

NSX-T Administration Guide

VMware NSX-T Data Center 1.1



vmware®

You can find the most up-to-date technical documentation on the VMware website at:

<https://docs.vmware.com/>

If you have comments about this documentation, submit your feedback to

docfeedback@vmware.com

VMware, Inc.
3401 Hillview Ave.
Palo Alto, CA 94304
www.vmware.com

Copyright © 2017 VMware, Inc. All rights reserved. [Copyright and trademark information.](#)

Contents

About Administering VMware NSX-T	6
1 Overview of NSX-T	7
Data Plane	9
Control Plane	9
Management Plane	10
NSX Manager	10
NSX Controller	11
Logical Switches	11
Logical Routers	12
NSX Edge	13
Transport Zones	13
Key Concepts	14
2 Creating Logical Switches and Configuring VM Attachment	17
Understanding BUM Frame Replication Modes	18
Create a Logical Switch	19
Layer 2 Bridging	20
Create a VLAN Logical Switch for the NSX Edge Uplink	24
Connecting a VM to a Logical Switch	25
Test Layer 2 Connectivity	34
3 Configuring Switching Profiles for Logical Switches and Logical Ports	38
Understanding QoS Switching Profile	39
Understanding Port Mirroring Switching Profile	41
Understanding IP Discovery Switching Profile	44
Understanding SpoofGuard	45
Understanding Switch Security Switching Profile	48
Understanding MAC Management Switching Profile	49
Associate a Custom Profile with a Logical Switch	50
Associate a Custom Profile with a Logical Switch Port	51
4 Configuring the Tier-1 Logical Router	53
Create a Tier-1 Logical Router	54
Add Downlink Ports for the Tier-1 Logical Router	54
Configure Route Advertisement on a Tier-1 Logical Router	56
Configure a Tier-1 Logical Router Static Route	57

- 5 Configuring a Tier-0 Logical Router 60**
 - [Create a Tier-0 Logical Router 61](#)
 - [Attach Tier-0 and Tier-1 62](#)
 - [Connect a Tier-0 Logical Router to a VLAN Logical Switch 65](#)
 - [Configure a Static Route 68](#)
 - [BGP Configuration Options 72](#)
 - [Configure BFD on a Tier-0 Logical Router 77](#)
 - [Enable Route Redistribution on the Tier-0 Logical Router 77](#)
 - [Understanding ECMP Routing 80](#)
 - [Create an IP Prefix List 85](#)
 - [Create a Route Map 86](#)

- 6 Network Address Translation 87**
 - [Tier-1 NAT 88](#)
 - [Tier-0 NAT 94](#)

- 7 Firewall Sections and Firewall Rules 98**
 - [Add a Firewall Rule Section 99](#)
 - [Delete a Firewall Rule Section 100](#)
 - [Enable and Disable Section Rules 100](#)
 - [Disable and Enable Section Logs 100](#)
 - [About Firewall Rules 100](#)
 - [Add a Firewall Rule 102](#)
 - [Delete a Firewall Rule 105](#)
 - [Edit the Default Distributed Firewall Rule 106](#)
 - [Change the Order of a Firewall Rule 107](#)
 - [Filter Firewall Rules 107](#)
 - [Exclude Objects from Firewall Enforcement 108](#)

- 8 Configuring Groups and Services 109**
 - [Create an IP Set 109](#)
 - [Create an IP Pool 110](#)
 - [Create a MAC Set 110](#)
 - [Create an NSGroup 111](#)
 - [Configuring Services and Service Groups 112](#)

- 9 DHCP 114**
 - [Create a DHCP Server Profile 114](#)
 - [Create a DHCP Server 115](#)
 - [Attach a DHCP Server to a Logical Switch 116](#)
 - [Detach a DHCP Server from a Logical Switch 116](#)
 - [Create a DHCP Relay Profile 116](#)

- [Create a DHCP Relay Service](#) 116
- [Add a DHCP Service to a Logical Router Port](#) 117

10 [Configuring Metadata Proxies](#) 118

- [Add a Metadata Proxy Server](#) 118
- [Attach a Metadata Proxy Server to a Logical Switch](#) 119
- [Detach a Metadata Proxy Server from a Logical Switch](#) 120

11 [Operations and Management](#) 121

- [Add a License Key](#) 121
- [Managing User Accounts](#) 122
- [Setting Up Certificates](#) 123
- [Configuring Appliances](#) 128
- [Manage Tags](#) 129
- [Search for Objects](#) 129
- [Find the SSH Fingerprint of a Remote Server](#) 130
- [Backing Up and Restoring the NSX Manager](#) 131
- [Managing Appliances and Appliance Clusters](#) 142
- [Logging System Messages](#) 154
- [Configure IPFIX](#) 157
- [Trace the Path of a Packet with Traceflow](#) 159
- [View Port Connection Information](#) 160
- [Monitor a Logical Switch Port Activity](#) 160
- [Monitor Port Mirroring Sessions](#) 161
- [Monitor Fabric Nodes](#) 163
- [Collect Support Bundles](#) 163

About Administering VMware NSX-T

The *NSX-T Administration Guide* provides information about configuring and managing networking for VMware NSX-T[®], including how to create logical switches and ports and how to set up networking for tiered logical routers. It also describes how to configure NAT, firewalls, SpoofGuard, grouping, and DHCP.

Intended Audience

This information is intended for anyone who wants to configure NSX-T. The information is written for experienced Windows or Linux system administrators who are familiar with virtual machine technology, networking, and security operations.

VMware Technical Publications Glossary

VMware Technical Publications provides a glossary of terms that might be unfamiliar to you. For definitions of terms as they are used in VMware technical documentation, go to <http://www.vmware.com/support/pubs>.

Overview of NSX-T

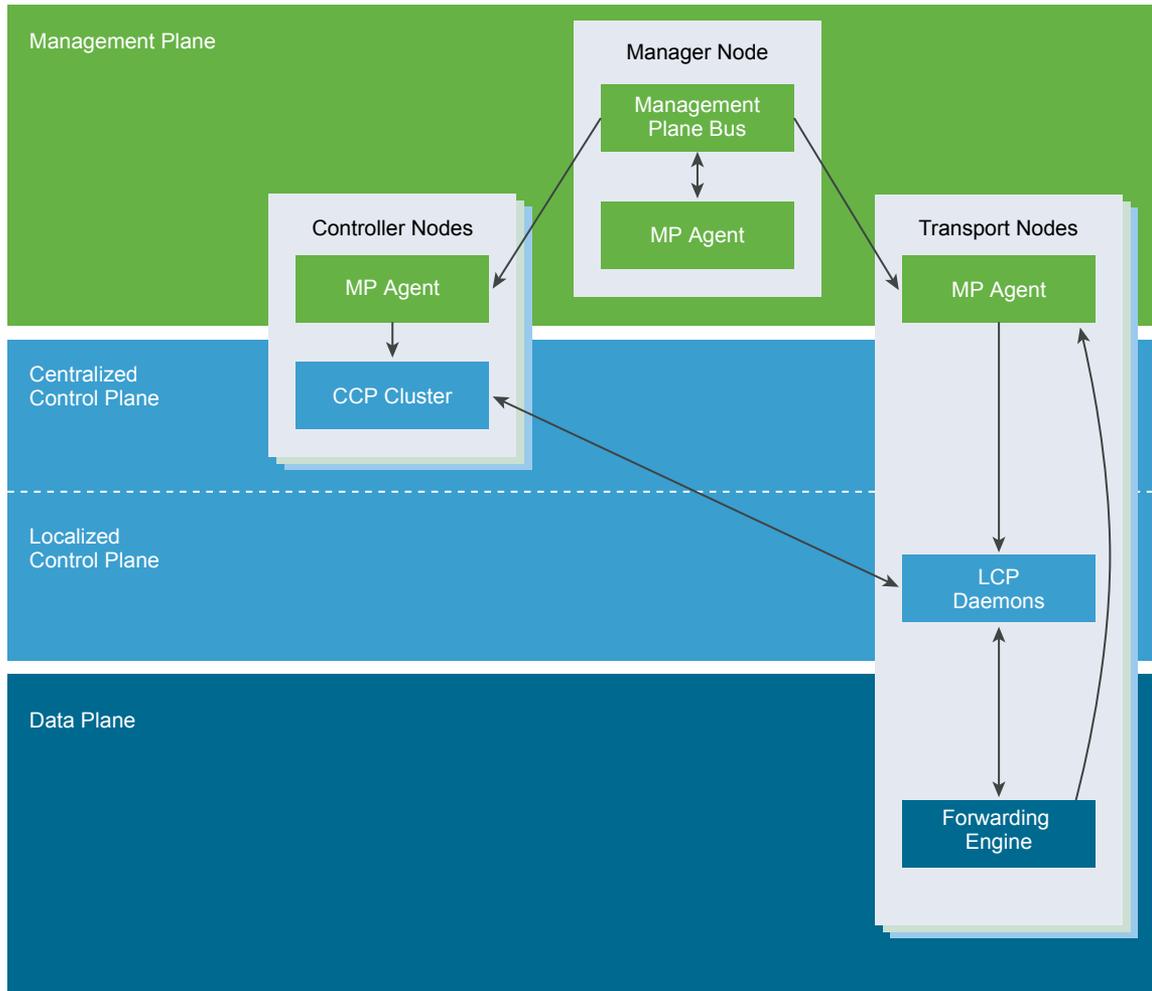
In much the same way that server virtualization programmatically creates, snapshots, deletes and restores software-based virtual machines (VMs), NSX-T network virtualization programmatically creates, snapshots, deletes, and restores software-based virtual networks.

With network virtualization, the functional equivalent of a network hypervisor reproduces the complete set of Layer 2 through Layer 7 networking services (for example, switching, routing, access control, firewalling, QoS) in software. As a result, these services can be programmatically assembled in any arbitrary combination, to produce unique, isolated virtual networks in a matter of seconds.

NSX-T works by implementing three separate but integrated planes: management, control, and data. The three planes are implemented as a set of processes, modules, and agents residing on three types of nodes: manager, controller, and transport nodes.

- Every node hosts a management plane agent.
- The NSX Manager node hosts API services. Each NSX-T installation supports a single NSX Manager node and does not support an NSX Manager cluster.
- NSX Controller nodes host the central control plane cluster daemons.
- NSX Manager and NSX Controller nodes may be co-hosted on the same physical server.

- Transport nodes host local control plane daemons and forwarding engines.



This chapter includes the following topics:

- [Data Plane](#)
- [Control Plane](#)
- [Management Plane](#)
- [NSX Manager](#)
- [NSX Controller](#)
- [Logical Switches](#)
- [Logical Routers](#)
- [NSX Edge](#)
- [Transport Zones](#)
- [Key Concepts](#)

Data Plane

Performs stateless forwarding/transformation of packets based on tables populated by the control plane and reports topology information to the control plane, and maintains packet level statistics.

The data plane is the source of truth for the physical topology and status for example, VIF location, tunnel status, and so on. If you are dealing with moving packets from one place to another, you are in the data plane. The data plane also maintains status of and handles failover between multiple links/tunnels. Per-packet performance is paramount with very strict latency or jitter requirements. Data plane is not necessarily fully contained in kernel, drivers, userspace, or even specific userspace processes. Data plane is constrained to totally stateless forwarding based on tables/rules populated by control plane.

The data plane also may have components that maintain some amount of state for features such as TCP termination. This is different from the control plane managed state such as MAC:IP tunnel mappings, because the state managed by the control plane is about how to forward the packets, whereas state managed by the data plane is limited to how to manipulate payload.

Control Plane

Computes all ephemeral runtime state based on configuration from the management plane, disseminates topology information reported by the data plane elements, and pushes stateless configuration to forwarding engines.

The control plane is sometimes described as the signaling for the network. If you are dealing with processing messages in order to maintain the data plane in the presence of static user configuration, you are in the control plane (for example, responding to a vMotion of a virtual machine (VM) is a control plane responsibility, but connecting the VM to the logical network is a management plane responsibility) Often the control plane is acting as a reflector for topological info from the data plane elements to one another for example, MAC/Tunnel mappings for VTEPs. In other cases, the control plane is acting on data received from some data plane elements to (re)configure some data plane elements such as, using VIF locators to compute and establish the correct subset mesh of tunnels.

The set of objects that the control plane deals with include VIFs, logical networks, logical ports, logical routers, IP addresses, and so on.

The control plane is split into two parts in NSX-T, the central control plane (CCP), which runs on the NSX Controller cluster nodes, and the local control plane (LCP), which runs on the transport nodes, adjacent to the data plane it controls. The Central Control Plane computes some ephemeral runtime state based on configuration from the management plane and disseminates information reported by the data plane elements via the local control plane. The Local Control Plane monitors local link status, computes most ephemeral runtime state based on updates from data plane and CCP, and pushes stateless configuration to forwarding engines. The LCP shares fate with the data plane element which hosts it.

Management Plane

The management plane provides a single API entry point to the system, persists user configuration, handles user queries, and performs operational tasks on all management, control, and data plane nodes in the system.

For NSX-T anything dealing with querying, modifying, and persisting user configuration is a management plane responsibility, while dissemination of that configuration down to the correct subset of data plane elements is a control plane responsibility. This means that some data belongs to multiple planes depending on what stage of its existence it is in. The management plane also handles querying recent status and statistics from the control plane, and sometimes directly from the data plane.

The management plane is the one and only source-of-truth for the configured (logical) system, as managed by the user via configuration. Changes are made using either a RESTful API or the NSX-T UI.

In NSX there is also a management plane agent (MPA) running on all cluster and transport nodes. Example use cases are bootstrapping configurations such as central management node address(es) credentials, packages, statistics, and status. The MPA can run relatively independently of the control plane and data plane, and to be restarted independently if its process crashes or wedges, however, there are scenarios where fate is shared because they run on the same host. The MPA is both locally accessible and remotely accessible. MPA runs on transport nodes, control nodes, and management nodes for node management. On transport nodes it may perform data plane related tasks as well.

Tasks that happen on the management plan include:

- Configuration persistence (desired logical state)
- Input validation
- User management -- role assignments
- Policy management
- Background task tracking

NSX Manager

NSX Manager provides the graphical user interface (GUI) and the REST APIs for creating, configuring, and monitoring NSX-T components, such as controllers, logical switches, and edge services gateways.

NSX Manager is the management plane for the NSX-T eco-system. NSX Manager provides an aggregated system view and is the centralized network management component of NSX-T. It provides a method for monitoring and troubleshooting workloads attached to virtual networks created by NSX-T. It provides configuration and orchestration of:

- Logical networking components – logical switching and routing
- Networking and Edge services
- Security services and distributed firewall - Edge services and security services can be provided by either built-in components of NSX Manager or by integrated 3rd party vendors.

NSX Manager allows seamless orchestration of both built-in and external services. All security services, whether built-in or 3rd party, are deployed and configured by the NSX-T management plane. The management plane provides a single window for viewing services availability. It also facilitates policy based service chaining, context sharing, and inter-service events handling. This simplifies the auditing of the security posture, streamlining application of identity-based controls (for example, AD and mobility profiles).

NSX Manager also provides REST API entry-points to automate consumption. This flexible architecture allows for automation of all configuration and monitoring aspects via any cloud management platform, security vendor platform, or automation framework.

The NSX-T Management Plane Agent (MPA) is an NSX Manager component that lives on each and every node (hypervisor). The MPA is in charge of persisting the desired state of the system and for communicating non-flow-controlling (NFC) messages such as configuration, statistics, status and real time data between transport nodes and the management plane.

NSX Controller

NSX Controller is an advanced distributed state management system that controls virtual networks and overlay transport tunnels.

NSX Controller is deployed as a cluster of highly available virtual appliances that are responsible for the programmatic deployment of virtual networks across the entire NSX-T architecture. The NSX-T Central Control Plane (CCP) is logically separated from all data plane traffic, meaning any failure in the control plane does not affect existing data plane operations. Traffic doesn't pass through the controller; instead the controller is responsible for providing configuration to other NSX Controller components such as the logical switches, logical routers, and edge configuration. Stability and reliability of data transport are central concerns in networking. To further enhance high availability and scalability, the NSX Controller is deployed in a cluster of three instances.

Logical Switches

The logical switching capability in the NSX Edge platform provides the ability to spin up isolated logical L2 networks with the same flexibility and agility that exists for virtual machines.

A cloud deployment for a virtual data center has a variety of applications across multiple tenants. These applications and tenants require isolation from each other for security, fault isolation, and to avoid overlapping IP addressing issues. Endpoints, both virtual and physical, can connect to logical segments and establish connectivity independently from their physical location in the data center network. This is enabled through the decoupling of network infrastructure from logical network (i.e., underlay network from overlay network) provided by NSX-T network virtualization.

A logical switch provides a representation of Layer 2 switched connectivity across many hosts with Layer 3 IP reachability between them. If you plan to restrict some logical networks to a limited set of hosts or you have custom connectivity requirements, you may find it necessary to create additional logical switches.

Logical Routers

NSX-T logical routers provide North-South connectivity, thereby enabling tenants to access public networks, and East-West connectivity between different networks within the same tenants.

A logical router is a configured partition of a traditional network hardware router. It replicates the hardware's functionality, creating multiple routing domains within a single router. Logical routers perform a subset of the tasks that can be handled by the physical router, and each can contain multiple routing instances and routing tables. Using logical routers can be an effective way to maximize router usage, because a set of logical routers within a single physical router can perform the operations previously performed by several pieces of equipment.

With NSX-T it's possible to create two-tier logical router topology: the top-tier logical router is Tier 0 and the bottom-tier logical router is Tier 1. This structure gives both provider administrator and tenant administrators complete control over their services and policies. Administrators control and configure Tier-0 routing and services, and tenant administrators control and configure Tier-1. The north end of Tier-0 interfaces with the physical network, and is where dynamic routing protocols can be configured to exchange routing information with physical routers. The south end of Tier-0 connects to multiple Tier-1 routing layer(s) and receives routing information from them. To optimize resource usage, the Tier-0 layer does not push all the routes coming from the physical network towards Tier-1, but does provide default information.

Southbound, the Tier-1 routing layer interfaces with the logical switches defined by the tenant administrators, and provides one-hop routing function between them. For Tier-1 attached subnets to be reachable from the physical network, route redistribution towards Tier-0 layer must be enabled. However, there isn't a classical routing protocol (such as OSPF or BGP) running between Tier-1 layer and Tier-0 layer, and all the routes go through the NSX-T control plane. Note that the two-tier routing topology is not mandatory, if there is no need to separate provider and tenant, a single tier topology can be created and in this scenario the logical switches are connected directly to the Tier-0 layer and there is no Tier-1 layer.

A logical router consists of two optional parts: a distributed router (DR) and one or more service routers (SR).

A DR spans hypervisors whose VMs are connected to this logical router, as well as edge nodes the logical router is bound to. Functionally, the DR is responsible for one-hop distributed routing between logical switches and/or logical routers connected to this logical router. The SR is responsible for delivering services that are not currently implemented in a distributed fashion, such as stateful NAT.

A logical router always has a DR, and it has SRs if any of the following is true:

- The logical router is a Tier-0 router, even if no stateful services are configured
- The logical router is Tier-1 router linked to a Tier-0 router and has services configured that do not have a distributed implementation (such as NAT, LB, DHCP)

The NSX-T management plane (MP) is responsible for automatically creating the structure that connects the service router to the distributed router. The MP creates a transit logical switch and allocates it a VNI, then creates a port on each SR and DR, connects them to the transit logical switch, and allocates IP addresses for the SR and DR.

NSX Edge

NSX Edge provides routing services and connectivity to networks that are external to the NSX-T deployment.

With NSX Edge, virtual machines or workloads that reside on the same host on different subnets can communicate with one another without having to traverse a traditional routing interface.

NSX Edge is required for establishing external connectivity from the NSX-T domain, through a Tier-0 router via BGP or static routing. Additionally, an NSX Edge must be deployed if you require network address translation (NAT) services at either the Tier-0 or Tier-1 logical routers.

The NSX Edge gateway connects isolated, stub networks to shared (uplink) networks by providing common gateway services such as NAT, and dynamic routing. Common deployments of NSX Edge include in the DMZ and multi-tenant Cloud environments where the NSX Edge creates virtual boundaries for each tenant.

Transport Zones

A transport zone controls which hosts a logical switch can reach. It can span one or more host clusters. Transport zones dictate which hosts and, therefore, which VMs can participate in the use of a particular network.

A Transport Zone defines a collection of hosts that can communicate with each other across a physical network infrastructure. This communication happens over one or more interfaces defined as Virtual Tunnel Endpoints (VTEPs).

If two transport nodes are in the same transport zone, VMs hosted on those transport nodes can "see" and therefore be attached to NSX-T logical switches that are also in that transport zone. This attachment makes it possible for the VMs to communicate with each other, assuming that the VMs have Layer 2/Layer 3 reachability. If VMs are attached to switches that are in different transport zones, the VMs cannot communicate with each other. Transport zones do not replace Layer 2/Layer 3 reachability requirements, but they place a limit on reachability. Put another way, belonging to the same transport zone is a prerequisite for connectivity. After that prerequisite is met, reachability is possible but not automatic. To achieve actual reachability, Layer 2 and (for different subnets) Layer 3 networking must be operational.

A node can serve as a transport node if it contains at least one hostswitch. When you create a host transport node and then add the node to a transport zone, NSX-T installs a hostswitch on the host. For each transport zone that the host belongs to, a separate hostswitch is installed. The hostswitch is used for attaching VMs to NSX-T logical switches and for creating NSX-T logical router uplinks and downlinks.

Key Concepts

The common NSX-T concepts that are used in the documentation and user interface.

Control Plane	Computes runtime state based on configuration from the management plane. Control plane disseminates topology information reported by the data plane elements, and pushes stateless configuration to forwarding engines.
Data Plane	Performs stateless forwarding or transformation of packets based on tables populated by the control plane. Data plane reports topology information to the control plane and maintains packet level statistics.
External Network	A physical network or VLAN not managed by NSX-T. You can link your logical network or overlay network to an external network through an NSX Edge. For example, a physical network in a customer data center or a VLAN in a physical environment.
Fabric Node	Node that has been registered with the NSX-T management plane and has NSX-T modules installed. For a hypervisor host or NSX Edge to be part of the NSX-T overlay, it must be added to the NSX-T fabric.
Fabric Profile	Represents a specific configuration that can be associated with an NSX Edge cluster. For example, the fabric profile might contain the tunneling properties for dead peer detection.
Logical Port Egress	Inbound network traffic to the VM or logical network is called egress because traffic is leaving the data center network and entering the virtual space.
Logical Port Ingress	Outbound network traffic from the VM to the data center network is called ingress because traffic is entering the physical network.
Logical Router	NSX-T routing entity.
Logical Router Port	Logical network port to which you can attach a logical switch port or an uplink port to a physical network.
Logical Switch	API entity that provides virtual Layer 2 switching for VM interfaces and Gateway interfaces. A logical switch gives tenant network administrators the logical equivalent of a physical Layer 2 switch, allowing them to connect a set of VMs to a common broadcast domain. A logical switch is a logical entity independent of the physical hypervisor infrastructure and spans many hypervisors, connecting VMs regardless of their physical location. This allows VMs to migrate without requiring reconfiguration by the tenant network administrator.

In a multi-tenant cloud, many logical switches might exist side-by-side on the same hypervisor hardware, with each Layer 2 segment isolated from the others. Logical switches can be connected using logical routers, and logical routers can provide uplink ports connected to the external physical network.

Logical Switch Port

Logical switch attachment point to establish a connection to a virtual machine network interface or a logical router interface. The logical switch port reports applied switching profile, port state, and link status.

Management Plane

Provides single API entry point to the system, persists user configuration, handles user queries, and performs operational tasks on all of the management, control, and data plane nodes in the system. Management plane is also responsible for querying, modifying, and persisting use configuration.

NSX Controller Cluster

Deployed as a cluster of highly available virtual appliances that are responsible for the programmatic deployment of virtual networks across the entire NSX-T architecture.

NSX Edge Cluster

Collection of NSX Edge node appliances that have the same settings as protocols involved in high-availability monitoring.

NSX Edge Node

Component with the functional goal is to provide computational power to deliver the IP routing and the IP services functions.

NSX-T Hostswitch or KVM Open vSwitch

Software that runs on the hypervisor and provides physical traffic forwarding. The hostswitch or OVS is invisible to the tenant network administrator and provides the underlying forwarding service that each logical switch relies on. To achieve network virtualization, a network controller must configure the hypervisor hostswitches with network flow tables that form the logical broadcast domains the tenant administrators defined when they created and configured their logical switches.

Each logical broadcast domain is implemented by tunneling VM-to-VM traffic and VM-to-logical router traffic using the tunnel encapsulation mechanism Geneve. The network controller has the global view of the data center and ensures that the hypervisor hostswitch flow tables are updated as VMs are created, moved, or removed.

NSX Manager

Node that hosts the API services, the management plane, and the agent services.

Open vSwitch (OVS)

Open source software switch that acts as a hypervisor hostswitch within XenServer, Xen, KVM, and other Linux-based hypervisors. NSX Edge switching components are based on OVS.

Overlay Logical Network

Logical network implemented using Layer 2-in-Layer 3 tunneling such that the topology seen by VMs is decoupled from that of the physical network.

Physical Interface (pNIC)	Network interface on a physical server that a hypervisor is installed on.
Tier-0 Logical Router	Provider logical router is also known as Tier-0 logical router interfaces with the physical network. Tier-0 logical router is a top-tier router and can be realized as active-active or active-standby cluster of services router. The logical router runs BGP and peers with physical routers. In active-standby mode the logical router can also provide stateful services.
Tier-1 Logical Router	Tier-1 logical router is the second tier router that connects to one Tier-0 logical router for northbound connectivity and one or more overlay networks for southbound connectivity. Tier-1 logical router can be an active-standby cluster of services router providing stateful services.
Transport Zone	Collection of transport nodes that defines the maximum span for logical switches. A transport zone represents a set of similarly provisioned hypervisors and the logical switches that connect VMs on those hypervisors. NSX-T can deploy the required supporting software packages to the hosts because it knows what features are enabled on the logical switches.
VM Interface (vNIC)	Network interface on a virtual machine that provides connectivity between the virtual guest operating system and the standard vSwitch or vSphere distributed switch. The vNIC can be attached to a logical port. You can identify a vNIC based on its Unique ID (UUID).
VTEP	Virtual tunnel end point. Tunnel endpoints enable hypervisor hosts to participate in an NSX-T overlay. The NSX-T overlay deploys a Layer 2 network on top of an existing Layer 3 network fabric by encapsulating frames inside of packets and transferring the packets over an underlying transport network. The underlying transport network can be another Layer 2 networks or it can cross Layer 3 boundaries. The VTEP is the connection point at which the encapsulation and decapsulation takes place.

Creating Logical Switches and Configuring VM Attachment

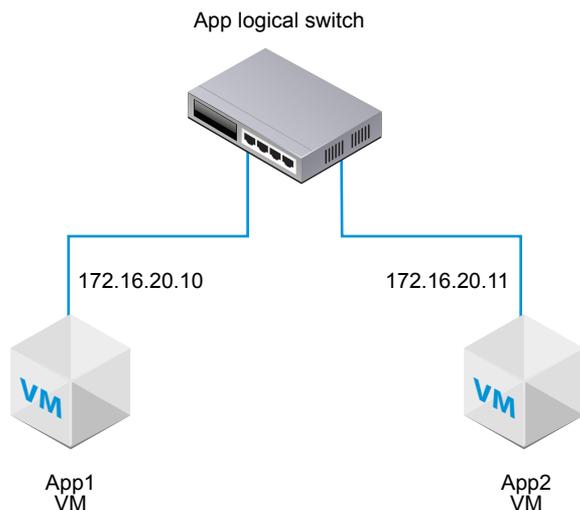
2

An NSX-T logical switch reproduces switching functionality, broadcast, unknown unicast, multicast (BUM) traffic, in a virtual environment completely decoupled from underlying hardware.

Logical switches are similar to VLANs, in that they provide network connections to which you can attach virtual machines. The VMs can then communicate with each other over tunnels between hypervisors if the VMs are connected to the same logical switch. Each logical switch has a virtual network identifier (VNI), like a VLAN ID. Unlike VLAN, VNIs scale well beyond the limits of VLAN IDs.

When you add logical switches, it is important that you map out the topology that you are building.

Figure 2-1. Logical Switch Topology



For example, the topology shows a single logical switch connected to two VMs. The two VMs can be on different hosts or the same host, in different host clusters or in the same host cluster. Because the VMs in the example are on the same virtual network, the underlying IP addresses configured on the VMs must be in the same subnet.

This chapter includes the following topics:

- [Understanding BUM Frame Replication Modes](#)
- [Create a Logical Switch](#)
- [Layer 2 Bridging](#)

- [Create a VLAN Logical Switch for the NSX Edge Uplink](#)
- [Connecting a VM to a Logical Switch](#)
- [Test Layer 2 Connectivity](#)

Understanding BUM Frame Replication Modes

Each host transport node is a tunnel endpoint. Each tunnel endpoint has an IP address. These IP addresses can be in the same subnet or in different subnets, depending on your configuration of IP pools or DHCP for your transport nodes.

When two VMs on different hosts communicate directly, unicast-encapsulated traffic is exchanged between the two tunnel endpoint IP addresses associated with the two hypervisors without any need for flooding.

However, as with any Layer 2 network, sometimes traffic that is originated by a VM needs to be flooded, meaning that it needs to be sent to all of the other VMs belonging to the same logical switch. This is the case with Layer 2 broadcast, unknown unicast, and multicast traffic (BUM traffic). Recall that a single NSX-T logical switch can span multiple hypervisors. BUM traffic originated by a VM on a given hypervisor needs to be replicated to remote hypervisors that host other VMs that are connected to the same logical switch. To enable this flooding, NSX-T supports two different replication modes:

- Hierarchical two-tier (sometimes called MTEP)
- Head (sometimes called source)

Hierarchical two-tier replication mode is illustrated by the following example. Say you have Host A, which has VMs connected to virtual network identifiers (VNIs) 5000, 5001, and 5002. Think of VNIs as being similar to VLANs, but each logical switch has a single VNI associated with it. For this reason, sometimes the terms VNI and logical switch are used interchangeably. When we say a host is on a VNI, we mean that it has VMs that are connected to a logical switch with that VNI.

A tunnel endpoint table shows the host-VNI connections. Host A examines the tunnel endpoint table for VNI 5000 and determines the tunnel endpoint IP addresses for other hosts on VNI 5000.

Some of these VNI connections will be on the same IP subnet, also called an IP segment, as the tunnel endpoint on Host A. For each of these, Host A creates a separate copy of every BUM frame and sends the copy directly to each host.

Other hosts' tunnel endpoints are on different subnets or IP segments. For each segment where there is more than one tunnel endpoint, Host A nominates one of these endpoints to be the replicator.

The replicator receives from Host A one copy of each BUM frame for VNI 5000. This copy is flagged as Replicate locally in the encapsulation header. Host A does not send copies to the other hosts in the same IP segment as the replicator. It becomes the responsibility of the replicator to create a copy of the BUM frame for each host it knows about that is on VNI 5000 and in the same IP segment as that replicator host.

The process is replicated for VNI 5001 and 5002. The list of tunnel endpoints and the resulting replicators might be different for different VNIs.

With head replication also known as headend replication, there are no replicators. Host A simply creates a copy of each BUM frame for each tunnel endpoint it knows about on VNI 5000 and sends it.

If all the host tunnel endpoints are on the same subnet, the choice of replication mode does not make any difference because the behaviour will not differ. If the host tunnel endpoints are on different subnets, hierarchical two-tier replication helps distribute the load among multiple hosts. Hierarchical two-tier is the default mode.

Create a Logical Switch

Logical switches attach to single or multiple VMs in the network. The VMs connected to a logical switch can communicate with each other using the tunnels between hypervisors.

Prerequisites

- Verify that a transport zone is configured. See the *NSX-T Installation Guide*.
- Verify that fabric nodes are successfully connected to NSX-T management plane agent (MPA) and NSX-T local control plane (LCP).

In the GET `https://<nsx-mgr>/api/v1/transport-nodes/<transport-node-id>/state` API call, the state must be success. See the *NSX-T Installation Guide*.

- Verify that transport nodes are added to the transport zone. See the *NSX-T Installation Guide*.
- Verify that the hypervisors are added to the NSX-T fabric and VMs are hosted on these hypervisors.
- Familiarize yourself with the logical switch topology and BUM frame replication concepts. See [Chapter 2 Creating Logical Switches and Configuring VM Attachment](#) and [Understanding BUM Frame Replication Modes](#).
- Verify that your NSX Controller cluster is stable.

Procedure

- 1 From your browser, log in to an NSX Manager at `https://nsx-manager-ip-address`.
- 2 Select **Switching > Switches**.
- 3 Click **Add**.
- 4 Assign a name for the logical switch.
- 5 Select a transport zone for the logical switch.

VMs that are attached to logical switches that are in the same transport zone can communicate with each other.

6 Select a replication mode for the logical switch.

The replication mode (hierarchical two-tier or head) is required for overlay logical switches, but not for VLAN-based logical switches.

Replication Mode	Description
Hierarchical two-tier	The replicator is a host that performs replication of BUM traffic to other hosts within the same VNI. Each host nominates one host tunnel endpoint in every VNI to be the replicator. This is done for each VNI.
Head	Hosts create a copy of each BUM frame and send the copy to each tunnel endpoint it knows about for each VNI.

7 (Optional) Click the **Switching Profiles** tab and select switching profiles.

8 Click **Save**.

In the NSX Manager UI, the new logical switch is a clickable link.

What to do next

Attach VMs to your logical switch. See [Connecting a VM to a Logical Switch](#).

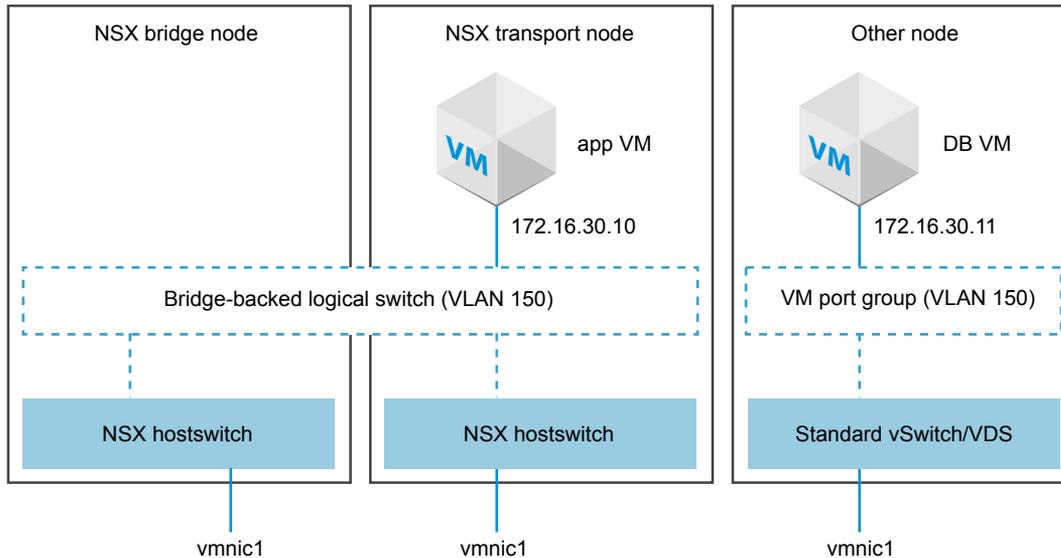
Layer 2 Bridging

When an NSX-T logical switch requires a Layer 2 connection to a VLAN-backed port group or needs to reach another device, such as a gateway, that resides outside of an NSX-T deployment, you can use an NSX-T Layer 2 bridge. This is especially useful in a migration scenario, in which you need to split a subnet across physical and virtual workloads.

The NSX-T concepts involved in Layer 2 bridging are bridge clusters, bridge endpoints, and bridge nodes. A bridge cluster is an high-availability (HA) collection of bridge nodes. A bridge node is a transport node that does bridging. Each logical switch that is used for bridging a virtual and the physical deployment has an associated VLAN ID. A bridge endpoint identifies the physical attributes of the bridge, such as the bridge cluster ID and the associated VLAN ID.

In this release of NSX-T, Layer 2 bridging is provided by ESXi hosts serving as bridge nodes. A bridge node is an ESXi host transport node that has been added to a bridge cluster.

In the following example, two NSX-T transport nodes are part of the same overlay transport zone. This makes it possible for their NSX-T host switches (sometimes called NSX-T vSwitches, as shown in the figure) to be attached to the same bridge-backed logical switch.

Figure 2-2. Bridge Topology

The transport node on the left belongs to a bridge cluster and is therefore a bridge node.

Because the logical switch is attached to a bridge cluster, it is called a bridge-backed logical switch. To be eligible for bridge backing, a logical switch must be in an overlay transport zone, not in a VLAN transport zone.

The middle transport node is not part of the bridge cluster. It is a normal transport node. It can be a KVM or ESXi host. In the diagram, a VM on this node called "app VM" is attached to the bridge-backed logical switch.

The node on the right is not part of the NSX-T overlay. It might be any hypervisor with a VM (as shown in the diagram) or it might be a physical network node. If the non-NSX-T node is an ESXi host, you can use a standard vSwitch or a vSphere distributed switch for the port attachment. One requirement is that the VLAN ID associated with the port attachment must match the VLAN ID on the bridge-backed logical switch. Also, the communication occurs over Layer 2, so the two end devices must have IP addresses in the same subnet.

As stated, the purpose of the bridge is to enable Layer 2 communication between the two VMs. When traffic is transmitted between the two VMs, the traffic traverses the bridge node.

Create a Bridge Cluster

A bridge cluster is a collection of transport nodes that do bridging and participate in high availability (HA). Only one transport node is active at a time. Having a multi-node cluster of NSX-T bridge nodes helps ensure that at least one NSX-T bridge node is always available. To create a bridge-backed logical switch, you must associate it with a bridge cluster. Therefore, even if you have only one bridge node, it must belong to a bridge cluster to be useful.

After creating the bridge cluster, you can later edit it to add additional bridge nodes.

Prerequisites

- Create at least one NSX-T transport node for use as a bridge node.
- The transport node used as a bridge node must be an ESXi host. KVM is not supported for bridge nodes.
- It is recommended that bridge nodes not have any hosted VMs.
- A transport node can be added to only one bridge cluster. You cannot add the same transport node to multiple bridge clusters.

Procedure

- 1 In the NSX Manager UI, navigate to **Fabric > Configuration > Bridges**.
- 2 Give the bridge cluster a name.
- 3 Select a transport zone for the bridge cluster.
The transport zone must be of type overlay, not VLAN.
- 4 From the **Available** column, select transport nodes and click the right arrow to move them to the **Selected** column.

What to do next

You can now associate a logical switch with the bridge cluster.

Create a Layer 2 Bridge-Backed Logical Switch

When you have VMs that are connected to the NSX-T overlay, you might want them to have Layer 2 connectivity with other devices or VMs that are outside of your NSX-T deployment. In this case, you can use a bridge-backed logical switch.

For an example topology, see [Figure 2-2](#).

Prerequisites

- At least one ESXi host to serve as a bridge node. A bridge node is an ESXi transport node that only does bridging. This transport node must be added to a bridge cluster. See [Create a Bridge Cluster](#).
- At least one ESXi or KVM host to serve as a regular transport node. This node has hosted VMs that require connectivity with devices outside of a NSX-T deployment.
- A VM or another end device outside of the NSX-T deployment. This end device must be attached to a VLAN port matching the VLAN ID of the bridge-backed logical switch.
- One logical switch in an overlay transport zone to serve as the bridge-backed logical switch.

Procedure

- 1 From a browser, log in to an NSX Manager at `https://<nsx-mgr>`.
- 2 Select **Switching > Switches**.
- 3 From the list of switches, select an overlay switch (traffic type: overlay).

- 4 On the switch configuration page, select **Related > Bridge Clusters**.
- 5 Click **ATTACH**, select a bridge cluster, and enter a VLAN ID.

For example:

Summary Monitor Manage Related

All Bridge Clusters Attached to Switch

ATTACH EDIT DETACH COLUMNS

Bridge Cluster	VLAN	HA on VLAN	Active Node
bridge-cluster-1	150	Enabled	None

- 6 Connect VMs to the logical switch if they are not already connected.

The VMs must be on transport nodes in the same transport zone as the bridge cluster.

You can test the functionality of the bridge by sending a ping from the NSX-T-internal VM to a node that is external to NSX-T. For example, in [Figure 2-2](#), app VM on the NSX-T transport node should be able to ping DB VM on the external node, and the reverse.

You can monitor traffic on the bridge switch by navigating to **Switching > Switches > Monitor**.

You can view the bridge traffic with the GET `https://192.168.110.31/api/v1/bridge-endpoints/<endpoint-id>/statistics` API call:

```
{
  "tx_packets": {
    "total": 134416,
    "dropped": 0,
    "multicast_broadcast": 0
  },
  "rx_bytes": {
    "total": 22164,
    "multicast_broadcast": 0
  },
  "tx_bytes": {
    "total": 8610134,
    "multicast_broadcast": 0
  },
  "rx_packets": {
    "total": 230,
    "dropped": 0,
    "multicast_broadcast": 0
  }
}
```

```

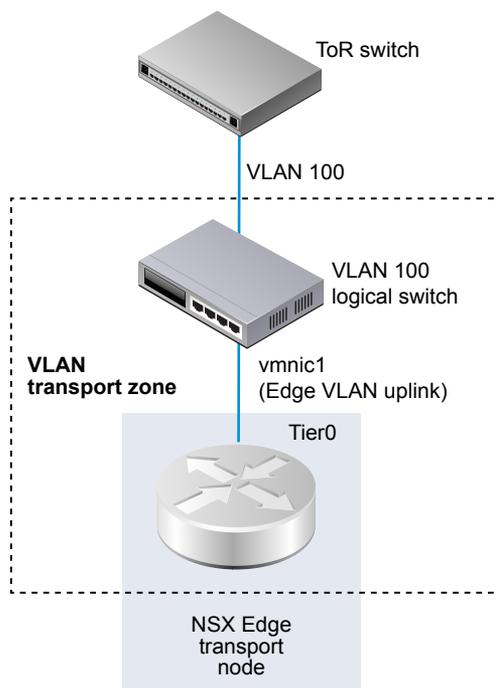
},
"last_update_timestamp": 1454979822860,
"endpoint_id": "ba5ba59d-22f1-4a02-b6a0-18ef0e37ef31"
}

```

Create a VLAN Logical Switch for the NSX Edge Uplink

Edge uplinks go out through VLAN logical switches.

When you are creating a VLAN logical switch, it is important to have in mind a particular topology that you are building. For example, the following simple topology shows a single VLAN logical switch inside of a VLAN transport zone. The VLAN logical switch has VLAN ID 100. This matches the VLAN ID on the TOR port connected to the hypervisor host port used for the Edge's VLAN uplink.



Prerequisites

- To create a VLAN logical switch, you must first create a VLAN transport zone.
- An NSX-T vSwitch must be added to the NSX Edge. To confirm on an Edge, run the `get host-switch` command. For example:

```

nsx-edge1> get host-switch

Host Switch       : c0a78378-1c20-432a-9e23-ddb34f1c80c9
Switch Name      : hs1
Transport Zone   : c46dcd72-808a-423d-b4cc-8752c33f6b2c
Transport Zone   : 73def985-d122-4b7b-ab6a-a58176dfc32d
Physical Port    : fp-eth0
Uplink Name      : uplink-1
Transport VLAN   : 4096

```

```

Default Gateway      : 192.168.150.1
Subnet Mask          : 255.255.255.0
Local VTEP Device    : fp-eth0
Local VTEP IP        : 192.168.150.102

```

- Verify that your NSX Controller cluster is stable.
- Verify that fabric nodes are successfully connected to the NSX-T management plane agent (MPA) and the NSX-T local control plane (LCP).

In the GET `https://<nsx-mgr>/api/v1/transport-nodes/<transport-node-id>/state` API call, the state must be success. See the *NSX-T Installation Guide*.

Procedure

- 1 From a browser, log in to an NSX Manager at `https://<nsx-mgr>`.
- 2 Select **Switching > Switches**.
- 3 Click **Add**.
- 4 Type a name for the logical switch.
- 5 Select a transport zone for the logical switch.

When you select a VLAN transport zone, the VLAN ID field appears.

- 6 Type a VLAN ID.
Enter 0 in the VLAN field if there is no VLAN ID for the uplink to the physical TOR.
- 7 (Optional) Click the **Switching Profiles** tab and select switching profiles.

What to do next

Add a logical router.

Connecting a VM to a Logical Switch

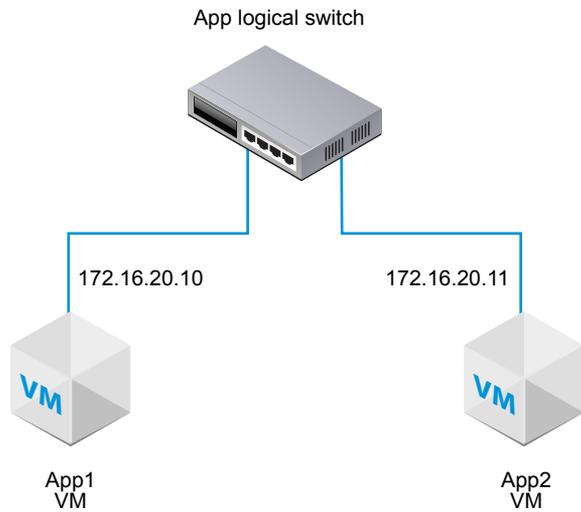
Depending on your host, the configuration for connecting a VM to a logical switch can vary.

The supported hosts that can connect to a logical switch are; an ESXi host that is managed in vCenter Server, a standalone ESXi host, and a KVM host.

Attach a VM Hosted on vCenter Server to an NSX-T Logical Switch

If you have a ESXi host that is managed in vCenter Server, you can access the host VMs through the Web-based vSphere Web Client. In this case, you can use this procedure to attach VMs to NSX-T logical switches.

The example shown in this procedure shows how to attach a VM called app-vm to a logical switch called app-switch.



The installation-based vSphere Client application does not support attaching a VM to an NSX-T logical switch. If you do not have the (Web-based) vSphere Web Client, see [Attach a VM Hosted on Standalone ESXi to an NSX-T Logical Switch](#).

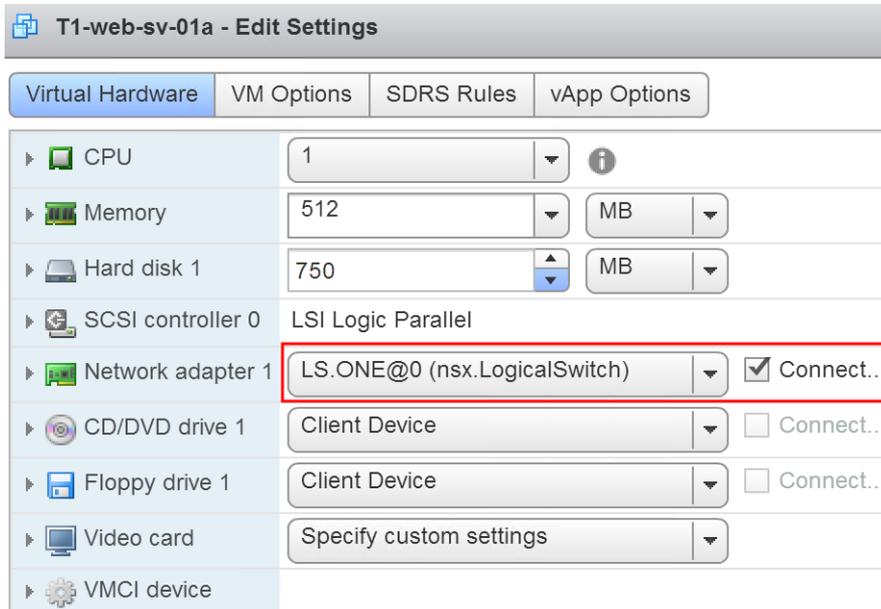
Prerequisites

- The VMs must be hosted on hypervisors that have been added to the NSX-T fabric.
- The fabric nodes must have NSX-T management plane (MPA) and NSX-T control plane (LCP) connectivity.
- The fabric nodes must be added to a transport zone.
- A logical switch must be created.

Procedure

- 1 In the vSphere Web Client, edit the VM settings, and attach the VM to the NSX-T logical switch.

For example:



- 2 Click **OK**.

After attaching a VM to a logical switch, logical switch ports are added to the logical switch. You can view logical switch ports on the NSX Manager in **Switching > Ports**.

In the NSX-T API, you can view NSX-T-attached VMs with the GET `https://<nsx-mgr>/api/v1/fabric/virtual-machines` API call

In the NSX-T Manager UI under **Switching > Ports**, the VIF attachment ID matches the ExternalID found in the API call. Find the VIF attachment ID matching the VM's externalId and make sure that the Admin and Operational status are Up/Up.

If two VMs are attached to the same logical switch and have IP addresses configured in the same subnet, they should be able to ping each other.

What to do next

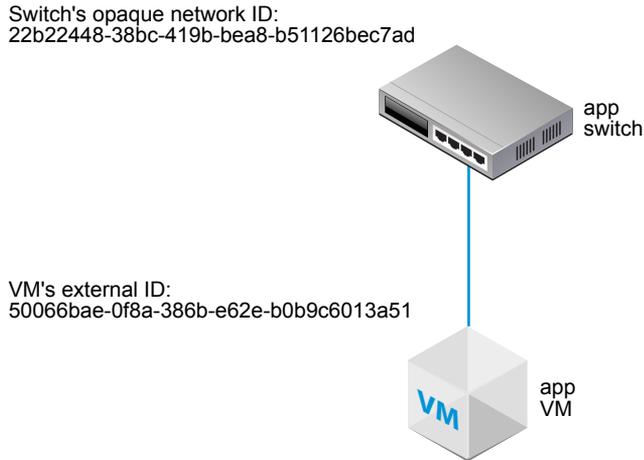
Add a logical router.

You can monitor the activity on the logical switch port to troubleshoot problems. See [Monitor a Logical Switch Port Activity](#).

Attach a VM Hosted on Standalone ESXi to an NSX-T Logical Switch

If you have a standalone ESXi host, you cannot access the host VMs through the web-based vSphere Web Client. In this case, you can use this procedure to attach VMs to NSX-T logical switches.

The example shown in this procedure shows how to attach a VM called app-vm to a logical switch called app-switch.



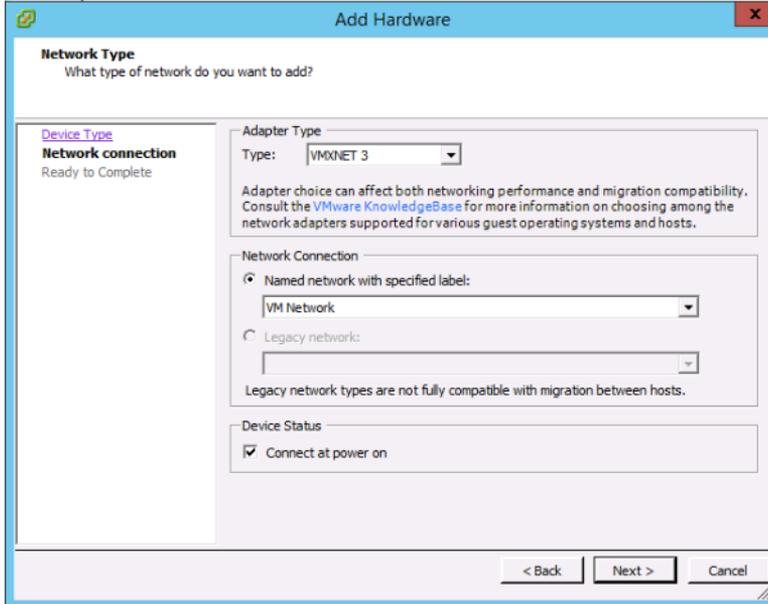
Prerequisites

- The VM must be hosted on hypervisors that have been added to the NSX-T fabric.
- The fabric nodes must have NSX-T management plane (MPA) and NSX-T control plane (LCP) connectivity.
- The fabric nodes must be added to a transport zone.
- A logical switch must be created.
- You must have access to the NSX Manager API.
- You must have write access to the VM's VMX file.

Procedure

- 1 Using the (install-based) vSphere Client application or some other VM management tool, edit the VM and add a VMXNET 3 Ethernet adapter.

Select any named network. You will change the network connection in a later step.



- 2 Use the NSX-T API to issue the GET `https://<nsx-mgr>/api/v1/fabric/virtual-machines/<VM-ID>` API call.

In the results, find the VM's externalId.

For example:

```
GET https://<nsx-mgr>/api/v1/fabric/virtual-machines/60a5a5d5-ea2b-407e-a806-4fdc8468f735

{
  "resource_type": "VirtualMachine",
  "id": "60a5a5d5-ea2b-407e-a806-4fdc8468f735",
  "display_name": "app-vm",
  "compute_ids": [
    "instanceUuid:50066bae-0f8a-386b-e62e-b0b9c6013a51",
    "moIdOnHost:5",
    "externalId:50066bae-0f8a-386b-e62e-b0b9c6013a51",
    "hostLocalId:5",
    "locationId:564dc020-1565-e3f4-f591-ee3953eef3ff",
    "biosUuid:4206f47d-fef7-08c5-5bf7-ea26a4c6b18d"
  ],
  "external_id": "50066bae-0f8a-386b-e62e-b0b9c6013a51",
  "type": "REGULAR",
  "host_id": "cb82b0fa-a8f1-11e5-92a9-6b7d1f8661fa",
```

```
"local_id_on_host": "5"
}
```

3 Power off and unregister the VM from the host.

You can use your VM management tool or the ESXi CLI, as shown here.

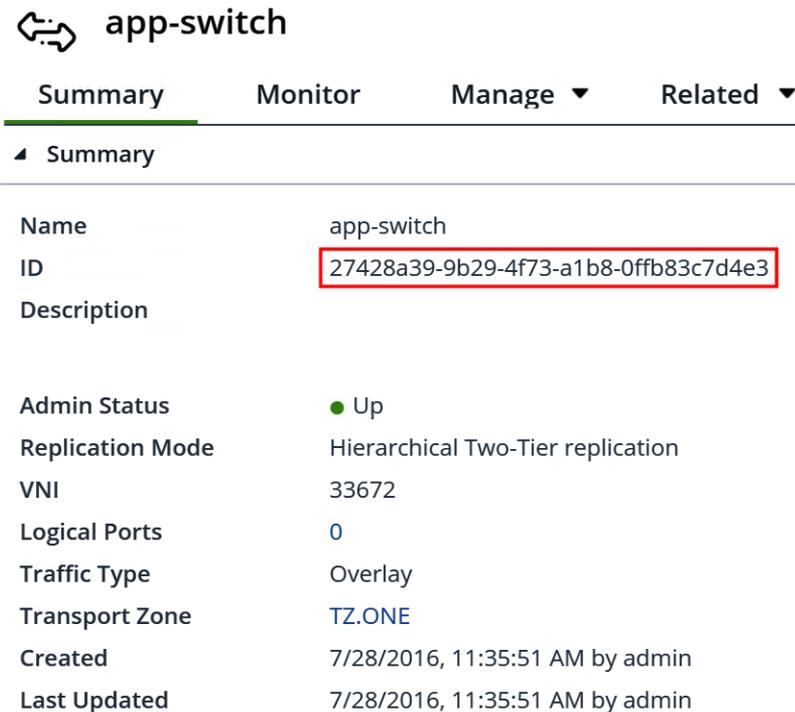
```
[user@host:~] vim-cmd /vmsvc/getallvms
Vmid   Name      File           Guest OS      Version  Annotation
5      app-vm    [ds2] app-vm/app-vm.vmx  ubuntuGuest  vmx-08
8      web-vm    [ds2] web-vm/web-vm.vmx  ubuntu64Guest vmx-08

[user@host:~] vim-cmd /vmsvc/power.off 5
Powering off VM:

[user@host:~] vim-cmd /vmsvc/unregister 5
```

4 From the NSX Manager UI, get the logical switch ID.

For example:



The screenshot shows the NSX Manager UI for a logical switch named "app-switch". The "Summary" tab is selected. The ID field is highlighted with a red box, showing the value "27428a39-9b29-4f73-a1b8-0ffb83c7d4e3".

Summary	Monitor	Manage ▼	Related ▼																						
<p>Summary</p> <table border="1"> <tr> <td>Name</td> <td>app-switch</td> </tr> <tr> <td>ID</td> <td>27428a39-9b29-4f73-a1b8-0ffb83c7d4e3</td> </tr> <tr> <td>Description</td> <td></td> </tr> <tr> <td>Admin Status</td> <td>● Up</td> </tr> <tr> <td>Replication Mode</td> <td>Hierarchical Two-Tier replication</td> </tr> <tr> <td>VNI</td> <td>33672</td> </tr> <tr> <td>Logical Ports</td> <td>0</td> </tr> <tr> <td>Traffic Type</td> <td>Overlay</td> </tr> <tr> <td>Transport Zone</td> <td>TZ.ONE</td> </tr> <tr> <td>Created</td> <td>7/28/2016, 11:35:51 AM by admin</td> </tr> <tr> <td>Last Updated</td> <td>7/28/2016, 11:35:51 AM by admin</td> </tr> </table>				Name	app-switch	ID	27428a39-9b29-4f73-a1b8-0ffb83c7d4e3	Description		Admin Status	● Up	Replication Mode	Hierarchical Two-Tier replication	VNI	33672	Logical Ports	0	Traffic Type	Overlay	Transport Zone	TZ.ONE	Created	7/28/2016, 11:35:51 AM by admin	Last Updated	7/28/2016, 11:35:51 AM by admin
Name	app-switch																								
ID	27428a39-9b29-4f73-a1b8-0ffb83c7d4e3																								
Description																									
Admin Status	● Up																								
Replication Mode	Hierarchical Two-Tier replication																								
VNI	33672																								
Logical Ports	0																								
Traffic Type	Overlay																								
Transport Zone	TZ.ONE																								
Created	7/28/2016, 11:35:51 AM by admin																								
Last Updated	7/28/2016, 11:35:51 AM by admin																								

5 Modify the VM's VMX file.

Delete the **ethernet1.networkName = "<name>"** field and add the following fields:

- ethernet1.opaqueNetwork.id = "<logical switch's ID>"
- ethernet1.opaqueNetwork.type = "nsx.LogicalSwitch"

- ethernet1.externalId = "<VM's externalId>"
- ethernet1.connected = "TRUE"
- ethernet1.startConnected = "TRUE"

For example:

OLD

```
ethernet1.pciSlotNumber = "224"  
ethernet1.virtualDev = "vmxnet3"  
ethernet1.networkName = "VM Network"  
ethernet1.addressType = "vpx"  
ethernet1.generatedAddress = "00:50:56:86:7b:d7"  
ethernet1.uptCompatibility = "true"  
ethernet1.present = "TRUE"
```

NEW

```
ethernet1.pciSlotNumber = "224"  
ethernet1.virtualDev = "vmxnet3"  
ethernet1.addressType = "vpx"  
ethernet1.generatedAddress = "00:50:56:86:7b:d7"  
ethernet1.uptCompatibility = "true"  
ethernet1.present = "TRUE"  
ethernet1.opaqueNetwork.id = "22b22448-38bc-419b-bea8-b51126bec7ad"  
ethernet1.opaqueNetwork.type = "nsx.LogicalSwitch"  
ethernet1.externalId = "50066bae-0f8a-386b-e62e-b0b9c6013a51"  
ethernet1.connected = "TRUE"  
ethernet1.startConnected = "TRUE"
```

- 6 In the NSX Manager UI, add a logical switch port, and use the VM's externalId for the VIF attachment.

For example:

New Logical Port
✕

Name: *

Description:

Logical Switch: *

Admin State: * Up

Attachment Type: *

Attachment ID:

Switching Profiles Type: *

Switching Profiles Id:

- 7 Reregister the VM and power it on.

You can use your VM management tool or the ESXi CLI, as shown here.

```
[user@host:~] vim-cmd /solo/register /path/to/file.vmx
```

For example:

```
[user@host:~] vim-cmd solo/registervm /vmfs/volumes/355f2049-6c704347/app-vm/app-vm.vmx
9
```

```
[user@host:~] vim-cmd /vmsvc/power.on 9
```

Powering on VM:

In the NSX Manager UI under **Switching > Ports**, find the VIF attachment ID matching the VM's externalId and make sure that the Admin and Operational status are Up/Up.

If two VMs are attached to the same logical switch and have IP addresses configured in the same subnet, they should be able to ping each other.

What to do next

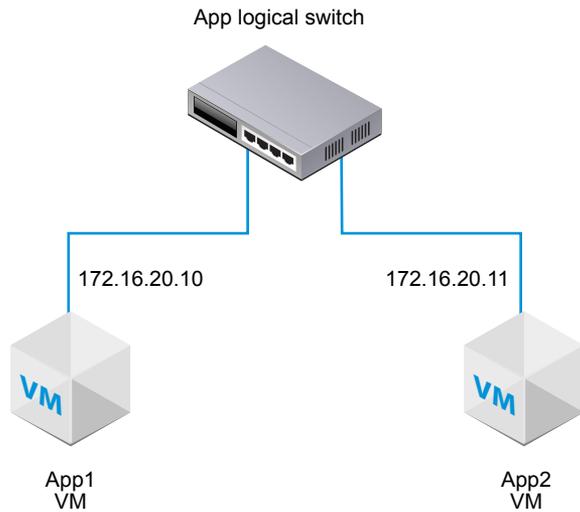
Add a logical router.

You can monitor the activity on the logical switch port to troubleshoot problems. See [Monitor a Logical Switch Port Activity](#).

Attach a VM Hosted on KVM to an NSX-T Logical Switch

If you have a KVM host, you can use this procedure to attach VMs to NSX-T logical switches.

The example shown in this procedure shows how to attach a VM called app-vm to a logical switch called app-switch.



Prerequisites

- The VM must be hosted on hypervisors that have been added to the NSX-T fabric.
- The fabric nodes must have NSX-T management plane (MPA) and NSX-T control plane (LCP) connectivity.
- The fabric nodes must be added to a transport zone.
- A logical switch must be created.

Procedure

- 1 From the KVM CLI, run the `virsh dumpxml <your vm> | grep interfaceid` command.

- In the NSX Manager UI, add a logical switch port, and use the VM's interface ID for the VIF attachment.

For example:

New Logical Port
✕

Name:*

Description:

Logical Switch:*

Admin State:* Up

Attachment Type:*

Attachment ID:

Switching Profiles Type:*

Switching Profiles Id:

In the NSX Manager UI under **Switching > Ports**, find the VIF attachment ID and make sure that the Admin and Operational status are Up/Up.

If two VMs are attached to the same logical switch and have IP addresses configured in the same subnet, they should be able to ping each other.

What to do next

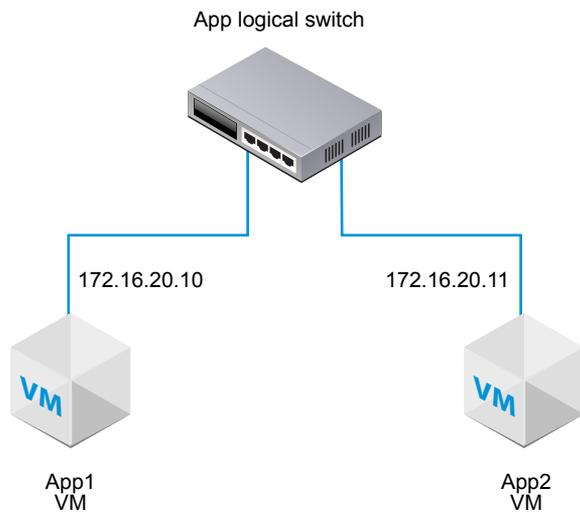
Add a logical router.

You can monitor the activity on the logical switch port to troubleshoot problems. See [Monitor a Logical Switch Port Activity](#).

Test Layer 2 Connectivity

After you successfully set up your logical switch and attach VMs to the logical switch, you can test the network connectivity of the attached VMs.

If your network environment is configured properly, based on the topology the App2 VM can ping the App1 VM.

Figure 2-3. Logical Switch Topology**Procedure**

- 1 Log in to one of the VMs attached to the logical switch using SSH or the VM console.
For example, App2 VM 172.16.20.11.
- 2 Ping the second VM attached to the logical switch to test connectivity.

```
$ ping -c 2 172.16.20.10
PING 172.16.20.10 (172.16.20.10) 56(84) bytes of data.
64 bytes from 172.16.20.10: icmp_seq=1 ttl=63 time=0.982 ms
64 bytes from 172.16.20.10: icmp_seq=2 ttl=63 time=0.654 ms
64 bytes from 172.16.20.10: icmp_seq=3 ttl=63 time=0.791 ms

--- 172.16.20.10 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1990ms
rtt min/avg/max/mdev = 0.654/0.809/0.902/0.104 ms
```

- 3 (Optional) Identify the problem that causes the ping to fail.
 - a Verify that the VM network settings are correct.
 - b Verify that the VM network adapter is connected to the correct logical switch.
 - c Verify that the logical switch Admin status is UP.
 - d From the NSX Manager, select **Switching > Switches**.

- e Click the logical switch and note the UUID and VNI information.
- f From the NSX Controller, run the following commands to troubleshoot the problem.

Command	Description
get logical-switch <vni-or-uuid> arp-table	Displays the ARP table for the specified logical switch. Sample output.
	<pre>nsx-controller1> get logical-switch 41866 arp-table VNI IP MAC Connection-ID 41866 172.16.20.11 00:50:56:b1:70:5e 295422</pre>
get logical-switch <vni-or-uuid> connection-table	Displays the connections for the specified logical switch. Sample output.
	<pre>nsx-controller1> get logical-switch 41866 connection-table Host-IP Port ID 192.168.110.37 36923 295420 192.168.210.53 37883 295421 192.168.210.54 57278 295422</pre>
get logical-switch <vni-or-uuid> mac-table	Displays the MAC table for the specified logical switch. Sample output.
	<pre>nsx-controller1> get logical-switch 41866 mac-table VNI MAC VTEP-IP Connection-ID 41866 00:50:56:86:f2:b2 192.168.250.102 295421 41866 00:50:56:b1:70:5e 192.168.250.101 295422</pre>
get logical-switch <vni-or-uuid> stats	Displays statistics information about the specified logical switch. Sample output.
	<pre>nsx-controller1> get logical-switch 41866 stats update.member 11 update.vtep 11 update.mac 4 update.mac.invalidate 0 update.arp 7 update.arp.duplicate 0 query.mac 2 query.mac.miss 0 query.arp 9 query.arp.miss 6</pre>
get logical-switch <vni-or-uuid> stats-sample	Displays a summary of all logical switch statistics over time. Sample output.
	<pre>nsx-controller1> get logical-switch 41866 stats-sample 21:00:00 21:10:00 21:20:00 21:30:00 21:40:00 update.member 0 0 0 0 0 update.vtep 0 0 0 0 0 update.mac 0 0 0 0 0 update.mac.invalidate 0 0 0 0 0 update.arp 0 0 0 0 0 update.arp.duplicate 0 0 0 0 0</pre>

Command	Description
	<pre>query.mac 0 0 0 0 0 query.mac.miss 0 0 0 0 0 query.arp 0 0 0 0 0 query.arp.miss 0 0 0 0 0</pre>
get logical-switch <vni-or-uuid> vtep	<p>Displays all virtual tunnel end points related to the specified logical switch. Sample output.</p> <pre>nsx-controller1> get logical-switch 41866 vtep VNI IP LABEL Segment MAC Connection-ID 41866 192.168.250.102 0x8801 192.168.250.0 00:50:56:65:f5:fc 295421 41866 192.168.250.100 0x1F801 192.168.250.0 02:50:56:00:00:00 295420 41866 192.168.250.101 0x16001 192.168.250.0 00:50:56:64:7c: 28 295422</pre>

The first VM attached to the logical switch is able to send packets to the second VM.

Configuring Switching Profiles for Logical Switches and Logical Ports

3

Switching profiles include Layer 2 networking configuration details for logical switches and logical ports. NSX Manager supports several types of switching profiles, and maintains one or more system-defined default switching profiles for each profile type.

The following types of switching profiles are available.

- QoS (Quality of Service)
- Port Monitoring
- IP Discovery
- SpoofGuard
- Switch Security
- MAC Management

Note You cannot edit or delete the default switching profiles in the NSX Manager. You can create custom switching profiles instead.

Each default or custom switching profile has a unique reserved identifier. You use this identifier to associate the switching profile to a logical switch or a logical port. For example, the default QoS switching profile ID is f313290b-eba8-4262-bd93-fab5026e9495.

A logical switch or logical port can be associated with one switching profile of each type. You cannot have for example, two QoS different switching profiles associated to a logical switch or logical port.

If you do not associate a switching profile type while creating or updating a logical switch, then the NSX Manager associates a corresponding default system-defined switching profile. The children logical ports inherit the default system-defined switching profile from the parent logical switch.

When you create or update a logical switch or logical port you can choose to associate either a default or a custom switching profile. When the switching profile is associated or disassociated from a logical switch the switching profile for the children logical ports is applied based on the following criteria.

- If the parent logical switch has a profile associated with it, the child logical port inherits the switching profile from the parent.
- If the parent logical switch does not have a switching profile associated with it, a default switching profile is assigned to the logical switch and the logical port inherits that default switching profile.

- If you explicitly associate a custom profile with a logical port, then this custom profile overrides the existing switching profile.

Note If you have associated a custom switching profile with a logical switch, but want to retain the default switching profile for one of the child logical port, then you must make a copy of the default switching profile and associate it with the specific logical port.

You cannot delete a custom switching profile if it is associated to a logical switch or a logical port. You can find out whether any logical switches and logical ports are associated with the custom switching profile by going to the Assigned To section of the Summary view and clicking on the listed logical switches and logical ports.

This chapter includes the following topics:

- [Understanding QoS Switching Profile](#)
- [Understanding Port Mirroring Switching Profile](#)
- [Understanding IP Discovery Switching Profile](#)
- [Understanding SpoofGuard](#)
- [Understanding Switch Security Switching Profile](#)
- [Understanding MAC Management Switching Profile](#)
- [Associate a Custom Profile with a Logical Switch](#)
- [Associate a Custom Profile with a Logical Switch Port](#)

Understanding QoS Switching Profile

QoS provides high-quality and dedicated network performance for preferred traffic that requires high bandwidth. The QoS mechanism does this by prioritizing sufficient bandwidth, controlling latency and jitter, and reducing data loss for preferred packets even when there is a network congestion. This level of network service is provided by using the existing network resources efficiently.

For this release, shaping and traffic marking namely, CoS and DSCP is supported. The Layer 2 Class of Service (CoS) allows you to specify priority for data packets when traffic is buffered in the logical switch due to congestion. The Layer 3 Differentiated Services Code Point (DSCP) detects packets based on their DSCP values. CoS is always applied to the data packet irrespective of the trusted mode.

NSX-T trusts the DSCP setting applied by a virtual machine or modifying and setting the DSCP value at the logical switch level. In each case, the DSCP value is propagated to the outer IP header of encapsulated frames. This enables the external physical network to prioritize the traffic based on the DSCP setting on the external header. When DSCP is in the trusted mode, the DSCP value is copied from the inner header. When in the untrusted mode, the DSCP value is not preserved for the inner header.

Note DSCP settings work only on tunneled traffic. These settings do not apply to traffic inside the same hypervisor.

You can use the QoS switching profile to configure the average ingress and egress bandwidth values to set the transmit limit rate. The peak bandwidth rate is used to support burst traffic a logical switch is allowed to prevent congestion on the northbound network links. These settings however, do not guarantee the bandwidth but help limit the use of network bandwidth.

The QoS switching profile settings are applied to the logical switch and inherited by the child logical switch port.

Configure a Custom QoS Switching Profile

You can define the DSCP value and configure the ingress and egress settings to create a custom QoS switching profile.

Prerequisites

- Familiarize yourself with the QoS switching profile concept. See [Understanding QoS Switching Profile](#).
- Identify the network traffic you want to prioritize.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Switching > Switching Profiles** from the navigation panel.
- 3 Click **Add**.
- 4 Complete the QoS switching profile details.

Option	Description
Name and Description	Assign a name to the custom QoS switching profile. You can optionally describe the setting that you modified in the profile.
Type	Select QoS from the drop-down menu.
DSCP	Select either a Trusted or Untrusted option from the Mode drop-down menu. When you select the Trusted mode the inner header DSCP value is applied to the outer IP header for IP/IPv6 traffic. For non IP/IPv6 traffic, the outer IP header takes the default value. Trusted mode is supported on an overlay-based logical port. The default value is 0. Untrusted mode is supported on overlay-based and VLAN-based logical port. For the overlay-based logical port, the DSCP value of the outbound IP header is set to the configured value irrespective to the inner packet type for the logical port. For the VLAN-based logical port, the DSCP value of IP/IPv6 packet will be set to the configured value. The DSCP values range for untrusted mode is between 0 to 63. Note DSCP settings work only on tunneled traffic. These settings do not apply to traffic inside the same hypervisor.

Option	Description
Class of Service	<p>Set the traffic priority level.</p> <p>CoS is supported on VLAN-based logical port. CoS groups similar types of traffic in the network and each type of traffic is treated as a class with its own level of service priority. The lower priority traffic is slowed down or in some cases dropped to provide better throughput for higher priority traffic. CoS can also be configured for the VLAN ID with zero packet.</p> <p>The CoS values range from 0 to 7, where 0 is the best effort service.</p>
Ingress	<p>Set custom values for the outbound network traffic from the VM to the logical network.</p> <p>You can use the average bandwidth to reduce network congestion. The peak bandwidth rate is used to support burst traffic and the burst duration is set in the burst size setting. You cannot guarantee the bandwidth. However, you can use the setting to limit network bandwidth. The default value 0, disables the ingress traffic.</p> <p>For example, when you set the average bandwidth for the logical switch to 30 Mbps the policy limits the bandwidth. You can cap the burst traffic at 100 Mbps for a duration 20 Bytes.</p>
Ingress Broadcast	<p>Set custom values for the outbound network traffic from the VM to the logical network based on broadcast.</p> <p>The default value 0, disables the ingress broadcast traffic.</p> <p>For example, when you set the average bandwidth for a logical switch to 50 Kbps the policy limits the bandwidth. You can cap the burst traffic to 400 Kbps for a duration of 60 Bytes.</p>
Egress	<p>Set custom values for the inbound network traffic from the logical network to the VM.</p> <p>The default value 0, disables the egress traffic.</p>

If the ingress, ingress broadcast, and egress options are not configured, the default values are used as protocol buffers.

5 Click **Save**.

A custom QoS switching profile appears as a link.

What to do next

Attach this QoS customized switching profile to a logical switch so that the modified parameters in the switching profile are applied to the network traffic. See [Associate a Custom Profile with a Logical Switch](#).

Understanding Port Mirroring Switching Profile

Logical port mirroring lets you replicate and redirect all of the traffic coming in or out of a logical switch port attached to a VM VIF port. The mirrored traffic is sent encapsulated within a Generic Routing Encapsulation (GRE) tunnel to a collector so that all of the original packet information is preserved while traversing the network to a remote destination.

Typically port mirroring is used in the following scenarios:

- Troubleshooting - Analyze the traffic to detect intrusion and debug and diagnose errors on a network.

- Compliance and monitoring - Forward all of the monitored traffic to a network appliance for analysis and remediation.

Compared to the physical port mirroring, logical port mirroring ensures that all of the VM network traffic is captured. If you implement port mirroring only in the physical network, some of the VM network traffic fails to be mirrored. This happens because communication between VMs residing on the same host never enters the physical network and therefore does not get mirrored. With logical port mirroring you can continue to mirror VM traffic even when that VM is migrated to another host.

The port mirroring process is similar for both VM ports in the NSX-T domain and ports of physical applications. You can forward the traffic captured by a workload connected to a logical network and mirror that traffic to a collector. The IP address should be reachable from the guest IP address on which the VM is hosted. This process is also true for physical applications connected to Gateway nodes.

Configure a Custom Port Mirroring Switching Profile

You can create a custom port mirroring switching profile with a different destination and key value.

Prerequisites

- Familiarize yourself with the port mirroring switching profile concept. See [Understanding Port Mirroring Switching Profile](#).
- Identify the IP address of the destination logical port ID you want to redirect network traffic to.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Switching > Switching Profiles** from the navigation panel.
- 3 Click **Add**.
- 4 Complete the port mirroring switching profile details.

Option	Description
Name and Description	Assign a name to the custom port mirroring switching profile. You can optionally describe the setting you modified to customize this profile.
Type	Select Port Mirroring from the drop-down menu.
Direction	Select an option from the drop-down menu to use this source for Ingress , Egress , or Bidirectional traffic. Ingress is the outbound network traffic from the VM to the logical network. Egress is the inbound network traffic from the logical network to the VM. Bidirectional is the two-way of traffic from the VM to the logical network and from the logical network to the VM. This is the default option.
Packet Truncation	Optional. The range is 60 - 65535.

Option	Description
Key	<p>Enter a random 32-bit value to identify mirrored packets from the logical port. This Key value is copied to the Key field in the GRE header of each mirror packet. If the Key value is set to 0, the default definition is copied to the Key field in the GRE header.</p> <p>The default 32-bit value is made of the following values.</p> <ul style="list-style-type: none"> ■ The first 24-bit is a VNI value. VNI is part of the IP header of encapsulated frames. ■ The 25th bit indicates if the first 24-bit is a valid VNI value. One represents a valid value and zero represents an invalid value. ■ The 26th bit indicates the direction of the mirrored traffic. One represents an ingress direction and zero represents an egress direction. ■ The remaining six bits are not used.
Destinations	<p>Enter the destination ID of the collector for the mirroring session. The destination IP address ID can only be an IPv4 address within the network or a remote IPv4 address not managed by NSX-T. You can add up to three destination IP addresses separated by a comma.</p>

5 Click **Save**.

A custom port mirroring switching profile appears as a link.

What to do next

Verify that the customized port mirroring switching profile works. See [Verify Custom Port Mirroring Switching Profile](#).

Verify Custom Port Mirroring Switching Profile

Before you start using the custom port mirroring switching profile, verify that the customization works properly.

Prerequisites

- Verify that the custom port mirroring switching profile is configured. See [Configure a Custom Port Mirroring Switching Profile](#).
- Verify that the customized port mirroring switching profile is attached to a logical switch. See [Associate a Custom Profile with a Logical Switch](#).

Procedure

- 1 Locate two VMs with VIF attachments to the logical port configured for port mirroring.

For example, VM1 10.70.1.1 and VM2 10.70.1.2 have VIF attachments and they are located in the same logical network.

- 2 Run the `tcpdump` command on a destination IP address.

```
sudo tcpdump -n -i eth0 dst host destination_IP_addres and proto gre
```

For example, the destination IP address is 10.24.123.196.

- 3 Log in to the first VM and ping the second VM to verify that the corresponding ECHO requests and replies are received at the destination address.

For example, the first VM 10.70.1.1 pings the second VM 10.70.1.2 to verify port mirroring.

No.	Time	Source	Destination	Protocol	Length	Info
8	0.748510	10.70.1.1	10.70.1.2	ICMP	140	Echo (ping) request id=0x650c, seq=57/14592, ttl=64
9	0.748521	10.70.1.2	10.70.1.1	ICMP	140	Echo (ping) reply id=0x650c, seq=57/14592, ttl=64
30	1.748345	10.70.1.1	10.70.1.2	ICMP	140	Echo (ping) request id=0x650c, seq=58/14848, ttl=64
31	1.748602	10.70.1.2	10.70.1.1	ICMP	140	Echo (ping) reply id=0x650c, seq=58/14848, ttl=64
59	2.748266	10.70.1.1	10.70.1.2	ICMP	140	Echo (ping) request id=0x650c, seq=59/15104, ttl=64
60	2.748515	10.70.1.2	10.70.1.1	ICMP	140	Echo (ping) reply id=0x650c, seq=59/15104, ttl=64
90	3.748306	10.70.1.1	10.70.1.2	ICMP	140	Echo (ping) request id=0x650c, seq=60/15360, ttl=64
91	3.748563	10.70.1.2	10.70.1.1	ICMP	140	Echo (ping) reply id=0x650c, seq=60/15360, ttl=64

What to do next

Attach this port mirroring customized switching profile to a logical switch so that the modified parameters in the switching profile are applied to the network traffic. See [Associate a Custom Profile with a Logical Switch](#).

Understanding IP Discovery Switching Profile

IP Discovery uses DHCP or ARP snooping to learn the VM MAC and IP addresses. After the MAC and IP addresses are learnt, the entries are shared with the NSX Controller to achieve ARP suppression. ARP suppression minimizes ARP traffic flooding within VMs connected to the same logical switch.

DHCP snooping inspects the DHCP packets exchanged between the VM DHCP client and the DHCP server to learn the VM IP and MAC addresses.

ARP snooping inspects the outgoing ARPs and GARPs of the VM to learn the IP and MAC addresses.

Configure IP Discovery Switching Profile

You can enable the ARP snooping or DHCP snooping to create a custom IP Discovery switching profile that learns the IP and MAC addresses to ensure the IP integrity of a logical switch.

Prerequisites

Familiarize yourself with the IP Discovery switching profile concept. See [Understanding IP Discovery Switching Profile](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Switching > Switching Profiles** from the navigation panel.
- 3 Click **Add**.

4 Complete the IP Discovery switching profile details.

Option	Description
Name and Description	Assign a name to the custom IP Discovery switching profile. You can optionally describe the setting you enabled in the profile.
Type	Select IP Discovering from the drop-down menu.
ARP Snooping	Toggle the ARP Snooping button to enable the feature. ARP snooping inspects the VM outgoing ARP and GARP to learn the VM MAC and IP addresses. ARP snooping is applicable if the VM uses a static IP address instead of DHCP.
DHCP Snooping	Toggle the DHCP Snooping button to enable the feature. DHCP snooping inspects the DHCP packets exchanged between the VM DHCP client and the DHCP server, to learn the VM MAC and IP addresses.

5 Click **Save**.

A custom IP Discovery switching profile appears as a link.

What to do next

Attach this IP Discovery customized switching profile to a logical switch so that the modified parameters in the switching profile are applied to the network traffic. See [Associate a Custom Profile with a Logical Switch](#).

Understanding SpoofGuard

SpoofGuard helps prevent a form of malicious attack called "web spoofing" or "phishing." A SpoofGuard policy blocks traffic determined to be spoofed.

SpoofGuard is a tool that is designed to prevent virtual machines in your environment from sending traffic with an IP address it is not authorized to end traffic from. In the instance that a virtual machine's IP address does not match the IP address on the corresponding logical port and switch address binding in SpoofGuard, the virtual machine's vNIC is prevented from accessing the network entirely. SpoofGuard can be configured at the port or switch level. There are several reasons SpoofGuard might be used in your environment:

- Preventing a rogue virtual machine from assuming the IP address of an existing VM.
- Ensuring the IP addresses of virtual machines cannot be altered without intervention – in some environments, it's preferable that virtual machines cannot alter their IP addresses without proper change control review. SpoofGuard facilitates this by ensuring that the virtual machine owner cannot simply alter the IP address and continue working unimpeded.
- Guaranteeing that distributed firewall (DFW) rules will not be inadvertently (or deliberately) bypassed – for DFW rules created utilizing IP sets as sources or destinations, the possibility always exists that a virtual machine could have its IP address forged in the packet header, thereby bypassing the rules in question.

NSX-T SpoofGuard configuration covers the following:

- MAC SpoofGuard - authenticates MAC address of packet
- IP SpoofGuard - authenticates MAC and IP addresses of packet
- Dynamic Address Resolution Protocol (ARP) inspection, that is, ARP and Gratuitous Address Resolution Protocol (GARP) SpoofGuard and Neighbor Discovery (ND) SpoofGuard validation are all against the MAC source, IP Source and IP-MAC source mapping in the ARP/GARP/ND payload.

At the port level, the allowed MAC/VLAN/IP whitelist is provided through the Address Bindings property of the port. When the virtual machine sends traffic, it is dropped if its IP/MAC/VLAN does not match the IP/MAC/VLAN properties of the port. The port level SpoofGuard deals with traffic authentication, i.e. is the traffic consistent with VIF configuration.

At the switch level, the allowed MAC/VLAN/IP whitelist is provided through the Address Bindings property of the switch. This is typically an allowed IP range/subnet for the switch and the switch level SpoofGuard deals with traffic authorization.

Traffic must be permitted by port level AND switch level SpoofGuard before it will be allowed into switch. Enabling or disabling port and switch level SpoofGuard, can be controlled using the SpoofGuard switch profile.

Configure Port Address Bindings

Address bindings specify the IP and MAC address of a logical port and are used to specify the port whitelist in SpoofGuard.

With port address bindings you'll specify the IP and MAC address, and VLAN if applicable, of the logical port. When SpoofGuard is enabled, it ensures that the specified address bindings are enforced in the data path. In addition to SpoofGuard, port address bindings are used for DFW rule translations.

Procedure

- 1 In NSX Manager, navigate to **Switching > Ports**.
- 2 Click the logical port to which you want apply address binding.
The logical port summary appears.
- 3 Under the Summary tab, expand **Address Bindings**.
- 4 Click **Add**.
The Add Address Binding dialogue box appears
- 5 Specify the IP and MAC address of the logical port to which you want to apply address binding. VLAN can also be optionally specified.
- 6 Click **Save**.

What to do next

Use the port address bindings when you [Configure a SpoofGuard Switching Profile](#).

Configure Switch Address Bindings

Address bindings allow a range of IP and MAC addresses and VLANs to be bound to switch.

In SpoofGuard, address bindings provide the allowed MAC/VLAN/IP whitelist. With the corresponding SpoofGuard enabled, SpoofGuard ensures that the specified address bindings are enforced in the data path.

Procedure

- 1 In NSX Manager, navigate to **Switching > Switches**.
- 2 Click the logical switch to which you want apply address binding.
In the right-hand window the switch summary appears.
- 3 Under the Summary tab, expand **Address Bindings**.
- 4 Click **Add**.
The Add Address Binding dialogue box appears.
- 5 Enter the MAC addresses and the IP range of the switch (and VLAN if applicable) in the switch address binding.
After the IP range/subnet is specified, the data path will apply bindings across all ports on the switch.
- 6 Click **Save**.

What to do next

Now you'll [Configure a SpoofGuard Switching Profile](#) and add the address bindings to the SpoofGuard whitelist.

Configure a SpoofGuard Switching Profile

When SpoofGuard is configured, if the IP address of a virtual machine changes, traffic from the virtual machine may be blocked until the corresponding configured port/switch address bindings are updated with the new IP address.

Enable SpoofGuard for the port group(s) containing the guests. When enabled for each network adapter, SpoofGuard inspects packets for the prescribed MAC and its corresponding IP address.

Prerequisites

Before configuring SpoofGuard, add address bindings or switch bindings on each logical switch. Address binding allows you to bind an IP address and MAC address to a port or switch. [Configure Port Address Bindings](#)[Configure Switch Address Bindings](#)

Procedure

- 1 In NSX Manager, navigate to **Switching > Switching Profiles**.

2 Click **Add**.

The New Switching Profile window appears.

3 Name the profile and select **SpoofGuard** as the type. You can also add a profile description.

4 To enable port level SpoofGuard, choose **port bindings**, and to enable switch level SpoofGuard select **switch bindings**.

Address bindings are the allowed whitelist for port and switch SpoofGuard.

5 Click **Save**.

A new switching profile has been created with a SpoofGuard Profile.

What to do next

Associate the SpoofGuard profile with a logical switch. [Associate a Custom Profile with a Logical Switch](#)

Understanding Switch Security Switching Profile

Switch security provides stateless Layer2 and Layer 3 security by checking the ingress traffic to the logical switch and dropping unauthorized packets sent from VMs by matching the IP address, MAC address, and protocols to a set of allowed addresses and protocols. You can use switch security to protect the logical switch integrity by filtering out malicious attacks from the VMs in the network.

You can configure the Bridge Protocol Data Unit (BPDU) filter, DHCP Snooping, DHCP server block, and rate limiting options to customize the switch security switching profile on a logical switch.

Configure a Custom Switch Security Switching Profile

You can create a custom switch security switching profile with MAC destination addresses from the allowed BPDU list and configure rate limiting.

Prerequisites

Familiarize yourself with the switch security switching profile concept. See [Understanding Switch Security Switching Profile](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Switching > Switching Profiles** from the navigation panel.
- 3 Click **Add**.
- 4 Complete the switch security profile details.

Option	Description
Name and Description	Assign a name to the custom switch security profile. You can optionally describe the setting that you modified in the profile.
Type	Select Switch Security from the drop-down menu.

Option	Description
BPDU Filter	<p>Toggle the BPDU filter button to enable BPDU filtering.</p> <p>When the BPDU filter is enabled, all of the traffic to BPDU destination MAC address is blocked. The BPDU filter when enabled also disables STP on the logical switch ports because these ports are not expected to take part in STP.</p>
BPDU Filter Allow List	<p>Click the destination MAC address from the BPDU destination MAC addresses list to allow traffic to the permitted destination.</p>
DHCP Filter	<p>Toggle the Server Block button and Client Block button to enable DHCP filtering.</p> <p>DHCP Server Block blocks traffic from a DHCP server to a DHCP client. Note that it does not block traffic from a DHCP server to a DHCP relay agent.</p> <p>DHCP Client Block prevents a VM from acquiring a DHCP IP address by blocking DHCP requests.</p>
Block Non-IP Traffic	<p>Toggle the Block Non-IP Traffic button to allow only IPv4, IPv6, ARP, GARP and BPDU traffic.</p> <p>The rest of the non-IP traffic is blocked. The permitted IPv4, IPv6, ARP, GARP and BPDU traffic is based on other policies set in address binding and SpoofGuard configuration.</p> <p>By default, this option is disabled to allow non-IP traffic to be handled as regular traffic.</p>
Rate Limits	<p>Set a rate limit for the ingress or egress Broadcast and Multicast traffic.</p> <p>Rate limits are configured to protect the logical switch or the VM from for example, broadcast traffic storms.</p> <p>To avoid any connectivity problems, the minimum rate limit value must be ≥ 10 pps.</p>

5 Click **Save**.

A custom switch security profile appears as a link.

What to do next

Attach this switch security customized switching profile to a logical switch so that the modified parameters in the switching profile are applied to the network traffic. See [Associate a Custom Profile with a Logical Switch](#).

Understanding MAC Management Switching Profile

The MAC management switching profile supports two functionalities: MAC learning and MAC address change.

MAC learning provides network connectivity to deployments where multiple MAC addresses are configured behind one vNIC, for example, in a nested hypervisor deployment where an ESXi VM runs on an ESXi host and multiple VMs run inside the ESXi VM. Without MAC learning, when the ESXi VM's vNIC connects to a switch port, its MAC address is static. VMs running inside the ESXi VM do not have network connectivity because their packets have different source MAC addresses. With MAC learning, the vSwitch inspects the source MAC address of every packet coming from the vNIC, learns the MAC address and allows the packet to go through. If a MAC address that is learned is not used for a certain period of time, it is removed. This aging property is not configurable.

MAC learning also supports unknown unicast flooding. Normally, when a packet that is received by a port has an unknown destination MAC address, the packet is dropped. With unknown unicast flooding enabled, the port floods unknown unicast traffic to every port on the switch that has MAC learning and unicast flooding enabled. This property is enabled by default, but only if MAC learning is enabled.

The MAC management switching profile also supports the ability of a VM to change its MAC address. A VM connected to a port with the MAC address change property enabled can run an administrative command to change the MAC address of its vNIC and still send and receive traffic on that vNIC. This feature is supported on ESXi only and not on KVM. This property is disabled by default.

If you enable MAC learning or MAC address change, to improve security, configure SpoofGuard as well.

For more information about creating a MAC management switching profile and associating the profile with a switch or port, see the *NSX-T API Guide*.

Note In this release, the MAC management switching profile feature is only available through the NSX API. It is not available from the NSX Manager UI.

Associate a Custom Profile with a Logical Switch

To apply your custom switching profile to your network you must associate it with a logical switch.

When custom switching profiles are attached to a logical switch they override existing default switching profiles. The custom switching profile is inherited by children logical switch ports.

Note If you have associated a custom switching profile with a logical switch, but want to retain the default switching profile for one of the child logical switch port, then you must make a copy of the default switching profile and associate it with the specific logical switch port.

Prerequisites

- Verify that a logical switch is configured. See [Create a Logical Switch](#).
- Verify that a custom switching profile is configured. See [Chapter 3 Configuring Switching Profiles for Logical Switches and Logical Ports](#).

Procedure

- 1 From your browser, log in to an NSX Manager at `https://nsx-manager-ip-address`.
- 2 Select **Switching > Switches** from the navigation panel.
- 3 Double-click the logical switch to apply the custom switching profile.
- 4 Click the **Manage** tab.
- 5 Select the custom switching profile type from the drop-down menu.
 - **QoS**
 - **Port Mirroring**
 - **IP Discovering**

- **SpoofGuard**
- **Switch Security**

6 Click **Change**.

7 Select the previously created custom switching profile from the drop-down menu.

8 Click **Save**.

The logical switch is now associated with the custom switching profile.

9 Verify that the new custom switching profile with the modified configuration appears under the **Manage** tab.

10 (Optional) Click the **Related** tab and select **Ports** from the drop-down menu to verify that the custom switching profile is applied to child logical ports.

What to do next

If you do not want to use the inherited switching profile from a logical switch, you can apply a custom switching profile to the child logical switch port. See [Associate a Custom Profile with a Logical Switch Port](#).

Associate a Custom Profile with a Logical Switch Port

A logical switch port provides a logical connection point for a VIF, a patch connection to a router, or a Layer 2 gateway connection to an external network. Logical switch ports also expose switching profiles, port statistics counters, and a logical link status.

You can change the inherited switching profile from the logical switch to a different custom switching profile for the child logical switch port.

Prerequisites

- Verify that a logical switch port is configured. See [Connecting a VM to a Logical Switch](#).
- Verify that a custom switching profile is configured. See [Chapter 3 Configuring Switching Profiles for Logical Switches and Logical Ports](#).

Procedure

1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.

2 Select **Switching > Port** from the navigation panel.

3 Double-click the logical switch port to apply the custom switching profile.

4 Click the **Manage** tab.

5 Select the custom switching profile type from the drop-down menu.

- **QoS**
- **Port Mirroring**
- **IP Discovering**

- **SpoofGuard**
- **Switch Security**

6 Click **Change**.

7 Select the previously created custom switching profile from the drop-down menu.

8 Click **Save**.

The logical switch port is now associated with the custom switching profile.

9 Verify that the new custom switching profile with the modified configuration appears under the **Manage** tab.

What to do next

You can monitor the activity on the logical switch port to troubleshoot problems. See [Monitor a Logical Switch Port Activity](#).

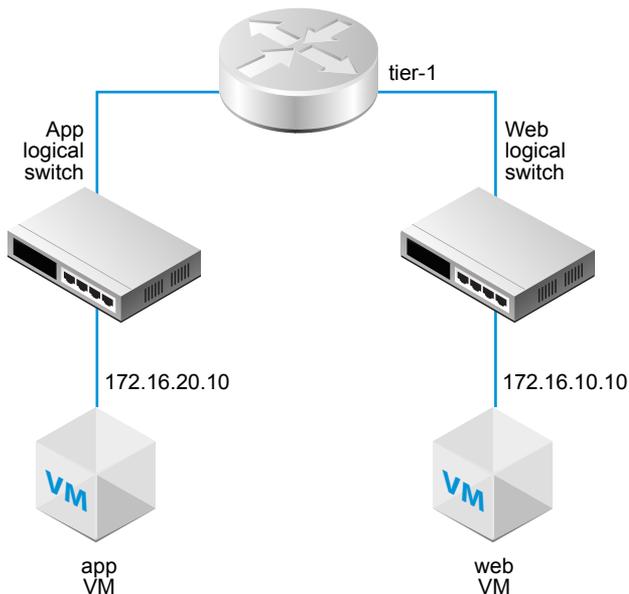
Configuring the Tier-1 Logical Router

4

An NSX-T logical router reproduces routing functionality in a virtual environment completely decoupled from underlying hardware. Tier-1 logical routers have downlink ports to connect to NSX-T logical switches and uplink ports to connect to NSX-T tier-0 logical routers.

When you add a logical router, it is important that you plan the networking topology you are building.

Figure 4-1. Tier-1 Logical Router Topology



For example, this simple topology shows two logical switches connected to a tier-1 logical router. Each logical switch has a single VM connected. The two VMs can be on different hosts or the same host, in different host clusters or in the same host cluster. If a logical router does not separate the VMs, the underlying IP addresses configured on the VMs must be in the same subnet. If a logical router does separate them, the IP addresses on the VMs must be in different subnets.

This chapter includes the following topics:

- [Create a Tier-1 Logical Router](#)
- [Add Downlink Ports for the Tier-1 Logical Router](#)
- [Configure Route Advertisement on a Tier-1 Logical Router](#)

- [Configure a Tier-1 Logical Router Static Route](#)

Create a Tier-1 Logical Router

The tier-1 logical router must be connected to the tier-0 logical router to get the northbound physical router access.

Prerequisites

- Verify that the logical switches are configured. See [Create a Logical Switch](#).
- Verify that an NSX Edge cluster is deployed to perform network address translation (NAT) configuration. See the *NSX-T Installation Guide*.
- Familiarize yourself with the tier-1 logical router topology. See [Chapter 4 Configuring the Tier-1 Logical Router](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Click **Add** and select **Tier-1 Router**.
- 4 Assign a name for the logical router.
- 5 (Optional) Select a tier-0 logical router to connect to this tier-1 logical router.

If you do not yet have any tier-0 logical routers configured, you can leave this field blank for now and edit the router configuration later.

- 6 (Optional) Select an edge cluster to connect to this tier-1 logical router.

If the tier-1 logical router is going to be used for NAT configuration, it must be connected to an NSX Edge cluster. If you do not yet have any edge clusters configured, you can leave this field blank for now and edit the router configuration later.

- 7 Click **Save**.

In the NSX Manager UI, the new logical router is a clickable link.

What to do next

Create downlink ports for your tier-1 logical router. See [Add Downlink Ports for the Tier-1 Logical Router](#).

Add Downlink Ports for the Tier-1 Logical Router

When you create a downlink port on a tier-1 logical router, the port serves as a default gateway for the VMs that are in the same subnet.

Prerequisites

Verify that a tier-1 logical router is configured. See [Create a Tier-1 Logical Router](#).

Procedure

- 1 Click the tier-1 logical router link to create ports.
- 2 Click the **Configuration** tab.
- 3 Click **Add** under the Logical Router Ports section.
- 4 Assign a name for the logical router port.
- 5 Select whether this attachment creates a switch port or updates an existing switch port.
If the attachment is for an existing switch port, select the port from the drop-down menu.
- 6 Enter the router port IP address in CIDR notation.
For example, the IP address can be 172.16.10.1/24.
You can also enter a preconfigured DHCP service IP address.
- 7 Click **Save**.
- 8 (Optional) Repeat steps 1-7 for creating additional tier-1 logical router ports.
- 9 Verify that the tier-1 logical router can route East-West VM traffic.

In this example, the tier-1 logical router has two downlink ports that connect to two logical switches. Each logical switch has a VM attached. The VMs can ping each other.

```
web-virtual-machine$ ping 172.16.20.10
PING 172.16.20.10 (172.16.20.10): 56(84) data bytes
64 bytes from 172.16.20.10: icmp_req=0 ttl=64 time=178 ms
^C
--- 172.16.20.10 ping statistics ---
1 packets transmitted, 1 packets received, 0.0% packet loss
round-trip min/avg/max/stddev = 178 ms/178 ms/178 ms/0.000 ms
```

```
app-virtual-machine$ ping 172.16.10.10
PING 172.16.10.10 (172.16.10.10): 56(84) data bytes
64 bytes from 172.16.10.10: icmp_req=0 ttl=64 time=178 ms
^C
--- 172.16.10.10 ping statistics ---
1 packets transmitted, 1 packets received, 0.0% packet loss
round-trip min/avg/max/stddev = 178 ms/178 ms/178 ms/0.000 ms
```

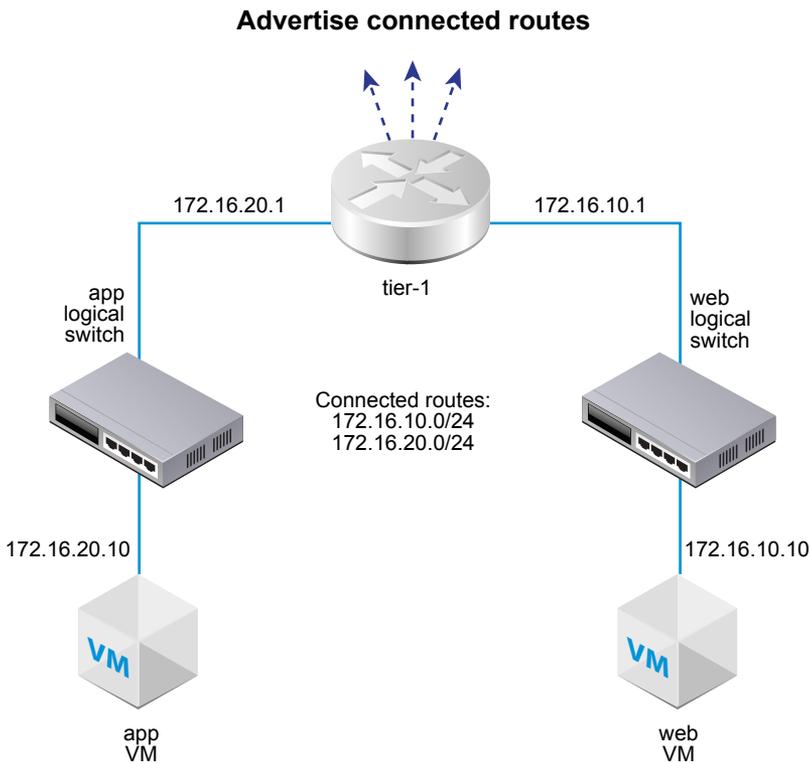
What to do next

Enable route advertisement to provide North-South connectivity between VMs and external physical networks or between different tier-1 logical routers that are connected to the same tier-0 logical router. See [Configure Route Advertisement on a Tier-1 Logical Router](#).

Configure Route Advertisement on a Tier-1 Logical Router

To provide Layer 3 connectivity between VMs connected to logical switches that are attached to different tier-1 logical routers, it is necessary to enable tier-1 route advertisement towards tier-0. You do not need to configure a routing protocol or static routes between tier-1 and tier-0 logical routers. NSX-T creates NSX-T static routes automatically when you enable route advertisement.

For example, to provide connectivity to and from the VMs through other peer routers, the tier-1 logical router must have route advertisement configured for connected routes. If you don't want to advertise all connected routes, you can specify which routes to advertise.



Prerequisites

- Verify that VMs are attached to logical switches. See [Chapter 2 Creating Logical Switches and Configuring VM Attachment](#).
- Verify that downlink ports for the tier-1 logical router are configured. See [Add Downlink Ports for the Tier-1 Logical Router](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing**.
- 3 Click a tier-1 logical router.

- 4 Select **Route Advertisement** from the Routing drop-down menu.
- 5 Enable route advertisement by clicking **Edit** and making sure the Status button is Enabled.
- 6 Specify which routes to advertise, either all routes or selected routes.
 - Click **Edit** and select **Advertise All NSX Connected Routes**.
 - Click **Add** and enter information about the routes to be advertised. For each route, you can enter a name and a route prefix in CIDR format.
- 7 Click the **Status** toggle button to enable Route Advertisement.

For example:

The screenshot shows the NSX-T interface for configuring a Tier-1 logical router. The 'ROUTING' section is active, and the 'Routing' tab is selected. The 'Route Advertisement' section is visible, showing the following settings:

Setting	Status
Status	Enabled
Advertise All NSX Connected Routes	Yes
Advertise NAT Routes	Yes
Advertise Static Routes	No

- 8 Click **Save**.

What to do next

Familiarize yourself with the tier-0 logical router topology and create the tier-0 logical router. See [Chapter 5 Configuring a Tier-0 Logical Router](#).

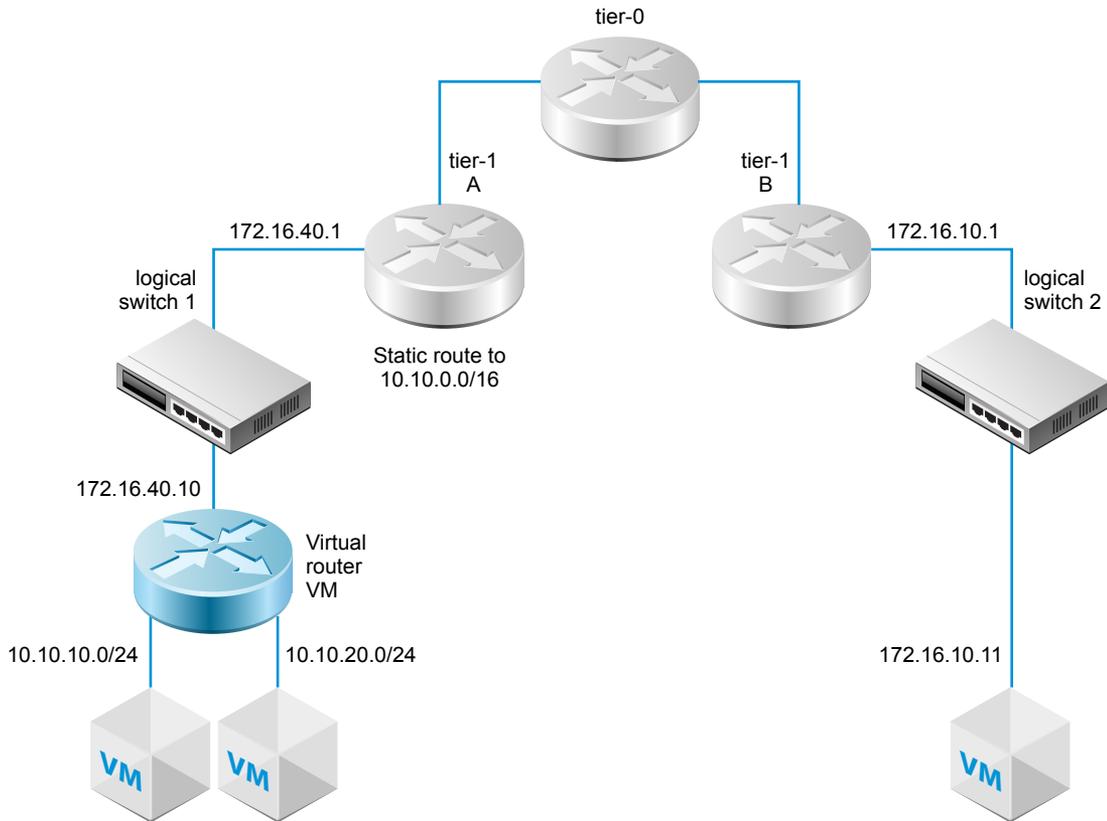
If you already have a tier-0 logical router connected to the tier-1 logical router, you can verify that the tier-0 router is learning the tier-1 router connected routes. See [Verify that a Tier-0 Router Has Learned Routes from a Tier-1 Router](#).

Configure a Tier-1 Logical Router Static Route

You can configure a static route on a tier-1 logical router to provide connectivity from NSX-T to a set of networks that are accessible through a virtual router.

For example, in the following diagram, the tier-1 A logical router has a downlink port to an NSX-T logical switch. This downlink port (172.16.40.1) serves the default gateway for the virtual router VM. The virtual router VM and tier-1 A are connected through the same NSX-T logical switch. The tier-1 logical router has a static route 10.10.0.0/16 that summarizes the networks available through the virtual router. Tier-1 A then has route advertisement configured to advertise the static route to tier-1 B.

Figure 4-2. Tier-1 Logical Router Static Route Topology



Prerequisites

Verify that a downlink port is configured. See [Add Downlink Ports for the Tier-1 Logical Router](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Select the tier-1 logical router.
- 4 Click the **Routing** tab and select **Static Route** from the drop-down menu.
- 5 Select **Add**.
- 6 Enter a network address in the CIDR format.
For example, 10.10.10.0/16.
- 7 Click **Insert Row** to add a next-hop IP address.
For example, 172.16.40.10.
- 8 Click **Save**.
The newly created static route network address appears in the row.
- 9 From the tier-1 logical router, select **Routing > Route Advertisement**.

10 Click **Edit** and select **Advertise Static Routes**.

11 Click **Save**.

The static route is propagated across the NSX-T overlay.

Configuring a Tier-0 Logical Router

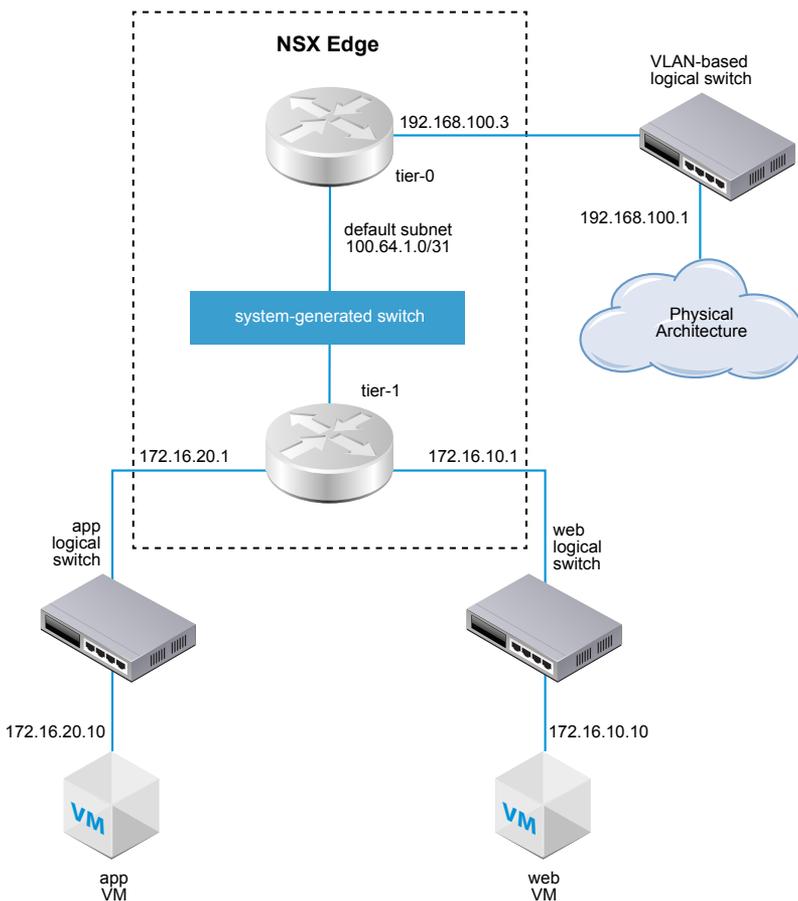
5

An NSX-T logical router reproduces routing functionality in a virtual environment completely decoupled from underlying hardware. The tier-0 logical router provides an on and off gateway service between the logical and physical network.

An NSX Edge cluster can back multiple tier-0 logical routers. Tier-0 routers support the BGP dynamic routing protocol and ECMP.

When you add a tier-0 logical router, it is important that you map out the networking topology you are building.

Figure 5-1. Tier-0 Logical Router Topology



For simplicity, the sample topology shows a single tier-1 logical router connected to a single tier-0 logical router hosted on a single NSX Edge node. Keep in mind that this is not a recommended topology. Ideally, you should have a minimum of two NSX Edge nodes to take full advantage of the logical router design.

The tier-1 logical router has a web logical switch and an app logical switch with respective VMs attached. The router-link switch between the tier-1 router and the tier-0 router is created automatically when you attach the tier-1 router to the tier-0 router. Thus, this switch is labeled as system generated.

This chapter includes the following topics:

- [Create a Tier-0 Logical Router](#)
- [Attach Tier-0 and Tier-1](#)
- [Connect a Tier-0 Logical Router to a VLAN Logical Switch](#)
- [Configure a Static Route](#)
- [BGP Configuration Options](#)
- [Configure BFD on a Tier-0 Logical Router](#)
- [Enable Route Redistribution on the Tier-0 Logical Router](#)
- [Understanding ECMP Routing](#)
- [Create an IP Prefix List](#)
- [Create a Route Map](#)

Create a Tier-0 Logical Router

Tier-0 logical routers have downlink ports to connect to NSX-T tier-1 logical routers and uplink ports to connect to external networks.

Prerequisites

- Verify that at least one NSX Edge is installed. See the *NSX-T Installation Guide*
- Verify that your NSX Controller cluster is stable.
- Verify that an edge cluster is configured. See the *NSX-T Installation Guide*.
- Familiarize yourself with the networking topology of the tier-0 logical router. See [Chapter 5 Configuring a Tier-0 Logical Router](#).

Procedure

- 1 From your browser, log in to an NSX Manager at `https://nsx-manager-ip-address`.
- 2 Select **Routing** from the navigation panel.
- 3 Click **Add** to create a tier-0 logical router.
- 4 Select **Tier-0 Router** from the drop-down menu.
- 5 Assign a name for the tier-0 logical router.

6 Select an existing edge cluster from the drop-down menu to back this tier-0 logical router.

7 (Optional) Select a high-availability mode.

By default, the active-active mode is used. In the active-active mode, traffic is load balanced across all members. In active-standby mode, all traffic is processed by an elected active member. If the active member fails, a new member is elected to be active.

8 (Optional) Click the **Advanced** tab to enter a subnet for the intra-tier 0 transit subnet.

This is the subnet that connects to the tier-0 services router to its distributed router. If you leave this blank, the default 169.0.0.0/28 subnet is used.

9 (Optional) Click the **Advanced** tab to enter a subnet for the tier-0-tier-1 transit subnet.

This is the subnet that connects the tier-0 router to any tier-1 routers that connect to this tier-0 router. If you leave this blank, the default address space assigned for these tier-0-to-tier-1 connections is 100.64.0.0/10. Each tier-0-to-tier-1 peer connection is provided a /31 subnet within the 100.64.0.0/10 address space.

10 Click **Save**.

The new tier-0 logical router appears as a link.

11 (Optional) Click the tier-0 logical router link to review the summary.

What to do next

Attach tier-1 logical routers to this tier-0 logical router.

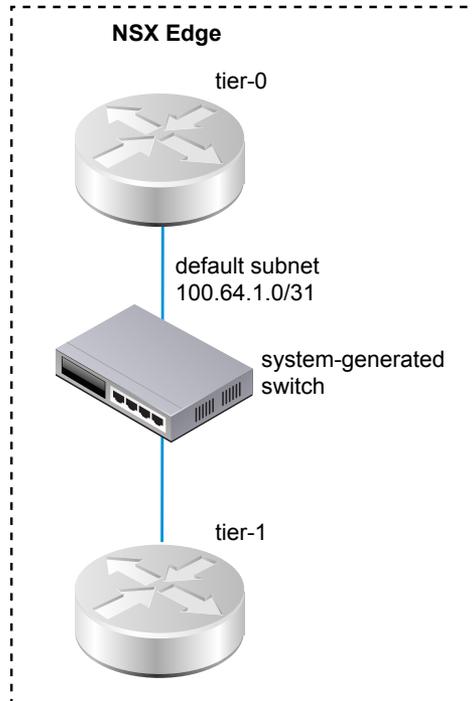
Configure the tier-0 logical router to connect it to a VLAN logical switch to create an uplink to an external network. See [Connect a Tier-0 Logical Router to a VLAN Logical Switch](#).

Attach Tier-0 and Tier-1

You can attach the tier-0 logical router to the tier-1 logical router so that the tier-1 logical router gets northbound and east-west network connectivity.

When you attach a tier-1 logical router to a tier-0 logical router, a router-link switch between the two routers is created. This switch is labeled as system-generated in the topology. The default address space assigned for these tier-0-to-tier-1 connections is 100.64.0.0/10. Each tier-0-to-tier-1 peer connection is provided a /31 subnet within the 100.64.0.0/10 address space. Optionally, you can configure the address space in the tier-0 **Summary > Advanced** configuration.

The following figure shows a sample topology.



Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Select the tier-1 logical router.
- 4 From the **Summary** tab, click **Edit**.
- 5 Select the tier-0 logical router from the drop-down menu.
- 6 (Optional) Select an edge cluster from the drop-down menu.

The tier-1 router needs to be backed by an edge device if the router is going to be used for services, such as NAT. If you do not select an edge cluster, the tier-1 router cannot perform NAT.

- 7 Specify members and a preferred member.

If you select an edge cluster and leave the members and preferred member fields blank, NSX-T sets the backing edge device from the specified cluster for you.

- 8 Click **Save**.
- 9 Click the **Configuration** tab of the tier-1 router to verify that a new point-to-point linked port IP address is created.

For example, the IP address of the linked port can be 100.64.1.1/31.

- 10 Select the tier-0 logical router from the navigation panel.

- Click the **Configuration** tab of the tier-0 router to verify that a new point-to-point linked port IP address is created.

For example, the IP address of the linked port can be 100.64.1.1/31.

What to do next

Verify that the tier-0 router is learning routes that are advertised by the tier-1 routers.

Verify that a Tier-0 Router Has Learned Routes from a Tier-1 Router

When a tier-1 logical router advertises routes to a tier-0 logical router, the routes are listed in the tier-0 router's routing table as NSX-T static routes.

Procedure

- On the NSX Edge, run the `get logical-routers` command to find the VRF number of the tier-0 service router.

```

nsx-edge-1> get logical-routers
Logical Router
UUID       : 736a80e3-23f6-5a2d-81d6-bbefb2786666
vrf        : 0
type       : TUNNEL

Logical Router
UUID       : 421a2d0d-f423-46f1-93a1-2f9e366176c8
vrf        : 5
type       : SERVICE_ROUTER_TIER0

Logical Router
UUID       : f3ce9d7d-7123-47d6-aba6-45cf1388ca7b
vrf        : 6
type       : DISTRIBUTED_ROUTER

Logical Router
UUID       : c8e64eff-02b2-4462-94ff-89f3788f1a61
vrf        : 7
type       : SERVICE_ROUTER_TIER1

Logical Router
UUID       : fb6c3f1f-599f-4421-af8a-99692dff3dd4
vrf        : 8
type       : DISTRIBUTED_ROUTER

```

- Run the `vrf <number>` command to enter the tier-0 service router context.

```

nsx-edge-1> vrf 5
nsx-edge1(tier0_sr)>

```

- 3 On the tier-0 service router, run the `get route` command and make sure the expected routes appear in the routing table.

Notice that the NSX-T static routes (ns) are learned by the tier-0 router because the tier-1 router is advertising routes.

```
nsx-edge1(tier0_sr)> get route

Flags: c - connected, s - static, b - BGP, ns - nsx_static
nc - nsx_connected, rl - router_link, t0n: Tier0-NAT, t1n: Tier1-NAT

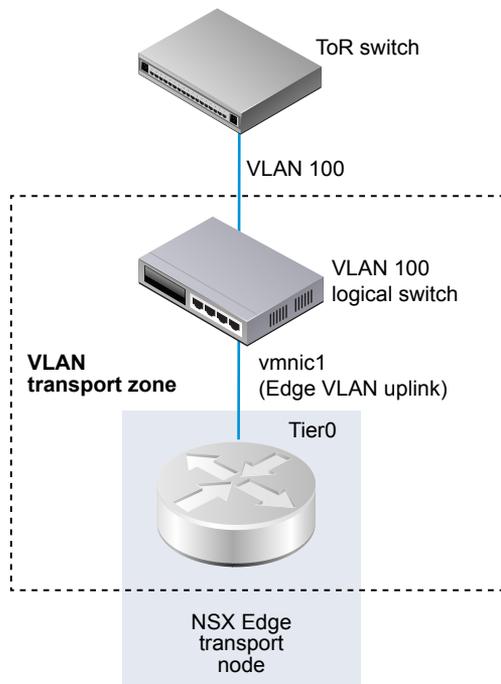
Total number of routes: 7

b  10.10.10.0/24      [20/0]      via 192.168.100.254
rl 100.91.176.0/31   [0/0]      via 169.254.0.1
c  169.254.0.0/28    [0/0]      via 169.254.0.2
ns 172.16.10.0/24    [3/3]      via 169.254.0.1
ns 172.16.20.0/24   [3/3]      via 169.254.0.1
c  192.168.100.0/24 [0/0]      via 192.168.100.2
```

Connect a Tier-0 Logical Router to a VLAN Logical Switch

To create the Edge uplink, you connect a tier-0 router to the VLAN switch.

The following simple topology shows a VLAN logical switch inside of a VLAN transport zone. The VLAN logical switch has a VLAN ID that matches the VLAN ID on the TOR port for the Edge's VLAN uplink.



Prerequisites

Create a VLAN logical switch. See [Create a VLAN Logical Switch for the NSX Edge Uplink](#).

Create a tier-0 router.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Select the tier-0 logical router.
- 4 From the **Configuration** tab, add a new logical router port.
- 5 Type a name for the port, such as uplink.
- 6 Select the **Uplink** type.
- 7 Select an edge transport node.
- 8 Select a VLAN logical switch.
- 9 Type an IP address in CIDR format in the same subnet as the connected port on the TOR switch.

For example:

New Router Port
✕

Name: *

Description:

Type: Uplink
 Downlink

Transport Node: *

Logical Switch: ✕ ▼
OR Create a New Switch

Logical Switch Port: Attach to new switch port
Switch Port Name:
 Attach to existing switch port

IP Address/mask: *

A new uplink port is added for the tier-0 router.

What to do next

Configure BGP or a static route.

Verify the Tier-0 Logical Router and TOR Connection

For routing to work on the uplink from the tier-0 router, connectivity with the top-of-rack device must be in place.

Prerequisites

- Verify that the tier-0 logical router is connected to a VLAN logical switch. See [Connect a Tier-0 Logical Router to a VLAN Logical Switch](#).

Procedure

- 1 Log in to the NSX Manager CLI.
- 2 On the NSX Edge, run the `get logical-routers` command to find the VRF number of the tier-0 service router.

```

nsx-edge-1> get logical-routers
Logical Router
UUID       : 736a80e3-23f6-5a2d-81d6-bbefb2786666
vrf        : 0
type       : TUNNEL

Logical Router
UUID       : 421a2d0d-f423-46f1-93a1-2f9e366176c8
vrf        : 5
type       : SERVICE_ROUTER_TIER0

Logical Router
UUID       : f3ce9d7d-7123-47d6-aba6-45cf1388ca7b
vrf        : 6
type       : DISTRIBUTED_ROUTER

Logical Router
UUID       : c8e64eff-02b2-4462-94ff-89f3788f1a61
vrf        : 7
type       : SERVICE_ROUTER_TIER1

Logical Router
UUID       : fb6c3f1f-599f-4421-af8a-99692dff3dd4
vrf        : 8
type       : DISTRIBUTED_ROUTER

```

- 3 Run the `vrf <number>` command to enter the tier-0 service router context.

```
nsx-edge-1> vrf 5
nsx-edge1(tier0_sr)>
```

- 4 On the tier-0 service router, run the `get route` command and make sure the expected route appears in the routing table.

Notice that the route to the TOR appears as connected (c).

```
nsx-edge1(tier0_sr)> get route

Flags: c - connected, s - static, b - BGP, ns - nsx_static
nc - nsx_connected, rl - router_link, t0n: Tier0-NAT, t1n: Tier1-NAT

Total number of routes: 7

b   10.10.10.0/24      [20/0]      via 192.168.100.254
rl  100.91.176.0/31    [0/0]      via 169.254.0.1
c   169.254.0.0/28     [0/0]      via 169.254.0.2
ns  172.16.10.0/24     [3/3]      via 169.254.0.1
ns  172.16.20.0/24    [3/3]      via 169.254.0.1
c   192.168.100.0/24  [0/0]      via 192.168.100.2
```

- 5 Ping the TOR.

```
nsx-edge1(tier0_sr)> ping 192.168.100.254
PING 192.168.100.254 (192.168.100.254): 56 data bytes
64 bytes from 192.168.100.254: icmp_seq=0 ttl=64 time=2.822 ms
64 bytes from 192.168.100.254: icmp_seq=1 ttl=64 time=1.393 ms
^C
nsx-edge1>
--- 192.168.100.254 ping statistics ---
3 packets transmitted, 2 packets received, 33.3% packet loss
round-trip min/avg/max/stddev = 1.393/2.107/2.822/0.715 ms
```

Packets are sent between the tier-0 logical router and physical router to verify a connection.

What to do next

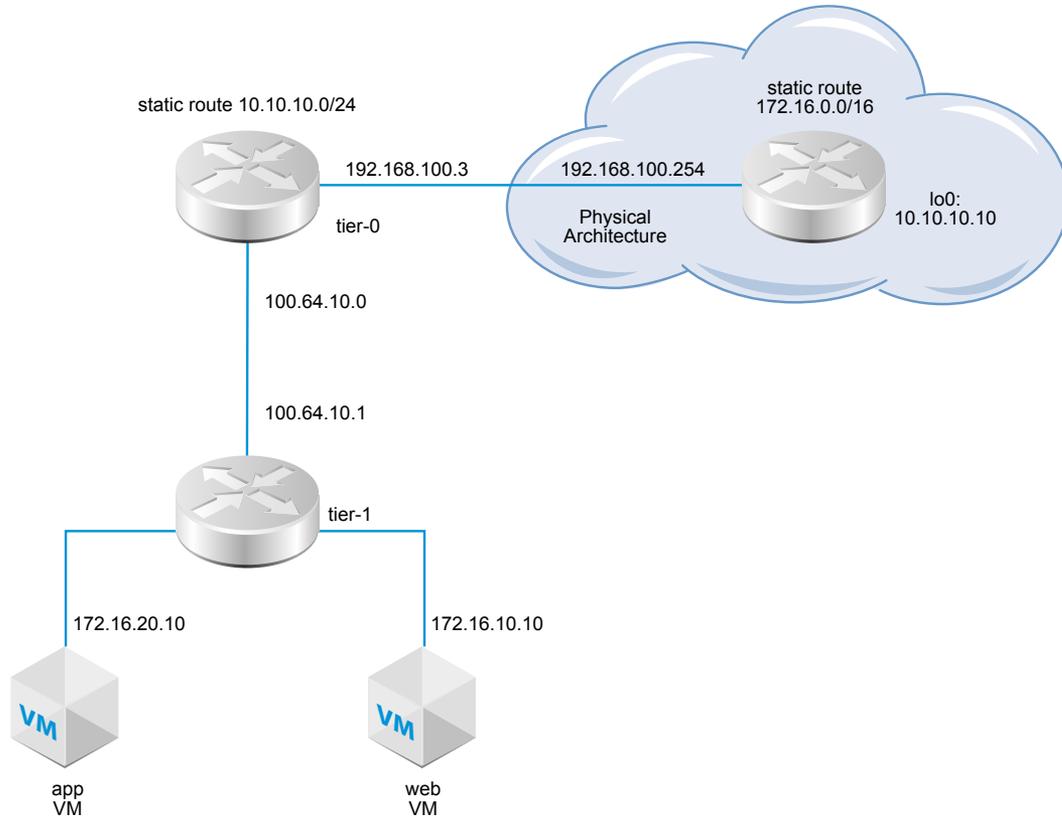
Depending on your networking requirements, you can configure a static route or BGP. See [Configure a Static Route](#) or [Configure BGP on a Tier-0 Logical Router](#).

Configure a Static Route

You can configure a static route on the tier-0 router to external networks. After you configure a static route, there is no need to advertise the route from tier-0 to tier-1, because tier-1 routers automatically have a static default route towards their connected tier-0 router.

The static route topology shows a tier-0 logical router with a static route to the 10.10.10.0/24 prefix in the physical architecture. For test purposes, the 10.10.10.10/32 address is configured on the external router loopback interface. The external router has a static route to the 172.16.0.0/16 prefix to reach the app and web VMs.

Figure 5-2. Static Route Topology



Prerequisites

- Verify that the physical router and tier-0 logical router are connected. See [Verify the Tier-0 Logical Router and TOR Connection](#).
- Verify that the tier-1 router is configured to advertise connected routes. See [Create a Tier-1 Logical Router](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Select the tier-0 logical router.
- 4 Click the **Routing** tab and select **Static Route** from the drop-down menu.
- 5 Select **Add**.

- 6 Enter a network address in the CIDR format.

For example, 10.10.10.0/24.

- 7 Click **Insert Row** to add a next-hop IP address.

For example, 192.168.100.254.

- 8 Click **Save**.

The newly created static route network address appears in the row.

What to do next

Check that the static route is configured properly. See [Verify the Static Route](#).

Verify the Static Route

Use the CLI to verify that the static route is connected. You must also verify the external router can ping the internal VMs and the internal VMs can ping the external router.

Prerequisites

Verify that a static route is configured. See [Configure a Static Route](#).

Procedure

- 1 Log in to the NSX Manager CLI.

2 Confirm the static route.

- a Get the service router UUID information.

```
get logical-routers
```

```
nsx-edge1> get logical-routers
Logical Router
UUID       : 736a80e3-23f6-5a2d-81d6-bbefb2786666
vrf        : 2
type       : TUNNEL

Logical Router
UUID       : d40bbfa4-3e3d-4178-8615-6f42ea335037
vrf        : 4
type       : SERVICE_ROUTER_TIER0

Logical Router
UUID       : d0289ba4-250e-41b4-8ffc-7cab4a46c3e4
vrf        : 5
type       : DISTRIBUTED_ROUTER

Logical Router
UUID       : a6ee6316-2212-4171-99cc-930c98bcad7f
vrf        : 6
type       : DISTRIBUTED_ROUTER
```

- b Locate the UUID information from the output.

```
Logical Router
UUID       : d40bbfa4-3e3d-4178-8615-6f42ea335037
vrf        : 4
type       : SERVICE_ROUTER_TIER0
```

- c Verify that the static route works.

```
get logical-router d40bbfa4-3e3d-4178-8615-6f42ea335037 static
```

```
Flags: c - connected, s - static, b - BGP, ns - nsx_static
nc - nsx_connected, rl - router_link, t0n: Tier0-NAT, t1n: Tier1-NAT

s   10.10.10.0/24      [1/1]      via 192.168.100.254
rl  100.64.1.0/31     [0/0]      via 169.0.0.1
ns  172.16.10.0/24    [3/3]      via 169.0.0.1
ns  172.16.20.0/24    [3/3]      via 169.0.0.1
```

- 3 From the external router, ping the internal VMs to confirm that they are reachable through the NSX-T overlay.

- a Connect to the external router.

```
ping 172.16.10.10
```

```
PING 172.16.10.10 (172.16.10.10) 56(84) bytes of data.
64 bytes from 172.16.10.10: icmp_req=1 ttl=62 time=127 ms
64 bytes from 172.16.10.10: icmp_req=2 ttl=62 time=1.96 ms
^C
--- 172.16.10.10 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1001ms
rtt min/avg/max/mdev = 1.966/64.793/127.620/62.827 ms
```

- b Test the network connectivity.

```
tracert 172.16.10.10
```

```
tracert to 172.16.10.10 (172.16.10.10), 30 hops max, 60 byte packets
 1 192.168.100.3 (192.168.100.3) 0.640 ms 0.575 ms 0.696 ms
 2 100.64.1.1 (100.64.1.1) 0.656 ms 0.604 ms 0.578 ms
 3 172.16.10.10 (172.16.10.10) 3.397 ms 3.703 ms 3.790 ms
```

- 4 From the VMs, ping the external IP address.

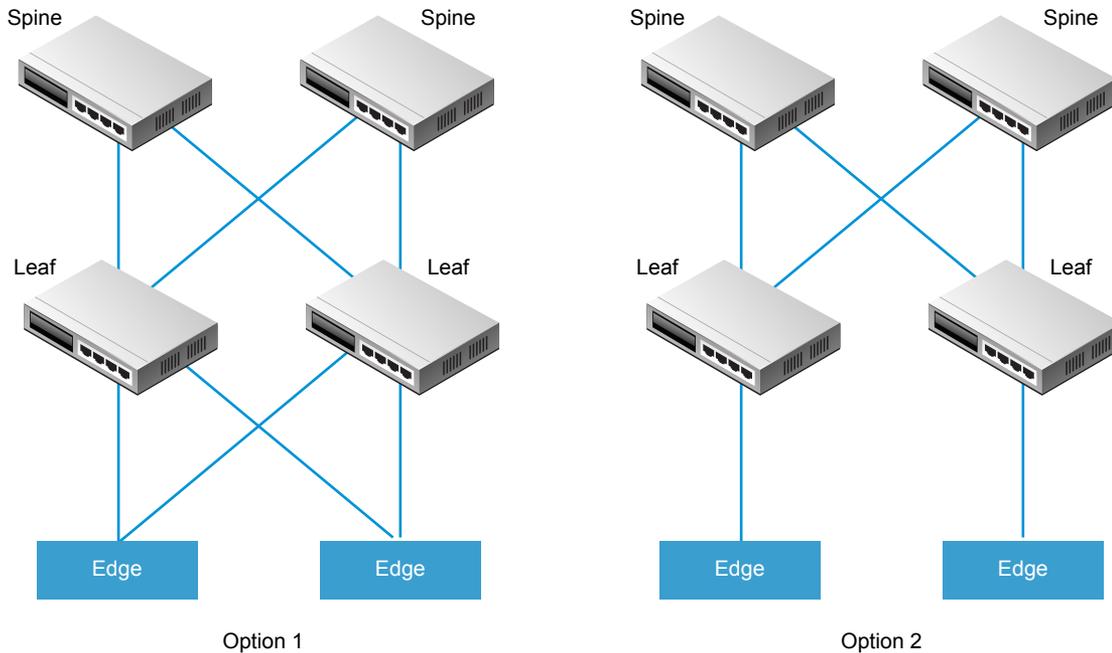
```
ping 10.10.10.10
```

```
PING 10.10.10.10 (10.10.10.10) 56(84) bytes of data.
64 bytes from 10.10.10.10: icmp_req=1 ttl=62 time=119 ms
64 bytes from 10.10.10.10: icmp_req=2 ttl=62 time=1.93 ms
^C
--- 10.10.10.10 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1001ms
rtt min/avg/max/mdev = 1.936/60.865/119.795/58.930 ms
```

BGP Configuration Options

To take full advantage of the tier-0 logical router, the topology must be configured with redundancy and symmetry with BGP between the tier-0 routers and the external top-of-rack peers. This design helps to ensure connectivity in the event of link and node failures.

There are two modes of configuration: active-active and active-standby. The following diagram shows two options for symmetric configuration. There are two NSX Edge nodes shown in each topology. In the case of an active-active configuration, when you create tier-0 uplink ports, you can associate each uplink port with up to eight NSX Edge transport nodes. Each NSX Edge node can have two uplinks.



For option 1, when the physical leaf-node routers are configured, they should have BGP neighborships with the NSX Edges. Route redistribution should include the same network prefixes with equal BGP metrics to all of the BGP neighbors. In the tier-0 logical router configuration, all leaf-node routers should be configured as BGP neighbors.

When you are configuring the tier-0 router's BGP neighbors, if you do not specify a local address (the source IP address), the BGP neighbor configuration is sent to all NSX Edge nodes associated with the tier-0 logical router uplinks. If you do configure a local address, the configuration goes to the NSX Edge node with the uplink owning that IP address.

In the case of option1, if the uplinks are on the same subnet on the NSX Edge nodes, it makes sense to omit the local address. If the uplinks on the NSX Edge nodes are in different subnets, the local address should be specified in the tier-0 router's BGP neighbor configuration to prevent the configuration from going to all associated NSX Edge nodes.

For option 2, ensure that the tier-0 logical router configuration includes the tier-0 services router's local IP address. The leaf-node routers are configured with only the NSX Edges that they are directly connected to as the BGP neighbor.

Configure BGP on a Tier-0 Logical Router

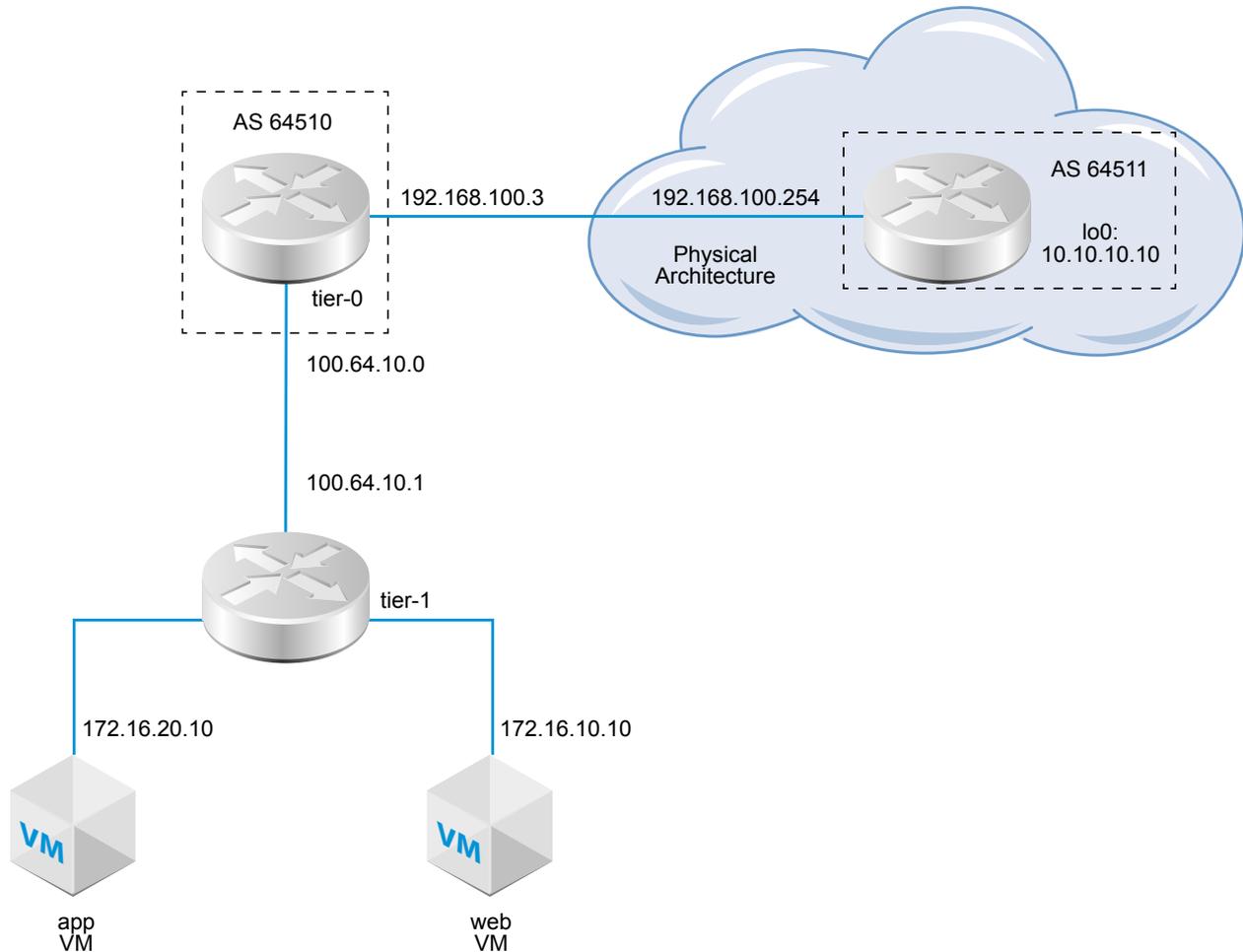
To enable access between your VMs and the outside world, you can configure an external BGP (eBGP) connection between a tier-0 logical router and a router in your physical infrastructure.

When configuring BGP, you must configure a local Autonomous System (AS) number for the tier-0 logical router. For example, the following topology shows the local AS number is 64510. You must also configure the remote AS number of the physical router. In this example, the remote AS number is 64511. The remote neighbor IP address is 192.168.100.254. The neighbor must be in the same IP subnet as the uplink on the tier-0 logical router. BGP multihop is not supported.

For test purposes, the 10.10.10.10/32 address is configured on the external router loopback interface.

Note Router ID used for forming BGP sessions on an edge node is auto-selected from the IP addresses configured on the uplinks of a tier-0 logical router. BGP sessions on an edge node can flap when router ID changes. This can happen when the IP address auto-selected for router ID is deleted or the logical router port on which this IP is assigned is deleted.

Figure 5-3. BGP Connection Topology



Prerequisites

- Verify that the tier-1 router is configured to advertise connected routes. See [Configure Route Advertisement on a Tier-1 Logical Router](#). This is not strictly a prerequisite for BGP configuration, but if you have a two-tier topology and you plan to redistribute your tier-1 networks into BGP, this step is required.
- Verify that a tier-0 router is configured. See [Create a Tier-0 Logical Router](#).
- Make sure the tier-0 logical router has learned routes from the tier-1 logical router. See [Verify that a Tier-0 Router Has Learned Routes from a Tier-1 Router](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Select the tier-0 logical router.
- 4 Click the **Routing** tab and select **BGP** from the drop-down menu.
- 5 Click **Edit** to configure the local AS number and click **Save**.
For example, 64510.
- 6 Click the **Status** toggle button to enable BGP.
The Status button must be appear as Enabled.
- 7 (Optional) Configure route aggregation, enable graceful restart, and enable ECMP.
Graceful restart is only supported if the edge cluster associated with the tier-0 router has only one edge node.
- 8 Click **Save**.
- 9 Click **Add** under the Neighbors section to add a BGP neighbor.
- 10 Enter the neighbor IP address.
For example, 192.168.100.254.
- 11 (Optional) Select a local address from the drop-down menu.
- 12 Enter the remote AS number.
For example, 64511.
- 13 (Optional) Configure the timers (keep alive time and hold down time) and a password.
- 14 (Optional) Add an address family and configure route filtering and route maps.

What to do next

Test whether BGP is working properly. See [Verify BGP Connections from a Tier-0 Service Router](#).

Verify BGP Connections from a Tier-0 Service Router

Use the CLI to verify from the tier-0 service router that a BGP connection to a neighbor is established.

Prerequisites

Verify that BGP is configured. See [Configure BGP on a Tier-0 Logical Router](#).

Procedure

- 1 Log in to the NSX Manager CLI.

- 2 On the NSX Edge, run the `get logical-routers` command to find the VRF number of the tier-0 service router.

```

nsx-edge-1> get logical-routers
Logical Router
UUID       : 736a80e3-23f6-5a2d-81d6-bbefb2786666
vrf        : 0
type       : TUNNEL

Logical Router
UUID       : 421a2d0d-f423-46f1-93a1-2f9e366176c8
vrf        : 5
type       : SERVICE_ROUTER_TIER0

Logical Router
UUID       : f3ce9d7d-7123-47d6-aba6-45cf1388ca7b
vrf        : 6
type       : DISTRIBUTED_ROUTER

Logical Router
UUID       : c8e64eff-02b2-4462-94ff-89f3788f1a61
vrf        : 7
type       : SERVICE_ROUTER_TIER1

Logical Router
UUID       : fb6c3f1f-599f-4421-af8a-99692dff3dd4
vrf        : 8
type       : DISTRIBUTED_ROUTER

```

- 3 Run the `vrf <number>` command to enter the tier-0 service router context.

```

nsx-edge-1> vrf 5
nsx-edge1(tier0_sr)>

```

- 4 Verify that the BGP state is Established, up.

```
get bgp neighbor
```

```

BGP neighbor: 192.168.100.254   Remote AS: 64511
BGP state: Established, up
Hold Time: 180s   Keepalive Interval: 60s
Capabilities:
    Route Refresh: advertised and received
    Address Family: IPv4 Unicast:advertised and received
    Graceful Restart: none
    Restart Remaining Time: 0
Messages: 28 received, 31 sent
Minimum time between advertisements: 30s (default)
For Address Family IPv4 Unicast:advertised and received
    Route Refresh: 0 received, 0 sent

```

```

Prefixes: 2 received, 2 sent, 2 advertised
1 Connections established, 2 dropped
Local host: 192.168.100.3, Local port: 179
Remote host: 192.168.100.254, Remote port: 33044

```

What to do next

Check the BGP connection from the external router. See [Verify North-South Connectivity and Route Redistribution](#).

Configure BFD on a Tier-0 Logical Router

BFD (Bidirectional Forwarding Detection) is a protocol that can detect forwarding path failures.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Select the tier-0 logical router.
- 4 Click the **Routing** tab and select **BFD** from the drop-down menu.
- 5 Click **Edit** to configure BFD.
- 6 Click the **Status** toggle button to enable BFD.

You can optionally change the global BFD properties **Receive interval**, **Transmit interval**, and **Declare dead interval**.

- 7 (Optional) Click **Add** under BFD Peers for Static Route Next Hops to add a BFD peer.

Specify the peer IP address and set the admin status to **Enabled**. Optionally, you can override the global BFD properties **Receive interval**, **Transmit interval**, and **Declare dead interval**.

Enable Route Redistribution on the Tier-0 Logical Router

When you enable route redistribution, the tier-0 logical router starts sharing specified routes with its northbound router.

Prerequisites

- Verify that the tier-0 and tier-1 logical routers are connected so that you can advertise the tier-1 logical router networks to redistribute them on the tier-0 logical router. See [Attach Tier-0 and Tier-1](#).
- If you want to filter specific IP addresses from route redistribution, verify that route maps are configured. See [Create a Route Map](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.

- 3 Select the tier-0 logical router.
- 4 Click the **Routing** tab and select **Route Redistribution** from the drop-down menu.
- 5 Click **Add** to complete the route redistribution criteria.

Option	Description
Name and Description	Assign a name to the route redistribution. You can optionally provide a description. An example name, advertise-to-bgp-neighbor.
Sources	Select the source route check boxes you want to redistribute. Static - Tier-0 static routes. NSX Connected - Tier-1 connected routes. NSX Static - Tier-1 static routes. These static routes are created automatically. Tier-0 NAT - Routes generated if NAT is configured on the tier-0 logical router. Tier-1 NAT - Routes generated if NAT is configured on the tier-1 logical router.
Route Map	(Optional) Assign a route map to filter a sequence of IP addresses from route redistribution.

- 6 Click **Save**.
- 7 Click the **Status** toggle button to enable route redistribution.

The Status button appears as Enabled.

Verify North-South Connectivity and Route Redistribution

Use the CLI to verify that the BGP routes are learned. You can also check from the external router that the NSX-T-connected VMs are reachable.

Prerequisites

- Verify that BGP is configured. See [Configure BGP on a Tier-0 Logical Router](#).
- Verify that NSX-T static routes are set to be redistributed. See [Enable Route Redistribution on the Tier-0 Logical Router](#).

Procedure

- 1 Log in to the NSX Manager CLI.
- 2 View the routes learned from the external BGP neighbor.

```
nsx-edge1(tier0_sr)> get route bgp

Flags: c - connected, s - static, b - BGP, ns - nsx_static
nc - nsx_connected, rl - router_link, t0n: Tier0-NAT, t1n: Tier1-NAT

b    10.10.10.0/24      [20/0]      via 192.168.100.254
```

- 3 From the external router, check that BGP routes are learned and that the VMs are reachable through the NSX-T overlay.

- a List the BGP routes.

```
user@router# run show ip route bgp
Codes: K - kernel route, C - connected, S - static, R - RIP, O - OSPF,
       I - ISIS, B - BGP, > - selected route, * - FIB route

B>* 172.16.10.0/24 [20/0] via 192.168.100.2, eth2, 00:00:48
B>* 172.16.20.0/24 [20/0] via 192.168.100.2, eth2, 00:00:48
B>* 172.16.30.0/24 [20/0] via 192.168.100.2, eth2, 00:00:48
```

- b From the external router, ping the NSX-T-connected VMs.

```
ping 172.16.10.10
```

```
PING 172.16.10.10 (172.16.10.10) 56(84) bytes of data.
64 bytes from 172.16.10.10: icmp_req=1 ttl=62 time=127 ms
64 bytes from 172.16.10.10: icmp_req=2 ttl=62 time=1.96 ms
^C
--- 172.16.10.10 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1001ms
rtt min/avg/max/mdev = 1.966/64.793/127.620/62.827 ms
```

- c Check the path through the NSX-T overlay.

```
traceroute 172.16.10.10
```

```
traceroute to 172.16.10.10 (172.16.10.10), 30 hops max, 60 byte packets
 1 192.168.100.3 (192.168.100.3) 0.640 ms 0.575 ms 0.696 ms
 2 100.91.176.1 (100.91.176.1) 0.656 ms 0.604 ms 0.578 ms
 3 172.16.10.10 (172.16.10.10) 3.397 ms 3.703 ms 3.790 ms
```

- 4 From the internal VMs, ping the external IP address.

```
ping 10.10.10.10
```

```
PING 10.10.10.10 (10.10.10.10) 56(84) bytes of data.
64 bytes from 10.10.10.10: icmp_req=1 ttl=62 time=119 ms
64 bytes from 10.10.10.10: icmp_req=2 ttl=62 time=1.93 ms
^C
--- 10.10.10.10 ping statistics ---
2 packets transmitted, 2 received, 0% packet loss, time 1001ms
rtt min/avg/max/mdev = 1.936/60.865/119.795/58.930 ms
```

What to do next

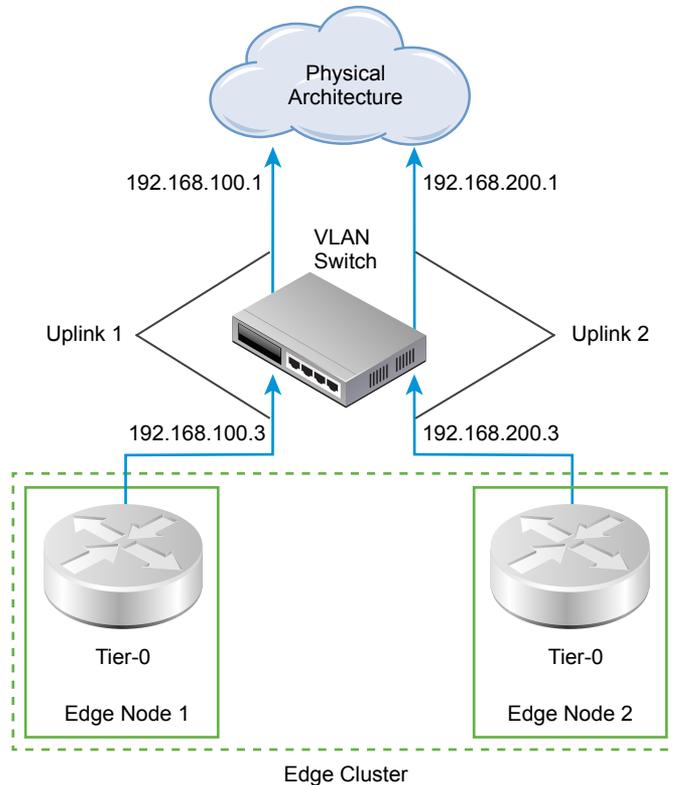
Configure additional routing functionality, such as ECMP.

Understanding ECMP Routing

Equal cost multi-path (ECMP) routing protocol increases the north and south communication bandwidth by adding an uplink to the tier-0 logical router and configure it for each Edge node in an Edge cluster. The ECMP routing paths are used to load balance traffic and provide fault tolerance for failed paths.

ECMP paths are automatically created from the VMs attached to logical switches to the Edge nodes on which the tier-0 logical router is instantiated. A maximum of eight ECMP paths are supported.

Figure 5-4. ECMP Routing Topology



For example, the topology shows two tier-0 logical routers in an edge cluster. Each tier-0 logical router is in an edge node and these nodes are part of the cluster. The uplink ports 192.168.100.3 and 198.168.200.3 define how the transport node connects to the logical switch to gain access to the physical network. When the ECMP routing paths are enabled these paths connect the VMs attached to logical switches and the two Edge nodes in the Edge cluster. The multiple ECMP routing paths increase the network throughput and resiliency.

Add an Uplink Port for the Second Edge Node

Before you enable ECMP, you must configure an uplink to connect the tier-0 logical router to the VLAN logical switch.

Prerequisites

- Verify that a transport zone and two transport nodes are configured. See the *NSX-T Installation Guide*.
- Verify that two Edge nodes and an Edge cluster are configured. See the *NSX-T Installation Guide*.
- Verify that a VLAN logical switch for uplink is available. See [Create a VLAN Logical Switch for the NSX Edge Uplink](#).
- Verify that a tier-0 logical router is configured. See [Create a Tier-0 Logical Router](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Select the tier-0 logical router.
- 4 Click the **Configuration** tab to add a router port.
- 5 Click **Add**.
- 6 Complete the router port details.

Option	Description
Name	Assign a name for the router port.
Description	Provide additional description that shows that the port is for ECMP configuration.
Type	Accept the default type Uplink .
Transport Node	Assign the host transport node from the drop-down menu.
Logical Switch	Assign the VLAN logical switch from the drop-down menu.
Logical Switch Port	Assign a new switch port name. You can also use an existing switch port.
IP Address/Mask	Enter an IP address that is in the same subnet as the connected port on the ToR switch.

Sample router port configuration.

New Router Port [X]

Name: *

Description:

Type: Uplink
 Downlink

Transport Node: *

Logical Switch: [OR Create a New Switch](#)

Logical Switch Port: Attach to new switch port
Switch Port Name:
 Attach to existing switch port

IP Address/mask: *

7 Click **Save**.

A new uplink port is added to the tier-0 router and the VLAN logical switch. The tier-0 logical router is configured on both of the edge nodes.

What to do next

Create a BGP connection for the second neighbor and enable the ECMP routing. See [Add a Second BGP Neighbor and Enable ECMP Routing](#).

Add a Second BGP Neighbor and Enable ECMP Routing

Before you enable ECMP routing, you must add a BGP neighbor and configure it with the newly added uplink information.

Prerequisites

Verify that the second edge node has an uplink port configured. See [Add an Uplink Port for the Second Edge Node](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Select the tier-0 logical router.
- 4 Click the **Routing** tab and select **BGP** from the drop-down menu.
- 5 Click **Add** under the Neighbors section to add a BGP neighbor.

- 6 Enter the neighbor IP address.
For example, 192.168.200.254.
- 7 Select the local address from the drop-down menu.
For example, uplink2 192.168.200.1.
- 8 Enter the remote AS number.
For example, 64511.
- 9 Click **Save**.
The newly added BGP neighbor appears.
- 10 Click **Edit** next to the BGP Configuration section.
- 11 Click the **ECMP** toggle button to enable ECMP.
The Status button must be appear as Enabled.
- 12 Click **Save**.

Multiple ECMP routing paths connect the VMs attached to logical switches and the two Edge nodes in the Edge cluster.

What to do next

Test whether the ECMP routing connections are working properly. See [Verify ECMP Routing Connectivity](#).

Verify ECMP Routing Connectivity

Use CLI to verify that the ECMP routing connection to neighbor is established.

Prerequisites

Verify that ECMP routing is configured. See [Add an Uplink Port for the Second Edge Node](#) and [Add a Second BGP Neighbor and Enable ECMP Routing](#).

Procedure

- 1 Log in to the NSX Manager CLI.
- 2 Get the distributed router UUID information.

```
get logical-routers
```

```
Logical Router
UUID       : 736a80e3-23f6-5a2d-81d6-bbefb2786666
vrf        : 2
type       : TUNNEL

Logical Router
UUID       : d40bbfa4-3e3d-4178-8615-6f42ea335037
vrf        : 4
type       : SERVICE_ROUTER_TIER0
```

```
Logical Router
UUID       : d0289ba4-250e-41b4-8ffc-7cab4a46c3e4
vrf        : 5
type       : DISTRIBUTED_ROUTER
```

```
Logical Router
UUID       : a6ee6316-2212-4171-99cc-930c98bcad7f
vrf        : 6
type       : DISTRIBUTED_ROUTER
```

- 3 Locate the UUID information from the output.

```
Logical Router
UUID       : d0289ba4-250e-41b4-8ffc-7cab4a46c3e4
vrf        : 5
type       : DISTRIBUTED_ROUTER
```

- 4 Type the VRF for the tier-0 distributed router.

```
vrf 5
```

- 5 Verify that the tier-0 distributed router is connected to the Edge nodes.

```
get forwarding
```

For example, edge-node-1 and edge-node-2.

- 6 Enter **exit** to leave the vrf context.
- 7 Open the active controller for the tier-0 logical router.
- 8 Verify that the tier-0 distributed router on the controller node is connected.

```
get logical-router <UUID> route
```

The route type for the UUID should appear as NSX_CONNECTED.

- 9 Start a SSH session on the two Edge nodes.

- 10 Start a session to capture packets.

```
set capture session 0 interface fp-eth1 dir tx
```

```
set capture session 0 expression src net <IP_Address>
```

- 11 Navigate to the Control Center and double click the httpdata11.bat and httpdata12.bat scripts.

A large number of HTTP requests to both Web VMs is sent and you see traffic hashed to both paths using the Edge nodes, which shows that ECMP is working.

- 12 Stop the capture session.

```
del capture session 0
```

- 13 Remove the bat scripts.

Create an IP Prefix List

An IP prefix list contains single or multiple IP addresses that are assigned access permissions for route advertisement. The IP addresses in this list are processed sequentially. IP prefix lists are referenced through BGP neighbor filters or route maps with in or out direction.

For example, you can add the IP address 192.168.100.3/27 to the IP prefix list and deny the route from being redistributed to the northbound router. This means that with the exception of the 192.168.100.3/24 IP address all other IP addresses are going to be shared the router.

You can also append an IP address with less-than-or-equal-to (le) and greater-than-or-equal-to (ge) modifiers to grant or limit route redistribution. For example, 192.168.100.3/27 ge 24 le 30 modifiers match subnet masks greater than or equal to 24-bits and less than or equal to 30-bits in length.

Note The default action for a route is **Deny**. When you create a prefix list to deny or permit specific routes, be sure to create an IP prefix with a blank network address and the **Permit** action if you want to permit all other routes.

Prerequisites

Verify that you have a tier-0 logical router configured. See [Create a Tier-0 Logical Router](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Select the tier-0 logical router.
- 4 Click the **Routing** tab and select **IP Prefix Lists** from the drop-down menu.
- 5 Select **Add**.
- 6 Assign a name for the IP prefix list.
- 7 Click **Insert Row** to add a network address in the CIDR format.
For example, 192.168.100.3/27.
- 8 Select **Deny** or **Permit** from the drop-down menu.
You grant or deny each IP address from being advertised, depending on your requirement.
- 9 (Optional) Set a range of IP address numbers in the le or ge modifiers.
For example, set le modifier to 30 and ge modifier to 24.
- 10 Click **Save**.

The newly created IP prefix list appears in the row.

Create a Route Map

A route map consists of a sequence of IP prefix lists, BGP path attributes, and an associated action. The router scans the sequence for an IP address match. If there is a match, the router performs the action and scans no further.

Route maps can be referenced at the BGP neighbor level and route redistribution. When IP prefix lists are referenced in route maps and the route map action of permitting or denying is applied, the action specified in the route map sequence overrides the specification within the IP prefix list.

Prerequisites

Verify that an IP prefix list is configured. See [Create an IP Prefix List](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Select the tier-0 logical router.
- 4 Select **Routing > Route Maps**.
- 5 Click **Add**.
- 6 Enter a name and an optional description for the route map.
- 7 Click **Add** to add an entry in the route map.
- 8 Select one or more IP prefix lists.
- 9 (Optional) Set BGP attributes.

BGP Attribute	Description
AS-path Prepend	Prepend a path with one or more AS (autonomous system) numbers to make the path longer and therefore less preferred.
MED	Multi-Exit Discriminator indicates to an external peer a preferred path to an AS.
Weight	Set a weight to influence path selection. The range is 0 - 65535.
Community	Specify a community using the aa:nn format, for example, 300:500. Or use the drop-down menu to select one of the following: <ul style="list-style-type: none"> ■ NO_EXPORT_SUBCONFED - Do not advertise to EBGP peers. ■ NO_ADVERTISE - Do not advertise to any peer. ■ NO_EXPORT - Do not advertise outside BGP confederation

- 10 In the Action column, select **Permit** or **Deny**.

You can permit or deny IP addresses in the IP prefix lists from advertising their addresses.

- 11 Click **Save**.

Network Address Translation

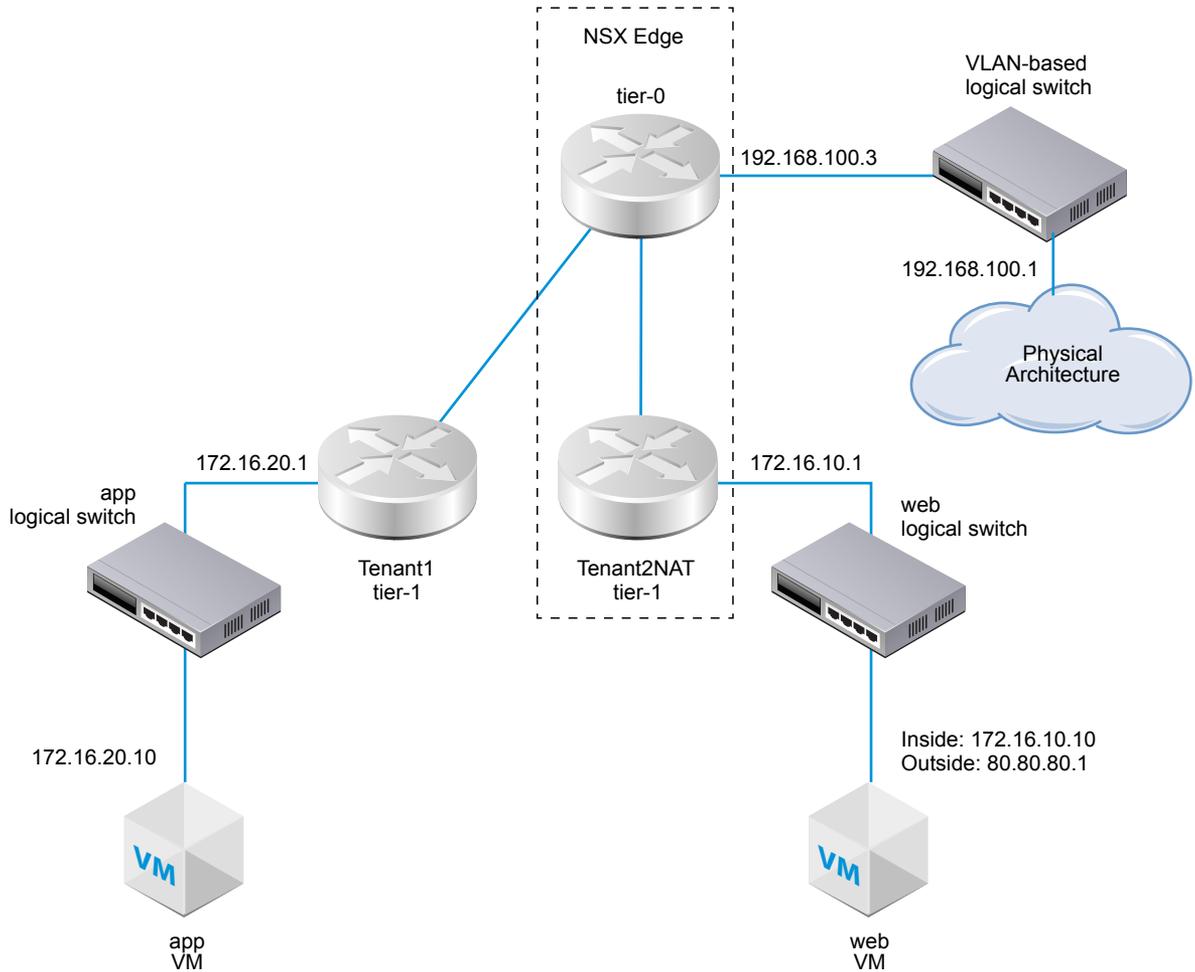
Network address translation (NAT) in NSX-T can be configured on tier-0 and tier-1 logical routers.

For example, the following diagram shows two tier-1 logical routers with NAT configured on Tenant2NAT. The web VM is simply configured to use 172.16.10.10 as its IP address and 172.16.10.1 as its default gateway.

NAT is enforced at the uplink of the Tenant2NAT logical router on its connection to the tier-0 logical router.

To enable NAT configuration, Tenant2NAT must have a service component on an NSX Edge cluster. Thus, Tenant2NAT is shown inside the NSX Edge. For comparison, Tenant1 can be outside of the NSX Edge because it is not using any Edge services.

Figure 6-1. NAT Topology



This chapter includes the following topics:

- [Tier-1 NAT](#)
- [Tier-0 NAT](#)

Tier-1 NAT

Tier-1 logical routers support source NAT and destination NAT.

Configure Source NAT on a Tier-1 Router

Source NAT (SNAT) changes the source address in the IP header of a packet. It can also change the source port in the TCP/UDP headers. The typical usage is to change a private (rfc1918) address/port into a public address/port for packets leaving your network.

In this example, as packets are received from the web VM, the Tenant2NAT tier-1 router changes the source port of the packets from 172.16.10.10 to 80.80.80.1. Having a public source address enables destinations outside of the private network to route back to the original source.

Prerequisites

- The tier-0 router must have an uplink connected to a VLAN-based logical switch. See [Connect a Tier-0 Logical Router to a VLAN Logical Switch](#).
- The tier-0 router must have routing (static or BGP) and route redistribution configured on its uplink to the physical architecture. See [Configure a Static Route](#), [Configure BGP on a Tier-0 Logical Router](#), and [Enable Route Redistribution on the Tier-0 Logical Router](#).
- The tier-1 routers must each have an uplink to a tier-0 router configured. Tenant2NAT must be backed by an edge cluster. See [Attach Tier-0 and Tier-1](#).
- The tier-1 routers must have downlink ports and route advertisement configured. See [Add Downlink Ports for the Tier-1 Logical Router](#) and [Configure Route Advertisement on a Tier-1 Logical Router](#).
- The VMs must be attached to the correct logical switches.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing**.
- 3 Click a tier-1 logical router on which you want to configure NAT.
- 4 Under NAT, click **Add**.
- 5 For the Action, select SNAT.
- 6 Select the protocol type.
By default, **Any Protocol** is selected.
- 7 For the Source IP address, enter the inside IP address of the VM.
If you leave the source IP blank, all sources on router's downlink ports are translated. In this example, the source IP is 172.16.10.10.
- 8 For the Translated IP address, enter the outside IP address for the VM.
Note that the outside/translated IP address does not need to be configured on the VM. Only the NAT router needs to know about the translated IP address.
In this example, the translated IP address is 80.80.80.1.
- 9 For the Destination IP address, you can leave it blank or enter an IP address.
If you leave Destination IP blank, the NAT applies to all destinations outside of the local subnet.
- 10 Enable the rule.
- 11 (Optional) Enable logging.

The new rule is listed under NAT. For example:

The screenshot shows the NSX Manager interface for a NAT rule named "Tenant2NAT". The interface has tabs for Summary, Configuration, Routing, and NAT. The NAT tab is active. Below the tabs, it says "NAT" and "No Statistics were collected". There are icons for ADD, EDIT, DELETE, and COLUMNS. A table lists the NAT rule configuration:

ID	Action	Match				Translated		Stats	
		Protocol	Source IP	Source Ports	Destination IP	Destination Ports	IP		Ports
4100	SNAT	Any	172.16.10.10	Any	Any	Any	80.80.80.1	Any	

Below the table, there is a "Priority: 1024" label and a small bar chart icon.

What to do next

Configure the tier-1 router to advertise NAT routes.

To advertise the NAT routes upstream from the tier-0 router to the physical architecture, configure the tier-0 router to advertise tier-1 NAT routes.

Configure Destination NAT on a Tier-1 Router

Destination NAT changes the destination address in IP header of a packet. It can also change the destination port in the TCP/UDP headers. The typical usage of this is to redirect incoming packets with a destination of a public address/port to a private IP address/port inside your network.

In this example, as packets are received from the app VM, the Tenant2NAT tier-1 router changes the destination port of the packets from 172.16.10.10 to 80.80.80.1. Having a public destination address enables a destination inside a private network to be contacted from outside of the private network.

Prerequisites

- The tier-0 router must have an uplink connected to a VLAN-based logical switch. See [Connect a Tier-0 Logical Router to a VLAN Logical Switch](#).
- The tier-0 router must have routing (static or BGP) and route redistribution configured on its uplink to the physical architecture. See [Configure a Static Route](#), [Configure BGP on a Tier-0 Logical Router](#), and [Enable Route Redistribution on the Tier-0 Logical Router](#).
- The tier-1 routers must each have an uplink to a tier-0 router configured. Tenant2NAT must be backed by an edge cluster. See [Attach Tier-0 and Tier-1](#).
- The tier-1 routers must have downlink ports and route advertisement configured. See [Add Downlink Ports for the Tier-1 Logical Router](#) and [Configure Route Advertisement on a Tier-1 Logical Router](#).
- The VMs must be attached to the correct logical switches.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.

- 2 Select **Routing**.
- 3 Click a tier-1 logical router on which you want to configure NAT.
- 4 Under NAT, click **Add**.
- 5 For the Action, select DNAT.
- 6 Select the protocol type.
By default, **Any Protocol** is selected.
- 7 For the Destination IP address, enter the outside IP address of the VM.
In this example, the destination IP address is 80.80.80.1. Note that the outside IP address does not need to be configured on the VM. Only the NAT router needs to know about the outside IP address.
- 8 For the Translated IP address, enter the inside IP address for the VM.
The inside IP address must be configured on the VM.
In this example, the inside/translated IP address is 172.16.10.10.
- 9 For the Source IP address, you can leave it blank or enter an IP address.
If you leave Source IP blank, the NAT applies to all sources outside of the local subnet.
- 10 Enable the rule.
- 11 (Optional) Enable logging.

The new rule is listed under NAT. For example:

 **Tenant2NAT**

Summary Configuration Routing NAT

NAT

No Statistics were collected

+ ADD EDIT DELETE COLUMNS

ID	Action	Match				Translated		Stats
		Protocol	Source IP	Source Ports	Destination IP	Destination Ports	IP	
▲ Priority: 1024								
4101	DNAT	Any	Any	Any	80.80.80.1	Any	172.16.10.10	Any

What to do next

Configure the tier-1 router to advertise NAT routes.

To advertise the NAT routes upstream from the tier-0 router to the physical architecture, configure the tier-0 router to advertise tier-1 NAT routes.

Advertise Tier-1 NAT Routes to the Upstream Tier-0 Router

Advertising tier-1 NAT routes enables the upstream tier-0 router to learn about these routes.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing**.
- 3 Click a tier-1 logical router on which you have configured NAT.
- 4 From the tier-1 router, select **Routing > Route Advertisement**.
- 5 Edit the route advertisement rules to enable NAT route advertisement.

**Tenant2NAT**

Summary

Configuration

Routing ▼

NAT

Route Advertisement

Status	● Enabled
Advertise All NSX Connected Routes	● Yes
Advertise NAT Routes	● Yes
Advertise Static Routes	● No

What to do next

Advertise tier-1 NAT routes from the tier-0 router to the upstream physical architecture.

Advertise Tier-1 NAT Routes to the Physical Architecture

Advertising tier-1 NAT routes from the tier-0 router enables the upstream physical architecture to learn about these routes.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing**.
- 3 Click a tier-0 logical router that is connected to a tier-1 router on which you have configured NAT.
- 4 From the tier-0 router, select **Routing > Route Redistribution**.
- 5 Edit the route advertisement rules to enable tier-1 NAT route advertisement.

Edit Redistribution Criteria - T1
✕

Name: *

Description:

Sources: *

- Static
- NSX Connected
- NSX Static
- Tier-0 NAT
- Tier-1 NAT

Route Map:

Save
Cancel

What to do next

Verify NAT is working as expected.

Verify Tier-1 NAT

Verify that SNAT and DNAT rules are working correctly.

Procedure

- 1 Log in the NSX Edge.
- 2 Run `get logical-routers` to determine the VRF number for the tier-0 services router.
- 3 Enter the tier-0 services router context by running the `vrf <number>` command.
- 4 Run the `show route` command and make sure that the tier-1 NAT address appears.

```

nsx-edge(tier0_sr)> get route

Flags: c - connected, s - static, b - BGP, ns - nsx_static
nc - nsx_connected, rl - router_link, t0n: Tier0-NAT, t1n: Tier1-NAT

Total number of routes: 8

t1n 80.80.80.1/32      [3/3]      via 169.0.0.1
...

```

- 5 If your Web VM is set up to serve Web pages, make sure you can open a Web page at `http://80.80.80.1`.
- 6 Make sure that the tier-0 router's upstream neighbor in the physical architecture can ping 80.80.80.1.
- 7 While the ping is still running, check the stats column for the DNAT rule.
There should be one active session.

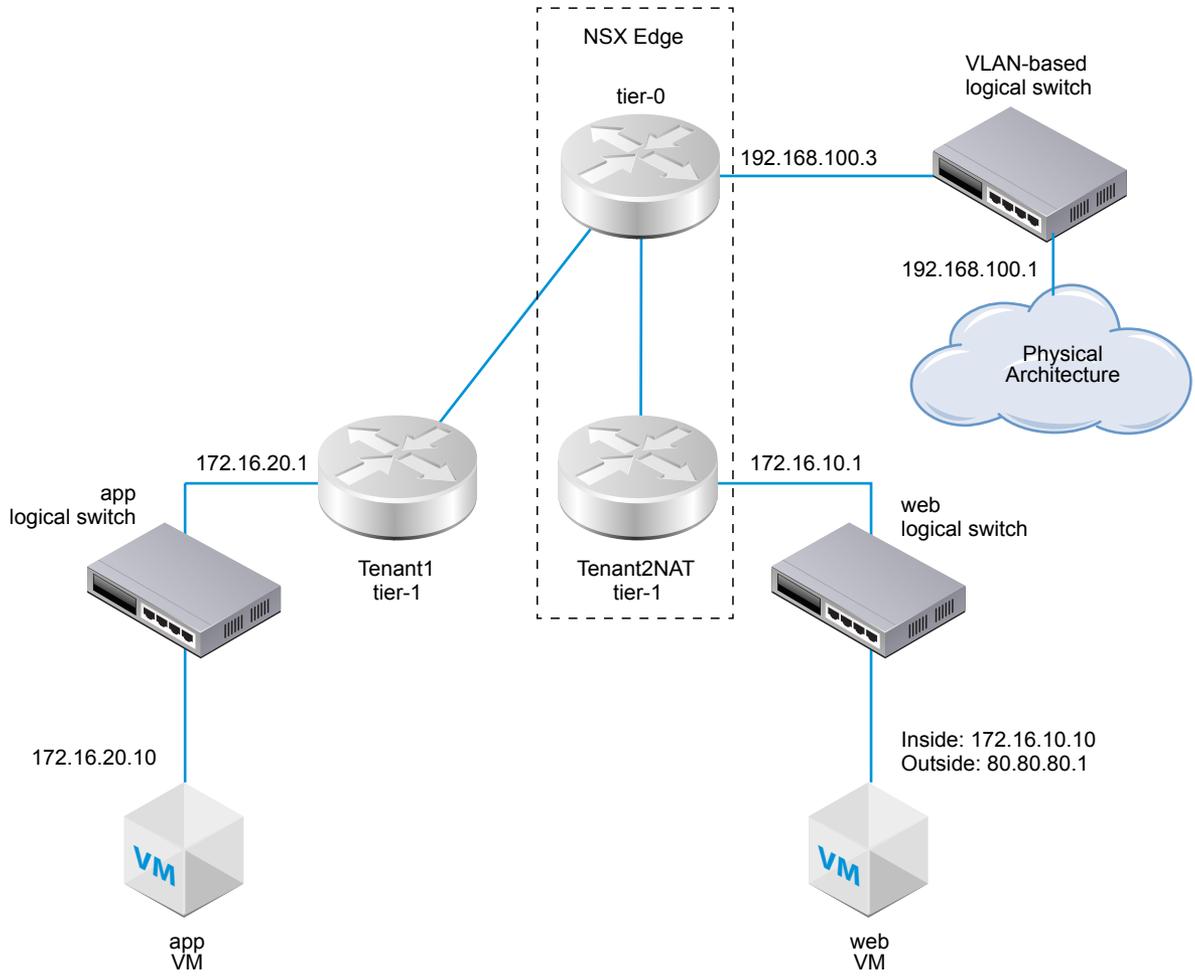
Tier-0 NAT

Tier-0 logical routers support reflexive NAT.

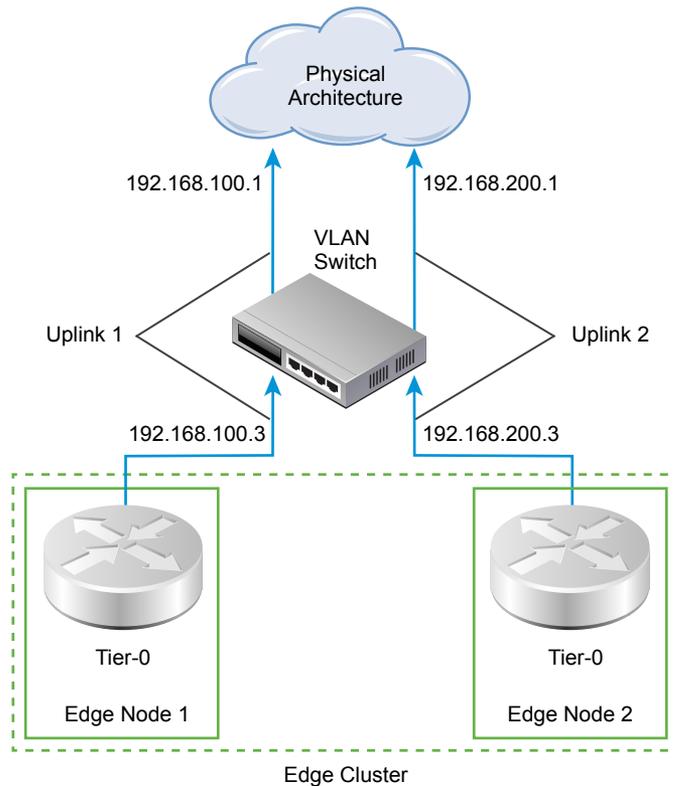
Reflexive NAT

When a tier-0 logical router is running in Active-Active ECMP mode, you cannot configure stateful NAT where asymmetrical paths might cause issues. For Active-Active ECMP routers, you can use reflexive NAT (sometimes called stateless NAT).

In this example, as packets are received from the web VM, the Tenant2NAT tier-1 router changes the source port of the packets from 172.16.10.10 to 80.80.80.1. Having a public source address enables destinations outside of the private network to route back to the original source.



However, when there are two Active-Active tier-0 routers involved, as shown here, reflexive NAT must be configured.



Configure Reflexive NAT on a Tier-0 Logical Router

When a tier-0 logical router is running in Active-Active ECMP mode, you cannot configure stateful NAT where asymmetrical paths might cause issues. For Active-Active ECMP routers, you can use reflexive NAT (sometimes called stateless NAT).

Prerequisites

- The tier-0 router must have two uplinks connected to a VLAN-based logical switch. See [Connect a Tier-0 Logical Router to a VLAN Logical Switch](#).
- The tier-0 router must have routing (static or BGP) and route redistribution configured on its uplinks to the physical architecture. See [Configure a Static Route](#), [Configure BGP on a Tier-0 Logical Router](#), and [Enable Route Redistribution on the Tier-0 Logical Router](#).
- The tier-1 routers must each have an uplink to a tier-0 router configured. Tenant2NAT must be backed by an edge cluster. See [Attach Tier-0 and Tier-1](#).
- The tier-1 routers must have downlink ports and route advertisement configured. See [Add Downlink Ports for the Tier-1 Logical Router](#) and [Configure Route Advertisement on a Tier-1 Logical Router](#).
- The VMs must be attached to the correct logical switches.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing**.
- 3 Click a tier-0 logical router on which you want to configure reflexive NAT.
- 4 Under NAT, click **Add**.
- 5 For the Action, select Reflexive.
- 6 For the Source IP address, enter the outside IP address of the VM.
In this example, the source IP is 80.80.80.1.
- 7 For the Translated IP address, enter the inside IP address for the VM.
In this example, the translated IP address is 172.16.10.10.
- 8 For the Destination IP address, you can leave it blank or enter an IP address.
If you leave Destination IP blank, the NAT applies to all destinations outside of the local subnet.
- 9 Enable the rule.
- 10 (Optional) Enable logging.

The new rule is listed under NAT. For example:

 **T0-router-1**

Summary Configuration Routing NAT

NAT

No Statistics were collected

+ ADD  EDIT  DELETE  COLUMNS 

ID	Action	Match				Translated		Stats
		Protocol	Source IP	Source Ports	Destination IP	Destination Ports	IP	
▲ Priority: 1024								
 4099	Reflexive	Any	80.80.80.1	Any	Any	Any	172.16.10.10	Any 

What to do next

Configure the tier-1 router to advertise NAT routes.

To advertise the NAT routes upstream from the tier-0 router to the physical architecture, configure the tier-0 router to advertise tier-1 NAT routes.

Firewall Sections and Firewall Rules



Firewall sections are used to group a set of firewall rules.

A firewall section is made up from one or more individual firewall rules. Each individual firewall rule contains instructions that determine whether a packet should be allowed or blocked; which protocols it is allowed to use; which ports it is allowed to use and so forth. Sections are used for multi-tenancy, such as specific rules for sales and engineering departments in separate sections.

A section can be defined as enforcing stateful or stateless rules. Stateless rules are treated as traditional stateless ACLs. Reflexive ACLs are not supported for stateless sections. A mix of stateless and stateful rules on a single logical switch port is not recommended and may cause undefined behavior.

Rules can be moved up and down within a section. For any traffic attempting to pass through the firewall, the packet information is subjected to the rules in the order shown in the section, beginning at the top and proceeding to the default rule at the bottom. The first rule that matches the packet has its configured action applied, and any processing specified in the rule's configured options is performed and all subsequent rules are ignored (even if a later rule is a better match). Thus, you should place specific rules above more general rules to ensure those rules are not ignored. The default rule, located at the bottom of the rule table, is a "catchall" rule; packets not matching any other rules will be enforced by the default rule.

This chapter includes the following topics:

- [Add a Firewall Rule Section](#)
- [Delete a Firewall Rule Section](#)
- [Enable and Disable Section Rules](#)
- [Disable and Enable Section Logs](#)
- [About Firewall Rules](#)
- [Add a Firewall Rule](#)
- [Delete a Firewall Rule](#)
- [Edit the Default Distributed Firewall Rule](#)
- [Change the Order of a Firewall Rule](#)
- [Filter Firewall Rules](#)
- [Exclude Objects from Firewall Enforcement](#)

Add a Firewall Rule Section

A firewall rule section is edited and saved independently and is used to apply separate firewall configuration to tenants.

Procedure

- 1 Select **Firewall** in the navigation panel.

Ensure that you are in the General tab to add an L3 rule. Click the Ethernet tab to add an L2 rule.

- 2 To add a section, in the first column, click the wheel (⚙) icon, or a rule and select either **Add Section Above** or **Add Section Below**.

Note For any traffic attempting to pass through the firewall, the packet information is subjected to the rules in the order shown in the Rules table, beginning at the top and proceeding to the default rules at the bottom. In some cases, the order of precedence of two or more rules might be important in determining the disposition of a packet.

- 3 Enter the section name and an optional description.
- 4 Select either **Stateful**, **False**, or **True**. This option is applicable only for L3.

Stateless firewalls watch network traffic, and restrict or block packets based on source and destination addresses or other static values. Stateful firewalls can watch traffic streams from end to end. Stateless firewalls are typically faster and perform better under heavier traffic loads. Stateful firewalls are better at identifying unauthorized and forged communications. There is no toggling between stateful and stateless once it is defined.

- 5 Select where you would like to apply the section.

Note If you have used **Applied To** in a section it will override any **Applied To** settings in the rules in that section.

Logical Port - Displays all of the logical ports

Logical Switch - Displays all of the logical switches

NSGroup - Displays all of the NSGroups

- 6 Click the checkbox next to the available port, switch, or group and then click the arrow.

The item moves to the Selected column.

- 7 Click **Save** to save the section.

The newly added Section appears in the **Firewall** window.

What to do next

Add Firewall rules to the section.

Delete a Firewall Rule Section

A firewall rule section can be deleted when it is no longer used.

When you delete a firewall rule section, all rules in that section are deleted. You cannot delete a section and add it again at a different place in the firewall table. To do so, you must delete the section and publish the configuration. Then add the deleted section to the firewall table and re-publish the configuration.

Procedure

- 1 Select **Firewall** in the navigation panel.
- 2 Make sure that you are in the General tab to add an L3 rule.
- 3 Click the Ethernet tab to add an L2 rule.
- 4 To delete a section, in the first column, right-click the wheel  next to the section you want to delete.
- 5 Click **Delete** to remove the section. The section and all the rules it contained are now deleted.

Enable and Disable Section Rules

You can enable or disable all rules in a firewall rule section.

Procedure

- 1 Select **Firewall** in the navigation panel.
- 2 In the first column, click the wheel icon and select **Disable Section Rules** or **Enable Section Rules**.
- 3 Click **Save**.

Disable and Enable Section Logs

Enabling logs for section rules records information on packets for all of the rules in a section. Depending on the number of rules in a section, a typical firewall section will generate large amounts of log information and can affect performance.

Logs are stored in the `/var/log/dfwptlogs.log` file on vSphere ESXi and KVM hosts.

Procedure

- 1 Select **Firewall** in the navigation panel.
- 2 In the first column, click the wheel  icon. Select **Disable Logs for Section Rules** or **Enable Logs for Section Rules**.
- 3 Click **Save**.

About Firewall Rules

NSX-T uses firewall rules to specify traffic handling in and out of the network.

Firewall offers multiple sets of configurable rules: Layer 3 rules (General tab) and Layer 2 rules (Ethernet tab). Layer 2 firewall rules are processed before Layer 3 rules. The Configuration tab contains the exclusion list, which contains logical switches, logical ports and groups that are to be excluded from firewall enforcement.

Firewall Rules are enforced as follows:

- Rules are processed in top-to-bottom ordering.
- Each packet is checked against the top rule in the rule table before moving down the subsequent rules in the table.
- The first rule in the table that matches the traffic parameters is enforced.

No subsequent rules can be enforced as the search is then terminated for that packet. Because of this behavior, it is always recommended to put the most granular policies at the top of the rule table. This will ensure they will be enforced before more specific rules.

The default rule, located at the bottom of the rule table, is a catchall rule; packets not matching any other rules will be enforced by the default rule. After the host preparation operation, the default rule is set to allow action. This ensures that VM-to-VM communication is not broken during staging or migration phases. It is a best practice to then change this default rule to block action and enforce access control through a positive control model (i.e., only traffic defined in the firewall rule is allowed onto the network).

Firewall rule options are accessible by clicking the drop down arrow next to Columns, and checking the columns you'd like to be included in the firewall rules window. The following options are available.

Table 7-1. Columns in the firewall rule screen

Column Name	Definition
Name	Name of the firewall rule.
Sources	The source of the rule can be either an IP or MAC address or an object other than an IP address. The source will match any if not defined. IPv6 is not supported for source or destination range.
ID	Unique system generated ID for each rule.
Direction	The direction rule element matches the direction a packet is traveling as it traverses the interface. A direction of In is for traffic ingressing through the firewall. A direction of out is for traffic egressing through a firewall. By default the direction will be In Out (both directions).
IP Protocol	This is applicable only for L3 rules. Both IPv4 and IPv6 are supported. The default value is both.
Destinations	The destination IP or MAC address/netmask of the connection that is affected by the rule. The destination will match any if not defined. IPv6 is not supported for source or destination range.
Services	The service can be a predefined port protocol combination for L3. For L2 it can be ether-type. For both L2 and L3 you can manually define a new service or service group. The service will match any, if it is not specified.
Action (required)	The action applied by the rule can be allow, block, or reject.
Applied To	Defines the scope at which this rule is applicable. If not defined the scope will be all logical ports. If you have added "applied to" in a section it will overwrite the rule.
Log	Logging can be turned off or on. Logs are stored at /var/log/dfwpktlogs.log file on ESX and KVM hosts.

Table 7-1. Columns in the firewall rule screen (Continued)

Column Name	Definition
Stats	Read-only field that displays the byte, packet count, and sessions.
Notes	Comments for the rule.

Below is the default firewall rule with a portion of the column options shown.

Figure 7-1. Firewall Rules Window

	Name	ID	Sources	Destinations	Services	Action	Applied To	Log	Stats
1	2b8f904d-b41e-454d-890e-54af...	3165	default - a3b0...	default - a3b0...	Any	Allow	All	No	packets: 0 bytes: 0 sessions: 0

Add a Firewall Rule

A firewall is a network security system that monitors and controls the incoming and outgoing network traffic based on predetermined firewall rules.

Firewall rules are added at the NSX Manager scope. Using the Applied To field, you can then narrow down the scope at which you want to apply the rule. You can add multiple objects at the source and destination levels for each rule, which helps reduce the total number of firewall rules to be added.

Note By default, a rule matches on the default of any source, destination, and service rule elements, matching all interfaces and traffic directions. If you want to restrict the effect of the rule to particular interfaces or traffic directions, you must specify the restriction in the rule.

Prerequisites

To use a group of addresses, first manually associate the IP and MAC address of each VM with their logical switch.

Procedure

- 1 Select **Firewall** in the navigation panel.
Ensure that you are in the General tab to add an L3 rule. Click the Ethernet tab to add an L2 rule.
- 2 To add a rule, in the first column, click the wheel (⚙️) icon, and at the bottom of the list select **Add Rule**.

A new row appears to define a firewall rule.

Note For any traffic attempting to pass through the firewall, the packet information is subjected to the rules in the order shown in the Rules table, beginning at the top and proceeding to the default rules at the bottom. In some cases, the order of precedence of two or more rules might be important in determining the disposition of a packet.

- 3 A new rule is added at the top of the section. If you want to add a rule at a specific place in a section, select a rule. In the first column, click the wheel (⦿) icon and select **Insert Rule Above** or **Insert Rule Below**.

A new row appears to define a firewall rule.

- 4 In the Upper right corner of the **Name** column, click the pencil icon. Enter the rule name in the Edit Name dialogue box.

A rule appears with the specified name.

- 5 Point to the **Sources** cell of the new rule, click the pencil icon, and select the source of the rule. The source will match any if not defined. The **Edit Sources** dialogue box appears.

Note When creating a new firewall rule, you can drag and drop objects to use for the Source, Destination, Service, and Applied To fields, instead of selecting these each time. This can help to speed up the rule creation process, especially when the same objects are often reused.

In order to do this, click **Objects** in the left hand corner of the firewall rules window, select the object type from the list, then drag and drop the object you need into the right field that is, Sources in your firewall rule.

Table 7-2. Edit Sources window

Option	Description
IP Address or MAC Address	Enter multiple IP or MAC addresses in a comma-separated list. The list can contain up to 255 characters. Both IPv4 and IPv6 formats are supported.
Objects	<p>Click the arrow and select the Object.</p> <ol style="list-style-type: none"> 1 Select IP Set, Logical Port, Logical Switch, or NS Group. <p>Available objects for the selected container are displayed.</p> <ol style="list-style-type: none"> 2 Select one or more objects and click the arrow. To select all of the available objects click the checkbox next to Available, then click the arrow. 3 The objects move to the selected column. 4 Click OK.

- 6 Point to the **Destinations** cell of the new rule. The destination will match any if not defined. The **Edit Destinations** dialogue box appears.

Table 7-3. Edit Destinations window

Option	Description
IP Address or MAC address	You can enter multiple IP or MAC addresses in a comma-separated list. The list can contain up to 255 characters. Both IPv4 and IPv6 formats are supported.
Objects	Click the arrow and select the Object. <ol style="list-style-type: none"> You can select IP Set, Logical Port, Logical Switch, or NS Group. Available objects for the selected container are displayed. Select one or more objects and click the arrow. To select all of the available objects click the checkbox next to Available, then click the arrow. The objects move to the selected column. Click OK.

- 7 Point to the **Service** cell of the new rule. The service will match any if not defined. The **Edit Services** dialogue box appears. The list already displays many predefined services, but you are not limited to these choices.

- 8 To select a predefined service, select one of more available objects, then click the arrow. Click **OK**.

- 9 To define a new service, click **New**. The NSService dialogue box appears.

Option	Description
Name	Name the new service.
Description	Describe the new service.
Type of Service	<ul style="list-style-type: none"> ■ ALG ■ ICMP ■ IP ■ L4 Port Set ■ IGMP
Protocol	Select one of the available protocols.
Source Ports	Enter the source port.
Destination Ports	Select the destination port.
Group existing services	Click the radio button to add an existing group service.

- 10 Point to the **Action** cell, and click the pencil icon. This parameter is required. The Edit Action dialogue box appears.

Option	Description
Allow	Allows all L3 or L2 traffic with the specified source, destination, and protocol to pass through the current firewall context. Packets that match the rule, and are accepted, traverse the system as if the firewall is not present
Drop	Drops packets with the specified source, destination, and protocol. Dropping a packet is a silent action with no notification to the source or destination systems. Dropping the packet causes the connection to be retried until the retry threshold is reached.
Reject	Rejects packets with the specified source, destination, and protocol. Rejecting a packet is a more graceful way to deny a packet, as it sends a destination unreachable message to the sender. If the protocol is TCP, a TCP RST message is sent. ICMP messages with administratively prohibited code are sent for UDP, ICMP, and other IP connections. One benefit of using Reject is that the sending application is notified after only one attempt that the connection cannot be established.

- 11 Point to the **Applied To** cell, and click the pencil icon. The Edit Applied To dialogue box appears. Select the type of object from the drop-down list. Click **OK**.
- 12 Point to the **Log** cell, and click the pencil icon. Logging is turned off by default. Select either **Yes** to enable logging, or **No**, to disable logging. Logs are stored at `/var/log/dfwpktlogs.log` file on ESX and KVM hosts. You can also write notes here. Note that selecting **Yes**, logs all sessions matching this rule. Enabling logging can affect performance.
- 13 For your rule or rules to take effect, click **Save**.
You can add multiple rules before clicking **Save**.

Delete a Firewall Rule

A firewall is a network security system that monitors and controls the incoming and outgoing network traffic based on predetermined firewall rules. Custom defined rules can be added and deleted.

Procedure

- 1 Select **Firewall** in the navigation panel.
Ensure that you are in the General tab to add an L3 rule. Click the Ethernet tab to add an L2 rule.
- 2 Right-click the number of the rule you want to move.
A drop-down list appears.
- 3 Select **Delete**.
The firewall rule is deleted.
- 4 Click **Save** for the changes to take effect.

The rule is deleted.

Edit the Default Distributed Firewall Rule

You can edit the default firewall settings that apply to traffic that does not match any of the user-defined firewall rules.

Default firewall settings apply to traffic that does not match any of the user-defined firewall rules. The Distributed Firewall default rule is displayed on the centralized firewall user interface. The default Layer 3 rule is shown under the General Tab and the default Layer 2 rule is shown under the Ethernet tab.

The default Distributed Firewall rules allow all L3 and L2 traffic to pass through all prepared clusters in your infrastructure. The default rule is always at the bottom of the rules table and cannot be deleted or added to. However, you can change the Action element of the rule from Allow to Drop or Reject (not recommended), and indicate whether traffic for that rule should be logged.

Procedure

- 1 Click **Firewall**.

The General Firewall screen appears.

- 2 Ensure that you are in the **General** tab to edit the default L3 rule. Click the **Ethernet** tab to edit an L2 rule.

- 3 Under the **Action** column, expand the section and select one of the options:

- **Allow** - Allows all L3 or L2 traffic with the specified source, destination, and protocol to pass through the current firewall context. Packets that match the rule, and are accepted, traverse the system as if the firewall is not present.
- **Drop** - Drops packets with the specified source, destination, and protocol. Dropping a packet is a silent action with no notification to the source or destination systems. Dropping the packet causes the connection to be retried until the retry threshold is reached.
- **Reject** - Rejects packets with the specified source, destination, and protocol. Rejecting a packet is a more graceful way to deny a packet, as it sends a destination unreachable message to the sender. If the protocol is TCP, a TCP RST message is sent. ICMP messages with administratively prohibited code are sent for UDP, ICMP, and other IP connections. One benefit of using Reject is that the sending application is notified after only one attempt that the connection cannot be established.

Note Selecting **Reject** as the action for the default rule is not recommended.

- 4 Under the **Log** column, expand the section and select either **Yes** to enable logging, or **No**, to disable logging. You can also write notes here. Note that selecting Yes, logs all sessions matching this rule. Enabling logging can affect performance.
- 5 Click **Save** and confirm your changes.

Change the Order of a Firewall Rule

Rules are processed in top-to-bottom ordering. You can change the order of the rules in the list.

For any traffic attempting to pass through the firewall, the packet information is subjected to the rules in the order shown in the rules table, beginning at the top and proceeding to the default rules at the bottom. In some cases, the order of precedence of two or more rules might be important in determining the traffic flow.

You can move a custom rule up or down in the table. The default rule is always at the bottom of the table and cannot be moved.

Procedure

- 1 Select **Firewall** in the navigation panel.
Ensure that you are in the General tab to add an L3 rule. Click the Ethernet tab to add an L2 rule.
- 2 Right-click the number of the rule you want to move.
- 3 Select **Move Up** or **Move Down**.
The rule will shift up or down one position.

Filter Firewall Rules

You can use a number of criteria to filter your rule set, which allows for easy rule modification.

Procedure

- 1 In the upper left hand corner of the Firewall window, click **Filter** FILTER .
The Filter dialogue box appears.
- 2 For filtering you can search in:
 - Sources - Searches inbound firewall rules.
 - Destinations - Searches outbound firewall rules.
 - Applied To - Searches rules on the Applied to criteria.
 - Services - From this list, select the application or service to be allowed or blocked. The list already displays many common services, but you are not limited to these choices. Use the Services cell to add any additional services or applications that do not already appear.
- 3 Click the criteria you want to search in and it appears at the top of the box.
- 4 Select the type of search:
 - IP Set - This option lists all of the IP address of either the rule source or destination.
 - Logical Port - Filters on logical ports.

- Logical Switch - Filters on Logical switches.
- NSGroup - Filters on NSGroups.

The filter result appears in the box.

Exclude Objects from Firewall Enforcement

A logical port, logical switch, or NSGroup can be excluded from a firewall rule.

After you've created a section with firewall rules you may want to exclude an NSX-T appliance port from the firewall rules.

Procedure

- 1 Select **Firewall** in the navigation panel. Select the **Configuration** tab.
The exclusion list screen appears.
- 2 Select **Objects** on the right hand corner of the window.
- 3 From the drop-down list, select **Logical Ports**, **Logical Switch**, or **NSGroup**.
- 4 Double-click the specific port, switch, or group you would like to exclude from the firewall rule. To close the Object dialogue box, click **Objects** again.
The Exclusion List is populated with the name and type of object you are excluding.
- 5 To remove an object from the exclusion list, click the **x**.
- 6 Click **Save**.

Configuring Groups and Services

8

You can configure groups to organize objects.

You can use the following groups in firewall rules:

- IP sets
- MAC sets
- Service groups
- NSGroups, which can include IP sets, MAC sets, logical ports, logical switches, and other NSGroups

In addition, you can create IP pools to assign IP addresses when you create transport nodes.

This chapter includes the following topics:

- [Create an IP Set](#)
- [Create an IP Pool](#)
- [Create a MAC Set](#)
- [Create an NSGroup](#)
- [Configuring Services and Service Groups](#)

Create an IP Set

An IP set is a group of IP addresses that you can use as sources and destinations in firewall rules.

An IP set can contain a combination of individual IP addresses, IP ranges, and subnets. You can specify IPv4 or IPv6 addresses, or both. An IP set can be a member of NSGroups.

Note IPv6 is not supported for source or destination ranges for firewall rules.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Inventory > Groups** from the navigation panel.
- 3 Select **IP Sets** at the top of the main panel.
- 4 Click **Add**.

- 5 Enter a name.
- 6 (Optional) Enter a description.
- 7 Enter individual addresses or a range of addresses.
- 8 Click **Save**.

Create an IP Pool

You can use an IP Pool to allocate IP addresses or subnets when you create L3 subnets.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Inventory > Groups** from the navigation panel.
- 3 Select **IP Pools** at the top of the main panel.
- 4 Click **Add**.
- 5 Enter a name.
- 6 (Optional) Enter a description.
- 7 Click **Add**.
- 8 Enter IP Ranges.
Mouse over the upper right corner of any cell and click the pencil icon to edit it.
- 9 (Optional) Enter a Gateway.
- 10 Enter a CIDR IP address with suffix.
- 11 (Optional) Enter DNS Servers.
- 12 (Optional) Enter a DNS Suffix.
- 13 Click **Save**.

Create a MAC Set

A MAC Set is a group of MAC addresses that you can use as sources and destinations in layer 2 firewall rules and as a member of an NS Group.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Inventory > Groups** from the navigation panel.
- 3 Select **MAC Sets** at the top of the main panel.
- 4 Click **Add**.
- 5 Enter a name.

- 6 (Optional) Enter a description.
- 7 Enter the MAC addresses.
- 8 Click **Save**.

Create an NSGroup

You can configure an NSGroup to contain a combination of IP sets, MAC sets, logical ports, logical switches, and other NSGroups, . You can specify NSGroups as sources and destinations, as well as in the `Applied To` field, in firewall rules.

An NSGroup has the following characteristics:

- You can specify direct members, which can be IP sets, MAC sets, logical switches, logical ports, and NSGroups.
- You can specify up to five membership criteria for logical switches and logical ports. For each criterion you specify a tag and optionally a scope.
- An NSGroup has direct members and effective members. Effective members include members that you specify using membership criteria, as well as all the direct and effective members that belong to this NSGroup's members. For example, assuming NSGroup-1 has direct member LogicalSwitch-1. You add NSGroup-2 and specify NSGroup-1 and LogicalSwitch-2 as members. Now NSGroup-2 has direct members NSGroup-1 and LogicalSwitch-2, as well as an effective member, LogicalSwitch-1. Next you add NSGroup-3 and specify NSGroup-2 as a member. NSGroup-3 now has direct member NSGroup-2 and effective members LogicalSwitch-1 and LogicalSwitch-2.
- An NSGroup can have a maximum of 500 direct members.
- The recommended limit for the number of effective members in an NSGroup is 5000. Exceeding this limit does not affect any functionality but might have a negative impact on performance. On the NSX Manager, when the number of effective members for an NSGroup exceeds 80% of 5000, the warning message `NSGroup xyz is about to exceed the maximum member limit. Total number in NSGroup is ...` appears in the log file, and when the number exceeds 5000, the warning message `NSGroup xyz has reached the maximum numbers limit. Total number in NSGroup = ...` appears. On the NSX Controller, when the number of translated VIFs/IPs/MACs in an NSGroup exceeds 5000, the warning message `Container xyz has reached the maximum IP/MAC/VIF translations limit. Current translations count in Container - IPs:..., MACs:..., VIFs:...` appears in the log file. The NSX Manager and NSX Controller check the NSGroups regarding the limit twice a day, at 7 AM and 7 PM.

For all the objects that you can add to an NSGroup as members, that is, logical switches, logical ports, IP sets, MAC sets, and NSGroups, you can navigate to the screen for any of the objects and select **Related > NSGroups** to see all the NSGroups that directly or indirectly has this object as a member. For example, in the example above, after you navigate to the screen for LogicalSwitch-1, selecting **Related > NSGroups** shows NSGroup-1, NSGroup-2, and NSGroup-3 because all three have LogicalSwitch-1 as a member, either directly or indirectly.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Inventory > Groups** from the navigation panel.
- 3 Select **NSGroups** at the top of the main panel.
- 4 Click **Add**.
- 5 Enter a name for the NSGroup.
- 6 (Optional) Enter a description.
- 7 (Optional) Click **Membership Criteria** to specify up to five criteria.

A criterion can apply to logical switches or logical ports.

A criterion can specify a tag value, a scope value, or both.

- 8 (Optional) Click **Members** to select members.
The available types are **IP Set**, **MAC Set**, **Logical Switch**, **Logical Port**, and **NSGroup**.
- 9 Click **Save**.

Configuring Services and Service Groups

You can configure an NSService to specify parameters for matching network traffic such as a port and protocol pairing. You can also use an NSService to allow or block certain types of traffic in firewall rules.

An NSService can be of the following types:

- Ether
- IP
- IGMP
- ICMP
- ALG
- L4 Port Set

An L4 Port Set supports the identification of source ports and destination ports. You can specify individual ports or a range of ports, up to a maximum of 15 ports.

An NSService can also be a group of other NSServices. An NSService that is a group can be of the following types:

- Layer 2
- Layer 3 and above

You cannot change the type after you create an NSService. Some NSServices are predefined. You cannot modify or delete them.

Create an NSService

You can create an NSService to specify the characteristics that network matching uses, or to define the type of traffic to block or allow in firewall rules.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Inventory > Services** from the navigation panel.
- 3 Click **Add**.
- 4 Enter a name.
- 5 (Optional) Enter a description.
- 6 Select **Specify a protocol** to configure an individual service, or select **Group existing services** to configure a group of NSServices.
- 7 For an individual service, select a type and a protocol.
The available types are **Ether**, **IP**, **IGMP**, **ICMP**, **ALG**, and **L4 Port Set**.
- 8 For a service group, select a type and members for the group.
The available types are **Layer 2** and **Layer 3 and above**.
- 9 Click **Save**.

DHCP

DHCP (Dynamic Host Configuration Protocol) allows clients to automatically obtain network configuration, such as IP address, subnet mask, default gateway, and DNS configuration, from a DHCP server.

You can create DHCP servers to handle DHCP requests or create DHCP relay services to relay DHCP traffic to external DHCP servers.

If you configure DHCP servers, to improve security, configure a DFW rule to allow traffic on UDP ports 67 and 68 only for valid DHCP server IP addresses.

Note A DFW rule that has Logical Switch/Logical Port/NSGroup as the source, Any as the destination, and is configured to drop DHCP packets for ports 67 and 68, will fail to block DHCP traffic. To block DHCP traffic, configure Any as the source as well as the destination.

This chapter includes the following topics:

- [Create a DHCP Server Profile](#)
- [Create a DHCP Server](#)
- [Attach a DHCP Server to a Logical Switch](#)
- [Detach a DHCP Server from a Logical Switch](#)
- [Create a DHCP Relay Profile](#)
- [Create a DHCP Relay Service](#)
- [Add a DHCP Service to a Logical Router Port](#)

Create a DHCP Server Profile

A DHCP server profile specifies an NSX Edge cluster or members of an NSX Edge cluster. A DHCP server with this profile services DHCP requests from VMs on logical switches that are connected to the NSX Edge nodes that are specified in the profile.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **DHCP** from the navigation panel.
- 3 Click **Server Profiles** and click **Add**.

- 4 Enter a name and optional description.
- 5 Select an NSX Edge cluster from the drop-down menu.
- 6 (Optional) Select members of the NSX Edge cluster.

You can specify up to 2 members.

What to do next

Create a DHCP server. See [Create a DHCP Server](#).

Create a DHCP Server

You can create DHCP servers to service DHCP requests from VMs that are connected to logical switches.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **DHCP** from the navigation panel.
- 3 Click **Servers** and click **Add**.
- 4 Enter a name and optional description.
- 5 Enter the IP address of the DHCP server and its subnet mask in CIDR format.
For example, enter 192.168.1.2/24.
- 6 Select a DHCP profile from the drop-down menu.
- 7 (Optional) Enter common options such as domain name, default gateway, DNS servers, and subnet mask.
- 8 (Optional) Enter classless static route options.
- 9 (Optional) Enter other options.
- 10 Click **Save**.
- 11 Select the newly created DHCP server.
- 12 Expand the IP Pools section.
- 13 Click **Add** to add IP ranges, default gateway, lease duration, warning threshold, error threshold, classless static route option, and other options.
- 14 Expand the Static Bindings section.
- 15 Click **Add** to add static bindings between MAC addresses and IP addresses, default gateway, hostname, lease duration, classless static route option, and other options.

What to do next

Attach a DHCP server to a logical switch. See [Attach a DHCP Server to a Logical Switch](#).

Attach a DHCP Server to a Logical Switch

You must attach a DHCP server to a logical switch before the DHCP server can process DHCP requests from VMs connected to the switch.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Switching > Switches** from the navigation panel.
- 3 Click the logical switch that you intend to attach a DHCP server to.
- 4 Click **Actions > Attach DHCP Server**.

Detach a DHCP Server from a Logical Switch

You can detach a DHCP server from a logical switch to reconfigure your environment.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Switching > Switches** from the navigation panel.
- 3 Click the logical switch that you intend to detach a DHCP server from.
- 4 Click **Actions > Detach DHCP Server**.

Create a DHCP Relay Profile

A DHCP relay profile specifies one or more external DHCP servers. When you create a DHCP relay service, you must specify a DHCP relay profile.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **DHCP** from the navigation panel.
- 3 Click **Relay Profiles** and click **Add**.
- 4 Enter a name and optional description.
- 5 Enter one or more external DHCP server addresses.

What to do next

Create a DHCP relay service. See [Create a DHCP Relay Service](#).

Create a DHCP Relay Service

You can create a DHCP relay service to relay traffic between DHCP clients and DHCP servers that are not created in NSX-T.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **DHCP** from the navigation panel.
- 3 Click **Relay Services** and click **Add**.
- 4 Enter a name and optional description.
- 5 Select a DHCP relay profile from the drop-down menu.

What to do next

Add a DHCP service to a logical router port. See [Add a DHCP Service to a Logical Router Port](#).

Add a DHCP Service to a Logical Router Port

When you add a DHCP relay service to a logical router port, VMs on the logical switch that is attached to that port can communicate with the DHCP servers that are configured in the relay service.

Prerequisites

- Verify you have a configured DHCP relay service. See [Create a DHCP Relay Service](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Routing** from the navigation panel.
- 3 Select the router that is connected to the desired logical switch and click the **Configuration** tab.
- 4 Select the router port that connects to the desired logical switch and click **Edit**.
- 5 Select a DHCP relay service from the **DHCP Service** drop-down list and click **Save**.

The logical router port displays the DHCP relay service in the **DHCP Service** column.

You can also select a DHCP relay service when you add a new logical router port.

Configuring Metadata Proxies

With a metadata proxy server, VM instances can retrieve instance-specific metadata from an OpenStack Nova API server.

The following steps describe how a metadata proxy works:

- 1 A VM sends an HTTP GET to `http://169.254.169.254:80` to request some metadata.
- 2 The metadata proxy server that is connected to the same logical switch as the VM reads the request, makes appropriate changes to the headers, and forwards the request to the Nova API server.
- 3 The Nova API server requests and receives information about the VM from the Neutron server.
- 4 The Nova API server finds the metadata and sends it to the metadata proxy server.
- 5 The metadata proxy server forwards the metadata to the VM.

A metadata proxy server runs on an NSX Edge node. For high availability, you can configure metadata proxy to run on two or more NSX Edge nodes in an NSX Edge cluster.

This chapter includes the following topics:

- [Add a Metadata Proxy Server](#)
- [Attach a Metadata Proxy Server to a Logical Switch](#)
- [Detach a Metadata Proxy Server from a Logical Switch](#)

Add a Metadata Proxy Server

A metadata proxy server enables VMs to retrieve metadata from an OpenStack Nova API server.

Prerequisites

Verify that you have created an edge cluster. For more information, see *NSX-T Installation Guide*.

Procedure

- 1 From your browser, log in to an NSX Manager at `https://nsx-manager-ip-address`.
- 2 Select **DHCP > Metadata Proxies**.
- 3 Click **Add**.
- 4 Enter a name for the metadata proxy server.

- 5 (Optional) Enter a description.
- 6 Enter the URL for the Nova server.
- 7 Enter the secret parameter.
- 8 Select an edge cluster from the drop-down list.
- 9 (Optional) Select members of the edge cluster.

For example:

New Metadata Proxy Server
✕

Name:

Description:

Nova Server URL: *

Secret: *

Edge Cluster: * ▾

Members: ✕ ✕ ▾

What to do next

Attach the metadata proxy server to a logical switch.

Attach a Metadata Proxy Server to a Logical Switch

To provide metadata proxy services to VMs that are connected to a logical switch, you must attach a metadata proxy server to the switch.

Prerequisites

Verify that you have created a logical switch. For more information, see [Create a Logical Switch](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.

- 2 Select **DHCP > Metadata Proxies**.
- 3 Select a metadata proxy server.
- 4 Select the menu option **Actions > Attach to Logical Switch**
- 5 Select a logical switch from the drop-down list.

You can also attach a metadata proxy server to a logical switch by navigating to **Switching > Switches**, selecting a switch, and selecting the menu option **Actions > Attach Metadata Proxy**.

Detach a Metadata Proxy Server from a Logical Switch

To stop providing metadata proxy services to VMs that are connected to a logical switch or use a different metadata proxy server, you can detach a metadata proxy server from a logical switch.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **DHCP > Metadata Proxies**.
- 3 Select a metadata proxy server.
- 4 Select the menu option **Actions > Detach from Logical Switch**
- 5 Select a logical switch from the drop-down list.

You can also detach a metadata proxy server from a logical switch by navigating to **Switching > Switches**, selecting a switch, and selecting the menu option **Actions > Detach Metadata Proxy**.

Operations and Management

You may need to change the configuration of the appliances you've installed, for example, adding licenses, certificates, and changing passwords. There are also routine maintenance tasks that you should perform, including running backups. Additionally, there are tools to help you find information about the appliances that are part of the NSX-T infrastructure and the logical networks created by NSX-T, including remote system logging, traceflow, and port connections.

This chapter includes the following topics:

- [Add a License Key](#)
- [Managing User Accounts](#)
- [Setting Up Certificates](#)
- [Configuring Appliances](#)
- [Manage Tags](#)
- [Search for Objects](#)
- [Find the SSH Fingerprint of a Remote Server](#)
- [Backing Up and Restoring the NSX Manager](#)
- [Managing Appliances and Appliance Clusters](#)
- [Logging System Messages](#)
- [Configure IPFIX](#)
- [Trace the Path of a Packet with Traceflow](#)
- [View Port Connection Information](#)
- [Monitor a Logical Switch Port Activity](#)
- [Monitor Port Mirroring Sessions](#)
- [Monitor Fabric Nodes](#)
- [Collect Support Bundles](#)

Add a License Key

You can use the NSX Manager UI to add one or more license keys.

The following non-evaluation license types are available:

- Standard
- Advanced
- Enterprise

When you install NSX Manager, a pre-installed evaluation license becomes active and is valid for 60 days. The evaluation license provides all the features of an enterprise license. You cannot install or unassign an evaluation license.

You can install one or more of the non-evaluation licenses, but for each type, you can only install one key. When you install a standard, advanced, or enterprise license, the evaluation license is no longer available. You can also unassign non-evaluation licenses. If you unassign all non-evaluation licenses, the evaluation license is restored.

If you have multiple keys of the same license type and want to combine the keys, you must go to <https://my.vmware.com> and use the Combine Keys functionality. The NSX Manager UI does not provide this functionality.

Procedure

- 1 From your browser, log in to an NSX Manager at `https://nsx-manager-ip-address`.
- 2 Select **System > Configuration > License** from the navigation panel.
- 3 Click **Add** to enter the license key.
- 4 Click **Save**.

Managing User Accounts

NSX-T appliances have a local administrative user, admin. You cannot create or delete users.

Change the Admin Password

You can change the password for the admin user on any NSX-T appliance.

Procedure

- 1 Log in to the NSX Manager CLI.
- 2 Run the `set user` command.

```
nsx> set user admin
Current password:
New password:
Confirm new password:
nsx>
```

The password must meet these password complexity requirements:

- At least eight characters in length

- At least one uppercase character
- At least one lowercase character
- At least one numeric character
- At least one special character

Account Lockout

After five consecutive failed login attempts, the administrator account is locked for 15 minutes.

For NSX Manager, NSX Controller, and NSX Edge nodes the administrator account is locked out for 15 minutes after the fifth consecutive failed login attempt. To reset the locked account, wait 15 minutes to login again. This behavior is intentional, in that it prevents an attacker from learning of the existence of an account by observing the change in login failure messages from "incorrect password" to "account locked."

Note This applies to administrator logins via SSH, or via the console.

Setting Up Certificates

You can generate a Certificate signing request (CSR) in the NSX Manager and send it to a Certificate Authority (CA) to get a server certificate.

The CSR can also be used generate self-signed certificates. If you have an existing certificate or a CA certificate you can import it for use. You can also import a Certificate Revocation List (CRL) that includes revoked certificates.

Create a Certificate Signing Request File

Certificate signing request (CSR) is an encrypted text that contains specific information such as, organization name, common name, locality, and country. You send the CSR file to a certificate authority (CA) to apply for a digital identity certificate.

Prerequisites

- Gather the information that you need to fill out the CSR file. You must know the FQDN of the server and the organizational unit, organization, city, state, and country.
- Verify that the public and private key pairs are available.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **System > Settings** from the navigation panel.
- 3 Click the **Certificates** tab.
- 4 Select **CSRs** from the drop-down menu.
- 5 Click **Generate CSR**.

6 Complete the CSR file details.

Option	Description
Name	Assign a name for your certificate.
Common Name	Enter the fully qualified domain name (FQDN) of your server. For example, test.vmware.com.
Organization Name	Enter your organization name with applicable suffixes. For example, VMware Inc.
Organization Unit	Enter the department in your organization that is handling this certificate For example, IT department.
Locality	Add the city in which your organization is located. For example, Palo Alto.
State	Add the state in which your organization is located. For example, California.
Country	Add the country in which your organization is located. For example, United States (US).
Message Algorithm	Set the encryption algorithm for your certificate. RSA encryption - is used for digital signatures and encryption of the message. Therefore, it is slower than DSA when creating an encrypted token but faster to analyze and validate this token. This encryption is slower to decrypt and faster to encrypt. DSA encryption - is used for digital signatures. Therefore, it is faster than RSA when creating an encrypted token but slower to analyze and validate this token. This encryption is faster to decrypt and slower to encrypt.
Key Size	Set the key bits size of the encryption algorithm. The default value, 2048, is adequate unless you specifically need a different Key size. Many CAs require a minimum value of 2048. Larger key sizes are more secure but have a greater impact on performance.
Description	Enter specific details to help you identify this certificate at a later date.

7 Click **Save**.

A custom CSR appears as a link.

8 Select the CSR and click **Actions**.

9 Select **Download CSR PEM** from the drop-down menu.

You can save the CSR PEM file for your records and CA submission.

10 Use the contents of the CSR file to submit a certificate request to the CA in accordance with the CA enrollment process.

The CA creates a server certificate based on the information in the CSR file, signs it with its private key, and sends you the certificate. The CA also sends you a root CA certificate.

Import a CA Certificate

You can import a signed CA certificate to become an interim CA for your company. After you import the certificate, you have the authority to sign your own certificates.

Prerequisites

Verify that a CA certificate is available.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **System > Settings** from the navigation panel.
- 3 Click **Import** on the Certificates tab.
- 4 Select **Import CA Certificate** from the drop-down menu and add the certificate details.

Option	Description
Name	Assign a name to the CA certificate.
Certificate Contents	Browse to the CA certificate file on your computer and add the file.
Description	Enter a summary of what is included in this CA certificate.

- 5 Click **Save**.

You can now sign your own certificates.

Import a Certificate

You can import a certificate with the private key to create self-signed certificates.

Prerequisites

Verify that a certificate is available.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **System > Settings** from the navigation panel.
- 3 Click **Import** on the Certificates tab.
- 4 Select **Import Certificate** from the drop-down menu and add the certificate details.

Option	Description
Name	Assign a name to the CA certificate.
Certificate Contents	Browse to the certificate file on your computer and add the file.
Private Key	Browse to the private key file on your computer and add the file.

Option	Description
Password	Add a password for this certificate.
Description	Enter a summary of what is included in this certificate.

- 5 Click **Save**.

You can now create your own self-signed certificates.

Create a Self-Signed Certificate

Using self-signed certificates might be less secure than using trusted certificates.

When you use a self-signed certificate the client user receives a warning message such as, *Invalid Security Certificate*. The client user must then accept the self-signed certificate when first connecting to the server in order to proceed. Allowing client users to select this option provides reduced security than other authorization methods.

Prerequisites

Verify that a CSR is available. See [Create a Certificate Signing Request File](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **System > Settings** from the navigation panel.
- 3 Click the **Certificates** tab.
- 4 Select **CSRs** from the drop-down menu.
- 5 Select the existing CSR.
- 6 Click **Actions** and select **Self Sign Certificate for CSR** from the drop-down menu.
- 7 Enter the number of days the self-sign certificate is valid.
The default time frame is 10 years.
- 8 Click **Save**.

The self-signed certificate appears in the **Certificate** list. The certificate type is designated as self-signed.

Replace a Certificate

If you need to replace a certificate, for example if your certificate is expiring, you can use an API request to replace the existing certificate.

Prerequisites

Verify that a certificate is available in the NSX Manager. See [Create a Self-Signed Certificate](#) and [Import a Certificate](#).

Procedure

- 1 Select **System > Settings** from the navigation panel.
- 2 Click the **Certificates** tab and select **Certificates** from the drop-down menu.
- 3 Click on the ID of the certificate you want to use and copy the certificate ID from the pop-up window.
- 4 Send a POST `/api/v1/node/services/http?action=apply_certificate&certificate_id=<CertificateID>` API request to replace the existing certificate.

```
POST https://192.168.110.201/api/v1/node/services/http?
action=apply_certificate&certificate_id=e61c7537-3090-4149-b2b6-19915c20504f
```

The API request restarts the HTTP service so that the service can begin using the new certificate. When the POST request succeeds, the response code is 200 Accepted.

Import a Certificate Revocation List

A Certificate revocation list (CRL) is a list of subscribers and their certificate status. When a potential user attempts to access a server, the server denies access based on the CRL entry for that particular user.

The list contains the following items:

- Revoked certificates and the reasons for revocation
- Dates the certificates are issued
- Entities that issued the certificates
- Proposed date for the next release

Prerequisites

Verify that a CRL is available.

Procedure

- 1 From your browser, log in to an NSX Manager at `https://nsx-manager-ip-address`.
- 2 Select **System > Settings** from the navigation panel.
- 3 Click the **Certificates** tab.
- 4 Select **CRLs** from the drop-down menu.

5 Click **Import** and add the CRL details.

Option	Description
Name	Assign a name to the CRL.
Certificate Contents	<p>Copy all of the items in the CRL and paste them in this section.</p> <p>A sample CRL.</p> <pre>-----BEGIN X509 CRL----- MIIBODCB4zANBgkqhkiG9w0BAQQFADBGMQswCQYDVQQGEwJBVTEMMaOGA1UECBM D UUxEMRkwFwYDVQQKEwBNaw5jb20gUHR5LiBMdGQuMQswCQYDVQQLEwJDUzEbMBk G A1UEAxMSU1NMZW5IGRlbW8gc2VydMvyFw0wMTAxMTUxNjI2NTdaFw0wMTAyMTQ x NjI2NTdaMFwIwEgIBARcNOTUxMDA5MjMzMjA1WjASAgEDFw05NTEyMDEwMTAwMD a MBMCAhI0Fw0wMTAxMTUxNjE5NDdaMBMCAhI1Fw0wMTAxMTUxNjIzNDZaMA0GCSq G SIb3DQEBAUAA0EAHPjQ3M93Q0j8Ufi +jZM7Y78TfAzG4jJn/E6MYBPFVQFYo/Gp UZexfjSVo5CIyyS0tYscz8o07avwBxTiMpDEQg== -----END X509 CRL--</pre>
Description	Enter a summary of what is included in this CRL.

6 Click **Save**.

The imported CRL appears as a link.

Configuring Appliances

Some system configuration tasks must be done using the command line or API.

For complete command line interface information, see the *NSX-T Command-Line Interface Reference*.

For complete API information, see the *NSX-T API Guide*.

Table 11-1. System configuration commands and API requests.

Task	Command Line (NSX Manager, NSX Controller, NSX Edge)	API Request (NSX Manager only)
Set system timezone	set timezone <timezone>	PUT https://<nsx-mgr>/api/v1/node
Set NTP Server	set ntp-server <ntp-server>	PUT https://<nsx-mgr>/api/v1/node/services/ntp
Set a DNS server	set name-servers <dns-server>	PUT https://<nsx-mgr>/api/v1/node/network/name-servers
Set DNS Search Domain	set search-domains <domain>	PUT https://<nsx-mgr>/api/v1/node/network/search-domains

Manage Tags

You can add tags to objects to make searching for objects easier. When you specify a tag for an object, you can also specify a scope.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Navigate to an object category.
For example, navigate to **Switching > Switches**.
- 3 Select an object.
- 4 Select the menu option **Actions > Manage Tags**.
- 5 Add or delete tags.

Option	Action
Add a tag	Click Add to specify a tag and optionally a scope.
Delete a tag	Select an existing tag and click Delete .

A logical port can have a maximum of 15 tags. All other objects can have a maximum of 10 tags.

- 6 Click **Save**.

Search for Objects

You can search for objects using various criteria.

The following criteria are available for searching:

- Resource Type
- Name
- Description
- Creation Time
- Modified Time
- Created by
- Modified by
- Tags

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Click the magnifying-glass icon in the upper-right corner of the main window pane.

- 3 Enter a regular expression search pattern for an object or object type.

The search pattern is anchored by default, meaning that the start of string anchor `^` and the end of string anchor `$` are assumed. Do not use these anchors in the search pattern. For example, if you want to search for resources that start with 'Logical', the search pattern can be `Logical.*`. If you want to search for resources that end with 'Switch', the search pattern can be `.*Switch`.

- 4 In the window showing the results, click the link **View ... results** at the bottom of the window to open the search pane where you can refine the search.
- 5 Specify one or more criteria to refine the search.

Find the SSH Fingerprint of a Remote Server

Some API requests that involve copying files to or from a remote server require that you provide the SSH fingerprint for the remote server in the request body. The SSH fingerprint is derived from a host key on the remote server.

To connect via SSH, the NSX Manager and the remote server must have a host key type in common. If there are multiple host keys types in common, whichever one is preferred according to the `HostKeyAlgorithm` configuration on the NSX Manager is used.

Having the fingerprint for a remote server helps you confirm you are connecting to the correct server, protecting you from man-in-the-middle attacks. You can ask the administrator of the remote server if they can provide the SSH fingerprint of the server. Or you can connect to the remote server to find the fingerprint. Connecting to the server over console is more secure than over the network.

The NSX Manager appliance is based on Ubuntu 14.04 and uses the default `HostKeyAlgorithm` order. This table lists the keys that are present on the NSX Manager by default, in order from most preferred to least preferred.

Table 11-2. NSX Manager Host Keys in Preferred Order

Host key types present on NSX Manager	Default Location for that Host Key Type
ECDSA (256 bit)	<code>/etc/ssh/ssh_host_ecdsa_key.pub</code>
ED25519	<code>/etc/ssh/ssh_host_ed25519_key.pub</code>
RSA	<code>/etc/ssh/ssh_host_rsa_key.pub</code>
DSA	<code>/etc/ssh/ssh_host_dsa_key.pub</code>

Procedure

- 1 Log in to the CLI of the remote server.

Logging in using a console is more secure than over the network.

- List the public key files in the `/etc/ssh` directory.

```
$ ls -al /etc/ssh/*pub
-rw-r--r-- 1 root root 601 Apr  8 18:10 ssh_host_dsa_key.pub
-rw-r--r-- 1 root root  93 Apr  8 18:10 ssh_host_ed25519_key.pub
-rw-r--r-- 1 root root 393 Apr  8 18:10 ssh_host_rsa_key.pub
```

- Compare the available keys to the HostKeyAlgorithm order.

In this example there are three SSH keys, DSA and RSA, and ED25519. ED25519 is highest in the preferred order, so that is the key that the NSX Manager will use when connecting to the remote server.

- Get the fingerprint of the preferred key.

```
$ ssh-keygen -lf /etc/ssh/ssh_host_ed25519_key.pub
256 d0:21:3e:ec:52:ff:19:a9:e7:71:b5:7f:63:23:57:f7  root@ubuntu (ED25519)
```

The fingerprint of the key is `d0:21:3e:ec:52:ff:19:a9:e7:71:b5:7f:63:23:57:f7`.

Note You must remove the colons from the SSH fingerprint in backup and restore API requests.

Backing Up and Restoring the NSX Manager

If your NSX Manager virtual appliance becomes inoperable, it can be restored from backup. The NSX Manager stores the desired state for the virtual network. If the NSX Manager appliance becomes inoperable, the data plane is not affected, but configuration changes cannot be made.

There are three types of backups created from the two backup methods:

Cluster backup	This backup includes the desired state of the virtual network.
Node backup	This backup includes the NSX Manager appliance configuration.
Inventory backup	This backup includes the set of ESX and KVM hosts and edges. This information is used during a restore operation to detect and fix discrepancies between the Management Plane's desired state and these hosts.

There are two backup methods:

Manual NSX Manager node backup and cluster backup	Manual node and cluster backups can be run anytime as needed.
Automated NSX Manager node backup, cluster backup and inventory backup	Automated backups run based on a schedule that you set. Automated backups are highly recommended. See Schedule Automated Backups .

To ensure you have a recent backup, you should configure automated backups. It is important to run cluster and inventory backups regularly.

You can restore an NSX-T configuration back to the state that is captured in any of the cluster backups. When restoring a backup, you must restore to a new NSX Manager appliance running the same NSX Manager version as the appliance that was backed up.

Backup and restore requires that hypervisors, the NSX Manager appliance, and NSX Controller appliances must have static management IP addresses. Changing management IP addresses is not supported. Using DHCP to assign management IP addresses for NSX Manager and NSX Controller appliances is not supported. Using DHCP to assign management IP addresses for hypervisors is supported only if the DHCP server is configured to always provide the same IP address to a given hypervisor.

Back Up the NSX Manager Configuration

The NSX Manager configuration backup consists of the NSX Manager node backup, cluster backup and inventory backup.

Procedure

1 [Configure Backup Location](#)

Backups are saved to a remote SFTP location that NSX Manager can access. The backup location must be configured before a backup can occur.

2 [Schedule Automated Backups](#)

Schedule frequent backups so you can restore an inoperable NSX Manager and its configuration data. Automated backups are disabled by default. You can schedule automated backups to occur on specific days of the week or at a specified interval. Scheduled backups are highly recommended.

Configure Backup Location

Backups are saved to a remote SFTP location that NSX Manager can access. The backup location must be configured before a backup can occur.

Procedure

- 1 Log in to the NSX Manager Virtual Appliance.
- 2 Click **System > Utilities > Backup**.
- 3 To provide access credentials to the backup location, click **Edit** in the upper right of the page.
- 4 Click the **Automatic Backup** toggle to enable automatic backups.
- 5 Enter the IP address or host name of the SFTP server.
- 6 Edit the default port if required.
- 7 Enter the username and password required to log in to the SFTP server.
- 8 In the **Destination Directory** field, enter the absolute directory path where the backups will be stored.

- 9 Enter the passphrase used to encrypt the backup data.

You will need this passphrase to restore a backup. If you forget the backup passphrase, you cannot restore any backups.

- 10 Enter the SSH fingerprint of the server that stores the backups. See [Find the SSH Fingerprint of a Remote Server](#).
- 11 Click **Save**.
- 12 Click **Backup Now** on the bottom of the page to confirm that files can be written to the SFTP server.

What to do next

Schedule automated backups.

Schedule Automated Backups

Schedule frequent backups so you can restore an inoperable NSX Manager and its configuration data. Automated backups are disabled by default. You can schedule automated backups to occur on specific days of the week or at a specified interval. Scheduled backups are highly recommended.

Prerequisites

- Determine an appropriate backup location. Select a location that provides protection against single points of failure. For example, do not place the backups on the same file store as the appliances. A failure on that file store could affect both the appliances and their backups.
- Find the ssh fingerprint of the server that stores the backups. See [Find the SSH Fingerprint of a Remote Server](#). The backup and restore API requests require that the SSH fingerprint does not contain colons.

Procedure

- 1 Log in to the NSX Manager Virtual Appliance.
- 2 Click **System > Utilities > Backup**.
- 3 Click **Edit** in the upper right corner of the page.
- 4 Click **File Server** and then verify that Automatic Backup is enabled.
- 5 Click **Schedule** at the top of the page.
- 6 For the Node/Cluster Backup, click **Weekly** and then set the day(s) and time of the backup to the SFTP server, or click **Interval** and then set a backup time.
- 7 Inventory backups are set to occur every 30 seconds by default and should occur frequently. Accept or change the default setting as necessary.
- 8 Click **Save**.

Note The first weekly-scheduled backup occurs at the specified weekday and time. The first interval-scheduled backup occurs immediately after saving the backup configuration with enabled automated backups.

The NSX Manager stores three separate backup files: node-level, cluster-level and inventory. The backup files are saved to the SFTP server in the directory specified in the backup configuration. Inside that directory, the files are saved in the following directories:

- /<user specified directory>/cluster-node-backups (cluster and node backups)
- /<user specified directory>/inventory-summary (inventory backups)

Restore the NSX Manager Configuration

If your NSX Manager appliance is inoperable and you have taken the recommended backups, you can restore your NSX Manager appliance. You will need the passphrase specified when the backup was created to restore a backup.

Procedure

1 Prepare to Restore the NSX Manager Backups

Before restoring the NSX Manager backups, you must install a new NSX Manager appliance. The new NSX Manager must be deployed with the same management IP address as the previous NSX Manager.

2 Restore the Cluster Backup

The cluster backup is used to restore the desired network state. You must restore a cluster backup before restoring a node backup.

3 Restore the NSX Manager Node Backup

The node backup restores the appliance configuration, which allows the NSX Controller cluster to connect to it. You must restore a cluster backup before restoring a node backup. The node backup file that you choose should have the same timestamp as the cluster backup file.

4 Download the Backup and Restore Helper Script

You must download the backup and restore helper script from the NSX Manager.

5 Revert Fabric Changes Since Last Cluster Backup

The backup and restore helper script compares the desired state after the backups have been restored with the most recent fabric state captured by the script, and provides instructions for making the fabric state match the desired state after restoring the backup.

6 Restore an NSX Controller Cluster

If an NSX Controller Cluster is unrecoverable, or if you need to replace one or more controllers due to changes to cluster membership, you should restore the entire cluster of controllers.

Prepare to Restore the NSX Manager Backups

Before restoring the NSX Manager backups, you must install a new NSX Manager appliance. The new NSX Manager must be deployed with the same management IP address as the previous NSX Manager.

Prerequisites

- Verify that you have node, cluster, and the latest inventory backup files available to restore.
- Verify that you have the passphrase of the node and cluster backup files.
- Verify that you know the version of the NSX Manager used to create the backups, and have an appropriate installation file (OVA, OVF, or QCOW2) of the same version available.
- Verify that you know the IP address that was assigned to the NSX Manager used to create the node backup.
- Verify that no one will attempt to make configuration changes to the NSX Manager until the restore process is completed.

Procedure

- 1 If the old NSX Manager appliance is still running (for example, if you are restoring to roll back an upgrade attempt), shut it down.
- 2 Install a new NSX Manager appliance.
 - The version of the new NSX Manager appliance must be the same as the version of the appliance used to create the backups.
 - You must configure this appliance with the IP address of the NSX Manager used to create the node backup.

See the *NSX-T Installation Guide* for information and instructions about these steps.

What to do next

Restore the cluster backup.

Restore the Cluster Backup

The cluster backup is used to restore the desired network state. You must restore a cluster backup before restoring a node backup.

Prerequisites

- Find the ssh fingerprint of the server that stores the backups. See [Find the SSH Fingerprint of a Remote Server](#). The backup and restore API requests require that the SSH fingerprint does not contain colons.

Procedure

- 1 Check the status of the NSX Manager is STABLE before restoring the backup.

```
GET https://192.168.110.201/api/v1/cluster/status
{
  "control_cluster_status" : {
    "status" : "NO_CONTROLLERS"
  },
  "mgmt_cluster_status" : {
    "online_nodes" : [ {
      "mgmt_cluster_listen_ip_address" : "192.168.110.201",
      "uuid" : "422E901F-B167-DA0A-951F-C0278CA8A4BA"
    } ],
    "status" : "STABLE"
  }
}
```

Note The control cluster status is NO_CONTROLLERS because the control cluster will not connect to the NSX Manager until the node backup has been restored.

- 2 Send a cluster backup restore API request, POST /api/v1/cluster/backups?action=restore, which copies the backup file from a remote location and restores it on the NSX Manager appliance. Specify the backup file and location information in the API request.

Restore Request

Fields:

passphrase	The passphrase specified when the backup was created. If you don't know this password, you cannot restore this backup.
server	The remote server where the backup files are stored.
uri	The backup file path on the remote server.
ssh_fingerprint	The SSH fingerprint of the remote server where the backup files are stored. See Find the SSH Fingerprint of a Remote Server .
username	The username used to log in to the remote server to copy the backup files.
password	The password used to log in to the remote server to copy the backup files.

Example cluster backup restore request:

```
POST https://192.168.110.201/api/v1/cluster/backups?action=restore
{
  "restore_file": {
    "passphrase" : "7Taspa5anecR",
    "file_store" : "remote",
    "server" : "192.168.120.151",
```

```

    "uri" : "/vol0/backups/backup-cluster-20160314.zip",
    "protocol" : {
      "name" : "scp",
      "ssh_fingerprint" : "b508dfc65562e46e95707c25baf246f1",
      "authentication_scheme" : {
        "scheme_name" : "password",
        "username" : "admin" ,
        "password" : "4uhasWak"
      }
    }
  }
}

```

- 3 Wait for the system to become stable again.
- 4 Disable automatic backups.
 - a Log in to the NSX Manager Virtual Appliance.
 - b Click **System > Utilities > Backup**.
 - c Click **Edit** in the upper right of the page.
 - d Click the **Automatic Backup** toggle to disable automatic backups.

You can enable automatic backups once again after completing the backup and restore helper script.

What to do next

Restart all NSX Controllers to remove any cached data before the node backup is restored, and before the NSX Manager and NSX Controllers synchronize. See [Reboot NSX Controller Cluster Members](#).

Restore the NSX Manager Node Backup

The node backup restores the appliance configuration, which allows the NSX Controller cluster to connect to it. You must restore a cluster backup before restoring a node backup. The node backup file that you choose should have the same timestamp as the cluster backup file.

Caution You must restore a cluster backup before you restore a node backup. When the node backup is restored, the controllers can now communicate with the NSX Manager, and they will update the realized network state to match the desired network state configured on the NSX Manager. If the cluster backup has not been restored, there is no desired network state configured and the current realized network state will be destroyed.

Prerequisites

- Complete the restore of the cluster backup on the NSX Manager. See [Restore the Cluster Backup](#).
- Verify you have a backup of NSX Manager. See [Back Up the NSX Manager Configuration](#).
- Find the ssh fingerprint of the server that stores the backups. See [Find the SSH Fingerprint of a Remote Server](#). The backup and restore API requests require that the SSH fingerprint does not contain colons.

Procedure

- 1 Check the status of the NSX Manager is STABLE before restoring the backup.

```
GET https://192.168.110.201/api/v1/cluster/status
{
  "control_cluster_status" : {
    "status" : "NO_CONTROLLERS"
  },
  "mgmt_cluster_status" : {
    "online_nodes" : [ {
      "mgmt_cluster_listen_ip_address" : "192.168.110.201",
      "uuid" : "422E901F-B167-DA0A-951F-C0278CA8A4BA"
    } ],
    "status" : "STABLE"
  }
}
```

Note The control cluster status is NO_CONTROLLERS because the control cluster will not connect to the NSX Manager until the node backup has been restored.

- 2 Send a node backup restore API request, POST `/api/v1/node/backups?action=restore`, which copies the backup file from a remote location and restores it on the NSX Manager appliance. Specify the backup file and location information in the API request.

Restore Request

Fields:

passphrase	The passphrase specified when the backup was created. If you don't know this password, you cannot restore this backup.
server	The remote server where the backup files are stored.
uri	The backup file path on the remote server.
ssh_fingerprint	The SSH fingerprint of the remote server where the backup files are stored. See Find the SSH Fingerprint of a Remote Server .
username	The username used to log in to the remote server to copy the backup files.
password	The password used to log in to the remote server to copy the backup files.

Example node backup restore request:

```
POST https://192.168.110.201/api/v1/node/backups?action=restore
{
  "restore_file": {
    "passphrase" : "7Taspa5anecR",
    "file_store" : "remote",
    "server" : "192.168.120.151",
```

```

    "uri" : "/vol0/backups/backup-node-192.168.110.201-20160314.bak",
    "protocol" : {
      "name" : "scp",
      "ssh_fingerprint" : "b508dfc65562e46e95707c25baf246f1",
      "authentication_scheme" : {
        "scheme_name" : "password",
        "username" : "admin" ,
        "password" : "4uhasWak"
      }
    }
  }
}

```

What to do next

Download the backup and restore helper script.

Download the Backup and Restore Helper Script

You must download the backup and restore helper script from the NSX Manager.

Prerequisites

- Verify that the machine used to run the helper script meets the system requirements. The helper script requires python 2 and TLS 1.2, and has been verified on Ubuntu 14.04.

Procedure

- ◆ Download the backup and restore helper script. You can do this from the command line or the API.
 - From the command line:

Run the `copy file` command to copy the script to the remote server. The `url` argument specifies the destination for the script, using standard URL syntax, e.g. `scp://user@server/home/path/to/destination`.

```

nsx-manager-1> copy file backup_restore_helper.py url
scp://backups@192.168.120.151/vol0/backups/scripts/

```

- From the API:

Send this API request, and save the output to a file called `backup_restore_helper.py`.

```

GET https://nsx-manager-1/api/v1/node/file-store/backup_restore_helper.py/data

```

What to do next

Revert fabric changes since the last cluster backup.

Revert Fabric Changes Since Last Cluster Backup

The backup and restore helper script compares the desired state after the backups have been restored with the most recent fabric state captured by the script, and provides instructions for making the fabric state match the desired state after restoring the backup.

Prerequisites

- Verify that you have downloaded the backup and restore helper script.
- Verify that you have downloaded the latest inventory backup from the SFTP server.

Procedure

- 1 Log in to the machine to which the backup and restore helper script was downloaded or copied.
- 2 Run the backup and restore helper script, using the `-d` option to specify which checkpoint (inventory) file to use.

Provide the following information:

<code>-m</code>	NSX Manager IP address
<code>-u</code>	NSX Manager username
<code>-p</code>	NSX Manager password
<code>-d</code>	checkpoint (latest inventory backup) file name

```
$ python backup_restore_helper.py -m 192.168.110.201 -u admin -p <password> -d
backups/backup_restore_checkpoint_20160318_013354.json
```

- 3 Follow the instructions in the `backup_restore_helper.py` script output to update the fabric state to match the desired state.

Restore an NSX Controller Cluster

If an NSX Controller Cluster is unrecoverable, or if you need to replace one or more controllers due to changes to cluster membership, you should restore the entire cluster of controllers.

Before restoring a cluster of controllers, you first determine if control cluster membership has changed between what is known by the management plane and the actual membership as known by the controllers themselves. Membership can differ if changes were made after a backup.

- If the entire cluster is unrecoverable, see [Redeploy the NSX Controller Cluster](#).
- Follow the steps below to determine if cluster membership has changed, and if so, restore the cluster.

Prerequisites

- Verify that you have a recent cluster-level backup.
- Perform a cluster-level restore. See [Restore the Cluster Backup](#).

Procedure

- 1 Log in to the CLI of an NSX Manager and then run the `get management-cluster status` command.
- 2 Log in to the CLI of an NSX Controller and then run the `get managers` command to ensure that the controller is registered with the Manager.
- 3 Run the `get control-cluster status` command.
- 4 To determine if there are membership changes, compare the IP addresses from the output of the `get management-cluster status` command to the output from the `get control-cluster status` command.

No action is needed if the set of IP addresses is the same. If any IP address is different, continue with the remaining steps to restore the entire controller cluster.

- 5 Log in to the CLI of the NSX Controllers to determine which is the master controller by running the `get control-cluster status` command.
The master controller output will show `is master: true`.
- 6 Run the `stop service <controller>` command on one non-master controller.
- 7 Log in to the master controller and then run the `detach control-cluster <ip-address[:port]>` command to detach the non-master controller from the previous step.
- 8 (Optional) Run the `detach controller <uuid>` command on the NSX Manager to detach this controller only if the `get management-cluster status` command shows this controller on the NSX Manager.
- 9 Log in to the CLI of the NSX Controller and then run the `deactivate control-cluster` command.
- 10 Remove the bootstrap file and the uuid file with the following commands: `rm -r /opt/vmware/etc/bootstrap-config` and `rm -r /config/vmware/node-uuid`
- 11 Perform steps 6-10 for the remaining non-master controllers.
- 12 Log in to the CLI of the master controller and then run the `stop service <controller>` command.
- 13 Run the `detach controller <uuid>` command on the NSX Manager to detach this controller.
- 14 Log in to the CLI of the master controller and then run the `deactivate control-cluster` command.
- 15 Remove the bootstrap file and the uuid file with the following commands: `rm -r /opt/vmware/etc/bootstrap-config` and `rm -r /config/vmware/node-uuid`
- 16 Run the `get management-cluster status` command from the NSX Manager. If there are still controllers shown in the output, run the `detach controller <uuid>` command to detach any that remain.

What to do next

Complete the following tasks in the listed order.

- 1 Complete a node-level restore. See [Restore the NSX Manager Node Backup](#).
- 2 Join the NSX Controllers with the Management Plane, as documented in the *NSX-T Installation Guide*.
- 3 Redeploy the NSX Controller cluster, as documented in the *NSX-T Installation Guide*.

Managing Appliances and Appliance Clusters

Each installation of NSX-T requires and supports only one instance of NSX Manager. NSX Controller clusters should have three members. NSX Edge clusters should have at least two members.

If an appliance in a controller or edge cluster becomes inoperable, or you need to remove it for any reason, you can replace it with a new appliance.

Important If you make any changes to NSX Controller or NSX Edge cluster membership you must take a cluster backup afterwards to back up the new configuration. See [Backing Up and Restoring the NSX Manager](#).

Manage NSX Manager

You can check the status of NSX Manager with a CLI command. If NSX Manager is inoperable and unrecoverable, you can reboot the NSX Manager appliance.

Get NSX Manager Status

You can use a CLI command to get the status of NSX Manager.

Procedure

- 1 Log in to the CLI of NSX Manager.
- 2 Run the `get management-cluster status` command. For example,

```
nsx-manager> get management-cluster status
Number of nodes in management cluster: 1
-192.168.110.105
Management cluster status: STABLE

Number of nodes in control cluster: 3
- 192.168.110.52
- 192.168.110.53
- 192.168.110.51
Control cluster status: STABLE.
```

Note Even though the result says management cluster, there can be only one instance of NSX Manager.

Reboot NSX Manager

You can reboot NSX Manager with a CLI command to recover from critical errors.

Procedure

- 1 Log in to the CLI of NSX Manager.
- 2 Run the reboot command. For example,

```
nsx-manager> reboot
Are you sure you want to reboot (yes/no): y
```

Manage NSX Controller Cluster

The NSX Controller cluster should have three members. If after troubleshooting, you determine that one of your NSX Controller appliances are unrecoverable, you can add an appliance to the cluster to replace it, or if needed, redeploy the NSX Controller cluster.

An NSX Controller cluster must have majority to function normally. If two out of three members are online, the cluster still has majority. You should restore the three-member cluster by bringing up the offline NSX Controller. If you cannot bring it up, you can add another NSX Controller appliance to replace it, and regain majority. See [Replace a Member of the NSX Controller Cluster](#).

If one of out three members are online, the cluster does not have majority, and will not function normally. If you bring up the offline members, the cluster will regain majority. If you cannot bring up either of the offline members, you can redeploy the NSX Controller cluster. See [Redeploy the NSX Controller Cluster](#).

Prerequisites

Verify through troubleshooting that the appliances are not recoverable. For example, these steps may recover the appliances without having to replace them.

- Verify that the appliances have network connectivity, and resolve if not.
- Reboot the appliances.

What to do next

Get the NSX Controller cluster status. See [Get the NSX Controller Cluster Status](#).

Get the NSX Controller Cluster Status

You can find out the status of the NSX Controller cluster from the NSX Manager. You can also check the status of each NSX Controller from its command-line interface.

Getting the status of the NSX Controller cluster and cluster members can help you determine the source of a problem with the NSX Controller cluster.

Table 11-3. NSX Controller Cluster Status

	Is at least one controller registered with the NSX Manager?	Does the NSX Controller cluster have majority?	Are any NSX Controller cluster members down?
NO_CONTROLLERS	No	N/A	N/A
UNAVAILABLE	Unknown	Unknown	Unknown
STABLE	Yes	Yes	No
DEGRADED	Yes	Yes	Yes
UNSTABLE	Yes	No	No

Procedure

- 1 Log in to the NSX Manager CLI.
- 2 Run the `get management-cluster status` command.

```

nsx-manager-1> get management-cluster status
Number of nodes in management cluster: 3
- 192.168.110.203 (UUID 564DDA9E-8E84-E374-1F12-C69FAAE6A698) Online
- 192.168.110.201 (UUID 564D2E9C-A521-6C27-104F-76BBB5FCAC7F) Online
- 192.168.110.202 (UUID 564DC1B0-259A-9D6C-AF1F-12AEB6951882) Online

Management cluster status: STABLE

Number of nodes in control cluster: 3
- 192.168.110.52 (UUID 1228c336-3932-4b5b-b87e-9f66259cebcd)
- 192.168.110.53 (UUID f5348a2e-2d59-4edc-9618-2c05ac073fd8)
- 192.168.110.51 (UUID 03fad907-612f-4068-8109-efdf73002038)

Control cluster status: STABLE

```

- 3 Log in to the NSX Controller CLI.
- 4 Run the `get control-cluster status` command.

```

nsx-controller-1> get control-cluster status
uuid: 03fad907-612f-4068-8109-efdf73002038
is master: true
in majority: true

```

uuid	address	status
03fad907-612f-4068-8109-efdf73002038	192.168.110.51	active
1228c336-3932-4b5b-b87e-9f66259cebcd	192.168.110.52	active
f5348a2e-2d59-4edc-9618-2c05ac073fd8	192.168.110.53	active

Reboot NSX Controller Cluster Members

If you need to reboot multiple members of your NSX Controller cluster, you must reboot one member at a time. A three-member cluster can have majority if one member is offline. If two members are offline, the cluster will lose majority and will not function normally.

Procedure

- 1 Log in to the CLI of an NSX Manager.
- 2 Get the status of the management and control clusters.

```

nsx-manager-1> get management-cluster status
Number of nodes in management cluster: 3
- 192.168.110.201 (UUID 564D2E9C-A521-6C27-104F-76BBB5FCAC7F) Online
- 192.168.110.202 (UUID 564D0B9E-DEBD-A19E-233C-C13432CB23FB) Online
- 192.168.110.203 (UUID 564D666C-EB23-CDC1-8101-95155E9EB916) Online

Management cluster status: STABLE

Number of nodes in control cluster: 3
- 192.168.110.53 (UUID f5348a2e-2d59-4edc-9618-2c05ac073fd8)
- 192.168.110.51 (UUID 03fad907-612f-4068-8109-efdf73002038)
- 192.168.110.52 (UUID 1228c336-3932-4b5b-b87e-9f66259cebcd)

Control cluster status: STABLE

```

- 3 Log in to the CLI of an NSX Controller you need to reboot, and reboot it.

```

nsx-controller-2> reboot
Are you sure you want to reboot (yes/no): y

```

- 4 Get the status of the management and control cluster again. Wait until the control cluster status is STABLE before rebooting any additional members.

In this example, the NSX Controller 192.168.110.53 is rebooting, and the control cluster has a status of DEGRADED. This means the cluster is in majority, but one of the members is down. See [Get the NSX Controller Cluster Status](#) for more information about NSX Controller cluster status.

```

nsx-manager-1> get management-cluster status
Number of nodes in management cluster: 3
- 192.168.110.201 (UUID 564D2E9C-A521-6C27-104F-76BBB5FCAC7F) Online
- 192.168.110.202 (UUID 564D0B9E-DEBD-A19E-233C-C13432CB23FB) Online
- 192.168.110.203 (UUID 564D666C-EB23-CDC1-8101-95155E9EB916) Online

Management cluster status: STABLE

Number of nodes in control cluster: 3
- 192.168.110.53 (UUID f5348a2e-2d59-4edc-9618-2c05ac073fd8)
- 192.168.110.52 (UUID 1228c336-3932-4b5b-b87e-9f66259cebcd)
- 192.168.110.51 (UUID 03fad907-612f-4068-8109-efdf73002038)

Control cluster status: DEGRADED

```

Once the NSX Controller cluster has status of STABLE, it is safe to reboot any additional members.

```

nsx-manager-1> get management-cluster status
Number of nodes in management cluster: 3
- 192.168.110.201 (UUID 564D2E9C-A521-6C27-104F-76BBB5FCAC7F) Online
- 192.168.110.202 (UUID 564D0B9E-DEBD-A19E-233C-C13432CB23FB) Online
- 192.168.110.203 (UUID 564D666C-EB23-CDC1-8101-95155E9EB916) Online

Management cluster status: STABLE

Number of nodes in control cluster: 3
- 192.168.110.53 (UUID f5348a2e-2d59-4edc-9618-2c05ac073fd8)
- 192.168.110.51 (UUID 03fad907-612f-4068-8109-efdf73002038)
- 192.168.110.52 (UUID 1228c336-3932-4b5b-b87e-9f66259cebcd)

Control cluster status: STABLE

```

- 5 If you need information about individual NSX Controller appliance statuses, you can log into an NSX Controller and run the `get control-cluster status` command.

```

nsx-controller-1> get control-cluster status
uuid: 03fad907-612f-4068-8109-efdf73002038
is master: true
in majority: true

```

uuid	address	status
03fad907-612f-4068-8109-efdf73002038	192.168.110.51	active
1228c336-3932-4b5b-b87e-9f66259cebcd	192.168.110.52	active
f5348a2e-2d59-4edc-9618-2c05ac073fd8	192.168.110.53	not active

- 6 Repeat the steps to reboot additional NSX Controller appliances if needed.

Replace a Member of the NSX Controller Cluster

An NSX Controller cluster must have at least three members. If an NSX Controller appliance becomes inoperable and you need to remove it from the cluster, you must first add a new NSX Controller appliance to make a four-member cluster. Once the fourth member is added, you can remove an NSX Controller appliance from the cluster.

Prerequisites

- Verify through troubleshooting that the appliances are not recoverable. For example, these steps may recover the appliances without having to replace them.
 - Verify that the appliances have network connectivity, and resolve if not.
 - Reboot the appliances.
- Verify that you know the version of the NSX Controller that you are replacing and have an appropriate installation file (OVA, OVF, or QCOW2) of the same version available.

Procedure

- 1 Install and configure a new NSX Controller.

See the *NSX-T Installation Guide* for information and instructions about these steps.

- a Install a new NSX Controller appliance.

The version of the new NSX Controller must be the same as the NSX Controller it is replacing.

- b Join the new NSX Controller with the management plane.
- c Join the new NSX Controller with the control cluster.

- 2 Shut down the NSX Controller you want to remove from the cluster.

- 3 Log in to another NSX Controller and check that the NSX Controller you want to remove has a status of not active.

```
nsx-controller-1> get control-cluster status
uuid: e075cf44-0d49-4eb2-9e4f-d8b10ca97a3b
is master: true
in majority: true
  uuid                               address                status
  ----                               -
06996547-f50c-43c0-95c1-8bb644dea498 192.168.110.53         active
471e5ac0-194b-437c-9359-564cea845333 192.168.110.54         active
e075cf44-0d49-4eb2-9e4f-d8b10ca97a3b 192.168.110.51         active
863f9669-509f-4eba-b0ac-61a9702a242b 192.168.110.52         not active
```

- 4 Detach the controller from the cluster.

```
nsx-controller-1> detach control-cluster 192.168.110.52
Successfully detached node from the control cluster.
```

- 5 Detach the controller from the management plane.

```
nsx-manager-1> detach controller 863f9669-509f-4eba-b0ac-61a9702a242b
The detach operation completed successfully
```

- 6 Verify the controllers are active and the control cluster is stable.

From an NSX Controller:

```
nsx-controller-1> get control-cluster status
uuid: e075cf44-0d49-4eb2-9e4f-d8b10ca97a3b
is master: true
in majority: true
  uuid                               address                status
  ----                               -
06996547-f50c-43c0-95c1-8bb644dea498 192.168.110.53         active
471e5ac0-194b-437c-9359-564cea845333 192.168.110.54         active
e075cf44-0d49-4eb2-9e4f-d8b10ca97a3b 192.168.110.51         active
```

From an NSX Manager:

```

nsx-manager-1> get management-cluster status
Number of nodes in management cluster: 1
- 192.168.110.201 (UUID 4213216E-F93A-71B2-DA20-AFE5E714644F) Online
- 192.168.110.202 (UUID 4227F3D2-B7FE-8925-EA45-95ECD829C3E2) Online
- 192.168.110.203 (UUID 4227824A-1BDD-3A72-3EB3-8D306FEAE42D) Online

Management cluster status: STABLE

Number of nodes in control cluster: 3
- 192.168.110.51 (UUID e075cf44-0d49-4eb2-9e4f-d8b10ca97a3b)
- 192.168.110.53 (UUID 06996547-f50c-43c0-95c1-8bb644dea498)
- 192.168.110.54 (UUID 471e5ac0-194b-437c-9359-564cea845333)

Control cluster status: STABLE

```

Redeploy the NSX Controller Cluster

If replacing one controller has not resolved NSX Controller cluster issues, or if multiple NSX Controller appliances are unrecoverable, you can redeploy the whole cluster. The NSX Manager contains all desired configuration state, and can be used to re-create your NSX Controller cluster.

Data path connections will not be disrupted during the restore of the NSX Controller cluster.

Prerequisites

- Verify through troubleshooting that the appliances are not recoverable. For example, these steps may recover the appliances without having to replace them.
 - Verify that the appliances have network connectivity, and resolve if not.
 - Reboot the appliances.
- Verify that you know the version of the NSX Controller that you are replacing and have an appropriate installation file (OVA, OVF, or QCOW2) of the same version available.
- Verify that you know the IP addresses that were assigned to the NSX Controller appliances.

Procedure

- 1 Shut down all controllers in the NSX Controller cluster.

2 Detach the controllers from the NSX Manager.

- a Log in to the NSX Manager CLI.
- b Get a list of controllers with the `get management-cluster status` command.

```
nsx-manager-1> get management-cluster status
Number of nodes in management cluster: 1
- 192.168.110.201 (UUID 422EC8D8-B43F-D206-5048-781A5AECDC6) Online

Management cluster status: STABLE

Number of nodes in control cluster: 3
- 192.168.110.53 (UUID c28d0ac7-3107-4548-817a-50d76db007ab)
- 192.168.110.51 (UUID 4a0916c7-2f4d-48c2-81b6-29b7b3758ef4)
- 192.168.110.52 (UUID 1a409f24-9b9a-431e-a03a-1929db74bf00)

Control cluster status: UNSTABLE
```

- c Detach the controllers with the `detach controller` command.

```
nsx-manager-1> detach controller 1a409f24-9b9a-431e-a03a-1929db74bf00
The detach operation completed successfully
nsx-manager-1> detach controller 4a0916c7-2f4d-48c2-81b6-29b7b3758ef4
The detach operation completed successfully
nsx-manager-1> detach controller c28d0ac7-3107-4548-817a-50d76db007ab
The detach operation completed successfully
```

3 Install three NSX Controller appliances and create a new NSX Controller cluster.

See the *NSX-T Installation Guide* for information and instructions about these steps.

- a Install three NSX Controller appliances.
 - The version of the new NSX Controller appliances must be the same as the NSX Controller appliances they are replacing.
 - Assign the new controllers the same IP addresses that were used on the old controllers.
- b Join the NSX Controller appliances with the management plane.
- c On one of the NSX Controller appliances, initialize the control cluster.
- d Join the other two controllers with the control cluster.

Manage NSX Edge Cluster

You can replace an NSX Edge if, for example, it has become inoperable, or if you need to change hardware. After you install a new NSX Edge and create a new transport node, you can modify the edge cluster to replace the old transport node with the new transport node.

Note Removing a tier-1 edge cluster will cause the tier-1 distributed router (DR) instance to be out of service briefly.

Procedure

- 1 If the NSX Edge you want to replace is still operating, you can put it in to maintenance mode to minimize downtime. If high availability is enabled on the associated logical routers, entering maintenance mode will cause the logical routers to use a different edge cluster member. You do not need to do this if the NSX Edge is inoperable.

- a Get the fabric node ID of the failed fabric node.

```
https://192.168.110.201/api/v1/fabric/nodes
...
  "resource_type": "EdgeNode",
  "id": "a0f4fa74-e77c-11e5-8701-005056aeed61",
  "display_name": "edgenode-02a",
...

```

- b Put the failed NSX Edge node into maintenance mode.

```
POST https://192.168.110.201/api/v1/fabric/nodes/a0f4fa74-e77c-11e5-8701-005056aeed61?
action=enter_maintenance_mode

```

- 2 Install a new NSX Edge.

See the *NSX-T Installation Guide* for information and instructions about these steps.

- 3 Join the new NSX Edge with the management plane with the `join management-plane` command.

See the *NSX-T Installation Guide* for information and instructions about these steps.

4 Configure the NSX Edge as a transport node.

See the *NSX-T Installation Guide* for information and instructions about these steps.

You can get the transport node configuration of the failed NSX Edge appliance from the API, and use this information to create the new transport node.

a Get the fabric node ID of the new fabric node.

```
https://192.168.110.201/api/v1/fabric/nodes
...
  "resource_type": "EdgeNode",
  "id": "d61c8d86-f4b8-11e5-b1b2-005056ae3c10",
  "display_name": "edgenode-03a",
...
```

b Get the transport node ID of the failed transport node.

```
GET https://192.168.110.201/api/v1/transport-nodes
...
{
  "resource_type": "TransportNode",
  "description": "",
  "id": "73cb00c9-70d0-4808-abfe-a12a43251133",
  "display_name": "TN-edgenode-01a",
...
}
```

- c Get the transport node configuration of the failed transport node.

```
GET https://192.168.110.201/api/v1/transport-nodes/73cb00c9-70d0-4808-abfe-a12a43251133
{
  "resource_type": "TransportNode",
  "description": "",
  "id": "73cb00c9-70d0-4808-abfe-a12a43251133",
  "display_name": "TN-edgenode-01a",
  "tags": [],
  "transport_zone_endpoints": [
    ...
  ],
  "host_switches": [
    ...
  ],
  "node_id": "a0f4fa74-e77c-11e5-8701-005056aeed61",
  "_create_time": 1457696199196,
  "_last_modified_user": "admin",
  "_last_modified_time": 1457696225606,
  "_create_user": "admin",
  "_revision": 2
}
```

- d Create the new transport node with POST /api/v1/transport-nodes.

In the request body, provide the following information for the new transport node:

- description for the new transport node (optional)
- display_name for the new transport node
- node_id of the fabric node used to create the new transport node

In the request body, copy the following information from the failed transport node:

- transport_zone_endpoints
- host_switches
- tags (optional)

```
POST https://192.168.110.201/api/v1/transport-nodes
{
  "description": "",
  "display_name": "TN-edgenode-03a",
  "tags": [
    ...
  ],
  "transport_zone_endpoints": [
    ...
  ],
  "host_switches": [
    ...
  ],
  "node_id": "d61c8d86-f4b8-11e5-b1b2-005056ae3c10"
}
```

- 5 Edit the edge cluster to replace the failed transport node with the new transport node.
- Get the ID of the new transport node and the failed transport node. The `id` field contains the transport node ID.

```
GET https://192.168.110.201/api/v1/transport-nodes
...
{
  "resource_type": "TransportNode",
  "description": "",
  "id": "73cb00c9-70d0-4808-abfe-a12a43251133",
  "display_name": "TN-edgenode-01a",
  ...
}
...
{
  "resource_type": "TransportNode",
  "description": "",
  "id": "890f0e3c-aa81-46aa-843b-8ac25fe30bd3",
  "display_name": "TN-edgenode-03a",
  ...
}
```

- Get the ID of the edge cluster. The `id` field contains the edge cluster ID. Get the members of the edge cluster from the `members` array.

```
GET https://192.168.110.201/api/v1/edge-clusters
....
{
  "resource_type": "EdgeCluster",
  "description": "",
  "id": "9a302df7-0833-4237-af1f-4d826c25ad78",
  "display_name": "Edge-Cluster-1",
  ...
  "members": [
    {
      "member_index": 0,
      "transport_node_id": "73cb00c9-70d0-4808-abfe-a12a43251133"
    },
    {

```

```

    "member_index": 1,
    "transport_node_id": "e5d17b14-cdeb-4e63-b798-b23a0757463b"
  }
],

```

- c Edit the edge cluster to replace the failed transport node with the new transport node. The `member_index` must match the index of the failed transport node.

Caution If the NSX Edge is still operating, this is a disruptive action. This will move all the logical router ports from the failed transport node to the new transport node.

In this example, the transport node TN-edgenode-01a (73cb00c9-70d0-4808-abfe-a12a43251133) has failed, and is replaced by transport node TN-edgenode-03a (890f0e3c-aa81-46aa-843b-8ac25fe30bd3) in edge cluster Edge-Cluster-1 (9a302df7-0833-4237-af1f-4d826c25ad78).

```

POST http://192.168.110.201/api/v1/edge-clusters/9a302df7-0833-4237-af1f-4d826c25ad78?
action=replace_transport_node
{
  "member_index": 0,
  "transport_node_id" : "890f0e3c-aa81-46aa-843b-8ac25fe30bd3"
}

```

- 6 (Optional) Delete the failed transport node and NSX Edge node.

Logging System Messages

Log messages from all NSX-T components except the ones running on ESXi conform to the RFC 5424 format. You can configure a remote logging server to receive log messages.

For more information about RFC 5424, see <https://tools.ietf.org/html/rfc5424>.

RFC 5424 defines the following format for log messages:

```
<facility * 8 + severity> version UTC-TZ hostname APP-NAME procid MSGID [structured-data] msg
```

A sample log message from NSX Manager:

```
<187>1 2016-03-15T22:53:00.114Z nsx-manager NSX - SYSTEM [nsx@6876 comp="nsx-manager"
errorCode="MP4039" subcomp="manager"] Connection verification failed for broker '10.160.108.196'.
Marking broker unhealthy.
```

NSX-T produces regular logs (facility local6, which has a numerical value of 22) and audit logs (facility local7, which has a numerical value of 23). All API calls trigger an audit log.

RFC 5424 defines the following severity levels:

Severity Value	Description
0	Emergency: system is unusable
1	Alert: action must be taken immediately

Severity Value	Description
2	Critical: critical conditions
3	Error: error conditions
4	Warning: warning conditions
5	Notice: normal but significant condition
6	Informational: informational messages
7	Debug: debug-level messages

All logs with a severity of emergency, alert, critical, or error contain a unique error code in the structured data portion of the log message. The error code consists of a string and a decimal number. The string represents a specific module.

The MSGID field indicates a category for the log message. For a list of the categories, see [Log Message Categories](#).

Configure Remote Logging

You can configure NSX-T appliances and hypervisors to send log messages to a remote logging server. Remote logging is supported on NSX Manager, NSX Controller, NSX Edge appliances, and hypervisors. You can filter which log messages are sent to the logging server, based on these criteria:

- level: emerg, alert, crit, err, warning, notice, info, debug
- facility: codes are defined in RFC 5424. Facility local7 is used for audit messages, and local6 is used for non-audit messages.
- message ID or category: categories and examples are listed here: [Log Message Categories](#)

See the *NSX-T Command-Line Reference* and *NSX-T API Guide* for information about related commands and requests.

Prerequisites

- Configure a remote logging server to receive the logs from NSX-T appliances.
- Determine what log messages you want to send to the logging server.

Procedure

- 1 Log into the NSX-T appliance you want to configure with remote logging.
- 2 Configure a logging server with the `set logging-server` command using the following syntax. Multiple facilities or message IDs can be specified as a comma delimited list, without spaces.

```
set logging-server <hostname-or-ip-address[:port]> proto <proto> level <level> [facility
<facility>] [messageid <messageid>] [certificate <filename>]
```

You can run the command multiple times to add multiple logging server configurations.

```
nsx> set logging-server 192.168.110.60 proto udp level info facility syslog messageid SYSTEM,FABRIC
nsx> set logging-server 192.168.110.60 proto udp level info facility auth,user
```

3 (Optional) View the logging configuration with the `get logging-server` command.

```
nsx> get logging-servers
192.168.110.60 proto udp level info facility syslog messageid SYSTEM,FABRIC
192.168.110.60 proto udp level info facility auth,user
```

Log Message Categories

Log messages belong to a category. These categories can be used in the `set logging-server` command to filter which log messages are sent to a logging server.

Table 11-4. Log Message Categories

Log Message Category	Examples
FABRIC	Host node Host preparation Edge node Transport zone Transport node Uplink profiles Cluster profiles Edge cluster Bridge clusters and endpoints
SWITCHING	Logical switch Logical switch ports Switching profiles switch security features
ROUTING	Logical router Logical router ports Static routing Dynamic routing NAT
FIREWALL	Firewall rules Firewall rule sections
FIREWALL_PKTLOG	Firewall connection logs Firewall packet logs

Table 11-4. Log Message Categories (Continued)

Log Message Category	Examples
GROUPING	IP sets Mac sets NSGroups NSServices NSService groups VNI Pool IP Pool
DHCP	DHCP relay
SYSTEM	Appliance management (remote syslog, ntp, etc) Cluster management Trust management Licensing User and roles Task management Install (NSX Manager, NSX Controller) Upgrade (NSX Manager, NSX Controller, NSX Edge and host-packages upgrades) Realization Tags
MONITORING	SNMP Port connection Traceflow
-	All other log messages.

Configure IPFIX

IPFIX (Internet Protocol Flow Information Export) is a standard for the format and export of network flow information. When you enable IPFIX, all configured host transport nodes will send IPFIX messages to the IPFIX collectors using port 4739.

In the case of ESXi, NSX-T automatically opens port 4739. In the case of KVM, if firewall is not enabled, port 4739 is open, but if firewall is enabled, you must ensure that the port is open because NSX-T does not automatically open the port.

Prerequisites

- Install at least one IPFIX collector.
- Verify that the IPFIX collectors have network connectivity to the hypervisors.
- Verify that any relevant firewalls, including ESXi firewall, allow traffic on the IPFIX collector ports.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Tools > IPFIX** from the navigation panel.

3 Click the **Collectors** tab if it is not already selected.

4 Click **Configure Collectors**.

5 Click **Add** and enter the collector IP Address and Port.

You can add up to 8 collectors.

6 (Optional) In the Collection Options section, click **Edit** to specify the observation domain ID.

The observation domain ID identifies which observation domain the network flows originated from. The default value is 0, which indicates no specific observation domain.

7 Click the **Switch IPFIX Profiles** tab.

8 Click **Add** to add a profile.

Setting	Description
Active Timeout (seconds)	The length of time after which a flow will time out, even if more packets associated with the flow are received. Default is 300.
Idle Timeout (seconds)	The length of time after which a flow will time out, if no more packets associated with the flow are received (ESXi only, KVM times out all flows based on active timeout). Default is 300.
Max Flows	The maximum flows cached on a bridge (KVM only, not configurable on ESXi). Default is 16384.
Sampling Probability (%)	The percentage of packets that will be sampled (approximately). Increasing this setting may have a performance impact on the hypervisors and collectors. If all hypervisors are sending more IPFIX packets to the collector, the collector may not be able to collect all packets. Setting the probability at the default value of 0.1% will keep the performance impact low.

9 Click **Applied To** to apply the profile to one or more objects.

The types of object are logical ports and logical switches.

IPFIX on ESXi and KVM sample tunnel packets in different ways. On ESXi the tunnel packet is sampled as two records:

- Outer packet record with some inner packet information
 - SrcAddr, DstAddr, SrcPort, DstPort, and Protocol refer to the outer packet.
 - Contains some enterprise entries to describe the inner packet.
- Inner packet record
 - SrcAddr, DstAddr, SrcPort, DstPort, and Protocol refer to the inner packet.

On KVM the tunnel packet is sampled as one record:

- The inner packet record with some outer tunnel information
 - SrcAddr, DstAddr, SrcPort, DstPort, and Protocol refer to the inner packet.
 - Contains some enterprise entries to describe the outer packet.

Trace the Path of a Packet with Traceflow

Use Traceflow to inspect the path of a packet as it travels from one logical port on the logical network to another logical port on the same network. Traceflow traces the transport node-level path of a packet injected at a logical port. The trace packet traverses the logical switch overlay, but is not visible to interfaces attached to the logical switch. In other words, no packet is actually delivered to the test packet's intended recipients.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Navigate to the Traceflow screen. You have two options.
 - Select **Tools > Traceflow** from the navigation panel.
 - Select **Switching** from the navigation panel, click the **Ports** tab, select a VIF-attached port and click **Actions > Traceflow**
- 3 Select a traffic type.
The choices are Unicast, Multicast, and Broadcast.
- 4 Specify the source and destination information according to the traffic type.

Traffic Type	Specify Source Information	Specify Destination Information
Unicast	Select a VM and a virtual interface. The IP address and MAC address are displayed if VMtools is installed in the VM or if the VM is deployed using OpenStack plug-in (address bindings will be used in this case). If the VM has more than one IP address, select one from the drop-down menu. If the IP address and MAC address are not displayed, enter the IP address and MAC address in the text boxes. This will also apply to Multicast and Broadcast.	Select either VM Name or IP-MAC from the "Type" drop-down menu. <ul style="list-style-type: none"> ■ If VM Name is selected, select a VM and virtual interface. Select or enter an IP address and a MAC address ■ If IP-MAC is selected, select the trace type (Layer 2 or layer 3). If trace type is Layer 2, enter an IP address and a MAC address. If trace type is Layer 3, enter an IP address.
Multicast	Same as above.	Enter an IP Address. It must be a multicast address from 224.0.0.0 - 239.255.255.255.
Broadcast	Same as above.	Enter a subnet prefix length.

- 5 (Optional) Click **Advanced** to see the advanced options.
- 6 (Optional) In the left column, enter the desired values or input for the following fields:

Option	Description
Frame Size	e.g. 128
TTL	e.g. 64
Timeout (ms)	e.g. 10000

Option	Description
Ethertype	e.g. 2048
Payload Type	Select an option from the dropdown menu.
Payload Data	Payload formatted based on selected Payload Type (Base64, Hex, Plaintext, Binary, or Decimal)

- 7 (Optional) In the left column under "Protocol", select a protocol from the "Type" drop-down menu.
- 8 (Optional) Based on the protocol selected, complete the associated steps in the following table.

Protocol	Step 1	Step 2	Step 3
TCP	Enter a source port.	Enter a destination port.	Select the desired TCP Flags from the drop-down menu.
UDP	Enter a source port.	Enter a destination port.	N/A
ICMP	Enter an ICMP ID.	Enter a sequence value.	N/A

- 9 Click **Trace**.

Information about the connections, components, and layers is displayed. The output includes a table listing Observation Type (Delivered, Dropped, Received, Forwarded), Transport Node, and Component, and a graphical map of the topology if unicast and logical switch as a destination are selected. You can apply a filter (**All**, **Delivered**, **Dropped**) on the observations that are displayed. If there are dropped observations, the **Dropped** filter is applied by default. Otherwise, the **All** filter is applied.

View Port Connection Information

You can use the port connection tool to quickly visualize and troubleshoot the connection between two VMs.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Tools > Port Connection** from the navigation panel.
- 3 Select a VM from the **Source Virtual Machine** drop-down menu.
- 4 Select a VM from the **Destination Virtual Machine** drop-down menu.
- 5 Click **Go**.

A visual map of the port connection topology is displayed. You can click on any of the components in the visual output to reveal more information about that component.

Monitor a Logical Switch Port Activity

You can monitor the logical port activity for example, to troubleshoot network congestion and packets being dropped

Prerequisites

Verify that a logical switch port is configured. See [Connecting a VM to a Logical Switch](#).

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Switching > Port** from the navigation panel.
- 3 Double-click the logical switch port to monitor.
- 4 Click the **Monitor** tab.
- 5 Select **Begin Tracking**.

A port tracking page opens.

- 6 Start monitoring activity on the logical switch port.

You can view the bidirectional port traffic and identify dropped packets. The port tracker page also lists the switching profiles attached to the logical switch port.

For example, if you notice dropped packets because of network congestion you can configure a QoS switching profile for the logical switch port to prevent data loss on preferred packets. See [Understanding QoS Switching Profile](#).

Monitor Port Mirroring Sessions

You can monitor port mirroring sessions for troubleshooting and other purposes.

This feature has the following restrictions:

- A source mirror port cannot be in more than one mirror session.
- A destination port can only receive mirror traffic.
- With KVM, multiple NICs can be attached to the same OVS port. The mirroring happens at the OVS uplink port, meaning that traffic on all the pNICs attached to the OVS port is mirrored.
- Mirror session source and destination ports must be on the same host vSwitch. Therefore, if you vMotion the VM that has the source or destination port to another host, traffic on that port can no longer be mirrored.
- On ESXi, when mirroring is enabled on the uplink, raw production TCP packets are encapsulated using the Geneve protocol by VDL2 into UDP packets. A physical NIC that supports TSO (TCP segmentation offload) can change the packets and mark the packets with the MUST_TSO flag. On a monitor VM with VMXNET3 or E1000 vNICs, the driver treats the packets as regular UDP packets and cannot handle the MUST_TSO flag, and will drop the packets.

If a lot of traffic is mirrored to a monitor VM, there is a potential for the driver's buffer ring to become full and packets to be dropped. To alleviate the problem, you can take one or more of the following actions:

- Increase the rx buffer ring size.
- Assign more CPU resources to the VM.

- Use the Data Plane Development Kit (DPDK) to improve packet processing performance.

Note Make sure that the monitor VM's MTU setting (in the case of KVM, the hypervisor's virtual NIC device's MTU setting also) is large enough to handle the packets. This is especially important for encapsulated packets because encapsulation increases the size of packets. Otherwise, packets might be dropped. This is not an issue with ESXi VMs with VMXNET3 NICs, but is a potential issue with other types of NICs on both ESXi and KVM VMs.

Note In an L3 port mirroring session involving VMs on KVM hosts, you must set the MTU size to be large enough to handle the extra bytes required by encapsulation. The mirror traffic goes through an OVS interface and OVS uplink. You must set the OVS interface's MTU to be at least 100 bytes larger than the size of the original packet (before encapsulation and mirroring). If you see dropped packets, increase the MTU setting for the host's virtual NIC and the OVS interface. Use the following command to set the MTU for an OVS interface:

```
ovs-vsctl -- set interface <ovs_Interface> mtu_request=<MTU>
```

Note When you monitor the logical port of a VM and the uplink port of a host where the VM resides, you will see different behaviors depending on whether the host is ESXi or KVM. For ESXi, the logical-port mirror packets and the uplink mirror packets are tagged with the same VLAN ID and appear the same to the monitor VM. For KVM, the logical-port mirror packets are not tagged with a VLAN ID but the uplink mirror packets are tagged, and they appear different to the monitor VM.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Tools > Port Mirroring Session** from the navigation panel.
- 3 Enter a session name.
- 4 Select a transport node from the drop-down menu.
A port mirroring session must be between NICs on the same transport node.
- 5 Select a direction from the drop-down menu.
The choices are **Bidirectional**, **Ingress**, and **Egress**.
- 6 (Optional) Select a packet truncation value.
- 7 Click **Next**.
- 8 Select source PNICs.
- 9 (Optional) Toggle the **Encapsulated Packet** switch to disable the capturing of encapsulated traffic.
This switch is enabled by default.
- 10 Select source VNICs.

11 Select a destination.

You can select up to 3 VMs and up to 3 VNICs.

12 Click **Save**.

You cannot change the source and destination after saving the port mirroring session.

Monitor Fabric Nodes

You can monitor fabric nodes such as hosts, edges, edge clusters, bridges, and transport nodes from the NSX Manager UI.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **Fabric > Nodes** from the navigation panel.
- 3 Select one of the following tabs.
 - Hosts
 - Edges
 - Edge Clusters
 - Bridges
 - Transport Nodes

Note On the Hosts screen, if the MPA Connectivity status is Down or Unknown for a host, ignore the LCP Connectivity status because it might be inaccurate.

Collect Support Bundles

You can collect support bundles on registered cluster and fabric nodes and download the bundles to your machine or upload them to a file server.

If you choose to download the bundles to your machine, you get a single archive file consisting of a manifest file and support bundles for each node. If you choose to upload the bundles to a file server, the manifest file and the individual bundles are uploaded to the file server separately.

Procedure

- 1 From your browser, log in to an NSX Manager at <https://nsx-manager-ip-address>.
- 2 Select **System > Utilities** from the navigation panel.
- 3 Click the **Support Bundle** tab.
- 4 Select the target nodes.

The available types of nodes are management nodes, controller nodes, edges, and hosts.

- 5 (Optional) Specify log age in days to exclude logs that are older than the specified number of days.

- 6 (Optional) Toggle the switch that indicates whether to include or exclude core files and audit logs.
Core files and audit logs might contain sensitive information such as passwords or encryption keys.
- 7 (Optional) Select a check box to upload the bundles to a file server.
- 8 Click **Start Bundle Collection** to start collecting support bundles.
Depending on how many log files exist, each node might take several minutes.
- 9 Monitor the status of the collection process.
The status field shows the percentage of nodes that completed support bundle collection.
- 10 Click **Download** to download the bundle if the option to send the bundle to a file server was not set.