## Version History

<table>
<thead>
<tr>
<th>Date</th>
<th>Version</th>
<th>Author</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>December 2019</td>
<td>1.0</td>
<td>Ulises Alonso Camaró, F5 Networks</td>
<td>Initial public version.</td>
</tr>
</tbody>
</table>
| June 2020    | 2.0     | Ulises Alonso Camaró, F5 Networks| 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. |
INTRODUCTION ................................................................................................................................................. 5

NSX-T VERSIONS CONSIDERED IN THIS GUIDE ........................................................................................................ 6

BIG-IP VERSIONS CONSIDERED IN THIS GUIDE ........................................................................................................ 6

DEPLOYMENT GUIDE OVERVIEW .................................................................................................................................. 6

INLINE TOPOLOGIES OVERVIEW .................................................................................................................................. 7

PARALLEL TOPOLOGIES OVERVIEW ........................................................................................................................... 8

TOPOLOGIES’ MAIN CHARACTERISTICS SUMMARY .................................................................................................... 9

ADDITIONAL DEPLOYMENT POSSIBILITIES ............................................................................................................... 10

DESIGN CONSIDERATIONS ........................................................................................................................................... 11

Design consideration: Layer 2 networking ...................................................................................................................... 11

Design consideration: NAT .................................................................................................................................................. 11

Design consideration: Use of dynamic routing (BGP) with upstream networks ......................................................... 11

Design considerations: NSX-T’s distributed functions ................................................................................................ 12

Design consideration: Centralized management .............................................................................................................. 12

TOPOLOGY A: BIG-IPS INLINE-CONNECTED TO NSX-T’S TIER-0 GATEWAY .......................................................... 14

Implementation: Active/Standby NSX-T Edge with static routing .............................................................................. 16

Implementation: Active/Standby NSX-T Edge using BGP .......................................................................................... 24

Implementation: Active/Active NSX-T Edge using BGP ECMP ................................................................................ 36

TOPOLOGY B: BIG-IPS INLINE – CONNECTED LIKE AN NSX-T’S TIER-1 GATEWAY .......................................................... 48

Implementation: BIG-IPS inline-connected like an NSX-T’s Tier-1 Gateway ............................................................. 49

TOPOLOGY C: BIG-IPS PARALLEL-CONNECTED TO NSX-T’S TIER-0 GATEWAY .......................................................... 58

Implementation: BIG-IPS parallel-connected to NSX-T’s Tier-0 Gateway .............................................................. 60

TOPOLOGY D: BIG-IPS PARALLEL-CONNECTED TO NSX-T’S TIER-1 GATEWAY .......................................................... 70

Implementation: BIG-IPS parallel-connected to NSX-T’s Tier-1 Gateway .............................................................. 75
GENERAL NOTES..................................................................................................................................................79

BGP configuration details ......................................................................................................................................79

Best practices for BIG-IP in VMware NSX-T ........................................................................................................80

Considerations for container platforms ..............................................................................................................82
   General guidelines ...........................................................................................................................................82
   Exposing container services ...........................................................................................................................82
   Relevant configuration settings when using Red Hat OpenShift .................................................................84
   Relevant configuration settings when using Pivotal PKS .............................................................................84

VERIFYING THE DEPLOYMENT..........................................................................................................................86

Basic testing ......................................................................................................................................................86

Dynamic routing testing ......................................................................................................................................87

End to End testing: test egress forwarding connectivity through the BIG-IP ..................................................89

End to End testing: test egress forwarding connectivity without the BIG-IP ..................................................91

End to End testing: test Ingress connectivity through the BIG-IP ..................................................................91
Introduction

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

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

- For North-South traffic flows this is not an issue because the number of networks to which the F5 BIG-IP has to be connected is small and is not expected to change often.

- For East-West traffic this inhibits the possibility of using F5 BIG-IP hardware. Also, the number of network segments to which the F5 BIG-IP is expected to be connected for this use case is very high, but the VMware hypervisor only allows the VMs to be connected with up to 10 vNICs with one network segment per vNIC. In this guide this VMware limitation is overcome by creating multiple clusters of BIG-IPs. This allows higher distribution of the traffic and CPU utilization across the VMware cluster.

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

- F5 BIG-IPs can be connected to either Tier-0 (internally or externally) and to Tier-1 distributed routers while NSX-T’s load balancer can only be placed logically connected to Tier-1 Gateways.

- NSX-T’s load balancer is not a distributed function and runs centralized on NSX-T Edge’s nodes, which can represent a bottleneck. F5 BIG-IP can run in multiple hypervisors concurrently by either running Active-Active F5 Scale-N clusters or multiple F5 BIG-IP clusters.

- F5 BIG-IP provides proven, scalable, and world-class performance for ADC, NAT and Firewall capabilities, and provides additional functionalities such as Advanced WAF, SSL-VPN, Anti-DDoS protection, Secure Web Gateway with Identity Management and many other solutions with a unified management & visibility with F5 BIG-IQ.

\[\text{1}\] 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”.

---

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 2.4-3.0 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](https://clouddocs.f5.com) in clouddocs.f5.com for the most up to date information. When using BIG-IP Hardware platforms any BIG-IP version is supported.

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

Deployment guide overview

The document includes the 4 most common topologies:

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

- **Parallel topologies (these require SNAT):**
  - Topology C: BIG-IPs parallel-connected to NSX-T’s Tier-0 Gateway.
  - Topology D: BIG-IPs parallel-connected to NSX-T’s Tier-1 Gateway.

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

---

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

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

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

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

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

  For this topology three possible configurations are described:

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

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

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

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

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

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

- **Topology C – BIG-IPs parallel-connected to NSX-T’s Tier-0 Gateway.**
  
  Like Topology A which is also connected to a Tier-0 Gateway, this topology allows the use of either BIG-IP hardware or Virtual Editions. Other than the requirement of using SNAT, the main difference from Topology A is that each tenant can have their own BIG-IP instances with complete isolation. This is can be achieved either using BIG-IP hardware instantiating vCMP guests or using F5 BIG-IP Virtual Edition instances for each tenant.

- **Topology D – BIG-IPs parallel-connected to NSX-T’s Tier-1 Gateway.**
  
  This topology is similar to Topology C but with the BIG-IPs attached to the Tier-1 routers and would allow that Edge services could be applied at the NSX-T boundary for all traffic
flows without any traffic bypassing these Edge services. This is equivalent to the topology used by NSX-T Load Balancers.

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

Topologies’ main characteristics summary

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

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

<table>
<thead>
<tr>
<th>Topology</th>
<th>Type</th>
<th>Tier</th>
<th>VE</th>
<th>HW</th>
<th>Keeps client address</th>
<th>Inter-tenant distributed forwarding path</th>
<th>Enforcement point</th>
<th>Allows per-tenant VE clusters</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Inline</td>
<td>Tier-0</td>
<td>√</td>
<td>√</td>
<td>√</td>
<td>Not applicable</td>
<td>√</td>
<td>(for all tenants)</td>
</tr>
<tr>
<td>B</td>
<td>Inline</td>
<td>Tier-1</td>
<td>√</td>
<td></td>
<td>√</td>
<td></td>
<td>√</td>
<td>(per tenant)</td>
</tr>
<tr>
<td>C</td>
<td>Parallel</td>
<td>Tier-0</td>
<td>√</td>
<td>√</td>
<td></td>
<td>Not applicable</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Parallel</td>
<td>Tier-1</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
<td></td>
</tr>
</tbody>
</table>

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

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

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

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

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

**Keeps source address:** Ingress traffic doesn’t need to translate the source IP address of the clients. This avoids the need of using the X-Forwarded-For HTTP header.
**Inter-tenant distributed forwarding path:** when using plain routing between tenant workloads the processing path is fully distributed by only using NSX-T’s networking. In other words, this scenario is a path between Tier-1 workload to another Tier-1 workload and not using BIG-IP services. Note that when using NSX-T’s native LB the processing is done centralized in the NSX-T Edge nodes.

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

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

<table>
<thead>
<tr>
<th>Topology</th>
<th>Suitable for North-South</th>
<th>Suitable for intra-tenant East-West</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>✓</td>
<td>NA</td>
</tr>
<tr>
<td>B</td>
<td>✓</td>
<td>✓ (BIG-IP is a tenant enforcement point)</td>
</tr>
<tr>
<td>C</td>
<td>✓ (for non-routing traffic)</td>
<td>NA</td>
</tr>
<tr>
<td>D</td>
<td>✓ (for non-routing traffic)</td>
<td>✓ (for non-routing traffic)</td>
</tr>
</tbody>
</table>

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

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

### Additional deployment possibilities

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

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

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

Design consideration: Layer 2 networking

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

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

Design consideration: NAT

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

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

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

LACP - Link Aggregation Control Protocol is an IEEE standard.
Reflexive NAT - https://docs.vmware.com/en/VMware-NSX-T-Data-Center/2.5/administration/GUID-46900DFB-58EE-4E84-9873-357D91EFC854.html
NSX-T Edge’s Tier-0 routers exchange routes with upstream devices by means of eBGP. It is recommended the use of dynamic routing in the following use cases:

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

Design considerations: NSX-T’s distributed functions

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

Design consideration: Centralized management

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

- Centralized Management including self-service app-owner portal, application templates with security policies.
- Per-app analytics, performance metrics and dashboards.
- Multi-cloud capable and enabler for centralized CI/CD integrations.
- Fine grained RBAC where demarcation between the network, security, and app teams can be well established with their own specific views of a deployment.
Topology A: BIG-IPs inline-connected to NSX-T’s Tier-0 Gateway.

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

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

**Figure 4 - L1/L2 options for Topology A.**
Implementation: Active/Standby NSX-T Edge with static routing

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

![Diagram of Active/Standby NSX-T Edge with static routing]

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

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

1. Create the Tier-0 configuration.

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

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

   ![Figure 6 - Adding a Tier-0 Gateway/Gateway.](image)
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).

![Tier-0 Gateways](image)

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

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

Select the router created (T0-Topology-A in our example) and create two interfaces in the UI by first selecting the Edit option in the T0 Gateway, then scrolling down to the
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 8 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.

Figure 9 – Final Gateway Port configuration of the Tier-0 Gateway.
1.3. Create an HA VIP for the Tier-0 Gateway.

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

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

![Figure 10 - Adding an HA VIP to NSX-T's T0 Gateway.](image)

Selecting the two external interfaces just created.

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

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

![Figure 11 – Adding Tier-0 Gateway’s default route.](image)

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.**

![Tier-1 Gateways](image)

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

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

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

This configuration can be seen in the next figure:
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 13 - Adding a segment to the T1 Gateway.**

**Figure 14 – 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 15 – Static routes required in the BIG-IP units.

At this point, follow the testing steps described in the Verifying the deployment section.
Implementation: Active/Standby NSX-T Edge using BGP

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

---

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

1. Create the Tier-0 configuration.

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

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

   ![Figure 17 - Adding a Tier-0 Gateway.](image)
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).

![Tier-0 Gateways](image)

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

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

Select the router created (T0-Topology-A in our example) and create two interfaces in the UI by first selecting the Edit option in the T0 Gateway, then scrolling down to the
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 19 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.](image)

![Figure 20 – Final Gateway Port configuration of the Tier-0 Gateway.](image)
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.

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.

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

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

The remaining step is to redistribute the NSX-T routes into NSX-T’s BGP which then will be announced to the BGP peers (in this case the F5 BIG-IPs). This is done at Tier-0 Gateway level in the section shown in the next figure.
Create a redistribution entry which includes NSX connected networks as it can be seen in the next figure.

Figure 24 - Route redistribution settings at T0 Gateway
2. Create a Tier-1 Gateway.

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

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

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

![Tier-1 Gateways](image)

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

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

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

This configuration can be seen in the next figure:
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.

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

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

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

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

In order to enable BGP and routing protocols. Use the BIG-IPs UI and browse through Network > Route Domains > 0 (assuming that the default routing domain is the one being used). In this window enable BFD and BGP as seen in the next figure. Note that given this is
part of F5 BIG-IP’s base config it is not synchronized and must be done in all the F5 BIG-IP units.

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

The next step is to configure BFD and BGP itself. Log in through SSH into each BIG-IP unit and run the `imish` command which enters the ZebOS cli (ZebOS uses a typical router cli command set). The F5 BIG-IP must mimic NSX-T’s BGP configuration. This is shown in the next figure with embedded comments.
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
  neighbor 10.106.53.1 fall-over bfd
  no neighbor 10.106.53.1 capability graceful-restart
  neighbor 10.106.53.1 route-map default-route
  neighbor 10.106.53.2 remote-as 65001
  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 30 – ZebOS BGP without ECMP configuration in the BIG-IP.

At this point, follow the testing steps described in the Verifying the deployment section.
Implementation: Active/Active NSX-T Edge using BGP ECMP

For large / high performance deployments, NSX-T Edge nodes are typically configured in Active-Active. In this deployment guide it is assumed that when using NSX-T Active-Active the most likely scenario is that NSX-T Edge nodes are bare metal servers and the BIG-IPs are implemented in hardware. When using Active/Active NSX-T Edge it is likely to be used with ECMP\(^5\) 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.

\(^5\) Please note that NSX-T Edge Active-Active doesn’t imply the use ECMP or vice versa.
in order to keep the traffic symmetry outside the NSX-T environment. Instead, if ultimate scalability is required, the best option is adding blades with a chassis platform which provides ultimate scale-out performance without requiring any reconfiguration and keeps a simple architecture.

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

![Figure 32 – BGP peering detail with two uplink Layer 3 paths & transport zones for ECMP.](image)

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

1. Create a transport zone for each uplink
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 33 – Overall configuration of transport zones. The used ones by this topology are highlighted (red and blue for the uplinks).
2. Edit the Edge transport nodes to add the two uplink transport zones.

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

![Edit Edge Transport Node - edge-1-topology-a](image)

**Figure 34** – Adding the switches for each Uplink transport zone in each Edge transport nodes.
Besides each transport-zone, each associated N-VDS switch requires specific Uplink profile and Uplink interfaces. An example for Transport Zone tz-vlan-uplink-red is shown next.

![Figure 35 – N-VDS switch configuration for a sample Uplink transport zone.](image)

3. Create a Tier-0 configuration.

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

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

![Figure 36 - Adding a Tier-0 Gateway.](image)
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 37 - Filling the details of a Tier-0 Gateway in Active-Active mode.**

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

Select the just created Tier-0 Gateway and create 1 Gateway port for each peering address. This is one Gateway’s interface for the combination of each subnet (two in this example) and NSX-T Edge nodes (two in this example). In total 4 Gateway interfaces will be created as shown next. It is very important to correctly assign the right Edge Transport
node and switch. The ports and their configuration used in this example are shown next. The settings for each Gateway’s interfaces are analogous to the Active-Standby setup.

![Figure 38 – Adding the Gateway’s interfaces for the Uplink paths.](image)

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

![Figure 39 - Enable BGP with ECMP in the Tier-0 Gateway in Active-Active mode.](image)

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:

![BGP Neighbors](image)

**Figure 40 – BGP peerings for ECMP.**

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

![Route Redistribution](image)

**Figure 41 - Enabling Route redistribution at T0 Gateway**

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

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 ← as per VMware's recommendation NSXT-VI-SDN-038
  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
  neighbor 10.106.54.1 route-map default-route-uplink-red out
  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
!
```

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

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

Asymmetric traffic is allowed in the BIG-IP by unsetting the parameter VLAN-Keyed Connections as shown in the next figure. This must be configured in all the BIG-IP units.
**Figure 45 – Configuration required for ECMP which might generate asymmetric traffic.**

At this point, follow the testing steps described in the Verifying the deployment section.
Topology B: BIG-IPs inline – connected like an NSX-T’s Tier-1 Gateway.

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

![Overview of BIG-IPs inline-connected like an NSX-T’s Tier-1 Gateway.](image)

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

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

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

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

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

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

![Configuration Diagram](image)

**Figure 47 – Topology B example implementation.**

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

The transit network between the Tier-0 and the BIG-IPs uses a /24. Using a /24 prefix is larger than strictly necessary for an HA-pair (only 4 hosts address would be needed) but allows for more ingress VIP addresses and expanding the BIG-IP HA cluster into a Scale-N Active-Active cluster (up to 8 BIG-IPs per cluster) or multiple BIG-IP clusters.
From the figure above, it can be seen that this topology is only supported by BIG-IP VE. The configuration will be detailed next. As with all other topologies, this guide focuses in the configuration for the Layer 3 and higher layers that are specific to this topology.

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

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

![Figure 48 - Adding a Tier-0 Gateway.](image-url)
In the New Tier-0 Router dialog, complete the following:

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

![Tier-0 Gateways](image)

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

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

Select the router created (T0-Topology B in our example) and create two interfaces in the UI by first selecting the Edit option in the T0 Gateway, then scrolling down to the
Interfaces section clicking in the Set option of External and Service Interfaces. Enter the following parameters for each interface:

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

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

Figure 51 – Final Uplink interface configuration of the Tier-0 Gateway.
1.3. Create an HA VIP for the Tier-0 Gateway.

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

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

![Figure 52 - Adding an HA VIP to NSX-T’s T0 Gateway.](image)

Selecting the two external interfaces just created.

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

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

![Figure 53 – Adding Tier-0 Gateway’s default route.](image)

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

![Image of creating an overlay segment](image)

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

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

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

![Image of adding tenant routes](image)

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

2.2. Create tenant’s segments.

Follow the same steps as for creating the segment for the transit network, creating as many logical switches as networks are going to be used for the tenant. In this example we
will create only the ones for the tenant green, these will be:

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

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

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

![Figure 56 – Attaching the BIG-IP to NSX-T segments.](image)

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

![Figure 57 – Adding the NSX-T segment to the BIG-IP is just like a regular untagged VLAN.](image)
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.

![Self IPs and floating Self IPs required (shown in BIG-IP unit 1).](image)

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).

![Static route required in the BIG-IP units.](image)

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 be 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 but has two IPs.

These would be implemented as follows

![Figure 60 - Example of East-West Virtual Servers with multiple IP addresses](image)

It is important to differentiate the following Virtual Server Settings:

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

![Topology C overview](image)

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 latency added 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 will be done with a regular router port.

The main reason for choosing this topology should be that each tenant can have their own North-South BIG-IP VE, which they can manage independently. For the purpose of full isolation, this can be achieved for either Topology A or C using a hardware BIG-IP with vCMP technology. A multi-tenant setup with full isolation is shown in the Figure 62.
Figure 62 – Topology C with multiple tenants shown.
This topology has the following benefits:

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

On the other hand, it has the following drawbacks:

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

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

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

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

![Figure 63 – Topology C example implementation.](image)
In the example used for this topology BIG-IP VE is used which means that the segment between the BIG-IP and the Edge nodes uses the NSX-T overlay. This will be shown in the following configuration. Given the many possibilities of configuring NSX-T Edge nodes and their logical switch uplink ports, it is assumed that these have been already created. This guide is focused in the configuration for the Layer 3 and higher layers that are specific to this topology. See section Design consideration: Layer 2 networking for details.

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

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

![Tier-0 Gateways](Image)

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

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

Select the router created (T0-Topology C in our example) and create two interfaces in the UI by first selecting the Edit option in the T0 Gateway, then scrolling down to the
Interfaces section clicking in the Set option of External and Service Interfaces. Enter the following parameters for each interface:

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

![Figure 66 – Filling the details of a router port of one of the uplinks for the Tier-0 Gateway.](image)

![Figure 67 – Final Uplink interface configuration of the Tier-0 Gateway.](image)
1.3. Create an HA VIP for the Tier-0 Gateway.

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

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

![Figure 68 - Adding an HA VIP to NSX-T’s T0 Gateway.](image)

Select the two external interfaces just created.

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

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

![Figure 69 – Adding Tier-0 Gateway’s default route.](image)

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

1.4.1. Create a segment for the transit network.

Go to Networking > Segments > ADD SEGMENT and create a Segment within the Overlay or a VLAN Transport Zone, this will mainly depend if the BIG-IP is a VE or hardware. In this case we are using a VE and the transit network will be in the overlay Transport
Zone. The segment (we use segment-348 in this example) must be attached to the Tier-0 Gateway previously created. This configuration is shown next.

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

2. Create a Tier-1 Gateway.

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

![Tier-1 Gateways](image)

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

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

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

This configuration can be seen in the next figure:
3. Create the Layer 3 configuration on the BIG-IP.

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

![Figure 73 - Attaching the BIG-IP to an NSX-T overlay segment for the transit network.](image)

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

![Figure 74 – Adding the Layer 2 networks to the BIG-IPs.](image)

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.

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).
At this point, follow the testing steps described in the Verifying the deployment section.
Topology D: BIG-IPs parallel-connected to NSX-T’s Tier-1 Gateway.

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

- The purple flow shows a typical Virtual Server with a single IP address (VIP). The flow outlined is between segments orange and green. The VIP belongs to segment orange and the client is in the green segment. In order for the client to reach the VIP it has 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.
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 in Figure 81.

![Combined Topology A and D](image)

*Figure 81 - Combined Topology A and D.*
Implementation: BIG-IPs parallel-connected to NSX-T’s Tier-1 Gateway.

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

Figure 82 - Topology D implementation used through this section.

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

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

1. Create a Tier-1 Gateway.

This Tier-1 Gateway will have a transit network towards Tier-0 (automatically created) and in this example 3 user segments in the overlay transport zone (orange, green and blue).
In NSX-T manager, select Networking > Tier-1 Gateways > Add Tier-1 Gateway > Tier-1 Router filling the following parameters:

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

![Tier-1 Gateway Configuration](image)

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

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

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

This configuration can be seen in the next figure:
2. Create the Layer 3 configuration in the BIG-IP

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

Figure 85 – 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 Figure 86:

![Figure 86 – Static routes required in the BIG-IP units.](image)

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

BGP configuration details

The following configuration settings follow VMware Validated Design 5.0.1, 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 87 - VMware's baseline settings for BFD](image)

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.

---

Best practices for BIG-IP in VMware NSX-T

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

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

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

The above VM placement best practices can be achieved with the Dynamic Resource Scheduler (DRS). In the next picture, the creation of anti-affinity rules is shown to avoid two BIG-IPs of the same cluster running on the same hypervisor. Note: the anti-affinity rules should be “must” rather than “should” to guarantee anti-affinity and therefore high availability.
Figure 88 - Setting anti-affinity rules with VMware’s Dynamic Resource Scheduler.
Considerations for container platforms

General guidelines

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

As described in previous sections, for any of these container platforms the POD’s IP addresses should be routable from the BIG-IP. In other words, there cannot be any NAT between the BIG-IP and the PODs. Moreover, there are two ways which POD workers can be exposed with a resource of kind Service: via NodePort or via ClusterIP. Although both are supported it is highly recommended to use ClusterIP\(^7\). This is because when using NodePort mode the BIG-IP (or any other external host) cannot send traffic directly to the PODs which means for the BIG-IP that:

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

Exposing container services

Once the PODs that compose the workers of a given Service are defined, the BIG-IP must be automatically configured and updated when the PODs of the service are created, updated or deleted. This is performed by the F5 Container Ingress Services (CIS)\(^8\) which is installed as a Kubernetes POD that monitors configuration changes in Kubernetes. F5 CIS automatically updates the BIG-IP configuration by translating orchestration commands into F5 SDK/iControl REST calls. The overall architecture is shown in the next picture.

---

7 https://clouddocs.f5.com/containers/v2/kubernetes/kctlr-modes.html#kctlr-modes
8 https://clouddocs.f5.com/containers/v2/
Although in the diagram above only one CIS instance is shown, it is possible that a single instance of F5 BIG-IP can be managed by several CIS instances associating different container namespaces or projects to different partitions in the F5 BIG-IP.

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

<table>
<thead>
<tr>
<th></th>
<th>Red Hat OpenShift</th>
<th>Pivotal PKS</th>
<th>Vanilla Kubernetes</th>
</tr>
</thead>
<tbody>
<tr>
<td>OpenShift routes</td>
<td>Yes</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td>Ingress</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>AS3 declaration</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
These options can be combined in the same deployment. Note that in the above table, the LoadBalancer Service type is not mentioned. This is out of scope because it is meant to be implemented by a cloud provider’s load balancer. Also note that the LoadBalancer Service type is also not efficient in the use of IP address ranges because it requires an IP address for each instance.

Relevant configuration settings when using Red Hat OpenShift

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

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

```
use_native_loadbalancer = False
```

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

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

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

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

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

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

Relevant configuration settings when using Pivotal PKS

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

In the case of Pivotal PKS this is indicated with the PKS Ops Manager UI while performing PKS installation. Following the regular PKS configuration, it is needed to unset the NAT option in the Networking Tab as shown in the next screenshot.
Figure 90 - Indicating PKS networking options at installation time. The NAT option must be unset.
Verifying the deployment

Basic testing

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

<table>
<thead>
<tr>
<th>Topology</th>
<th>Adjacent next-nops</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Topology A</strong></td>
<td></td>
</tr>
<tr>
<td>Impl. static</td>
<td>Northbound – 10.105.217.1</td>
</tr>
<tr>
<td>Impl. static</td>
<td>Southbound – 10.106.63.1</td>
</tr>
<tr>
<td><strong>Topology A</strong></td>
<td>Impl. dynamic routing</td>
</tr>
<tr>
<td>Northbound</td>
<td>Northbound – 10.105.217.1</td>
</tr>
<tr>
<td>Southbound</td>
<td>Southbound – 10.106.53.{1,2}</td>
</tr>
<tr>
<td><strong>Topology A</strong></td>
<td>Impl. dynamic routing +ECMP</td>
</tr>
<tr>
<td>Northbound</td>
<td>Northbound – 10.105.217.1</td>
</tr>
<tr>
<td>Southbound Uplink Red</td>
<td>Southbound Uplink Red – 10.106.53.{1,2}</td>
</tr>
<tr>
<td>Southbound Uplink Blue</td>
<td>Southbound Uplink Blue – 10.106.54.{1,2}</td>
</tr>
<tr>
<td><strong>Topology B</strong></td>
<td>Northbound – 10.106.49.1</td>
</tr>
<tr>
<td>Southbound</td>
<td>Southbound – 10.106.{51,52}.10 (Servers)</td>
</tr>
<tr>
<td><strong>Topology C</strong></td>
<td>Northbound – 10.10.216.1</td>
</tr>
<tr>
<td>Southbound</td>
<td>Southbound – 10.106.48.1</td>
</tr>
<tr>
<td><strong>Topology D</strong></td>
<td>NorthBound – 10.106.32.1 (default route’s next-hop)</td>
</tr>
<tr>
<td></td>
<td>Southbound – 10.106.{32,33,34},100 (Servers)</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Topology</th>
<th>Segment / IP address</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Topology A</strong></td>
<td>10.106.32.10</td>
</tr>
<tr>
<td><strong>Topology B</strong></td>
<td>10.106.{51,52}.10</td>
</tr>
<tr>
<td><strong>Topology C</strong></td>
<td>10.106.51.10</td>
</tr>
<tr>
<td><strong>Topology D</strong></td>
<td>10.106.{32,33,34},100</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Topology</th>
<th>Test VM’s NSX-T next-hop</th>
<th>BIG-IP’s closest IP to test VM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topology A</td>
<td>Impl. static routing</td>
<td>10.106.32.1</td>
</tr>
<tr>
<td>Topology A</td>
<td>Impl. dynamic routing</td>
<td>10.106.32.1</td>
</tr>
<tr>
<td>Topology A</td>
<td>Impl. dynamic routing +ECMP</td>
<td>10.106.32.1</td>
</tr>
<tr>
<td>Topology B</td>
<td></td>
<td>10.106.{51,52}.1</td>
</tr>
<tr>
<td>Topology C</td>
<td></td>
<td>10.106.51.1</td>
</tr>
<tr>
<td>Topology D</td>
<td></td>
<td>10.106.{32,33,34}.1</td>
</tr>
</tbody>
</table>

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

**Dynamic routing testing**

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

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

```
bigipla.nsxt.bd.f5.com[0]#show bfd session
```

<table>
<thead>
<tr>
<th>Sess-Idx</th>
<th>Remote-Disc</th>
<th>Lower-Layer</th>
<th>Sess-Type</th>
<th>Sess-State</th>
<th>UP-Time</th>
<th>Remote-Addr</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>458133421</td>
<td>IPv4</td>
<td>Single-Hop</td>
<td>Up</td>
<td>2d19h49m</td>
<td>10.106.53.1/32</td>
</tr>
<tr>
<td>4</td>
<td>211353312</td>
<td>IPv4</td>
<td>Single-Hop</td>
<td>Up</td>
<td>2d19h49m</td>
<td>10.106.54.1/32</td>
</tr>
</tbody>
</table>

**Figure 91** - Verification of the NSX-T uplinks by checking the BFD sessions.

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

```
bigipla.nsxt.bd.f5.com[0]#show bgp neighbors | include BGP state
```

```
BGP state = Established, up for 2d19h50m
BGP state = Established, up for 2d19h50m
```

**Figure 92** - Verifying that the BGP sessions are Up.

As you can see in Figure 92, 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).
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
10.106.33.0/24  10.106.33.0/24  gw  10.106.53.1  dynamic ecmp
10.106.34.0/24  10.106.34.0/24  gw  10.106.53.1  dynamic ecmp

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  10.106.54.1  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  10.106.54.1  dynamic ecmp
```

Figure 94 - 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 95, where the parameters are in red are mandatory.

Figure 95 - 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 96 - Ping test using spine router's IP address and the well-known Internet address 8.8.8.8 for checking egress connectivity.](image)

<table>
<thead>
<tr>
<th>Topology</th>
<th>Closest's spine router IP address</th>
</tr>
</thead>
<tbody>
<tr>
<td>Topology A</td>
<td>10.105.217.1</td>
</tr>
<tr>
<td>Impl. static routing</td>
<td></td>
</tr>
<tr>
<td>Topology A</td>
<td>10.105.217.1</td>
</tr>
<tr>
<td>Impl. dynamic routing</td>
<td></td>
</tr>
<tr>
<td>Topology A</td>
<td>10.105.217.1</td>
</tr>
<tr>
<td>Impl. dynamic routing +ECMP</td>
<td></td>
</tr>
<tr>
<td>Topology B</td>
<td>10.105.216.1</td>
</tr>
</tbody>
</table>

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.

<table>
<thead>
<tr>
<th>Topology</th>
<th>Closest spine router’s IP address</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>10.105.216.1</td>
</tr>
<tr>
<td>D</td>
<td>10.105.217.1</td>
</tr>
</tbody>
</table>

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

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

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

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

![Creating a Standard Virtual Server for testing Ingress services' connectivity.](image)

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:

![Creating a new pool that will be used for the connectivity test with the Ingress Virtual Server.](image)

Then specifying the pool as shown in the next picture. Please note that the default HTTP health monitor is used.
Figure 99 - 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 100 - 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.