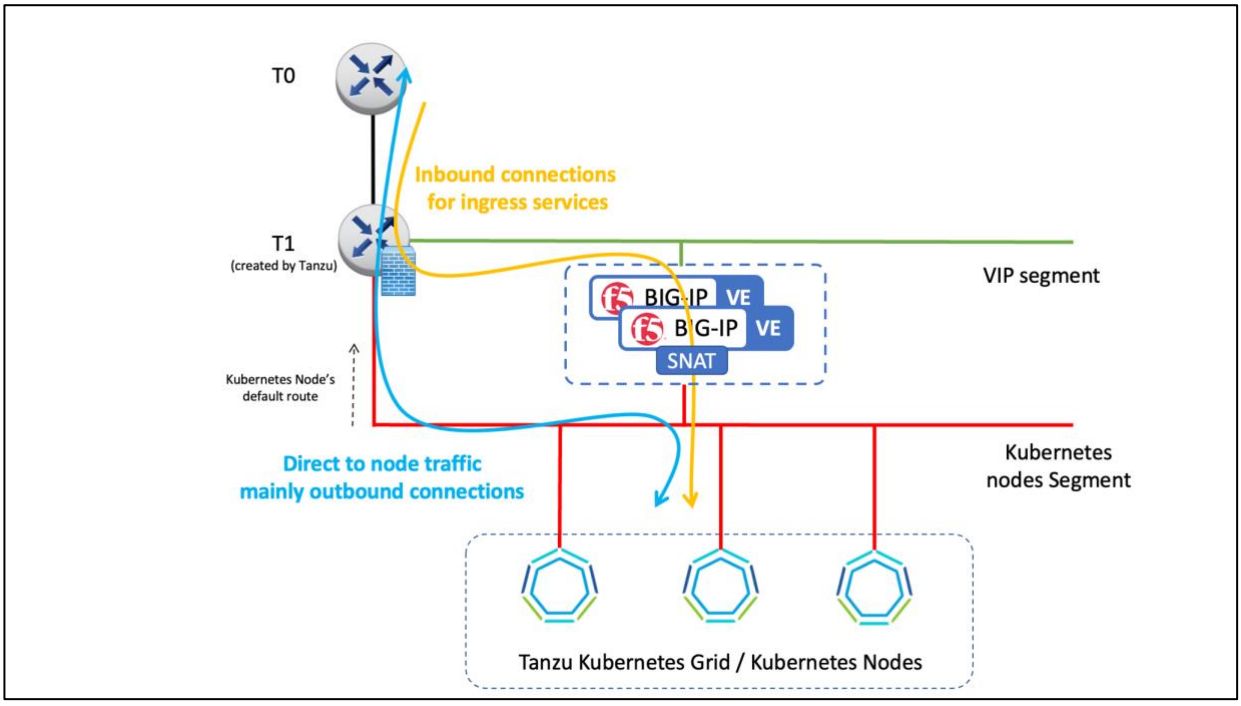


Figure 130 - A BIG-IP cluster per TKG workload cluster.

It can be observed that the return traffic is forced to go through the BIG-IP by means of using SNAT. The path of direct traffic to the nodes is unchanged.

Sometimes it is not required to have strict isolation or high scalability and a single BIG-IP can handle different TKG clusters in separate BIG-IP logical partitions. This is shown in the next figure.

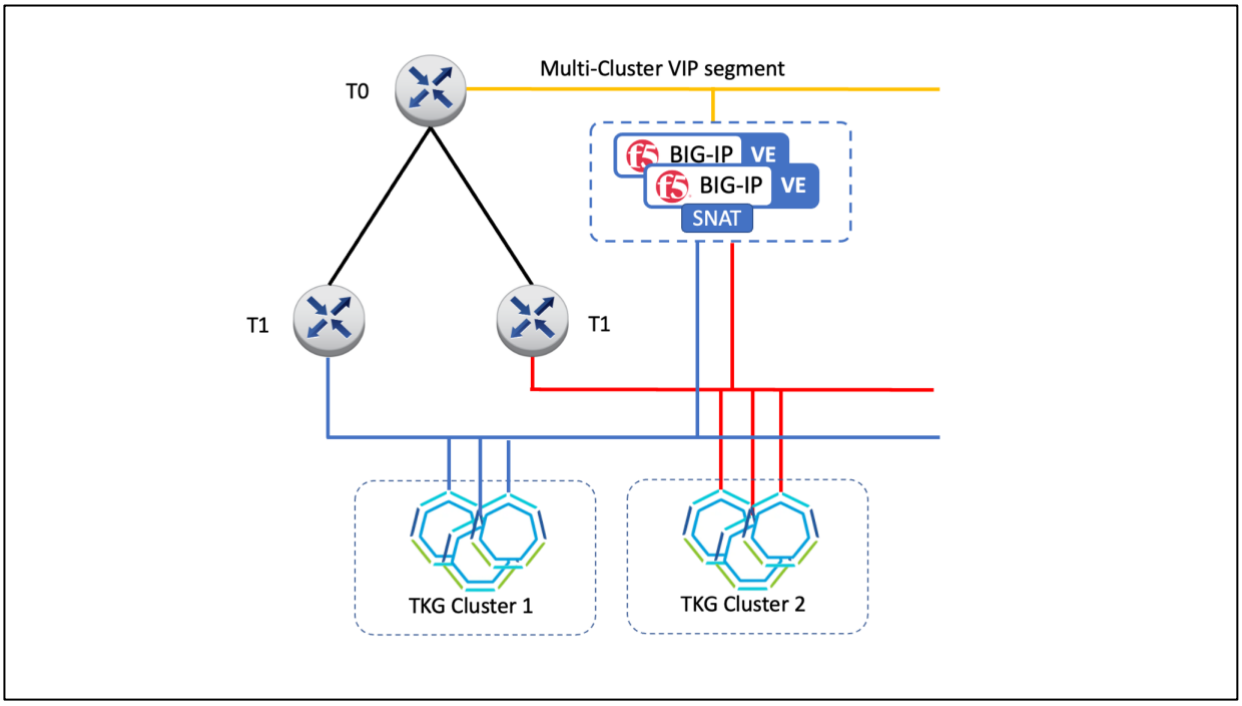


Figure 131 - A BIG-IP cluster for multiple TKG workload clusters.

It is considered that the use of separate routing domains for the different clusters is not needed in this setup because there are no forwarding virtual servers in these setups. If necessary, several VIP segments could be created. It should be considered ESXi imposes a limit of 10 vNICs per VM and hence this will limit the number of TKG clusters a BIG-IP can handle.

In either of these configurations the F5 BIG-IP can be used in a one-tier or a two-tier LB arrangement.

- In a one-tier LB arrangement, the BIG-IP performs the functions of both external Load Balancer and Ingress Controller. This reduces latency and troubleshooting is easier. BIG-IP Container Ingress Services (CIS) allows to expose most BIG-IP functionalities through the Kubernetes API.

- In a two-tier LB arrangement, the BIG-IP performs the function of external Load Balancer and distributes the requests amongst Ingress Controllers inside the TKG cluster. Typically, this Ingress Controller is NGINX plus which can implement Web Application and API protection functionalities. This arrangement allows for greater scalability and role separation where BIG-IPs are managed by the NetOps team and the Ingress controllers are managed by the DevOps teams.

CNI selection

At time of this writing (TKG 1.5) the two options offered by TKG are supported by BIG-IP's Container Ingress Services (CIS):

- Antrea – using the NodePortLocal feature allows an external load balancer to address the PODs individually. Antrea introduces an additional Geneve overlay with the additional overhead this imposes. On the other hand, Antrea is integrated with NSX-T 3.2 allowing POD traceflow and visibility of cluster security in NSX-T's UI.

- Calico – this doesn't create an additional overlay and POD addresses are exchanged by means of using BGP. In this scenario, the BIG-IPs are configured as additional Calico nodes.

For the must up to date information on how to configure CIS for these please check https://clouddocs.f5.com/containers/latest/

Allocating IP addresses for BIG-IPs in TKG's segment

When using either Antrea or Calico, the BIG-IP uses IP addresses in the same segment as the TKG nodes. These will be Self IPs, floating Self IPs and SNAT addresses. We have to reserve these IPs for the BIG-IPs so the TKG cluster doesn't make use of them at any time, typically at cluster scale-out.

The allocation of these addresses must be done within the IP pool pre-created by the TKG installer. The NSX-T API is used for this.

We will need to first find out the name of the IP pool. First login in the NSX-T manager > Networking > IP Address Pools where we will find the subnet allocated the cluster, type the namespace (in the case of TKGs) and cluster names to facilitate the search. In this screen we can obtain the API path to operate with the IP pool. This is shown in the upper part of the next figure.

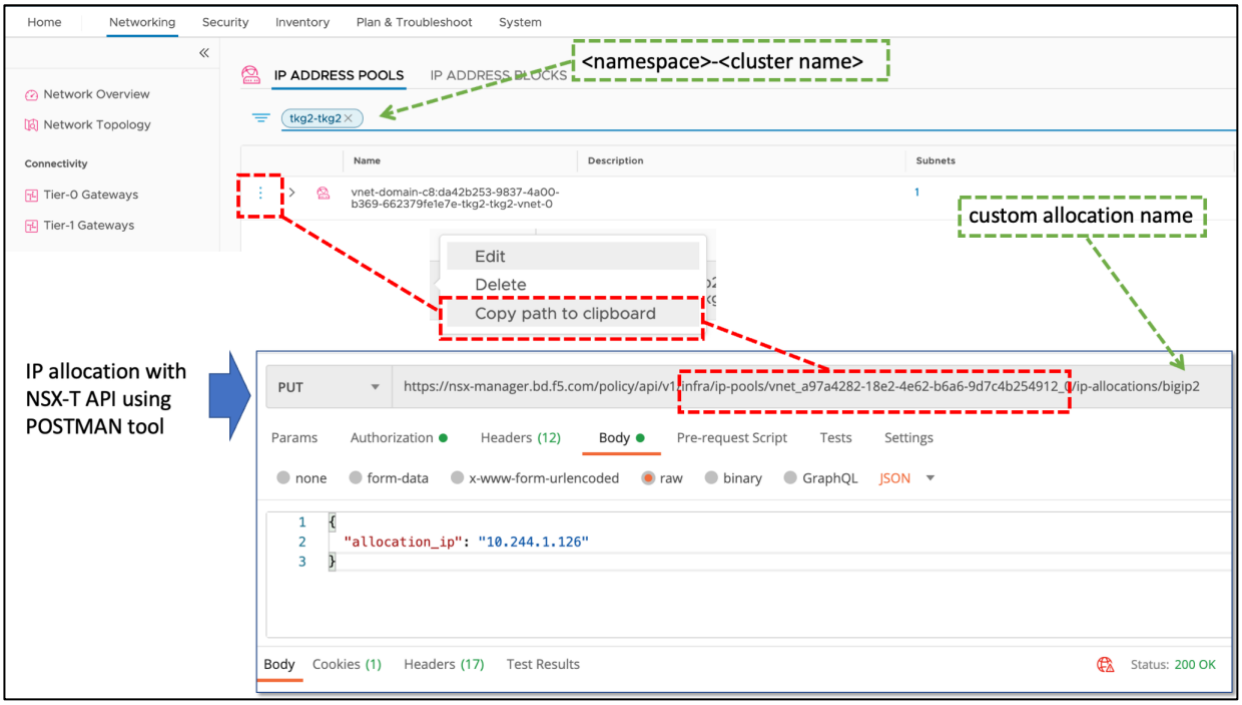


Figure 132 - Allocating IP addresses for the BIG-IPs.

The API path that we have obtained can be used in tools such as POSTMAN to make the IP address allocation with an API request. This is shown in the lower part of the previous figure.

More precisely, it is necessary to perform PUT policy/api/v1/<API path>/ip-allocations/<name of allocation> request indicating in the body the desired IP to allocate. The names of the IP allocations are not relevant. In this link at code.vmware.com you can find the full details of this API call and others to perform these allocations.

Deleting a TKG cluster

Before deleting a TKG cluster, the following actions should be done:

- Detaching of the ports used by the BIG-IP VEs in the TKG segment. Deleting of the BIG-IP VE Deletion should be equivalent, but the units should be unlicensed beforehand.

- If using the single tenant BIG-IP deployment type, the VIP segment should be deleted as well.

Otherwise, the TKG uninstall will fail because it cannot delete the associated T1 LR which has the above resources used for the BIG-IP.

BGP configuration details peering with NSX-T Edge nodes

The following configuration settings follow VMware Validated Design 5.0.1[15], 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.

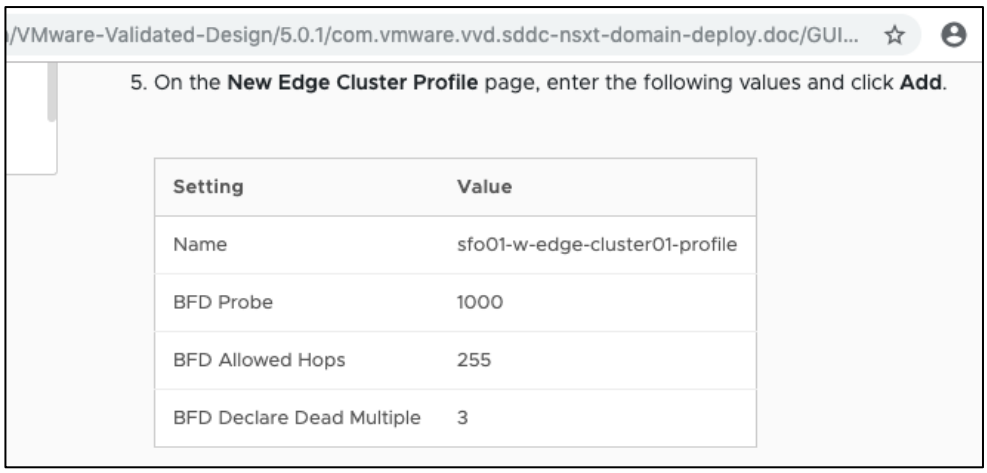


Figure 133 - VMware's baseline settings for BFD

[15] https://docs.vmware.com/en/VMware-Validated-Design/index.html

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.

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 and B extended	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) Southbouind – 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 and B extended	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 – 10106.54.10
Topology B and B extended	10.106.{51,52}.1	10.106.{51,52}.10
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):

```
bigip1a.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 134 - 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:

```
bigip1a.nsxt.bd.f5.com[0]#show bgp neighbors | include BGP state
  BGP state = Established, up for 2d19h50m
  BGP state = Established, up for 2d19h50m
```

Figure 135 - Verifying that the BGP sessions are Up.

As you can see in the previous figure, 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).

```
bigip1a.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                              32768 ?
*  10.106.32.0/24   10.106.54.1              0                    0 65001 ?
*>                  10.106.53.1              0                    0 65001 ?
*  10.106.33.0/24   10.106.54.1              0                    0 65001 ?
*>                  10.106.53.1              0                    0 65001 ?
*  10.106.34.0/24   10.106.54.1              0                    0 65001 ?
*>                  10.106.53.1              0                    0 65001 ?

Total number of prefixes 4
```

Figure 136 - 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@(bigip1a)(cfg-sync In Sync)(Active)(/Common)(tmos)# 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 137 - 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, where the parameters are in red are mandatory.

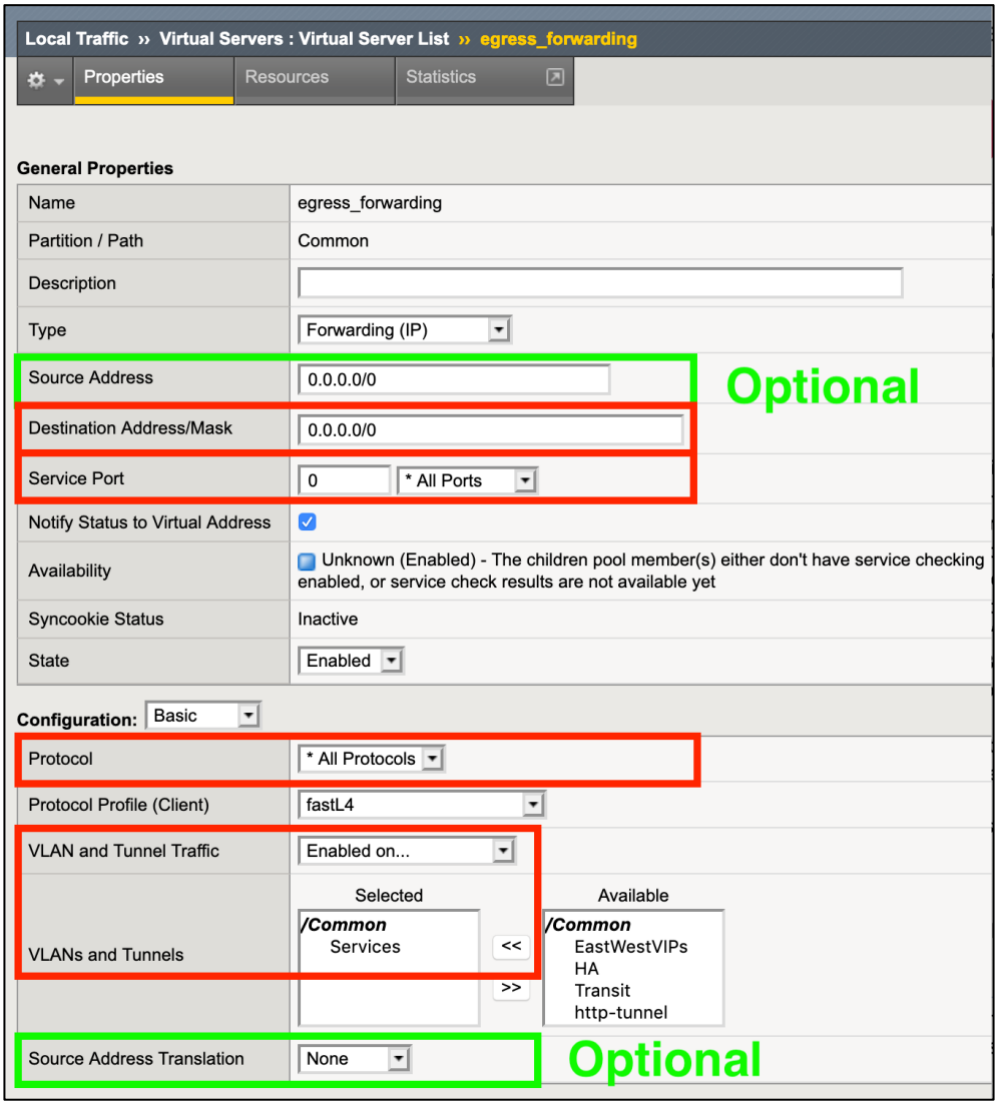


Figure 138 - Creating a Forwarding Virtual Server for testing egress traffic.

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.

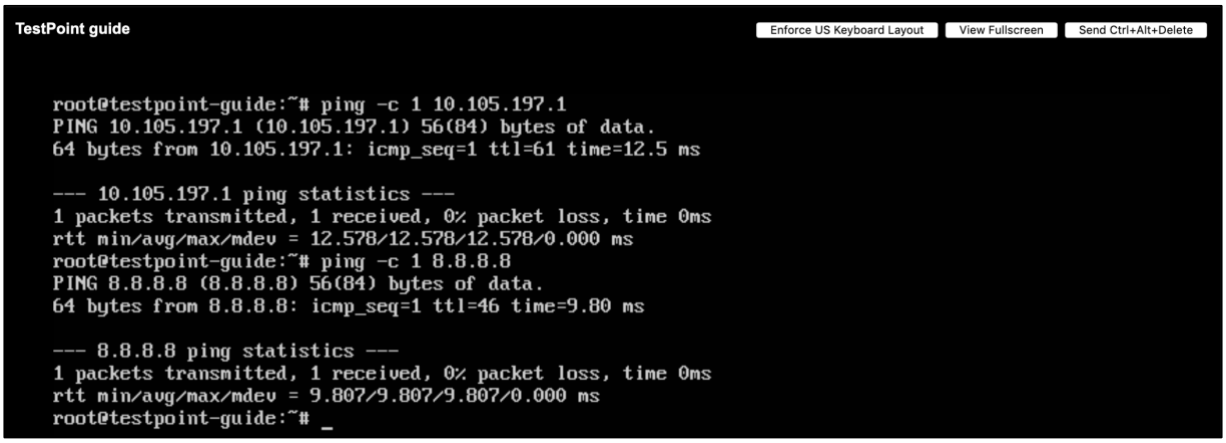


Figure 139 - Ping test using spine router's IP address and the well-known Internet address 8.8.8.8 for checking egress connectivity.

	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 and B extended	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 -nel -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 don'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 and B extended	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.

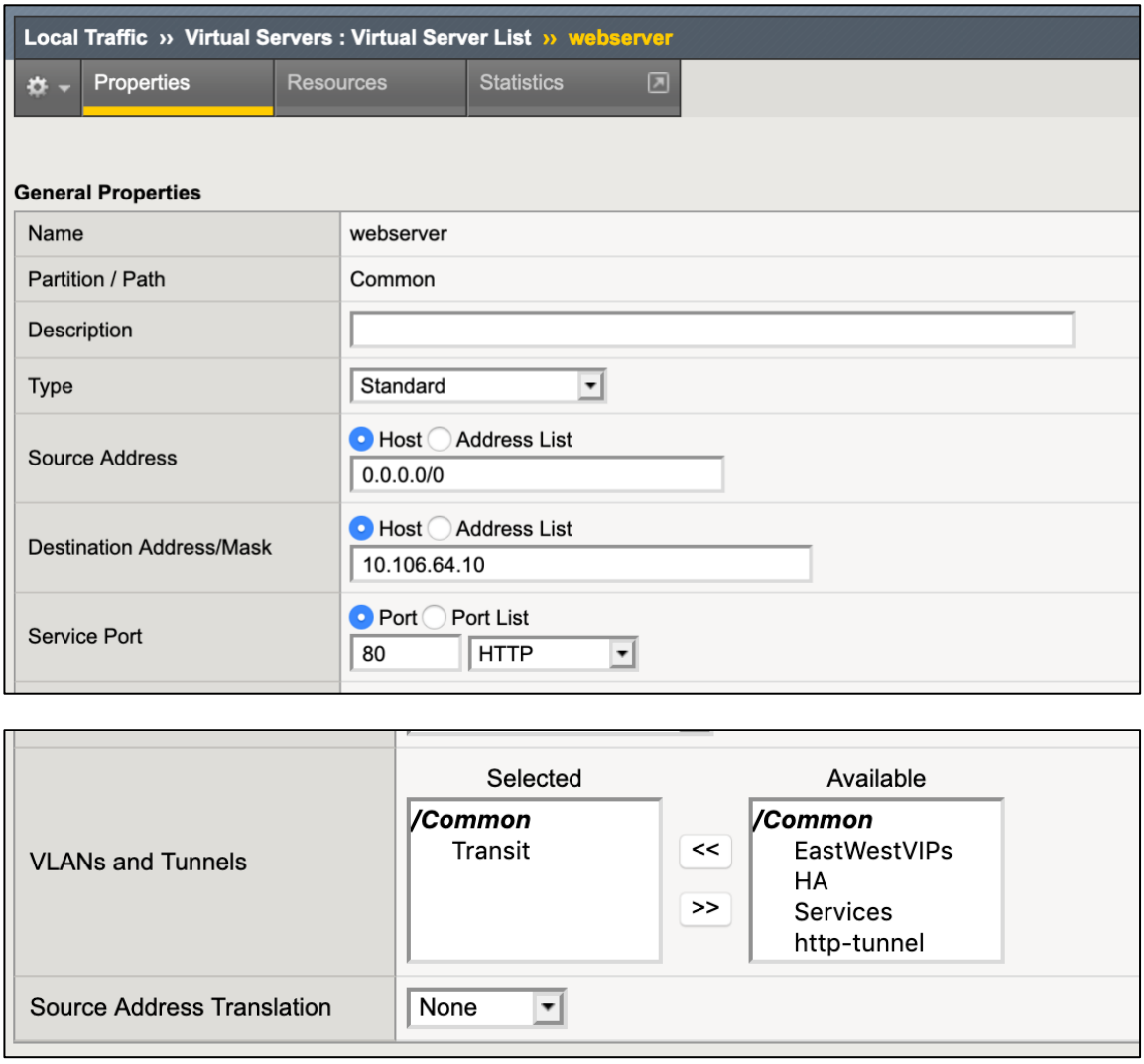


Figure 140 - Creating a Standard Virtual Server for testing Ingress services' connectivity.

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:

Figure 141 – 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.

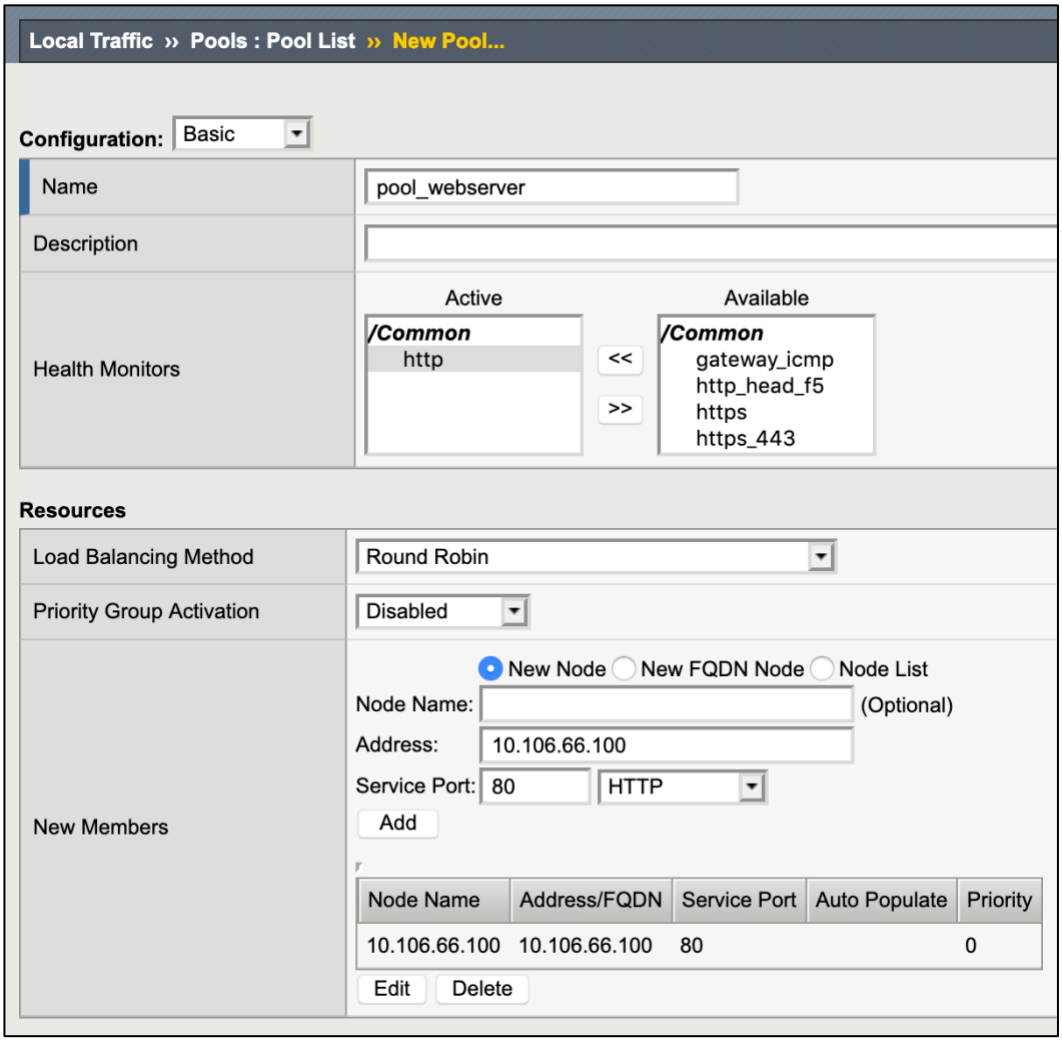


Figure 142 - Specifying pool member details for the test Ingress Virtual Server.

This pool health monitor already tests the connectivity form 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.

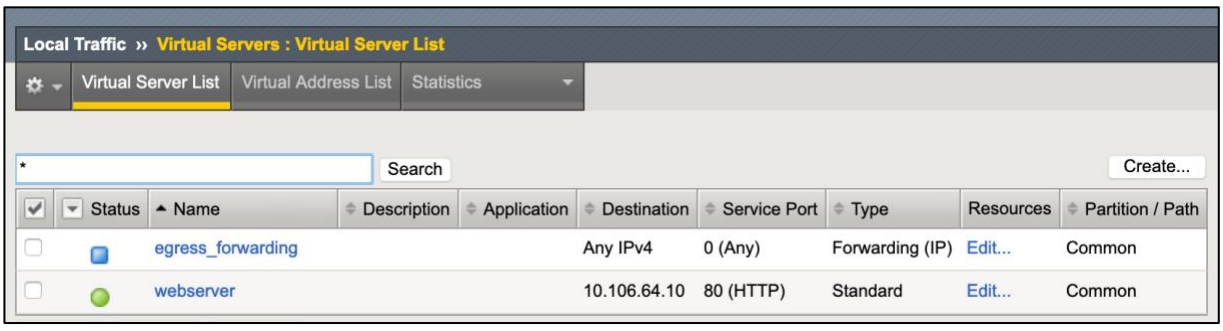


Figure 143 - virtual server status after creating the webserver VS for Ingress traffic.

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.