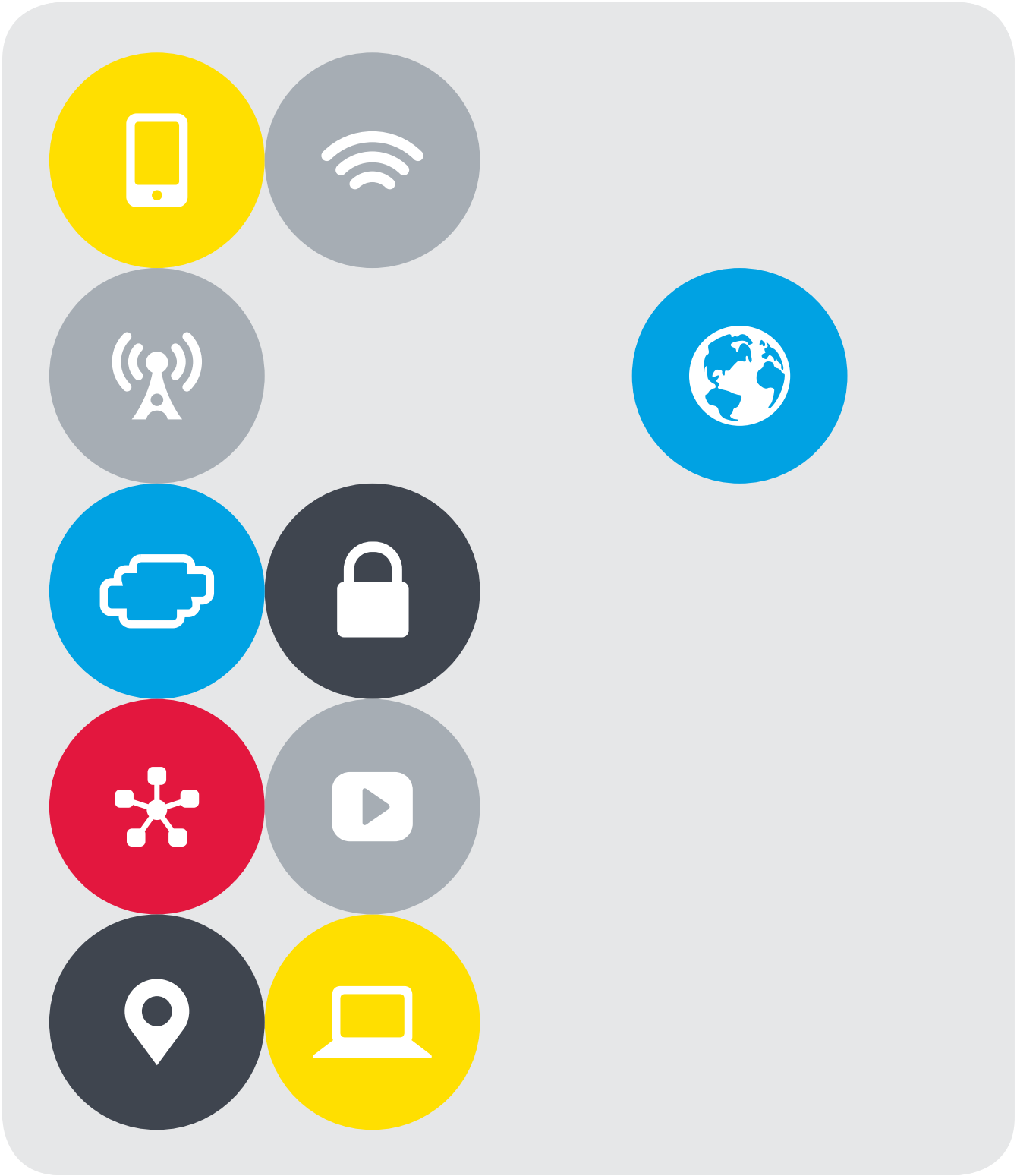
DESIGN GUIDE

VMware NSX and F5



Contents

Introduction

The purpose of this document is to provide a solution overview and design guidance for integrating F5 Application Delivery Controllers (ADCs) with VMware NSX network virtualization. The content is intended for network architects currently using or planning to use network virtualization and ADC/load balancing services in their environment. This guide will focus on the integration of F5® BIG-IQ® Cloud and F5 BIG-IP® virtual editions (VEs) with VMware NSX network virtualization.

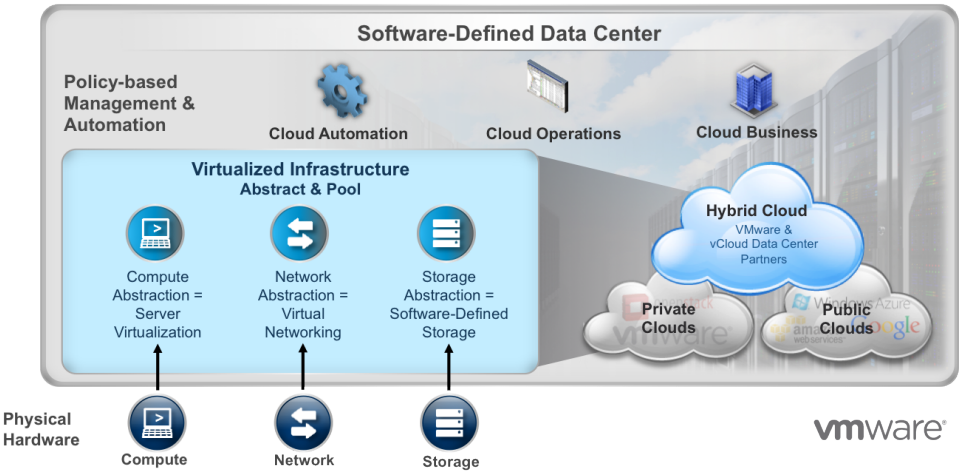
The Needs of the Business vs. IT Capabilities

Organizations want to increase the speed of innovation, reduce time to market, and drive the velocity of their businesses. Traditional IT architectures can be a barrier to these objectives. They can be costly to manage, and limited in configuration and scale. Legacy computing and IT operational models can be slow to respond to dynamic environments caused by shifting business needs or the rapid increase of users, devices, and applications. Furthermore, application architects and operations staff are experiencing difficultly orchestrating the lifecycle of applications and other IT services across the various silos of a typical IT organization. Society's increased reliance on business technology requires flexibility in how services are delivered and consumed.

To be competitive while remaining efficient with resources, data center and IT operational agility must be all-encompassing. A change in IT processes and architectures can drastically reduce lead times for deploying new applications and services, eliminate downtime due to unforeseen increases in workloads, and speed disaster recovery. Solving this challenge requires collaboration across all elements of the data center—including server infrastructure, networking and connectivity, and the overall application delivery architecture.

The Software-Defined Data Center

A software-defined data center (SDDC) architectural approach is required to meet many of today's business expectations, helping organizations transform data center economics and increase application deployment agility. The SDDC delivers elastic, "on-demand" IT infrastructure services through the use of resource virtualization and abstraction of traditional computing, network, and storage assets.



VMware's software-defined data center concept.

VMware vSphere and its data center management suite of products provide a foundational platform for storage and compute in the SDDC. VMware NSX network virtualization extends the same benefits to the data center network to accelerate network service provisioning, simplify network operations, and improve network economics.

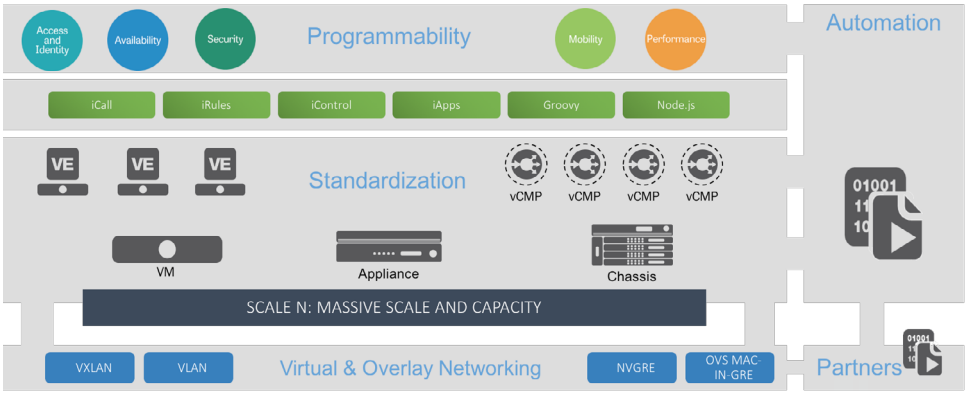
VMware NSX decouples the logical networking from the physical infrastructure, removing many of the operational and architectural barriers commonly seen with traditional network infrastructures and technology. Network virtualization with VMware NSX facilitates the automation and management of the logical network and network services lifecycle.

F5's Approach to Network Virtualization and SDDC

Network virtualization must be accompanied by changes in the provisioning of application delivery services. The deployment of ADC functions must be based on data center automation and the adoption of cloud computing principles. Modern architectures and data center models require a more flexible approach to application services such as availability—one that better aligns with trends toward micro-services and API-based architectures.

More broadly, given increased user mobility and an expanding "Internet of Things," service providers and organizations are reevaluating traditional architectural principles to determine how best to move forward with application delivery service provisioning that can keep up with, or at least catch up to, industry trends.

F5 Software-Defined Application Services™ (SDAS) is the next-generation model for delivering application services. SDAS takes advantage of F5 innovations in scalability models, programmability, and an intrinsic decoupling of the data and control planes to create a unique application services fabric capable of extending the benefits of F5 application delivery services to all applications, irrespective of location.



F5's Software-Defined Application Services framework.

SDAS is the first fabric-based application delivery and control system. It enables service injection, consumption, automation, and orchestration across a unified operating framework of pooled resources.

F5 Software-Defined Application Services deliver:

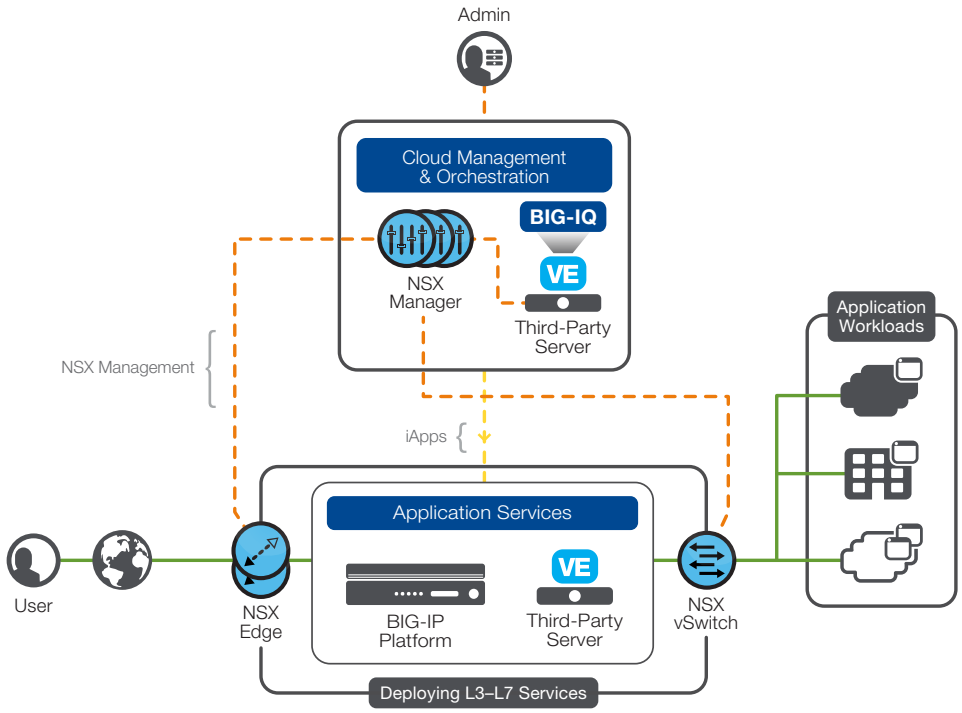
- **A fabric-based solution—**F5 ScaleN™ technology powers an elastic, all-active application service fabric that dramatically lowers the cost of delivering application services by increasing utilization and service densities.

- **Automation and orchestration—**Intelligent service automation and orchestration APIs reduce operational expenses and fill a critical gap in SDDC and network architectures. As a result, organizations with SDAS can streamline application deployment and support continuous delivery.

- **A unified operating framework—**A rich, extensible catalog of application services empowers application owners to address performance, security, and availability concerns in cloud, data center, service provider, and managed environments. The SDAS fabric provides a foundation for building elastic application services.

F5 BIG-IP Integration with VMware NSX

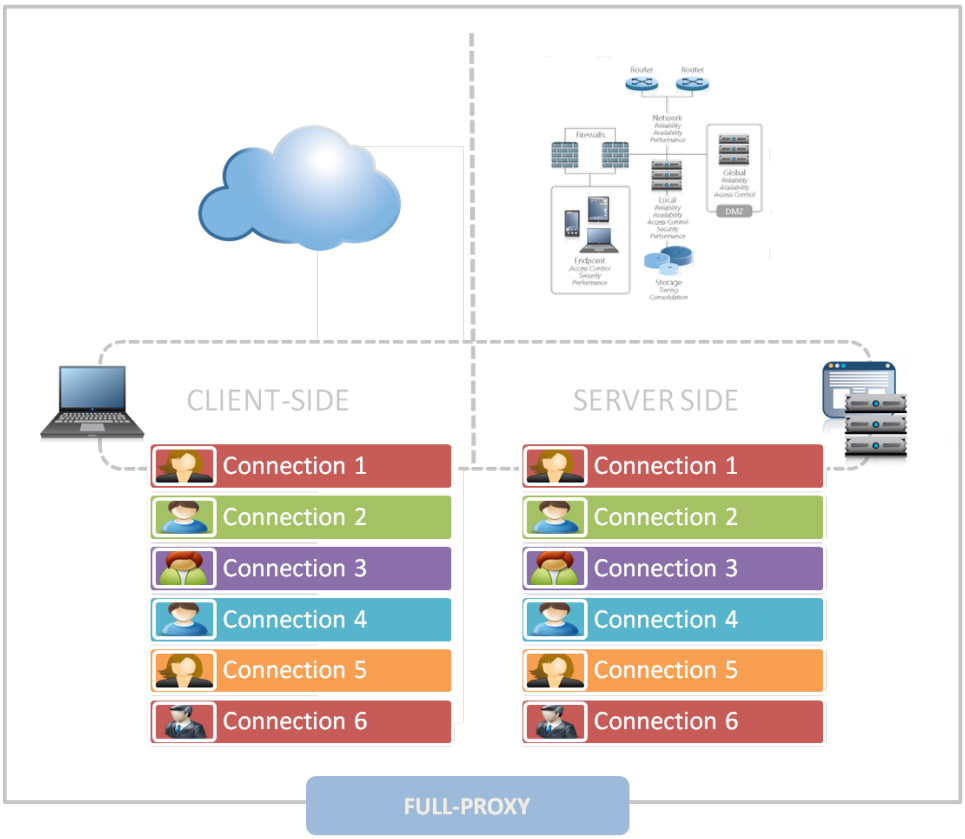
The integration between the F5 BIG-IP platform and VMware NSX extends VMware's SDDC strategy to include F5 SDAS—delivering interconnected automation for network and application layer services. Customers can rapidly deploy applications and other relevant IT services, creating an agile environment that can react swiftly to the demands of the business.



The F5 and VMware NSX deployment scenario.

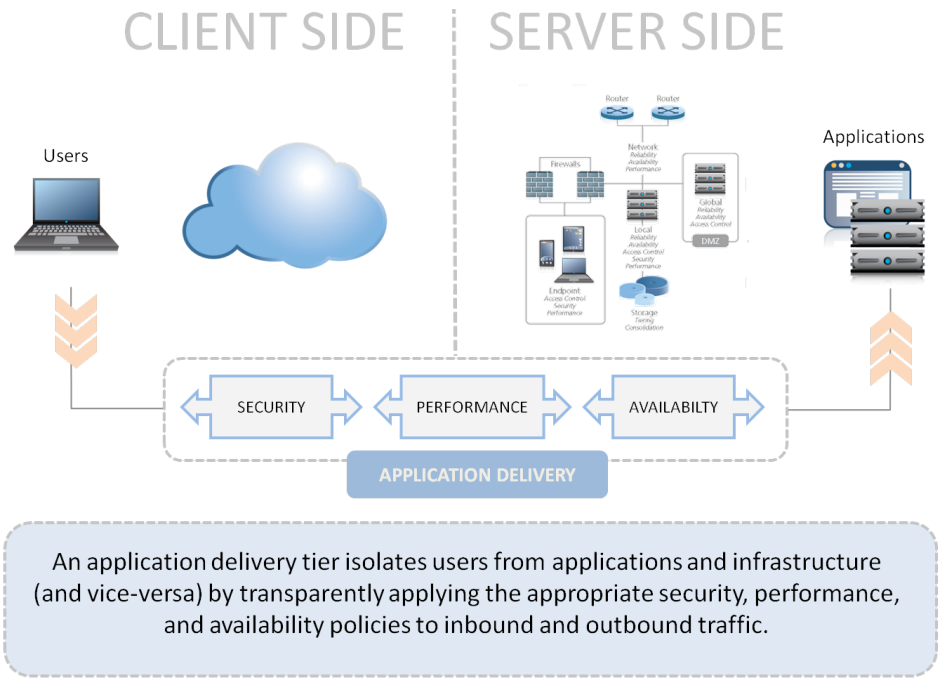
The F5 BIG-IP platform and its full-proxy architecture complements the native network virtualization capabilities of VMware NSX by introducing the ability to enhance the intelligence, security, availability, and performance of mission-critical applications.

A full proxy maintains two separate session tables—one on the client-side and one on the server-side. There is effectively an "air gap" isolation layer between the two internal to the proxy—enabling focused profiles to be applied specifically to address issues and manage traffic peculiar to each "side" of the proxy.



Full-proxy architecture.

Because all communication is funneled through virtualized applications and services at the application delivery tier, it serves as a strategic point of control at which delivery policies addressing traffic routing/intelligence and operational risk (performance, availability, security) can be enforced. Inserting an application delivery tier allows for an agile, flexible architecture—more supportive of the rapid changes today's IT organizations must deal with.



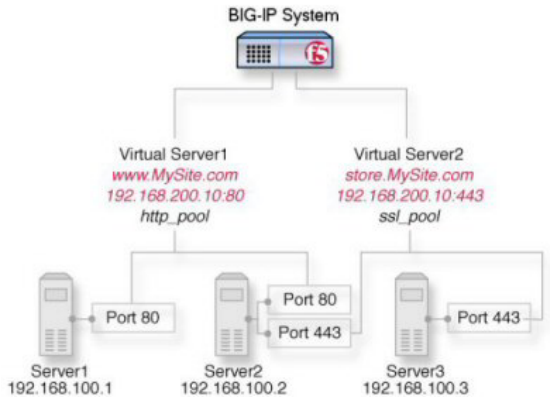
F5's application delivery tier using full proxy.

VMware NSX and BIG-IP virtual editions enable customers to build an adaptable, virtualized network infrastructure that delivers the scale, consolidation, and business continuity demanded by today's advanced application infrastructures. BIG-IP VEs deliver the same market-leading SDAS that run on F5 purpose-built hardware.

Combining the feature set of BIG-IP® Local Traffic Manager™ (LTM) with VMware NSX increases operational efficiency and ensures peak network performance by providing a flexible, high-performance application delivery system. With its application-centric perspective, BIG-IP LTM optimizes network infrastructure to deliver availability, security, and performance for critical business applications.

Key benefits of integrating BIG-IP LTM with VMware NSX include:

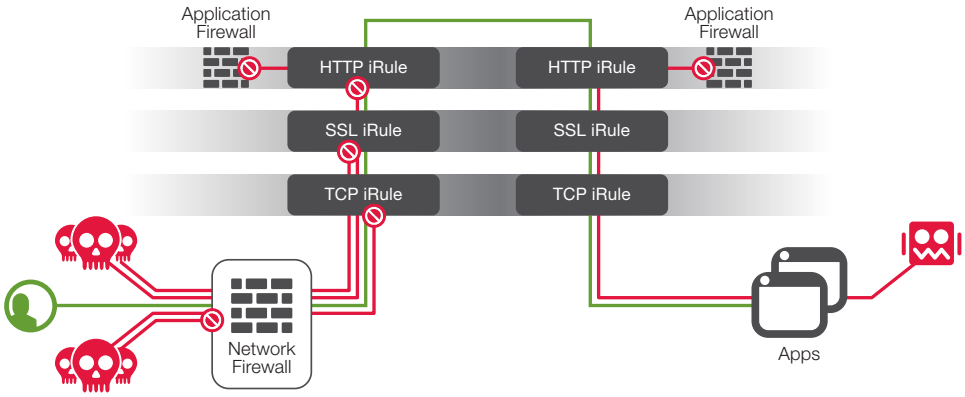
- **Application traffic management and optimization—**BIG IP LTM includes static and dynamic load balancing to eliminate single points of failure. Application proxies give protocol awareness allowing for the control traffic for the most important applications.



Load balancing web servers with F5 BIG-IP LTM.

BIG IP LTM also tracks the dynamic performance levels of servers in a group, ensuring that your applications are not just always on, but are easier to scale and manage. BIG-IP LTM makes real-time protocol and traffic management decisions based on application and server conditions, extensive connection management, and TCP and content offloading.

- **Out of the box security—**The BIG-IP system provides better security by actively terminating the flow of data. Traffic coming from the client can be examined before it is sent on its way to the application tier, ensuring that malicious traffic never passes the proxy barrier. Traffic returning from the server can be fully examined before it is deemed acceptable to pass back to the client, thereby ensuring that sensitive data such as credit card or social security numbers are never passed across the proxy barrier.

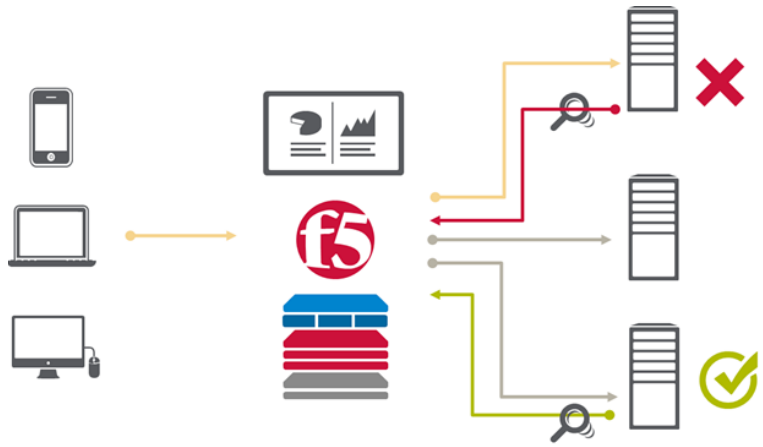


Full-proxy architecture.

The BIG-IP platform's full-proxy architecture also provides an effective means to combat modern attacks. Because of its ability to isolate applications, services, and even infrastructure resources, this full-proxy application delivery tier improves an organization's capability to withstand the onslaught of a concerted attack.

- **Secure application delivery—**The BIG-IP system delivers industry-leading SSL optimization, performance, and visibility for inbound and outbound traffic, so you can cost-effectively manage and protect your entire user experience by encrypting everything from the client to the server.

- **Intelligent application health monitoring—**Intelligently monitor your applications with the BIG-IP system's advanced application health monitoring capabilities. The BIG-IP system includes a number of built-in monitor templates for specific monitoring of applications, which may be used as is or customized to meet your requirements.



The BIG-IP platform's application health and intelligent load balancing.

BIG-IP LTM External Application Verification (EAV) monitors enhance application health monitoring by utilizing the BIG-IP system to perform complex and thorough service checks. They provide the ability to dive much deeper than merely performing a three-way handshake and neglecting the other layers of the application or service.

EAV also brings added application intelligence by performing application-aware monitoring, such as simulating an actual user login. This level of visibility into the application's health provides a deeper insight on whether the application is truly functioning as expected, rather than simply determining if a service is up/down or a network port is "listening."

- **Programmability—**For advanced configurations and traffic management scenarios, the F5 iRules® scripting language (F5's traffic scripting interface) enables programmatic analysis, manipulation, and detection of all aspects of the traffic in your networks. Customers can implement security mitigation rules, support new protocols, and fix application related errors in real time.

  Additional programmability options are available using iControl®, F5's open API that allows complete, dynamic, and programmatic control of F5 configuration objects. With iControl, you can work like a wizard—adding, modifying, or configuring your F5 device in real time.

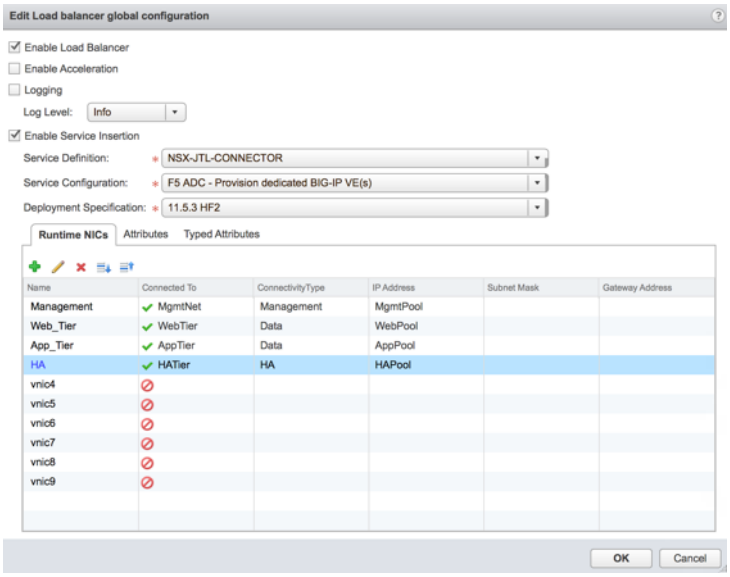
- **Wizard-driven configurations—**The BIG-IP platform includes F5 iApps™ Templates, a powerful feature that enables you to deploy and manage enterprise application services as a whole rather than individually managing configuration and objects. iApps gives you greater visibility into and control over application delivery—and helps you deploy in hours rather than weeks.

  iApps Templates are integrated with VMware NSX to provide an application-centric configuration approach, aligning the application delivery with your network and business needs.

- **Unprecedented application performance and availability—**Using real-time protocol and traffic management decisions based on application and server conditions, extensive connection management, and TCP and content offloading, BIG-IP LTM dramatically improves page load times and the user experience. It helps you adapt to shifting performance and application needs.

  BIG-IP LTM protects applications by removing single points of failure, giving you fine-grained bandwidth control and optimizing your most important applications. And since it tracks the dynamic performance levels of servers in a group, BIG-IP LTM ensures that all sites are not just always on, but are more scalable and easier to manage than ever.

- **Agile, automated, and simple deployment of BIG-IP ADCs with VMware NSX—**Quickly and easily spin-up application services in VMware NSX-enabled environments. BIG-IP virtual editions and VMware NSX offer on-demand deployment options for high-density and multi-tenant infrastructures, so you can make new applications available to users in minutes through the VMware NSX administration capabilities from within the vCenter Web Client.



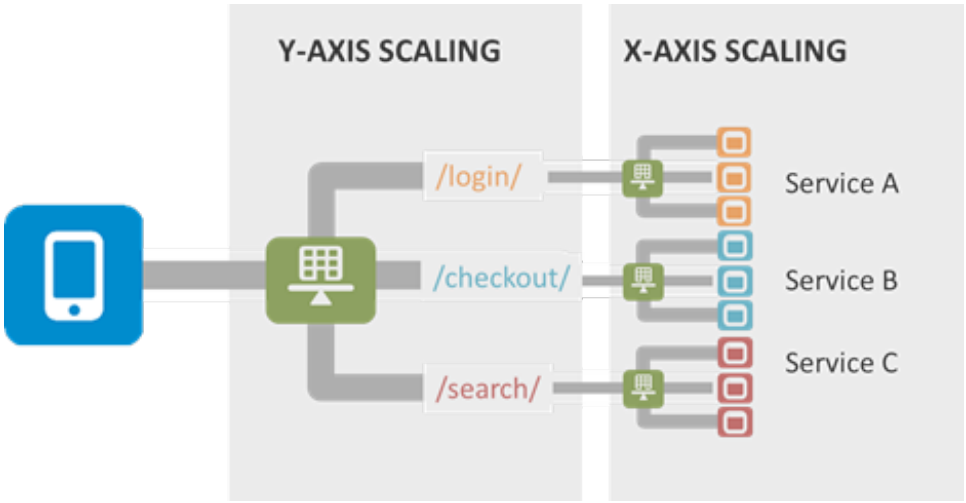
F5 BIG-IP integration using NSX Service Insertion.

Together, F5 and VMware have simplified the deployment of both network and application services within an SDDC. The combination of a unified deployment workflow for virtual machines (VMs) and services, along with the abstraction of complex application service configuration, simplifies and shortens the application deployment process.

F5 BIG-IP and VMware NSX Use Cases

By deploying BIG-IP ADCs and NSX together, organizations are able to achieve service provisioning automation and agility enabled by the SDDC. Here are some of the key use cases for BIG-IP and NSX integration:

- **Per-application, multi-application, and microservice load balancing—** Services that have a greater natural affinity to the application—load balancing, application security, performance—are migrating closer to the app not only in topology but in form factor, with VMs and software being the preferred method of delivery and control.



Load balancing based on the application's various services.

Whether it be load balancing multiple applications, a single monolithic application, or components that are part of microservice-based applications, the integration of VMware NXS and BIG-IP LTM offers customers the deployment flexibility, management simplicity, and scalability to meet their application delivery needs.

- **Development, testing, and lab environments—**Customers can ensure their application is truly tested "end-to-end" by ensuring VMware NSX and BIG-IP LTM are integrated into a testing/lab environment. Application network traffic patterns, load balancing behaviors, and scale/functionality can be analyzed and tested while isolated to a lab. Testing a single application and/or service independent of any other application ensures the testing environment and related results are not affected by any outside factors.

  Additionally, developers can deploy an Application Delivery Controller "on demand" to either validate their microservice or application against a load balancer without having to wait for the networking team to configure/provision the necessary services. Developers and testers can even tear down their load balancers when completed, returning their licenses and resources for other developers and testers to consume.

- **Automated data centers and self-service/cloud-provisioning environments—** Automating the provisioning of BIG-IP application delivery services in NSX-enabled data centers and cloud environments translates to increased virtual data center agility with fewer configuration errors, enhanced security, and policy/configuration quality and control.

  Application owners can get the application networking services they need faster, with total configuration time reduced from days or weeks to minutes. The result is greater operational efficiency, fewer consoles, and lower management costs, freeing more time for skilled personnel to work on more challenging tasks.

  The integration of BIG-IP ADCs with VMware NSX enables self-service and automated provisioning of application networking in data center and cloud environments—while retaining policy and configuration quality control for the teams responsible for security and application delivery services.

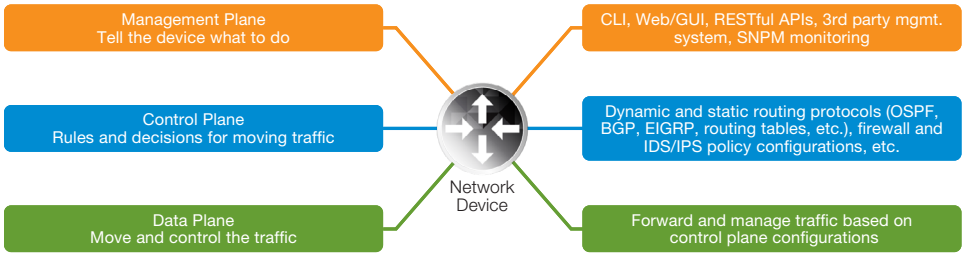
VMware NSX Overview

In order to clearly understand the components and architecture that make up the F5 BIG-IP and VMware NSX solution, it's important to first understand the major components of VMware NSX, including their roles and responsibilities within the NSX architecture and where they function in the network plane stack.

Network Planes

The primary responsibility of traditional and virtualized networking devices is to move and manage traffic across the network. It is important to understand how these network devices are programmed, the protocols and standards used to learn about network paths and other types of devices, and the way network traffic rules are enforced when processing network traffic.

You'll commonly hear about the three logical levels of network device functionality—the data plane, control plane, and management plane. The following diagram describes each network plane and its primary purpose.



Network planes—roles and functions in the data and protocol stack.

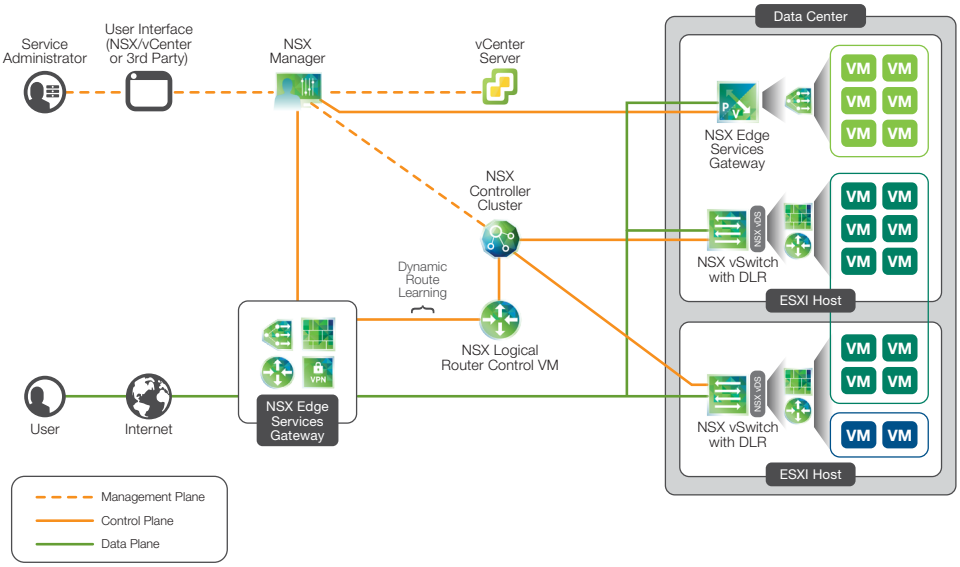
- **Management plane—**The management plane provides a method for configuring the network device. The configuration entered here will be placed into the control plane, and subsequently used by the data plane to properly ship the traffic around the network. Additionally, the management layer also uses protocols for proactive monitoring and device performance/statistics management.

- **Control plane—**The control plane maintains both static configurations from the management layer (i.e., routing protocol configurations, firewall policies, etc.) and dynamic configurations from network routing protocols. The control plane contains the "rules" for moving the traffic around the network.

- **Data plane—**The data plane is the workhorse of the network device. It is responsible for actually moving the packets across the network. It processes and forwards/blocks traffic based on the "rules" established in the control plane.

VMware NSX Architecture

Before designing and deploying the BIG-IP integration with VMware NSX, it is important to understand the core components of VMware's network virtualization solution.



Typical components for NSX deployment.

In this diagram, the service administrator can provision the entire lifecycle of either a network segment (VLAN or VXLAN) or network service (firewall, load balancing, VPN, etc.) using either a third-party management interface or through the vCenter Web Client. Any programmatic changes made by the service administrator to firewalls, networks, or routing are sent from the NSX Manager to either the NSX Edge Services Gateway (NSX Edge) or the NSX vSwitches through the NSX Controller.

NSX vSwitches provide efficient, in-hypervisor networking capabilities that allow for an optimized network path and enhanced security for "east-west" traffic, which uses logical network resources. The NSX vSwitch can perform distributed routing for VLAN and VXLAN-connected networks, firewall services at the vNIC level, and typical layer 2 switching services. The NSX vSwitch firewall capabilities control the ingress and egress of network traffic for each VM using traditional network configurations (IP address, IP port, etc.), by vCenter object type, or by the user's identity.

The NSX Edge Services Gateway on the internal network provides load balancing services to applications housed within the data center. The externally facing NSX Edge VM handles the "north-south" network traffic by providing routing, firewall, and VPN services to users

wishing to access resources within the secure network boundary. Additionally, the NSX Edge is controlling access to external network ports and protocols for systems that sit within the secure network boundary.

The following outlines the high-level roles and responsibilities of each NSX component function in the network virtualization stack.

Management Plane

NSX Manager Virtual Machine

- Provides the single point of configuration and the API entry point in a vSphere environment for NSX.
- Typically managed through the vCenter Web Client user interface or using third-party management interface.
- Option for one-to-one NSX Manager to vCenter server.

NOTE: Cross-vCenter NSX functionality does not support load balancing service insertion.

vCenter Server

- Provides a user interface (using vCenter Web Client) for NSX administration.
- Interfaces with NSX Manager to deliver and manage logical networking services, VMs, and relevant NSX configurations to ESXi hosts and VMs.

Control Plane

NSX Controller VM(s)

- Provides control plane functions for all NSX logical switches and NSX Edge Services Gateway(s) within a network and maintains information about all hosts, logical switches (VXLAN-based Port Group), firewall configurations, and Distributed Logical Routers (DLRs).
- Distributes learned routing updates received from NSX Logical Router Control VMs to DLRs on ESXi hosts.
- NSX Controller VM cluster can be deployed for high availability (HA) and scalability.

NSX Logical Router Control VM(s)

- Learns, maintains, and distributes combined static and dynamic routing information to the NSX Controller VMs.
- Combines Logical Interface Information from DLR, static routes, and dynamic routing info from external/NSX Edge sources.
- Supports various routing protocols (BGP, OSPF, static routing, etc.)
- One Logical Router VM (or pair if HA enabled, Active/Standby) for each DLR.

Data Plane

NSX vSwitch with Distributed Logical Router

- vSphere Virtual Distributed Switch (vDS)-based vSwitch with additional in-hypervisor functionality.
- Provides kernel-based switching, routing, advanced vNIC-level firewall/security, and VXLAN/VLAN networking.
- DLR routes inter-VM networking traffic (a.k.a., east-west traffic) within the hypervisor, rather than sending it to an external routing gateway.

NSX Edge Services Gateway

- Multi-purpose VM that can provide firewall services, VPN access, DNS/DHCP/IP services (DDI), load balancing, and routing/network address translation (NAT).
- Deployed in various VM sizes, based on enabled services and/or network throughput requirements.
- Deployed standalone, Active/Standby, or Active/Active depending on scalability or resiliency needs.

F5 BIG-IP and VMware NSX

Solution Overview

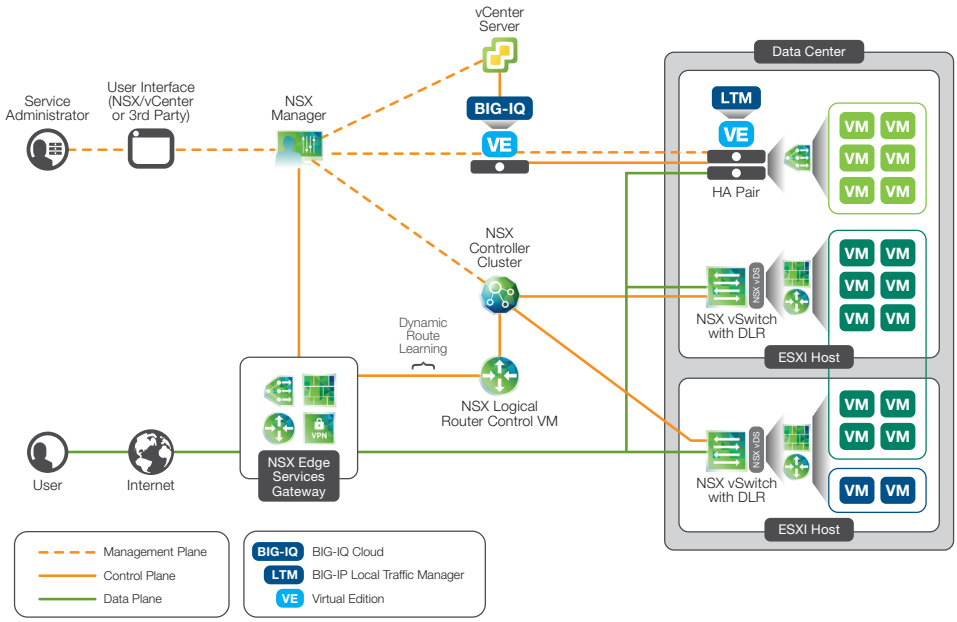
F5's BIG-IP application services platform integrates with VMware NSX to deliver your applications to users in a reliable, secure, and optimized way. By making intelligent traffic decisions that adapt to the changing demands of your applications, the BIG-IP system ensures your business critical applications are consistently available and scalable.

BIG-IP LTM services integrate with VMware NSX, giving network administrators and service consumers the ability to manage the entire lifecycle of both the BIG-IP virtual editions and the corresponding application delivery services they are providing.

BIG-IP and NSX Application Delivery Controller Architecture

The BIG-IP and VMware NSX integration is built from a service provider/consumer model. The service provider (also known as the BIG-IP administrator) manages the BIG-IP VE images, application configuration templates, licenses, and other BIG-IP service configurations. The service provider also controls what templates and services are made available to support the consumers.

Service consumers (also known as the vSpheres or NSX administrators) can deploy F5 ADCs in a standalone or HA configuration. They will be able to select the BIG-IP VE offered by the service provider. They can also select pre-configured templates to deliver intelligent traffic management capabilities for many industry-common applications. When the BIG-IP VE is no longer needed, the service consumer can deprovision it. The BIG-IP VE license is then reallocated for future use.



NSX and BIG-IP integration using F5 BIG-IQ Cloud and VMware NSX.

BIG-IQ Cloud is the main interface point between BIG-IP ADCs and VMware NSX. It is an intelligent framework for managing and orchestrating F5 application delivery solutions. Using BIG-IQ Cloud, service providers or BIG-IP administrators can manage BIG-IP VE images and application profile templates/configurations presented to VMware NSX.

See the following roles and responsibilities of each major F5/NSX-v solution component to better understand how they deliver a seamless integration experience.

Management Plane

NSX Manager Virtual Machine

- Receives instructions from the vCenter Web Client or other third-party management interface for orchestrating and managing BIG-IP VEs and application delivery services.
- Responsible for lifecycle management of BIG-IP VEs.
- Exchanges BIG-IP and application-related orchestration, configuration, and status information with BIG-IQ Cloud.

BIG-IQ Cloud Virtual Appliance

- Provides management interface for service provider functions, including BIG-IP server image repository management, integration with vCenter and NSX-v environments, application load balancing templates (iApps), etc.
- Manages pool of BIG-IP licenses used by NSX-integrated BIG-IP VEs.
- Options for standalone virtual machine or a highly available cluster of BIG-IQ Cloud virtual appliances.

vCenter Server

- Provides centralized management of vSphere-virtualized resources, including storage, networking, and virtual machines.
- Manages the virtualization platform where NSX-enabled BIG-IP VEs are deployed.

Control Plane

BIG-IQ Cloud Virtual Appliance

- Facilitates the configuration and automated deployment of BIG-IP VEs using VMware ESXi, vCenter, and NSX.

Data Plane

BIG-IP Virtual Edition

- Virtual machine that delivers BIG-IP LTM capabilities and functionality to NSX-enabled environments.
- Deployed and managed using VMware NSX's service insertion capabilities.
- Options for standalone virtual machine or high-availability pair in an active/standby configuration.

Integrating BIG-IP and VMware NSX

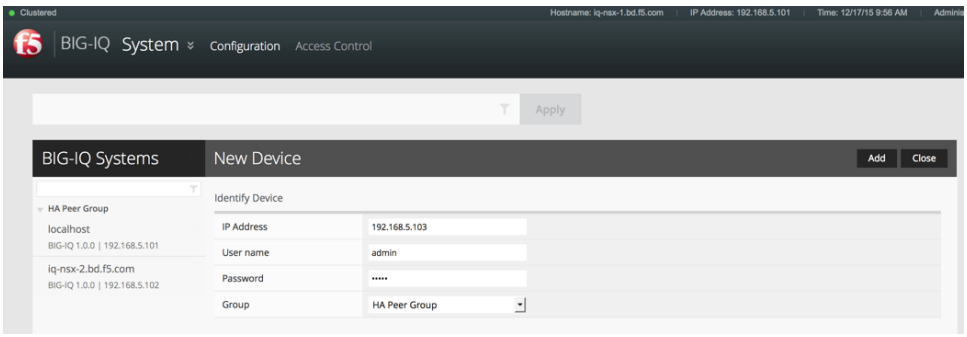
We've covered the basic components and topologies of VMware NSX, as well as the additional components to facilitate the BIG-IP service insertion. The next section provides a high-level overview of what it takes to integrate, deploy, and manage BIG-IP application delivery services through VMware NSX's administrative interface.

Prerequisites

Once VMware NSX is installed and configured, it's important to make sure you have all the prerequisites completed. For additional compatibility and interoperability information, refer to the BIG-IP and BIG-IQ Cloud product documentation for the supported versions of VMware NSX, vCenter/vSphere, and BIG-IP virtual editions.

Deploying and Configuring BIG-IQ Cloud

The first requirement needed to complete integration is the deployment and configuration of BIG-IQ Cloud.

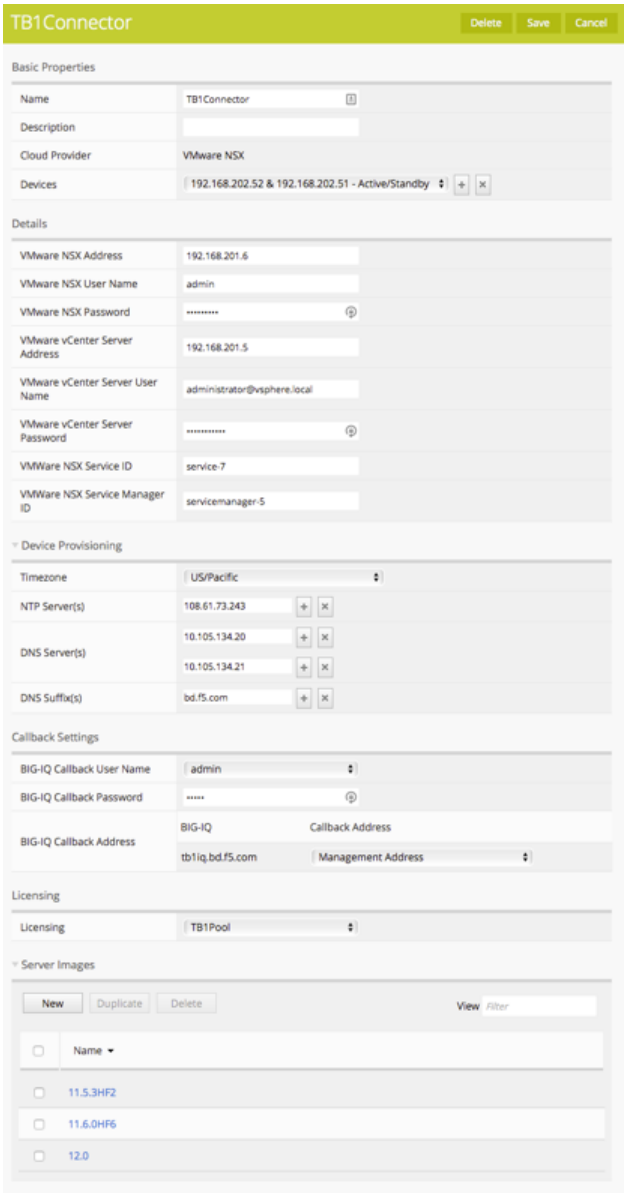


Adding additional BIG-IQ Cloud VE to HA Peer Group.

Additional BIG-IQ Cloud appliances can also be added and configured to support BIG-IQ Cloud resiliency. In the event a BIG-IQ Cloud appliance is rendered inoperable, the other BIG-IQ Cloud appliances will automatically assume the management and orchestration responsibilities with no additional configuration.

Configuring the NSX Connector

After the BIG-IQ Cloud is installed, the next step to completing the integration is to set up the NSX connector. The NSX connector establishes a connection between vCenter, NSX Manager, and BIG-IQ Cloud. Once this connection is established, BIG-IQ Cloud has the means to communicate with the NSX Manager and vCenter server to deploy and manage BIG-IP VEs and application delivery services.



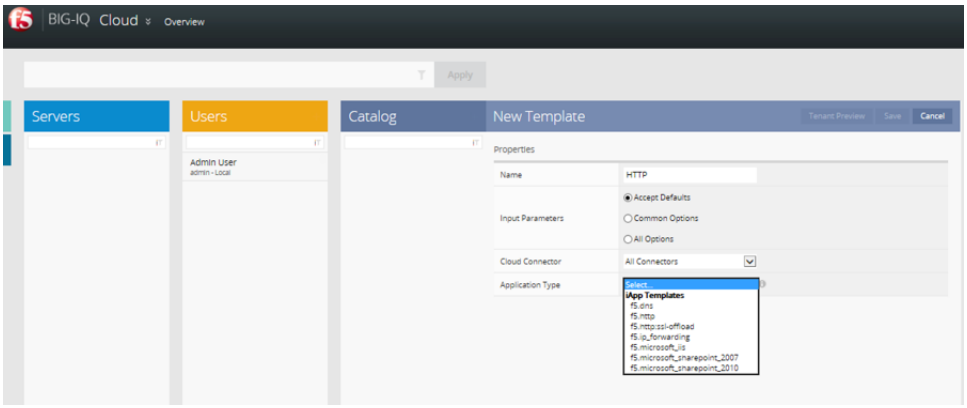
Configuring the NSX Connector in BIG-IQ Cloud

BIG-IP administrators can configure the available BIG-IP VE server images and assign a BIG-IP licensing pool used by a specific VMware NSX Manager. The NSX Connector can also be associated with the BIG-IQ Cloud catalog items to present specific application delivery configurations to NSX Manager. The BIG-IP VE images are stored on an internal web server or datastore for quick and easy access.

Importing iApps into BIG-IQ Cloud Catalog

F5 iApps provides the application-centric configuration capabilities for NSX-integrated BIG-IP ADCs. In order to populate the initial catalog of available iApp application templates within BIG-IQ Cloud, the administrator can deploy a new BIG-IP VE or discover an existing BIG-IP appliance containing the desired iApp(s).

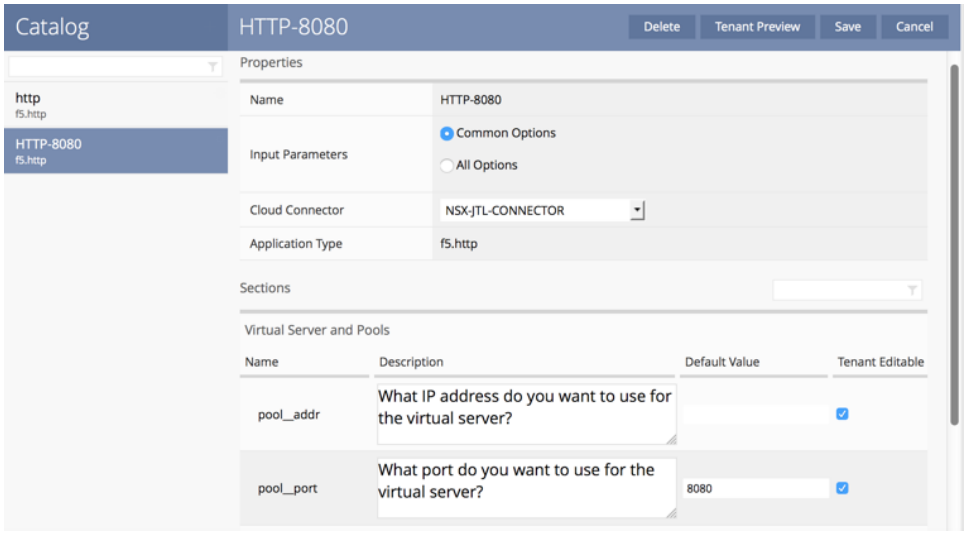
Once the BIG-IP system is discovered by BIG-IQ Cloud, any iApps on the deployed BIG-IP VE are automatically imported and available for customization and publishing through the BIG-IQ Cloud administrative console.



Imported iApp Templates appear in BIG-IQ Cloud catalog.

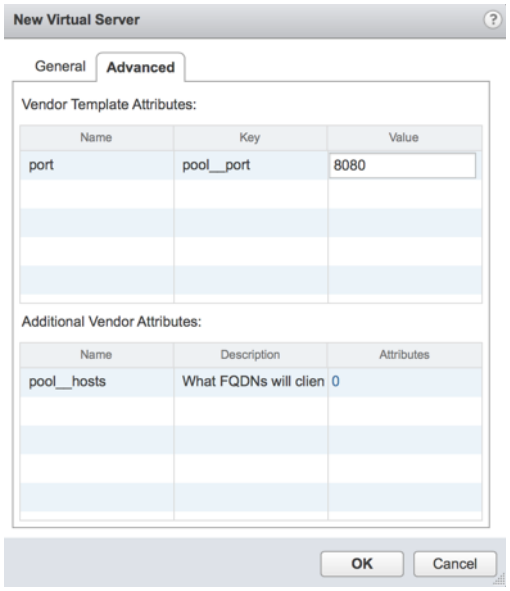
Creating Additional Catalog Items

Administrators can create customized catalog items that can be assigned to specific NSX Managers or available to all NSX Managers integrated with the BIG-IQ Cloud implementation. Configuration options can be marked as tenant-editable, allowing the consumer to choose the desired and permitted configuration settings for the application delivery service.



Setting tenant-editable options when configuring a catalog item in BIG-IQ Cloud.

BIG-IQ Cloud presents catalog items as an application profile through the NSX Administration User Interface. vSphere or NSX administrators can choose an application profile during the configuration of the NSX virtual server through the NSX Administration User Interface. NSX and vSphere administrators can also adjust any tenant-editable settings through the Advanced Vendor Template Attributes tab of the NSX virtual server.



Tenant-editable settings configurable from the advanced tab.

Deploying and Configuring an NSX-Managed BIG-IP VE

With the BIG-IQ Cloud catalog offerings in place to facilitate BIG-IP provisioning and configuration operations, it is time to deploy and configure NSX-managed load balancing services using a BIG-IP VE. There are three steps to deploy and configure a BIG-IP VE with the vCenter Web Client:

- Deploy NSX Edge
- Configure the NSX Edge for load balancing service insertion using a BIG-IP VE
- Configuration of the load balancing virtual server and pools

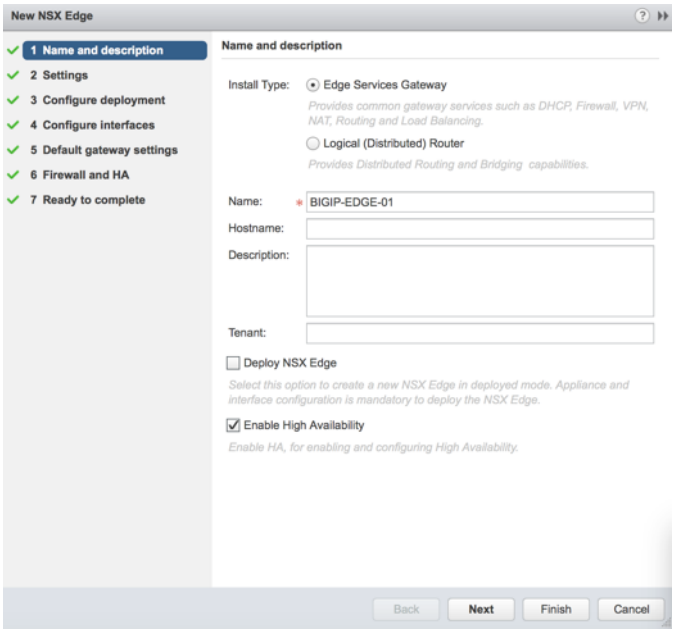
It's important to note that BIG-IQ Cloud catalog items supporting BIG-IP LTM functionality will work "out of the box" with an NSX-integrated BIG-IP VE. iApps that leverage other BIG-IP functionality require additional configuration outside of the vCenter Web Client and are not included in this design guide.

Deploying NSX Edge

The first step in deploying a new BIG-IP VE with VMware NSX is to deploy NSX Edge. There are two different configuration options when initially deploying BIG-IP load balancing services with NSX—"Deployed" and "Undeployed" modes.

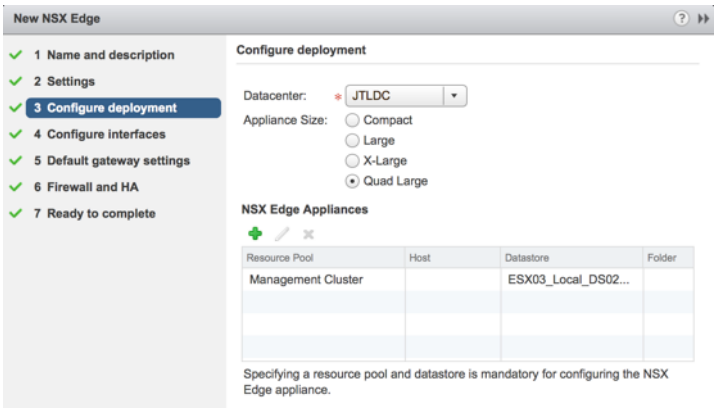
NSX Edge Undeployed Mode

When choosing Undeployed mode, the NSX Edge provides a "shell" for NSX to deploy and manage the BIG-IP using the vCenter Web Client. This configuration option can be used when there is a requirement to deploy only BIG-IP load balancing services; other NSX Edge services (i.e., Distributed Firewall) are not available.



Deploying NSX Edge for use with BIG-IP LTM integration in Undeployed mode.

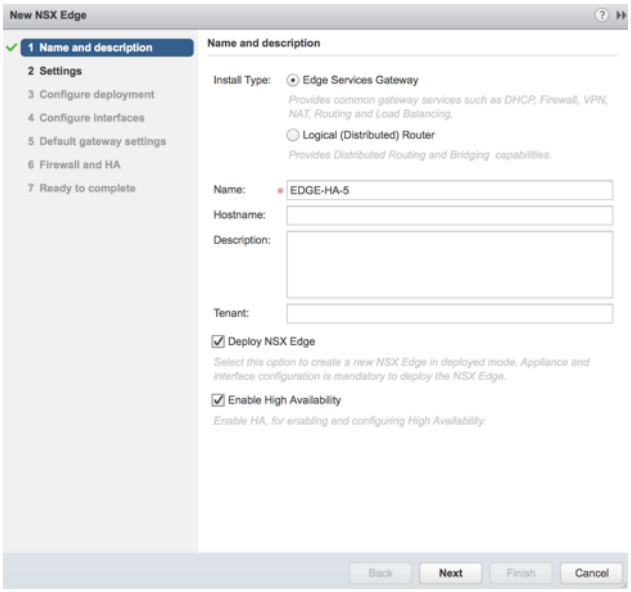
NSX administrators have the choice to deploy appliances in standalone or high-availability configurations, and what cluster resources will be used for the BIG-IP VEs. This step is repeated for each standalone or high-availability pair of BIG-IP VEs deployed through NSX.



Setting the deployment location for the BIG-IP VEs.

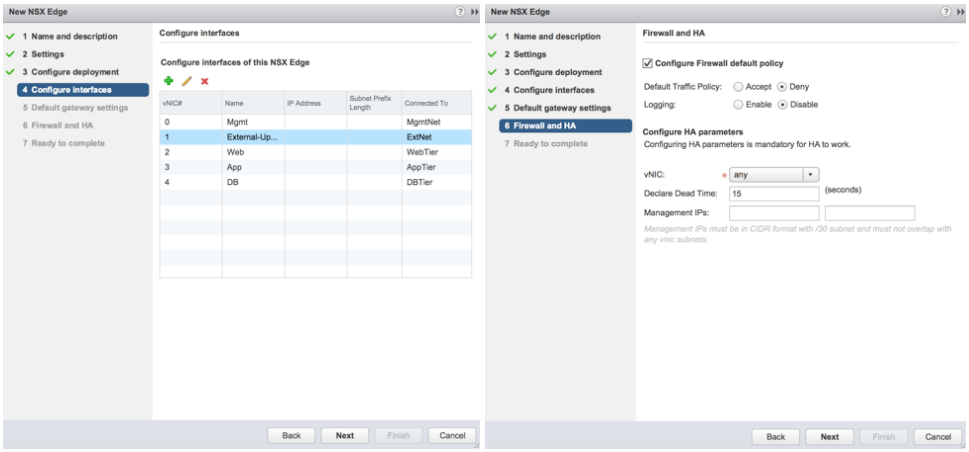
NSX Edge Deployed Mode

Using NSX Edge's Deployed mode will deploy NSX Edge Services Gateway appliances in addition to the BIG-IP VEs. This configuration option can be used when there is a requirement to leverage other NSX Edge Services (i.e., VPN, distributed firewall, etc.) in conjunction with BIG-IP load balancing services.



Deploying NSX Edge for use with BIG-IP integration in Deployed mode.

Similar to the Undeployed Mode, vSphere or NSX administrator can select whether both of the BIG-IP and NSX Edge appliances are deployed as standalone virtual machines or as a highly available pair of VMs.

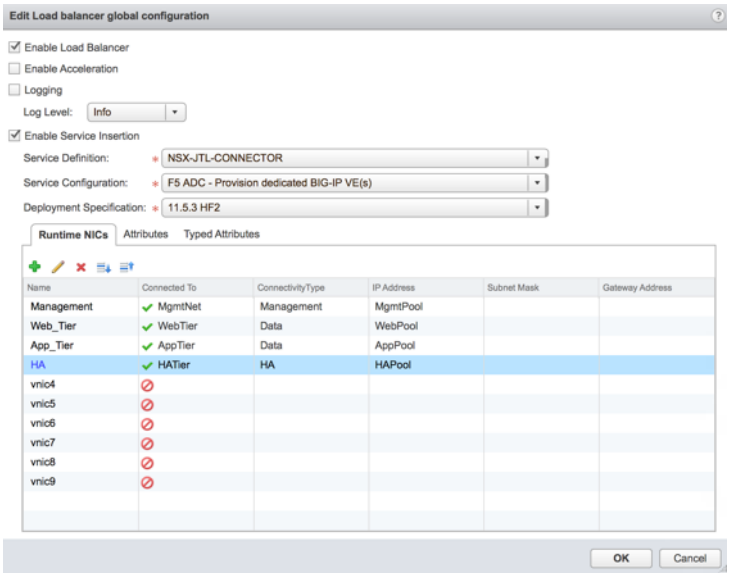


Configuring interfaces and firewall/HA options when deploying NSX Edge in Deployed mode.

Additional configurations are required for Deployed mode, including cluster resource selection, NSX Edge interface configuration, and whether other services (i.e., firewall and NSX Edge HA settings) should be enabled and used for the NSX Edge Services Gateway. This step is repeated for each standalone or high-availability pair of BIG-IP VEs deployed through NSX.

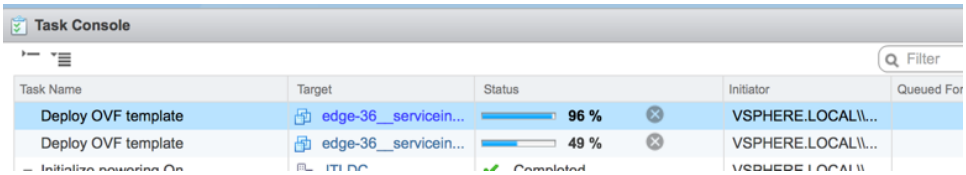
Configure and Deploy BIG-IP VEs for NSX Service Insertion

Once the NSX Edge has been configured, the next step is to configure the NSX Edge for service insertion. This step includes the configuration of network interfaces used by the BIG-IP VE, selecting the server image of the BIG-IP VE that will be deployed, and choosing the NSX Connector that will be used to provision and initially configure the BIG-IP VE.



Configuring service insertion properties for a BIG-IP VE (network interfaces, the BIG-IP VE image, etc.)

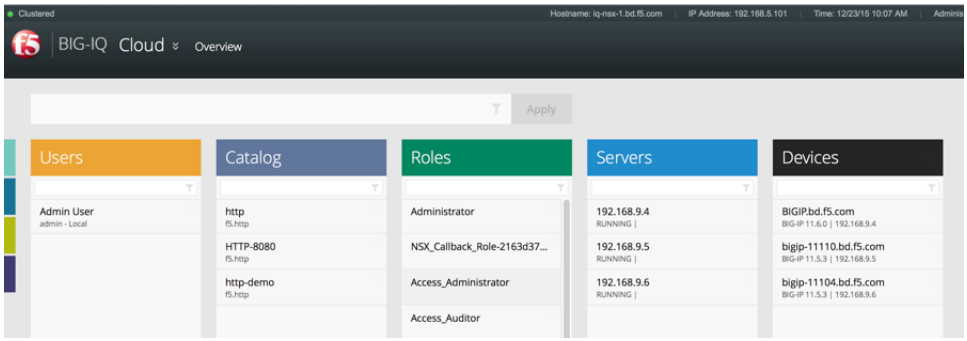
Once the consumer sets up the network interfaces and the remaining settings for service insertion, NSX Manager will deploy the BIG-IP VE, which will power on and begin initial configuration.



NSX deploying an HA Pair of BIG-IP VEs.

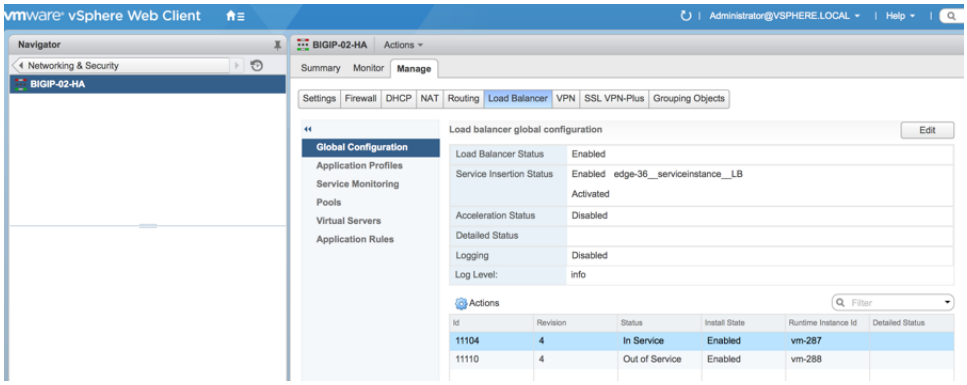
Upon initial boot, the BIG-IP VE management network interface will receive an IP address from DHCP. The DHCP address is used by BIG-IQ Cloud's device discovery process. Once the connection is established between BIG-IQ Cloud and the NSX-managed BIG-IP VE, BIG-IQ Cloud facilitates the initial setup and configuration by providing the designated IP addresses from the IP Pools, BIG-IP license allocation, and other necessary steps needed to complete BIG-IP device setup.

BIG-IQ Cloud will display the current status of the virtual appliances in the BIG-IQ Cloud administrator interface during the configuration process.



Viewing deployed BIG-IP VEs through BIG-IQ Cloud web interface.

The consumer can also view the status of the BIG-IP VEs through the vCenter Web Client. In this example, the administrator interface shows two BIG-IP VEs deployed in a high-availability pair, with the virtual machine marked "In Service" as the active BIG-IP node. The one marked "Out of Service" is the standby BIG-IP node.

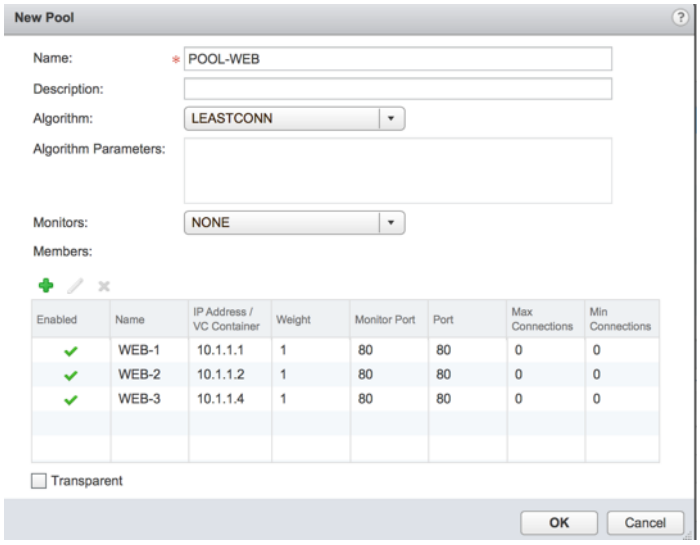


NSX Administration User Interface shows status of integrated BIG-IP VEs.

Configure Load Balancing Virtual Servers and Pools

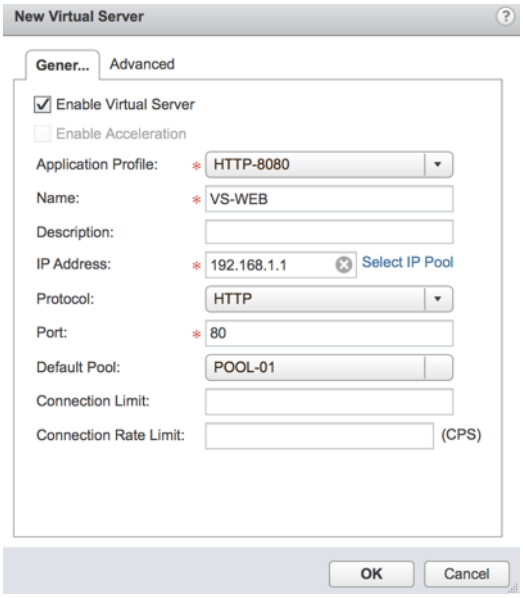
After the provisioning and configuration of BIG-IP VEs, it is time to configure the virtual servers and server resource pools in order to load balance application services. Service consumers will configure the BIG-IP load balancer settings (virtual servers, pool members, SSL certificates, advanced configurations, etc.) through a combination of standard NSX load balancing configuration options and advanced application profile settings provided by BIG-IQ Cloud catalog items.

Basic configuration settings, such as the virtual server's IP address, pool members, and the application profile are configured using the load balancing portion of the NSX Administration User Interface. In this example, we've configured a pool of three web servers.



Configuring pools within the NSX Administration User Interface.

This pool of web servers will be linked to a virtual server. Users will point their web browsers to the virtual server IP address and be routed to the appropriate web server based on the load balancing algorithm defined in the BIG-IQ Cloud catalog item.

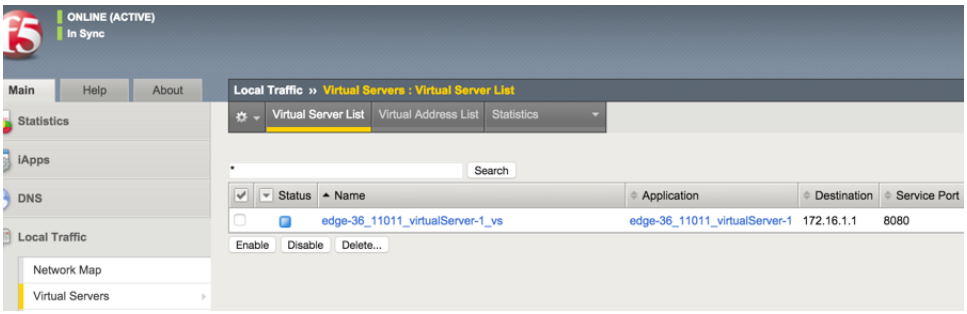


Configuring a virtual server in the NSX Administration User Interface.

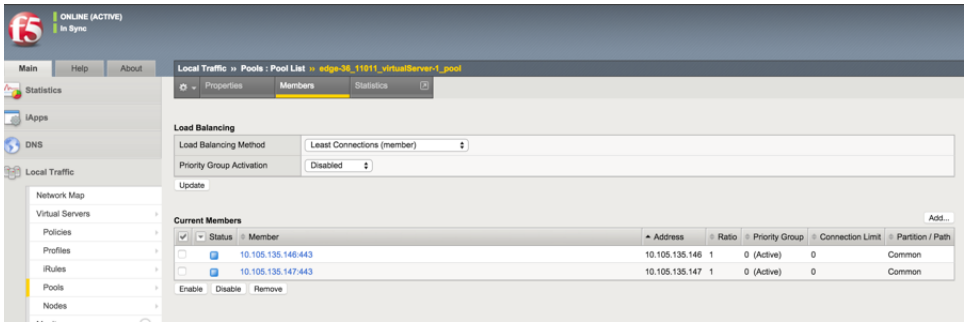
Creating a new virtual server and selecting the application profile means that the user is essentially selecting the BIG-IQ Cloud catalog item. Application profiles deliver application-centric configuration capabilities, including persistence, application performance/optimization, and SSL offload. BIG-IQ Cloud catalog settings are configured as advanced options using the Vendor Template Attributes.

Configuring application-specific properties used by BIG-IP LTM.

Once the pool and virtual server configurations are completed in the NSX Administration User Interface, the configuration is propagated to the BIG-IP VEs and is ready for use.



NSX-managed virtual server as it appears on the BIG-IP platform.



Virtual servers and pools configured through the NSX Administration User Interface displayed on the BIG-IP platform.

Any changes made to the pool or virtual server using the NSX Administration User Interface are automatically updated on the BIG-IP VEs. Changes made directly to the BIG-IP VEs are not synchronized or preserved by NSX Manager.

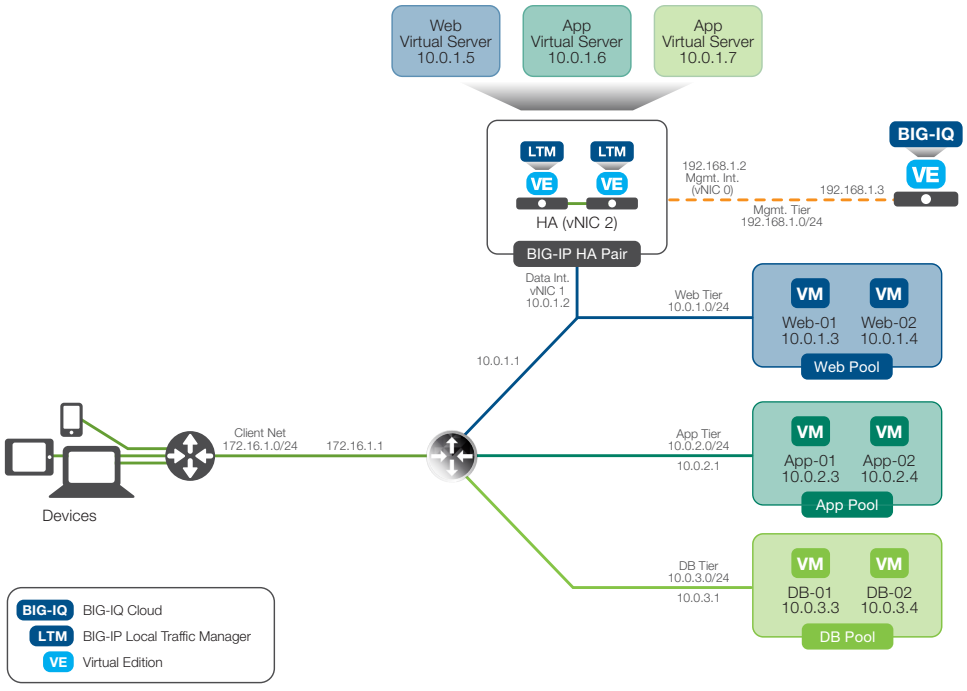
F5 BIG-IP/VMware NSX Configuration Topologies and Examples

BIG-IP ADCs and VMware NSX offer flexible configuration/topology options that help customers align with various network, security and/or perimeter zone configurations. Customers can “plug in” the available BIG-IP virtual edition network interfaces into any NSX-accessible port group or logical switch. This flexibility allows customers to place BIG-IP ADCs in their desired location on the network based on their topology, security posture, etc.

Three different configurations/topologies are available when using the integrated BIG-IP/VMware NSX solution. This next section outlines each baseline configuration and describes some common configurations typically used with BIG-IP and NSX deployments.

One Data Network Interface

The one data network interface configuration provides a single BIG-IP network interface that’s connected to an NSX-aware network and is used to pass production data between the application resources in the data center and the BIG-IP VEs.



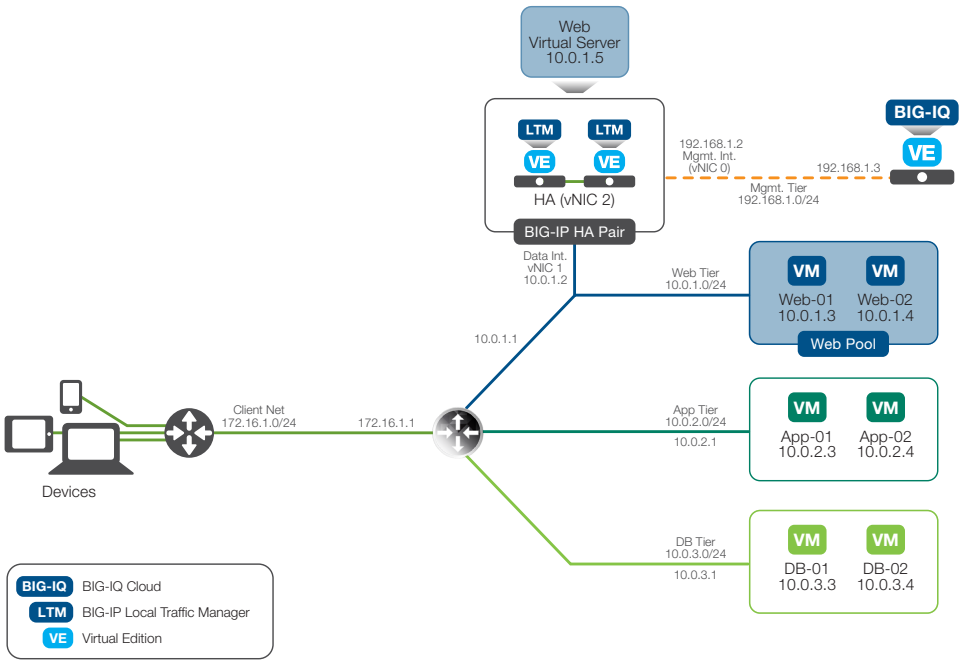
One interface deployment example.

The virtual server IP addresses exist on the same network where the BIG-IP VE data interface is connected. Back-end application resources that are front-ended by BIG-IP service management can exist either on the same segment as the BIG-IP data network or on another accessible network segment in the data center. This single interface configuration topology supports both redundant and standalone implementations of BIG-IP VEs.

This configuration is simple to set up and is typically seen in smaller network environments where back-end application services and resources exist on a single network segment or in data centers where multi-segment networks are accessible through a physical or virtual router. It can also be used to load balance internal application services when the BIG-IP VE is connected to a DMZ network interface or segment.

One Data Network Interface—Topology Example

This diagram depicts an example of a highly-available BIG-IP/VMware NSX-integrated deployment using a single network interface. The BIG-IP data interface is connected to the Web Tier internal network and provides BIG-IP LTM services for a pool of web servers on the internal network. The virtual server's IP is created on the Web Tier network. It manages a pool of web servers that are also part of the Web Tier network.



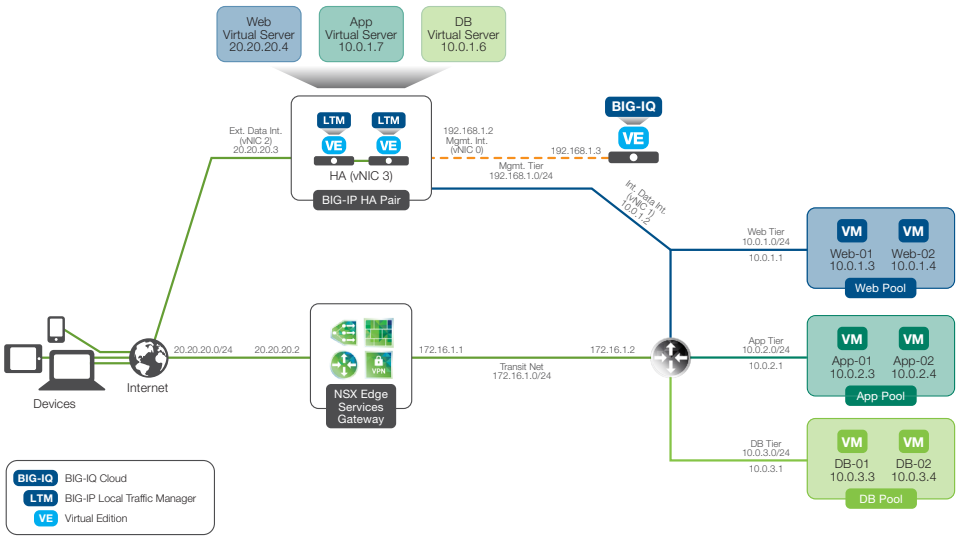
One interface off NSX Edge (DMZ) example—Web Tier from the DMZ.

Requests are made to the web servers from a client. The web server request is forwarded to the virtual server IP managed by the BIG-IP (10.0.1.5). The BIG-IP VEs then make a load balancing decision on which node to send the request to—based on the designated load balancing algorithm (10.0.1.3).

Once the decision is made, the BIG-IP VEs will make a request to the designated web server using BIG-IP full-proxy architecture. The request is sent to the appropriate web server node on the Web-Tier network segment (10.0.1.3). The web servers will then leverage existing NSX traffic routes available through the Distributed Logical Router to get to the App and Database server segments.

Two Data Network Interface

The two data network interface configuration provides two BIG-IP network interfaces that are connected to NSX-aware networks, passing production data between the application resources and the BIG-IP VEs. The two-interface configuration provides various options for connectivity, including direct connections to the Internet, back-end NSX networks, and/or an NSX Edge.



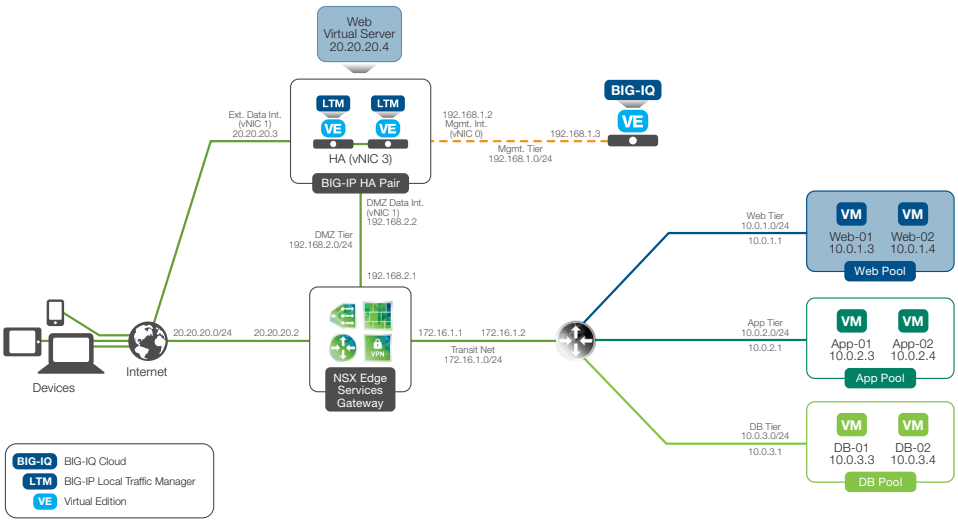
Two-interface topology.

As with the single interface configuration, the virtual server IP addresses for a two-interface configuration exist on the same networks where both BIG-IP data interfaces are connected. Back-end application servers that are front-ended by the BIG-IP VE can exist either on the same segments as the BIG-IP data network or on another accessible network segment in the data center. This dual interface configuration topology supports both redundant and standalone implementations of the BIG-IP VE.

Two Data Network Interface—Topology Example

The configuration shown below is typically seen in network environments where there is a need to load balance web services that exist on a trusted networks or in data centers where other networks are accessible through a physical or virtual router. It can also be used to load balance and secure internal application services when the BIG-IP is connected to a DMZ or Internet-attached network segment.

This diagram is an example of a high-availability BIG-IP/VMware NSX integrated deployment using two network interfaces. One of the BIG-IP VE data interfaces is connected to a DMZ network/interface from an NSX Edge, securing the network perimeter. A second network interface is directly connected to the Internet providing access to front-end web server resources.



Connecting the BIG-IP VE to NSX Edge and Internet using two network interfaces.

The BIG-IP VEs provide BIG-IP LTM services for a pool of Web servers that exist on a network segment inside the data center. The web server's virtual IP exists on the external Internet network. The NSX Edge controls routing and firewall services between the various internal application components and external clients.

In this example, requests are made to the web servers from a client. The BIG-IP virtual server (20.20.20.4) receives the request. Next, the BIG-IP VEs make a load balancing decision on which node to send the request to based on the designated load balancing algorithm (10.0.1.3). Once the decision is made, the BIG-IP VEs perform source address translation and make a request to the designated web server using BIG-IP full-proxy architecture.

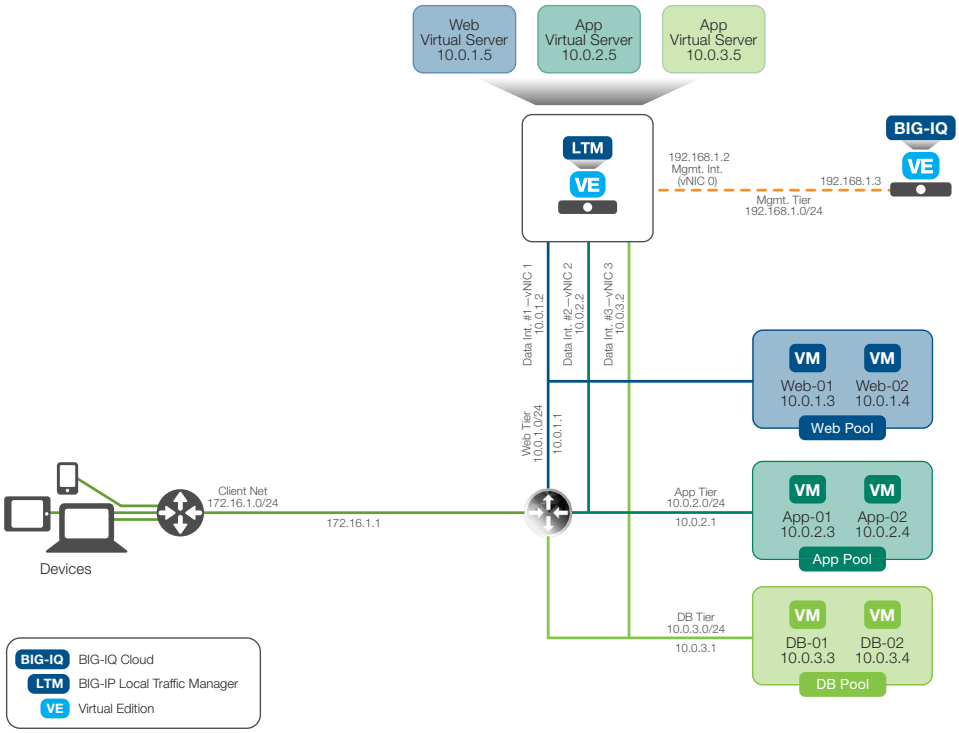
The request is routed back to the NSX Edge (192.168.2.1), where it is then sent to appropriate web server node on the Web-Tier network segment (10.0.1.3). The web servers then leverage existing NSX traffic routes to get to the App and Database server segments. All of the network traffic exchanged between endpoints, the BIG-IP VEs, and internal application servers can be secured using firewall rules enforced by the NSX Edge.

Three Data Network Interface

The three data network interface configuration provides multiple BIG-IP network interfaces that are connected to NSX-aware networks used to pass production data between application resources and the BIG-IP VEs.

The three-interface configuration provides various options for connectivity, including direct connections to the Internet, back-end NSX networks, and/or an NSX Edge. This three-interface configuration topology ***supports a standalone BIG-IP VE implementation only***, since the NSX/BIG-IP VE integration currently supports configurations with two to four network interfaces. With all four interfaces being used for data and management network traffic, there are no additional network interfaces for BIG-IP high availability.

This configuration is typically used for non-production environments (i.e., testing and validating multi-tier or microservice-based application load balancing functionality) and is ***not recommended*** for production implementations of VMware NSX/BIG-IP.



Three-interface topology.

With the three-interface configuration, the virtual server IP addresses exist on the same networks where all three of the BIG-IP VE data interfaces are connected. Back-end application servers that are front-ended by the BIG-IP VE can exist either on the same segments as the BIG-IP data network or on another accessible network segment in the data center. This configuration is typically seen in network environments that have a need to test/validate load balancing multi-tier or microservice-type applications.

Three Data Network Interface—Topology Example

Using the above diagram as an example, the BIG-IP virtual edition is deployed in a standalone configuration; the BIG-IP VE is directly connected to the Web-Tier, App-Tier and DB-Tier network segments. Client requests from the 172.16.1.0 network are made to the web servers. The BIG-IP web virtual server IP (10.0.1.5) receives a request. Next, the BIG-IP VE will make a load balancing decision on which node to send the request to based on the designated load balancing algorithm (10.0.1.3).

Once the decision is made, the BIG-IP VE makes a request to the designated web server using the BIG-IP system's full-proxy architecture and send it to appropriate web server node on the Web-Tier network segment (10.0.1.3). Requests from the web servers that are

destined for the application and database servers are routed to the BIG-IP App Tier virtual server IP (10.0.2.5) or the database virtual server IP (10.0.3.5). This communication does not need to cross the distributed logical router, since the BIG-IP VE is directly connected to the DB-Tier and App-Tier network segment.

Best Practices and Configuration Recommendations

In this section, we'll outline some best practices and recommendations for configuring and deploying BIG-IP application delivery services with VMware NSX. These guidelines help ensure a successful and seamless deployment of NSX-integrated load balancing services with BIG-IP systems.

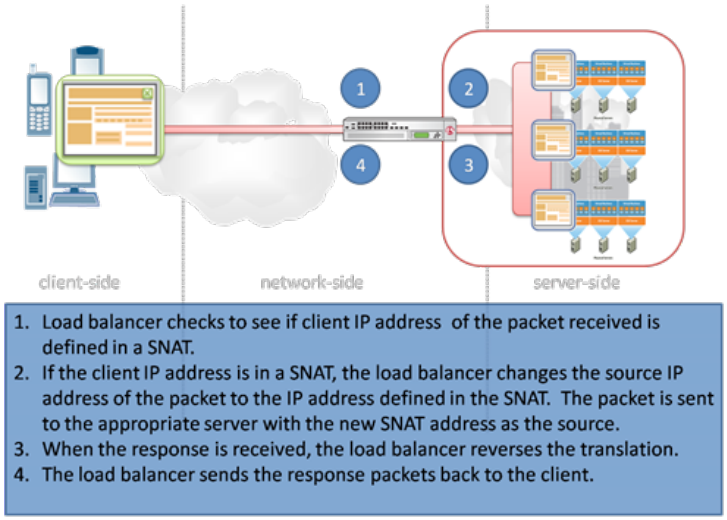
General Networking Considerations

Using BIG-IP VEs involves several networking considerations. Network virtualization supports many different networking segmentation and configuration standards—such as VXLANs, VLANs, and logical/virtual switching. Follow these simple configuration best practices when designing and deploying BIG-IP virtual editions.

- BIG-IP VEs require a minimum of two network interfaces and two unique port groups/logical switches for basic functionality. One interface and network is used for management network traffic and the other for data network traffic.
- It is not uncommon to see physical network configurations offering high-capacity, high-speed network interfaces that are used for consolidating all network communications. If using existing network links, make sure they are not oversubscribed and have ample bandwidth to handle the required network traffic volume.
- Use a dedicated network for BIG-IP VE high-availability services. It's important to keep network traffic that needs to be highly available segmented so production traffic or a busy internal network connection does not affect the device (or instance) heartbeat used to detect a failover.

Secure Network Address Translation

When you need to ensure that server responses always return through the BIG-IP system, or when you want to hide the source addresses of server-initiated requests from external devices, you can implement Secure Network Address Translation (SNAT). SNAT is a BIG-IP system feature that translates the source IP address within a connection to a BIG-IP system IP address that you define. The destination node then uses that new source address as its destination address when responding to the request.



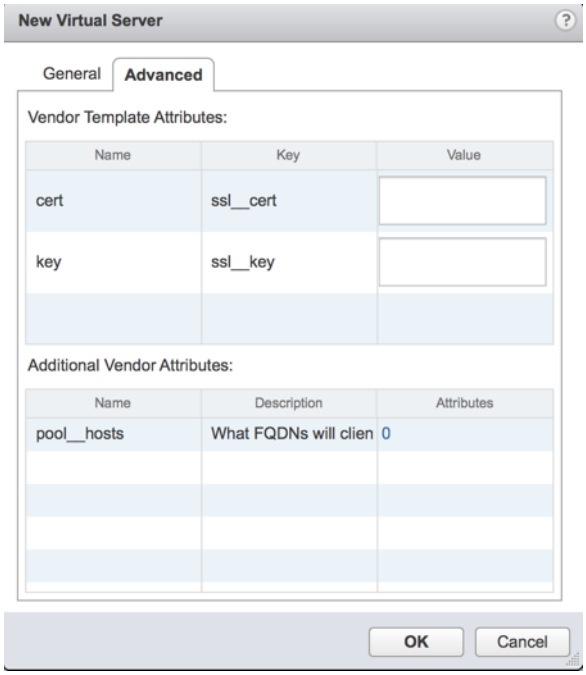
Secure Network Address Translation (SNAT) overview.

Simply put, if your default gateway of the pool member points to a device other than the BIG-IP ADC, you will need SNAT to ensure the return traffic goes through the BIG-IP ADC. If your default gateway of the pool member points to the BIG-IP ADC, SNAT is not required, since the traffic will be routed through the BIG-IP ADC.

SNAT can be enabled using a BIG-IQ Cloud catalog item. The BIG-IP administrator can enable ability to manage SNAT as a tenant-editable option in the BIG-IQ Cloud catalog item.

SSL Certificates

SSL certificate management is an important part of the NSX and BIG-IP integration, especially when performing traffic inspection or SSL offload with the BIG-IP virtual edition. NSX requires the certificate and private key to be imported separately as Advanced Vendor Template Attributes when configuring the virtual server. The private key's password must also be removed from the encrypted private key.



Importing the SSL certificate and private key.

Once the certificates and keys are uploaded to the BIG-IP VE, the application profile (iApp) takes care of creating the necessary SSL profiles and uploads the certificates and keys to the BIG-IP VE. If there is a requirement for SSL certificate chaining (i.e., Chaining a root and/or intermediate Certificate Authority certificate(s) with an SSL web certificate), all of the certificates must be combined together into a single file. This combined certificate file is then uploaded into the SSL Certificate field in the Advanced Vendor Settings of the NSX virtual server.

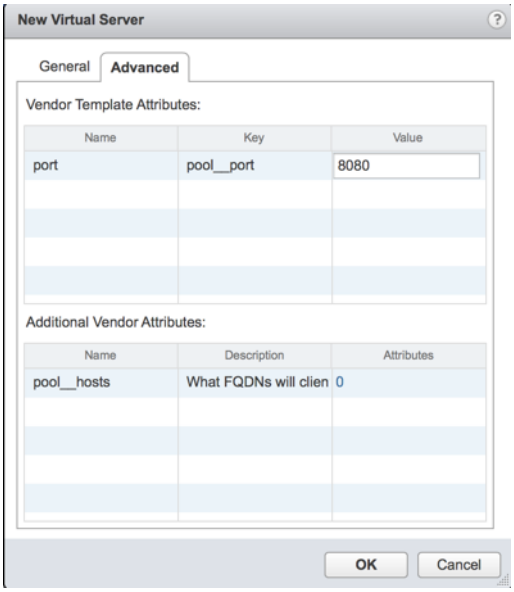
Virtual Server/Pool Management Considerations

The NSX Administration User Interface is where virtual server and server pool configuration occurs. This includes the configuration of virtual server IP addresses, pool members, the application profile, and other vendor-editable settings used to configure BIG-IP advanced load balancing options (i.e., SSL certificates, session persistence, etc.).

It is important to differentiate which configuration options are required/used by either BIG-IP LTM or NSX. The following list are fields that are used and required by both:

- **Virtual Server—**IP address (manually entered or from IP Pool)
- **Virtual Server—**Choosing application profile
- **Pool Member—**IP address for servers (manually entered or from vCenter object)

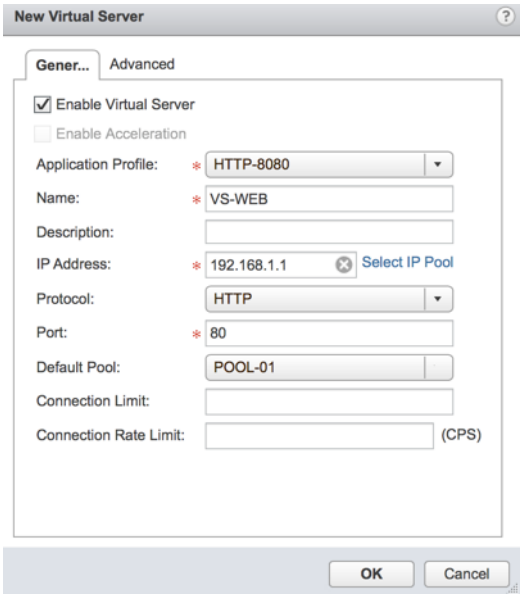
All other configuration options are managed through the New Virtual Server Advanced tab > Vendor Template Attributes (i.e., member server ports, virtual server ports, monitoring, persistence, SSL certificates, etc.).



Tenant-editable properties for a virtual server pool.

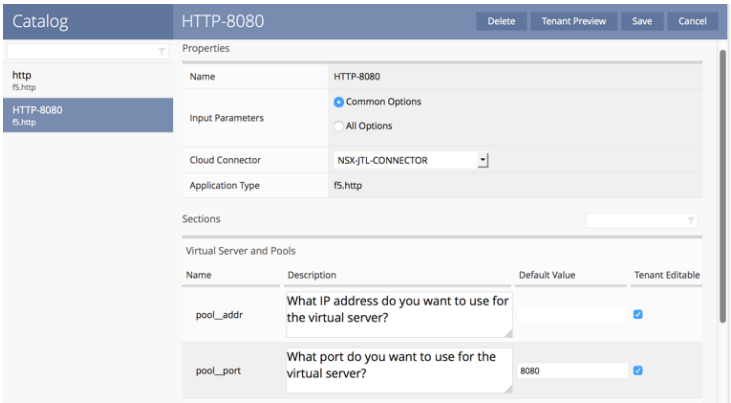
Although the vCenter/NSX Web Client UI requires configuration options beyond the three mentioned, these configuration options are not used by the BIG-IP ADC. For example, the New Virtual Server General tab requires a port to be specified when creating a new virtual server in the NSX Administration User Interface.

The value specified in the NSX Administration UI is not used by the BIG-IP ADC. Instead, the configuration value that is specified as part of the Application Profile, either as a default value, or as a "tenant editable" if the value used by the BIG-IP ADC.



Virtual server port setting not used by BIG-IP LTM but required by NSX.

In order to provide the vSphere or NSX administrator the ability to edit these settings, The BIG-IP administrator needs to set any customizable value as "tenant editable" in the BIG-IQ Cloud catalog item.



Configuring application profile properties to be "tenant-editable" in BIG-IQ Cloud.

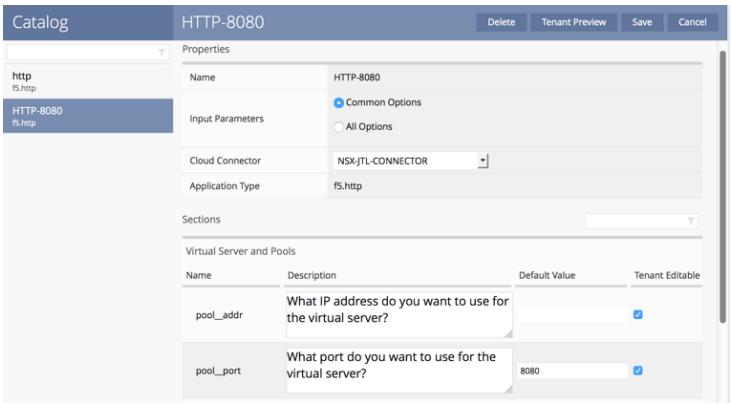
When the NSX or vSphere administrator selects the Application Profile on the virtual server, the tenant-editable settings will automatically populate in the Advanced Vendor Template Attributes tab of the virtual server.

Configuring Advanced Template Options

F5 iApps is the cornerstone for providing simple, application-centric configurations for BIG-IP systems. BIG-IQ Cloud functions as the translation point between iApps and VMware provider templates. BIG-IQ Cloud begins the process by translating an iApp and dividing it into two parts. The first part is filled out by the BIG-IP administrator or service provider as part of defining a BIG-IP catalog item. The remaining runtime portion of the iApp, once standardized, is reflected in the VMware NSX Administration User Interface.

Many of the necessary values used by the BIG-IP load balancing services are configured through the profile's advanced vendor settings. Therefore, the BIG-IP administrator must ensure all the necessary and authorized configuration options are properly configured and tenant-editable when adding the iApp to the BIG-IQ Cloud Catalog.

If the BIG-IP administrator needs to set default values (i.e., port set to 8080 for pool or virtual server), these types of settings must be set when adding an item to the catalog in BIG-IQ Cloud. The BIG-IP administrator can also mark any consumer-editable fields as "tenant editable" while adding the iApp configuration to the catalog. This provides an option for customers to change these values if necessary.



Customizing tenant-editable values (i.e., setting the virtual server's port to 8080).

This guide explains the deployment and integration of iApps that come pre-installed on the BIG-IP system. While other iApps can be utilized by adding them to an already-imported BIG-IP ADC, the BIG-IP administrator will need to manually import the iApp to all the relevant BIG-IP devices after those devices are deployed through NSX. It is recommended that the BIG-IP administrator manually add a BIG-IP device that is not managed by NSX as a platform for BIG-IQ Cloud's iApp discovery process.

If custom iApps (not shipped with BIG-IP LTM) are needed, these iApps may require additional compute/memory resources for any relevant BIG-IP VEs. In addition, these non-standard iApps must be manually copied to each BIG-IP VE and may require the modules to be enabled and licensed outside of the NSX Administration User Interface.

vSphere/vCenter/NSX General Design Considerations

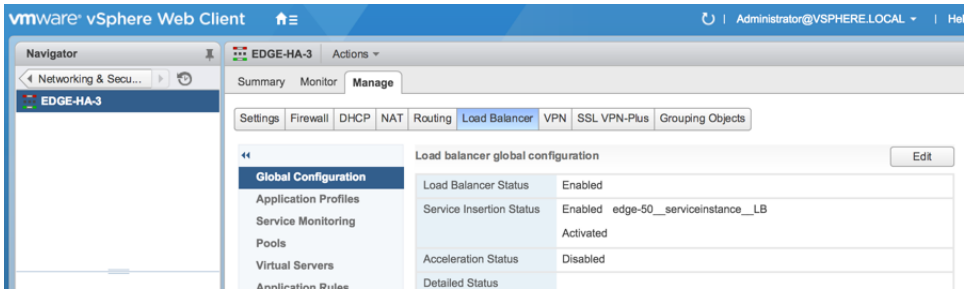
As part of the integration, important NSX and vSphere/vCenter elements should be considered when designing and deploying NSX/BIG-IP integrated solutions. The following table captures some of these design considerations and requirements.

Network Time Protocol (NTP) source	Required for BIG-IP and BIG-IQ Cloud.	Clock skew between BIG-IP and BIG-IQ Cloud beyond 5 minutes (30000 ms) may prevent the devices from operating correctly.
NSX-enabled IP Pools	Required for BIG-IP/NSX-integrated deployment.	NSX and BIG-IP integration sets the management network's default gateway based on IP Pool configuration.
Number of IP Pools required	1 IP Pool for management (Required) 1 HA Network Pool (Optional) 1 IP Pool for each data network (Required)	Each IP pool must contain a default gateway. IP address ranges must have enough capacity to support the desired number of BIG-IP and applicable services/applications.

DHCP services	Required on management network.	Required by the BIG-IP VE to establish initial connection with BIG-IQ Cloud during deployment. DHCP address no longer required after management network IP address issued from NSX IP Pool. Ensure a short DHCP lease time and ample DHCP IP addresses in the pool to support the BIG-IP VE lifecycle management.
VMware tools	Required on pool member virtual machines if vCenter object is used for selecting pool members.	Uses IP address from vCenter object integration for BIG-IP/NSX load balancing.

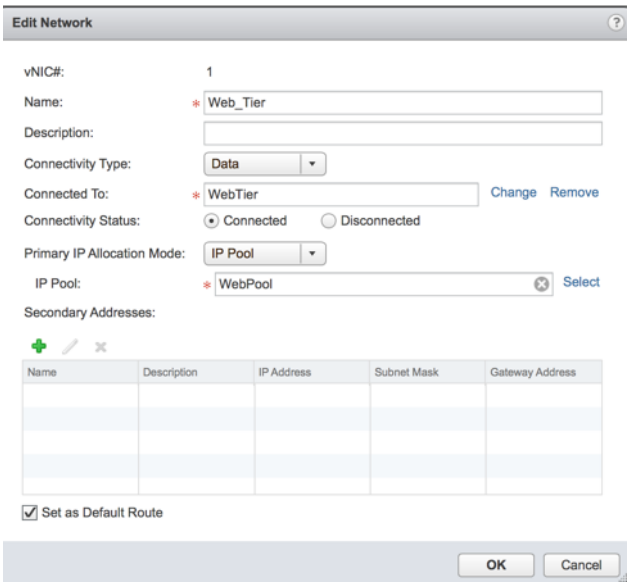
BIG-IP Load Balancing Service Insertion—Network Interface Considerations

Proper networking interface configuration is key when deploying and managing BIG-IP virtual editions managed through VMware NSX. The network interface configurations can be found by selecting the **NSX Edge Services Gateway** in the NSX Administration User Interface and clicking **Manage**. Click on **Load Balancing** and then click **Edit**.



Enabling load balancing service insertion for BIG-IP and BIG-IQ Cloud platforms.

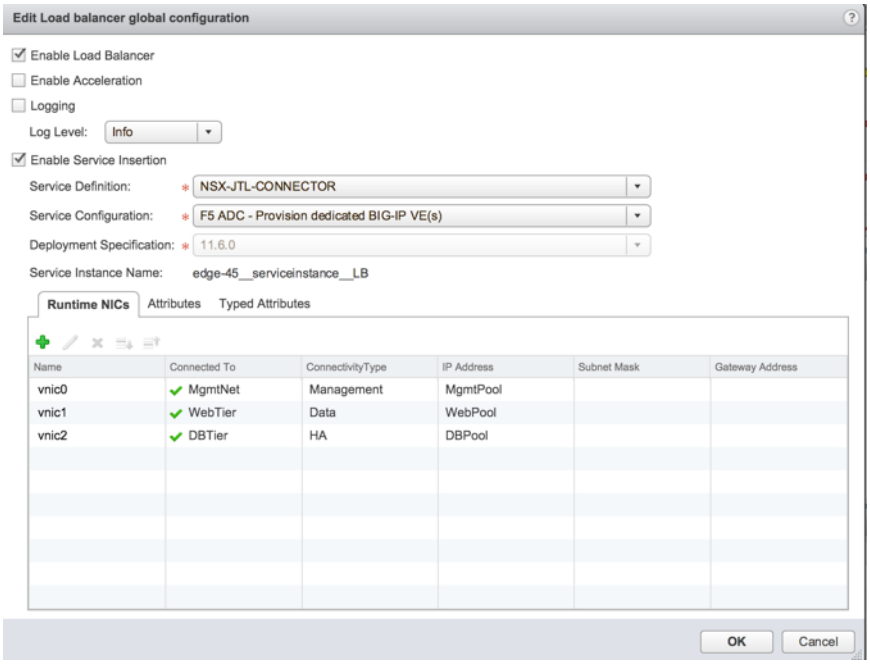
Follow these steps to set up BIG-IP networking using the NSX Administration User Interface:



Configuring data plane network interface used by the BIG-IP VE.

- Select the appropriate connectivity type (**Data** for data traffic, **HA** for high-availability traffic, and **Management** for management traffic).
- Select the appropriate port group or logical switch.
- Set the connectivity status to **Connected**.
- Select **IP Pool** as the Primary IP Address Allocation Mode.
- Choose the IP Pool that will allocate the BIG-IP VE's floating and non-floating IP addresses.
- Select **Set as Default Route** if the data interface will be used as a default gateway for the BIG-IP VE (the default route will be used if the **Set as Default Route** option is selected for the data interface).

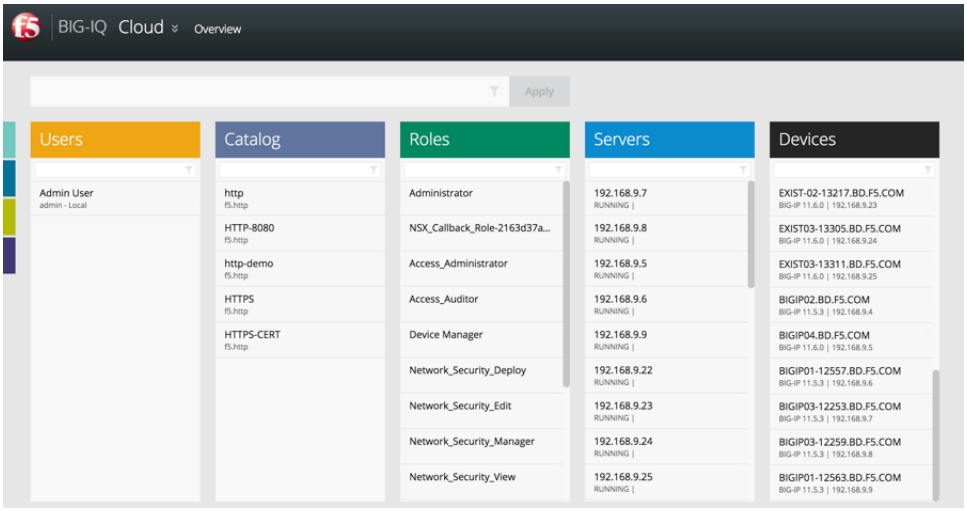
The following image is an example of a three-interface configuration, using one interface for data, one interface for management, and a single interface for HA. The management interface must also be listed first in the Runtime NICs list.



Configuration example for three-interface deployment.

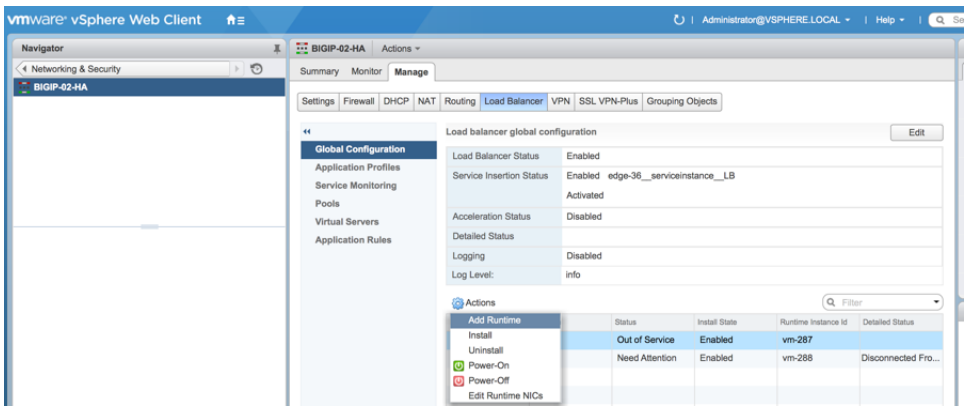
Finally, high-availability deployments of NSX-integrated BIG-IP VEs also require an FQDN to be specified in the Typed Attributes tab. BIG-IQ Cloud will append a unique ID number (from NSX) to the FQDN, and then use this for the BIG-IP hostname.

Setting the FQDN used by the BIG-IP VE.



Deployed BIG-IP VEs with FQDNs as show in the BIG-IQ Cloud management interface.

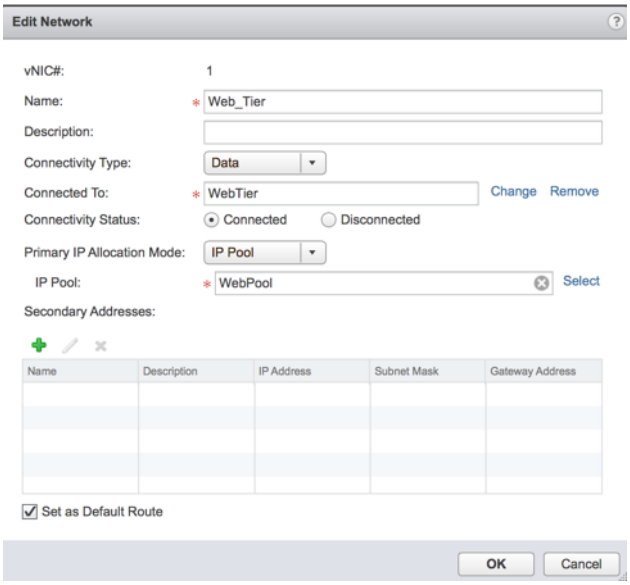
Sometimes, BIG-IP network configurations (i.e., logical switch assignment, default route, IP Pool assignment) may need to be changed. The NSX-integrated BIG-IP VE does not support network interface changes (i.e., port groups, IP Pools, etc.) after the BIG-IP VEs have been deployed. The BIG-IP VEs must first be uninstalled from the NSX Administration User Interface in the vCenter Web Client. Next, a new runtime must be used to reflect the new network configurations.



Deploying a new/replacement BIG-IP VE runtime.

BIG-IP Network Routing

When configuring NSX-integrated BIG-IP VEs, service consumers can specify one of the data interfaces as a default route. This is the only configurable routing option available through the NSX-integrated configuration of BIG-IP VEs. The management interface will automatically use the default gateway specified in the NSX IP Pool assigned to the interface.



Setting the default route used by the BIG-IP VE using the NSX Administration User Interface.

Additional static routes and VLANs can be added manually either through the BIG-IQ Cloud Device Administration interface or the BIG-IP Web Administration interfaces.

Administrator-Triggered and Automatic VM Migrations

One of the key benefits of vSphere is the ability to automatically or manually relocate virtual machines to other hosts or storage. While this feature works well for many workloads, it is not advisable for BIG-IP virtual editions. Moving an active workload that passes network traffic or demands a lot of CPU may cause an interruption in service and potentially drop any connections coming through the BIG-IP VE.

It is recommended that you also ensure controls are in place to prevent BIG-IP VEs from being automatically relocated to other hosts (via vCenter's Distributed Resource Scheduling) or low-performance storage media. If you have to move a BIG-IP VE to another host, you should move it when it is in an idle or stand-by state.

BIG-IP VE General Design Considerations

By default, the NSX-integrated BIG-IP VE deployment configures the virtual machine for two virtual CPUs and 4 GB of virtual memory. The amount of memory or CPU allocated is not configurable through the NSX Administration User Interface and is not impacted by the VM size selected during the provisioning of NSX Edge. Additional memory and CPU cannot be allocated to the BIG-IP VE through the vCenter Web Client once the BIG-IP VE is deployed using NSX/BIG-IQ Cloud integration.

Memory and CPU parameters can be adjusted within the BIG-IP VE image's OVF file itself. If additional CPU or memory is required to support additional BIG-IP modules or higher throughput requirements, the BIG-IP administrator will need to create an additional BIG-IP OVF server image with the adjusted CPU and memory parameters. Refer to the BIG-IQ Cloud product documentation for additional sizing and configuration information.

BIG-IP VEs are deployed with BIG-IP LTM enabled by default in "Nominal" mode. Nominal mode allows the BIG-IP VE to start with the minimal amount of memory, then dynamically allocate more memory to the BIG-IP LTM module as needed. The following table outlines the virtual server configuration maximums and design considerations when using BIG-IP and NSX integration.

Minimum number of network interfaces required for NSX/BIG-IP integration (without HA)	2	1 interface for management network traffic. 1 network interface for data traffic.
Minimum number of network interfaces required for NSX/BIG-IP integration (with HA)	3	1 interface for management network traffic. 1 network interface for data traffic. 1 network interface for HA traffic.
Maximum number of network interfaces and port groups/logical switches available for NSX/BIG-IP integration	4	1 interface for management network traffic. For HA—Up to 2 network interfaces for data traffic. No HA—Up to 3 network interfaces for data traffic if no HA is required. 1 interface for HA network traffic (if required).

Maximum number of virtual servers for compact VM deployment	64	BIG-IP VE vCPU and vRAM settings not affected when setting size of NSX Edge VM.
Maximum number of virtual servers for large VM deployment	64	BIG-IP VE vCPU and vRAM settings not affected when setting size of NSX Edge VM.
Maximum number of virtual servers for extra-large VM deployment	1,024	BIG-IP VE vCPU and vRAM Settings not affected when setting size of NSX Edge VM.
Maximum number of virtual servers for quad large VM deployment	1,024	BIG-IP VE vCPU and vRAM settings not affected when setting size of NSX Edge VM.

BIG-IQ Cloud General Design Considerations

BIG-IQ Cloud is required to facilitate any NSX-related operation initiated through the NSX Administration User Interface. BIG-IQ Cloud can be configured as a standalone appliance for environments that do not require high availability, or can be peered together to provide high availability.

In order to facilitate the integration between NSX and BIG-IP VEs, BIG-IQ Cloud (standalone or clustered) requires the NSX Connector to communicate with both vCenter and NSX Manager. BIG-IQ Cloud clusters and standalone appliances can be connected to VMware NSX/vCenter in the following ways:

- A single BIG-IQ Cloud standalone appliance or cluster can connect to multiple NSX Manager and vCenter Servers.
- A single BIG-IQ Cloud standalone appliance or cluster can connect to a single NSX Manager and vCenter Server.

The following table outlines other configuration maximums and design considerations that BIG-IQ Cloud administrators should consider when designing and deploying BIG-IQ Cloud with NSX-integrated BIG-IP VEs.

Minimum number of network interfaces required for BIG-IQ Cloud VE	1	BIG-IQ Cloud must be able to communicate with the network segment used by the BIG-IP VE management network interface, vCenter, and NSX Manager.
Minimum number of vCPUs and vRAM for each BIG-IQ Cloud VE	2 vCPUs 4 GB vRAM	Recommended for non-production or proof-of-concept environments, or environments that do not require high availability.
Recommended number of vCPUs and vRAM for each BIG-IQ Cloud VE	4 vCPUs 8 GB vRAM	Minimum resources recommended for production environments.
Minimum number of BIG-IQ Cloud VEs for high availability	2	HA requires BIG-IQ Cloud virtual appliances to be configured as HA Peer Group. BIG-IQ Cloud will automatically re-establish communication with NSX after 3 minutes if a BIG-IQ Cloud VE is rendered inoperable. Configuration changes initiated from NSX Manager require at least 1 BIG-IQ Cloud VE to be online.
Maximum number of BIG-IQ Cloud VEs in a BIG-IQ Cloud Cluster	3	N/A
Maximum number of managed BIG-IP VEs for each BIG-IQ Cloud Cluster	Up to 24	Total number of BIG-IP VEs (standalone and HA).
Maximum number of tenants for each BIG-IQ Cloud implementation	12	N/A

BIG-IP license type	BIG-IP Pool Licenses	Allows for dynamic and shared BIG-IP license management when VEs are provisioned/retired. Device-specific licensing not supported for BIG-IP VEs deployed by BIG-IQ Cloud.
BIG-IP VE server image format	OVF + necessary VMDK file(s)	OVA format is not supported. BIG-IP VE OVA files may need to be converted to OVF format if necessary.
BIG-IQ Cloud image repository	Web server using HTTP ESXi Datastore using HTTPS	Used to store NSX-deployable BIG-IP VE server images.
NSX Connector configuration requirements	NSX Manager IP address, username, and password vCenter IP address, username, and password NTP time source DNS server(s) DNS suffix	Required to ensure seamless deployment of BIG-IP VEs through NSX.

vSphere Host Placement and Resource Considerations

One of the benefits of virtualization is the effective management and sharing of compute and memory resources on a host or server. BIG-IP virtual editions depend on ample CPU and memory resources to efficiently process network traffic and application access requests. Any significant resource contention with other VMs may reduce traffic passing through BIG-IP VEs.

While oversubscribing ESXi hosts used for test/development deployments of BIG-IP VEs may be acceptable, oversubscription of memory and CPU for a BIG-IP production deployment is strongly discouraged. For this reason, F5 recommends that the BIG-IP VE instance be placed on a host with ample compute and storage resources. Placing a BIG-IP

VE instance on a busy host may reduce the performance of either the BIG-IP VE or the other applications collocated on the same hypervisor host.

The automated deployment of BIG-IP VEs will also set the appropriate resource reservations. To avoid resource contention, choose to deploy the BIG-IP VEs to a cluster that can meet the following general guidelines for resource management and planning:

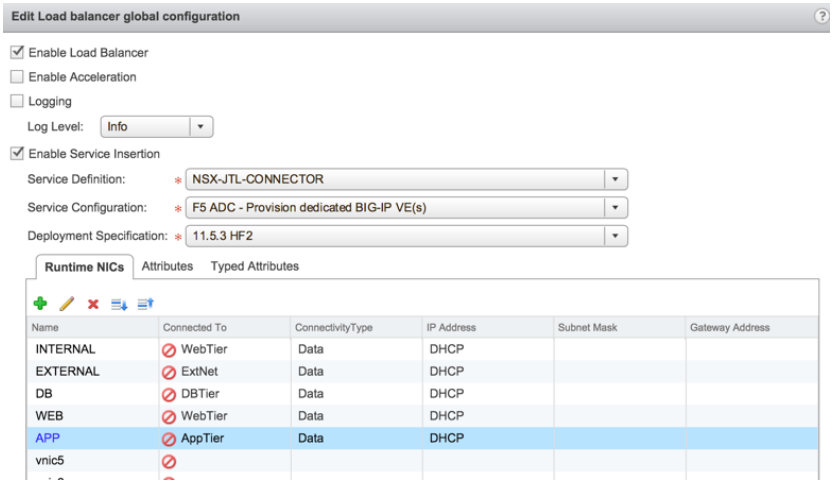
- Plan for 100 percent of memory and CPU capacity for the BIG-IP instance to be reserved.
- For each virtual CPU, plan for a single physical core to be reserved.

For example, if the processor speed of a server is 3 GHz and two virtual CPUs/4 GB of virtual RAM are allocated to the virtual machine, you will need to plan for up to 6000 MHz of CPU and 4 GB of RAM being used on the host for the BIG-IP VE.

NSX Edge Deployed Mode Considerations

vSphere and NSX administrators have the option to use BIG-IP LTM VE for load balancing in conjunction with other networking services offered by NSX Edge (i.e., firewall, VPN, etc.). The BIG-IP load balancing functionality is enabled using NSX load balancing service insertion. The NSX or vSphere administrator can also add BIG-IP load balancing functionality to either a new or existing NSX Edge deployment.

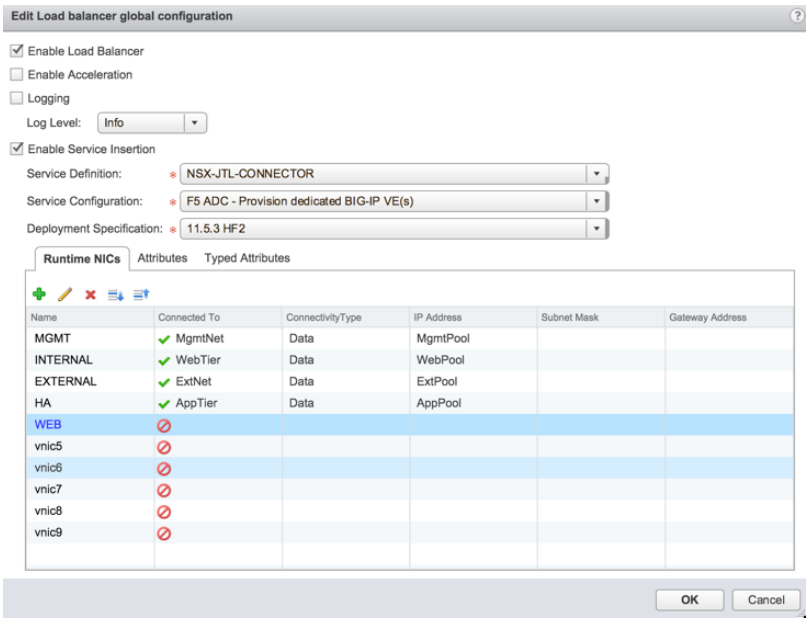
When the load balancing services are enabled on an NSX Edge configured for Deployed mode, the network interface configurations will be migrated to the load balancing configuration screen.



Configuring load balancer network interfaces with NSX Edge in Deployed mode.

NSX and vSphere administrators need to consider the following:

- The management interface must also be listed first in the Runtime NICs list.
- The network interfaces must be properly configured. Refer to the **BIG-IP Load Balancing Service Insertion—Network Interface Considerations** section of the document for the proper network interface settings when using BIG-IP load balancing with VMware NSX.
- The number of network interfaces cannot be more than four. Additional network interfaces can be reset by highlighting the desired network interface and clicking the red **X**.



Reconfigured load balancing network interfaces with NSX Edge in Deployed mode.

BIG-IP Redundancy and vSphere High Availability

The best way to get redundancy with BIG-IP ADCs is to have two instances of the BIG-IP VE deployed. A typical configuration would make one BIG-IP instance active while the other is in standby mode. In the event a host or BIG-IP instance fails, the passive node will immediately take over the work of the active node with minimal or no interruption in service. For BIG-IQ Cloud, redundancy is achieved by deploying multiple BIG-IQ Cloud appliances using clustered mode. BIG-IP changes initiated through the NSX Administration User Interface will continue to be processed if a BIG-IQ Cloud appliance is rendered inoperable.

vSphere offers a high-availability option to power virtual machines using other available hosts in the event of a host failure. This operation, including the reboot of the BIG-IP VEs, can take some time. Therefore, your BIG-IP and/or BIG-IQ Cloud instance will be down while it powers itself back on (which could be several minutes). Deploying another BIG-IP and/or BIG-IQ Cloud appliance will mitigate this risk, reducing the failover time to seconds versus minutes.

If a high-availability pair of BIG-IP VEs are deployed, NSX Manager will automatically prevent them from running on the same host to minimize an extended outage caused by an ESXi host failure. vSphere and NSX administrators will need to manually enforce the same logic for BIG-IQ Cloud VEs running on vSphere hypervisors.

BIG-IP Tuning and Optimization

Monitoring the health and performance of the system is critical to successfully implementing BIG-IP virtual editions. Critical elements such as CPU utilization, memory usage, and network performance should be constantly monitored for optimal performance and early problem detection. Also, some BIG-IP modules require more storage input/output (I/O) than others. Regardless, storage I/O performance should not be ignored.

Administrators should pay special attention to performance characteristics and statistics of the BIG-IP VE shortly after deployment to ensure its capacity is not exceeded. It may be necessary to consider separating services onto another BIG-IP instance to ensure optimal performance.

Physical vs. Virtual Considerations

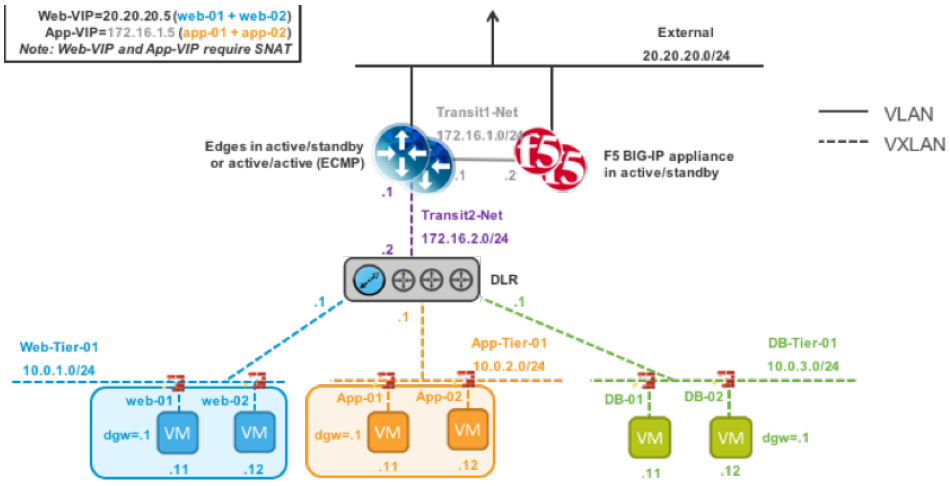
There may come a point where the benefits of a physical BIG-IP appliance outweigh the benefits of virtual editions. The number of access sessions, the volume of SSL transactions per second (TPS), and throughput are three typical elements that should receive attention. If use is encroaching on the upper end of the tested and published limits of BIG-IP VEs, F5 recommends considering a transition to a physical appliance to ensure acceptable performance and scalability of the ADC.

Alternative NSX-Aware Configurations

If native NSX/BIG-IP native integration is not feasible, customers still have options to integrate BIG-IP (virtual and physical) with the NSX platform. The following topologies are available to leverage NSX networking and existing BIG-IP physical and virtual appliances without the integrated deployment and configuration options offered by NSX Manager.

Parallel to the NSX Edge(s)

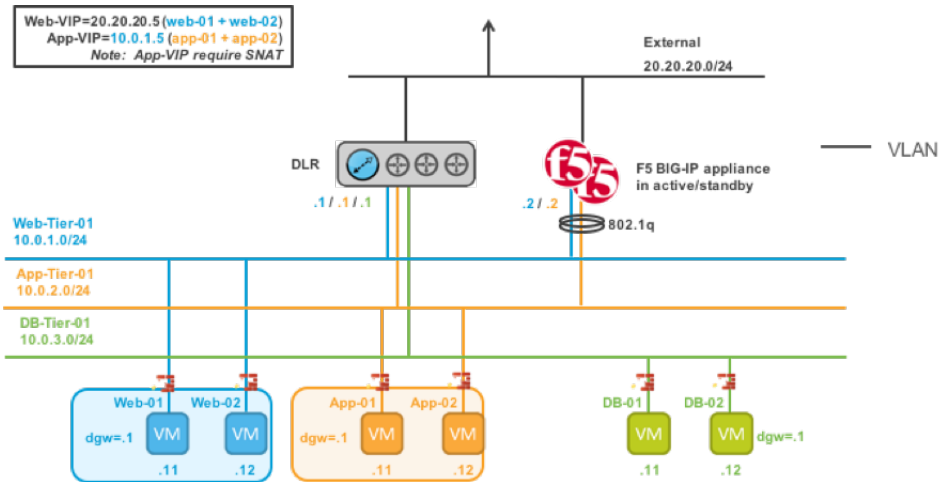
With this configuration, the BIG-IP appliances are logically installed parallel to the NSX Edges. The NSX Edge devices can be installed in active/standby mode or active/active mode. Below the Edges, a DLR provides connectivity to the different VLAN or VXLAN-connected applications tiers (Web, App, and DB).



Logical view “Parallel to NSX Edge” with BIG-IP appliances.

Parallel to DLR using VLANs

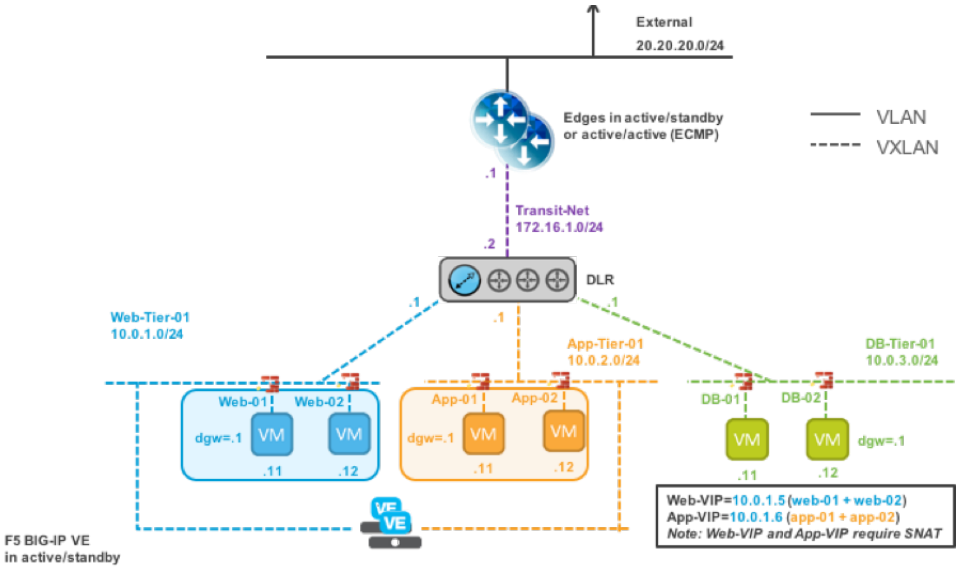
In this topology, the BIG-IP appliances are placed parallel to the DLR. The DLR provides connectivity to the different VLAN and VXLAN-connected applications tiers (Web, App, and DB). Standard 802.1q tagging can be used to connect the VLAN-backed virtual networking segments to the networking segments to which the BIG-IP devices are connected.



Logical view “Parallel to DLR” with BIG-IP appliances.

One-Arm Overlay with BIG-IP Virtual Editions

In this topology, the BIG-IP VEs are placed in one-arm mode in the load balanced server networks. This topology is popular on layer 3 physical fabrics, such as Leaf/Spine but also works on layer 2 physical fabrics. In this diagram, the NSX Edges can be installed in active/standby mode or active/active mode. Below the Edges, a DLR provides connectivity to the different VLAN or VXLAN-connected applications tiers (Web, App, and DB).



Logical View “One-Arm Connected” with BIG-IP virtual edition.

Conclusion

This document provides guidance for integrating F5 BIG-IP ADCs with VMware NSX network virtualization. The integration between F5 and VMware NSX extends VMware's software-defined data center strategy to include F5 Software-Defined Application Services, delivering interconnected automation for network and application layer services. NSX and vSphere administrators can rapidly deploy applications and other relevant IT services, creating an agile environment that can react swiftly to the demands of the business.

For more information about these solutions, please contact your local F5 or VMware representative.

References

BIG-IQ Cloud: VMware NSX Administration

BIG-IQ Cloud and Orchestration 1.0.0 Documentation

VMware for NSX Product Documentation

F5/VMware NSX Reference Architecture

F5 Networks, Inc. 401 Elliott Avenue West, Seattle, WA 98119 888-882-4447 f5.com

Americas info@f5.com
Asia-Pacific apacinfo@f5.com
Europe/Middle-East/Africa emeainfo@f5.com
Japan f5j-info@f5.com

Solutions for an application world.