# NSX-T 管理指南

VMware NSX-T Data Center 1.1

vmware®

您可以从 VMware 网站下载最新的技术文档：
https://docs.vmware.com/cn/。
VMware 网站还提供了最近的产品更新。
如果您对本文档有任何意见或建议，请将反馈信息发送至：
docfeedback@vmware.com

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

北京办公室
北京市
朝阳区新源南路 8 号
启皓北京东塔 8 层 801
www.vmware.com/cn

上海办公室
上海市
淮海中路 333 号
瑞安大厦 804-809 室
www.vmware.com/cn

广州办公室
广州市
天河路 385 号
太古汇一座 3502 室
www.vmware.com/cn

# 目录

# 关于管理 VMware NSX-T

NSX-T 管理指南 提供有关配置和管理 VMware NSX-T ® 网络的信息，包括如何创建逻辑交换机和端口以及如何为分层逻辑路由器设置网络。本文档还介绍了如何配置 NAT、防火墙、SpoofGuard、分组和 DHCP。

## 目标读者

本文档中的信息适用于任何要配置 NSX-T 的人员。本文档中的信息是为熟悉虚拟机技术、网络和安全操作且经验丰富的 Windows 或 Linux 系统管理员编写的。

## VMware 技术出版物术语表

VMware 技术出版物提供了一个术语表，其中包含一些您可能不熟悉的术语。有关 VMware 技术文档中使用的术语的定义，请访问 http://www.vmware.com/support/pubs。

# NSX-T 概述

1

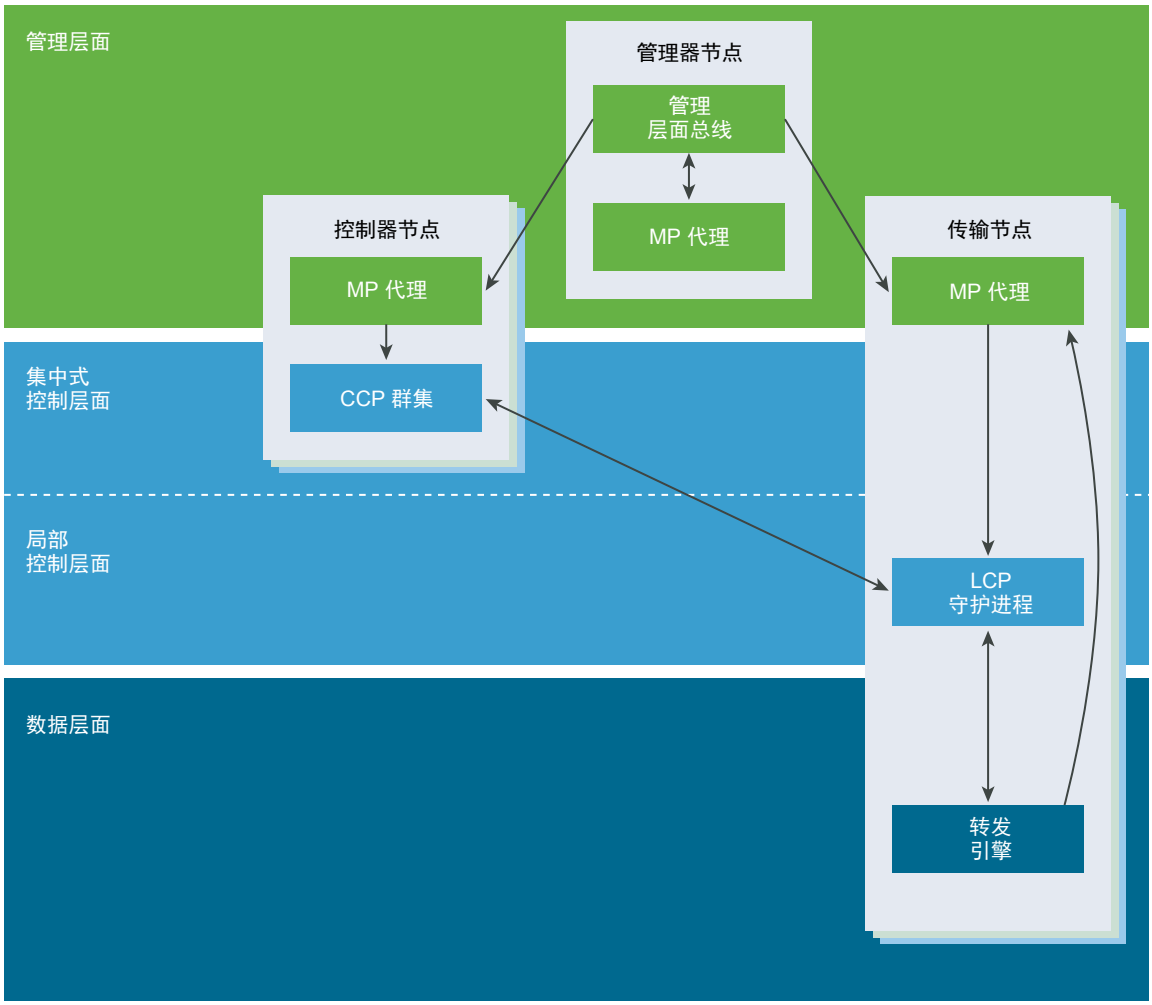
与服务器虚拟化以编程方式创建、删除和还原基于软件的虚拟机 (VM) 以及拍摄虚拟机快照的方式大致相同，NSX-T 网络虚拟化也是以编程方式创建、删除和还原基于软件的虚拟网络以及拍摄虚拟网络快照。

通过网络虚拟化，与网络管理程序功能等效的组件以软件形式再现一整套第 2 层至第 7 层网络服务（例如，交换、路由、访问控制、防火墙和 QoS）。因此，可通过编程方式任意组合这些服务，只需短短数秒，即可生成独一无二的独立虚拟网络。

NSX-T 的工作方式是实现三个单独但集成的层面：管理、控制和数据。这三个层面是作为位于三种类型的节点上的一组进程、模块和代理实现的：管理器、控制器和传输节点。

- 每个节点托管一个管理层面代理。
- NSX Manager 节点托管 API 服务。每个 NSX-T 安装支持单个 NSX Manager 节点，而不支持 NSX Manager 群集。
- NSX Controller 节点托管中央控制层面群集守护进程。
- 可以在同一物理服务器上托管 NSX Manager 和 NSX Controller 节点。

- 传输节点托管本地控制层面守护进程和转发引擎。




本章讨论了以下主题：

- 数据层面
- 控制层面
- 管理层面
- NSX Manager
- NSX Controller
- 逻辑交换机
- 逻辑路由器
- NSX Edge
- 传输区域
- 重要概念

# 数据层面

根据控制层面填充的表执行无状态数据包转发/转换，向控制层面报告拓扑信息以及维护数据包级别统计信息。

数据层面是物理拓扑和状态的真实数据源，例如，VIF 位置、隧道状态，等等。如果要将数据包从一个位置移动到另一个位置，则需要位于数据层面。数据层面还维护多个链路/隧道的状态并处理它们之间的故障切换。每个数据包的性能是至关重要的，并具有非常严格的延迟或抖动要求。数据层面并不一定完全包含在内核、驱动程序、用户空间甚至特定用户空间进程中。数据层面限制为基于控制层面填充的表/规则的完全无状态转发。

数据层面可能还具有维护一定数量的功能状态（如 TCP 终止）的组件。这与控制层面管理的状态（如 MAC:IP 隧道映射）不同，因为控制层面管理的状态与如何转发数据包有关，而数据层面管理的状态仅限于如何处理负载。

# 控制层面

根据管理层面中的配置计算所有瞬间运行时状态，传播数据层面元素报告的拓扑信息以及将无状态配置推送到转发引擎。

有时，将控制层面描述为网络信令。如果要处置处理消息以便将数据层面保持静态用户配置，则需要位于控制层面（例如，控制层面负责响应虚拟机 (VM) 的 vMotion，而管理层面负责将虚拟机连接到逻辑网络）。通常，控制层面作为在数据层面元素之间映射拓扑信息的反射器，例如，VTEP 的 MAC/隧道映射。在其他情况下，控制层面处理从某些数据层面元素中收到的数据以配置（或重新配置）某些数据层面元素，例如，使用 VIF 定位器计算和确定正确的隧道子集网格。

控制层面处理的对象集包括 VIF、逻辑网络、逻辑端口、逻辑路由器、IP 地址，等等。

控制层面在 NSX-T 中拆分成两个部分：中央控制层面 (Central Control Plane, CCP) 和本地控制层面 (Local Control Plane, LCP)，前者在 NSX Controller 群集节点上运行，后者在它控制的数据层面的相邻传输节点上运行。中央控制层面根据管理层面中的配置计算某种瞬间运行时状态，并通过本地控制层面传播数据层面元素报告的信息。本地控制层面监控本地链路状态，根据数据层面和 CCP 中的更新计算最新的运行时状态，并将无状态配置推送到转发引擎。LCP 与托管它的数据层面元素存在相同的风险。

# 管理层面

管理层面提供系统的单个 API 入口点，永久保留用户配置，处理用户查询，以及在系统中的所有管理、控制和数据层面节点上执行操作任务。

对于 NSX-T，管理层面负责处理查询、修改和永久保留用户配置，而控制层面负责将该配置向下传播到正确的数据层面元素子集。这意味着，某些数据属于多个层面，具体取决于它处于哪个阶段。管理层面还处理从控制层面中查询最近的状态和统计信息，有时直接从数据层面中进行查询。

管理层面是配置的（逻辑）系统的唯一真实数据源，这是用户通过配置管理的。可以使用 REST API 或 NSX-T UI 进行更改。

在 NSX 中，还会在所有群集和传输节点上运行管理层面代理 (Management Plane Agent, MPA)。引导配置就是一个示例用例，例如，中央管理节点地址凭据、软件包、统计信息和状态。MPA 可以相对独立于控制层面和数据层面运行，在其进程崩溃或停滞时可单独重新启动，但它们有时存在相同的风险，因为它们在相同的主机上运行。可以在本地和远程访问 MPA。MPA 在传输节点、控制节点和管理节点上运行以管理节点。在传输节点上，它还可以执行与数据层面相关的任务。

在管理层面上执行的任务包括：

- 永久保留配置（所需的逻辑状态）
- 输入验证
- 用户管理 - 角色分配
- 策略管理
- 后台任务跟踪

# NSX Manager

NSX Manager 提供了图形用户界面 (Graphical User Interface, GUI) 和 REST API 以创建、配置和监控 NSX-T 组件，例如，控制器、逻辑交换机和 Edge 服务网关。

NSX Manager 是 NSX-T 体系的管理层面。NSX Manager 提供了聚合系统视图并且是 NSX-T 的集中式网络管理组件。它提供了一种方法以监控连接到 NSX-T 创建的虚拟网络的工作负载以及进行故障排除。它提供了以下内容的配置和编排：

- 逻辑网络组件 - 逻辑交换和路由
- 网络和 Edge 服务
- 安全服务和分布式防火墙 - NSX Manager 的内置组件或集成的第三方供应商可以提供 Edge 服务和安全服务。

NSX Manager 允许无缝编排内置和外部服务。所有安全服务（无论是内置还是第三方服务）都是由 NSX-T 管理层面部署和配置的。管理层面提供了单个窗口以查看服务可用性。它还简化了基于策略的服务链、上下文共享和服务间事件处理。这会简化安全状态审核，简化了应用基于身份的控制（例如，AD 和移动性配置文件）。

NSX Manager 还提供 REST API 入口点以供自动化使用。这种灵活的架构允许通过任何云管理平台、安全供应商平台或自动化框架自动完成所有配置和监控操作。

NSX-T 管理层面代理 (MPA) 是位于每个和所有节点（管理程序）上的 NSX Manager 组件。MPA 负责永久保留所需的系统状态，以及在传输节点和管理层面之间传送非流量控制 (Non-Flow-Controlling, NFC) 消息，例如，配置、统计信息、状态和实时数据。

# NSX Controller

NSX Controller 是一种高级分布式状态管理系统，它控制虚拟网络和覆盖网络传输隧道。

NSX Controller 部署为一组高可用性的虚拟设备，它们负责在整个 NSX-T 架构中以编程方式部署虚拟网络。NSX-T 中央控制层面 (CCP) 在逻辑上与所有数据层面流量隔离，这意味着，控制层面中的任何故障不会影响现有的数据层面操作。流量不通过控制器传输；控制器负责为其他 NSX Controller 组件提供配置（如逻辑交换机、逻辑路由器和 Edge 配置）。数据传输的稳定性和可靠性是网络的核心问题。为了进一步提高高可用性和可扩展性，将在包含三个实例的群集中部署 NSX Controller。

## 逻辑交换机

NSX Edge 平台中的逻辑交换功能可以启动隔离的逻辑 L2 网络并提供虚拟机具有的相同灵活性和敏捷性。

虚拟数据中心的云部署具有跨多个租户的多种应用程序。出于安全和故障隔离的目的以及避免重叠的 IP 寻址问题，这些应用程序和租户需要互相隔离。端点（虚拟和物理）可以连接到逻辑分段并建立连接，而与它们在数据中心网络中的物理位置无关。这是通过将网络基础结构与 NSX-T 网络虚拟化提供的逻辑网络分离（即，将底层网络与覆盖网络分离）实现的。

逻辑交换机为第 2 层交换连接的多个主机提供了第 3 层 IP 访问能力。如果打算将某些逻辑网络限制为一组有限的主机，或者具有自定义连接要求，您可能会发现需要创建额外的逻辑交换机。

## 逻辑路由器

NSX-T 逻辑路由器提供南北向连接以允许租户访问公用网络，并在这些相同租户中的不同网络之间提供东西向连接。

逻辑路由器是为传统网络硬件路由器配置的一个分区。它复制硬件的功能，可在单个路由器中创建多个路由域。逻辑路由器执行物理路由器可以处理的一部分任务，每个逻辑路由器可以包含多个路由实例和路由表。使用逻辑路由器是一种有效的方法以最大限度提高路由器使用率，因为单个物理路由器中的一组逻辑路由器可以执行以前由很多设备执行的操作。

通过使用 NSX-T，可以创建两层逻辑路由器拓扑：顶层逻辑路由器是第 0 层，底层逻辑路由器是第 1 层。这种结构允许提供商管理员和租户管理员完全控制其服务和策略。提供商管理员控制和配置第 0 层路由和服务，租户管理员控制和配置第 1 层。在物理网络的第 0 层接口的北端，可以配置动态路由协议以便与物理路由器交换路由信息。第 0 层的南端连接到多个第 1 层路由层，并从这些层中接收路由信息。为了优化资源使用率，第 0 层不会将来自物理网络的所有路由推送到第 1 层，而是提供默认信息。

南向第 1 层路由层提供与租户管理员定义的逻辑交换机的接口，并在这些交换机之间提供单跃点路由功能。要从物理网络中访问第 1 层连接的子网，必须启用到第 0 层的路由重新分发。不过，不会在第 1 层和第 0 层之间运行传统路由协议（如 OSPF 或 BGP），所有路由将通过 NSX-T 控制层面。请注意，两层路由拓扑不是强制性的，如果不需要隔离提供商和租户，则可以创建单层拓扑，在这种情况下，逻辑交换机直接连接到第 0 层，而没有第 1 层。

逻辑路由器由两个可选的部分组成：一个分布式路由器 (Distributed Router, DR) 和一个或多个服务路由器 (Service Router, SR)。

DR 将跨虚拟机连接到该逻辑路由器的管理程序以及该逻辑路由器绑定到的 Edge 节点。从功能上讲，DR 负责连接到该逻辑路由器的逻辑交换机和/或逻辑路由器之间的单跃点分布式路由。SR 负责提供当前未以分布式方式实现的服务（如有状态 NAT）。

逻辑路由器始终具有 DR；如果满足任何以下条件，则还具有 SR：

- 逻辑路由器是第 0 层路由器，即使未配置有状态服务
- 逻辑路由器是链接到第 0 层路由器的第 1 层路由器，并且配置了没有分布式实现的服务（如 NAT、LB、DHCP）。

NSX-T 管理层面 (Management Plane, MP) 负责自动创建将服务路由器连接到分布式路由器的结构。MP 创建一个中转逻辑交换机并为其分配一个 VNI，然后在每个 SR 和 DR 上创建一个端口，将它们连接到中转逻辑交换机，并为 SR 和 DR 分配 IP 地址。

# NSX Edge

NSX Edge 为 NSX-T 部署外部的网络提供路由服务和连接。

通过使用 NSX Edge，位于不同子网中的同一主机上的虚拟机或工作负载可以相互通信，而无需通过传统路由接口。

需要使用 NSX Edge 以从 NSX-T 域建立外部连接（通过第 0 层路由器并经由 BGP 或静态路由）。此外，如果需要在第 0 层或第 1 层逻辑路由器中使用网络地址转换 (NAT) 服务，则必须部署 NSX Edge。

NSX Edge 提供了常见的网关服务（如 NAT）和动态路由以将隔离的末端网络连接到共享（上行链路）网络。DMZ 和多租户云环境中包含常见的 NSX Edge 部署，其中 NSX Edge 为每个租户创建虚拟边界。

# 传输区域

传输区域控制逻辑交换机可以访问的主机。它可以跨一个或多个主机群集。传输区域确定哪些主机可以参与使用特定的网络，进而确定哪些虚拟机可以参与使用该网络。

传输区域定义了一组可以通过物理网络基础结构相互通信的主机。该通信是通过一个或多个定义为虚拟隧道端点 (Virtual Tunnel Endpoint, VTEP) 的接口完成的。

如果两个传输节点位于相同的传输区域中，在这些传输节点上托管的虚拟机可以“看到”也位于该传输区域中的 NSX-T 逻辑交换机，从而可以连接到这些逻辑交换机。虚拟机可以通过该连接相互通信，并假定虚拟机具有第 2 层/第 3 层可访问性。如果虚拟机连接到位于不同传输区域中的交换机，则虚拟机无法相互通信。传输区域没有取代第 2 层/第 3 层可访问性要求，而是对该可访问性施加了一个限制。换句话说，属于同一传输区域是连接的一个必备条件。在满足该必备条件后，可以进行访问，但不会自动进行。要实现实际可访问性，第 2 层和（对于不同的子网）第 3 层网络必须正常运行。

如果某个节点包含至少一个主机交换机，则可以将其作为传输节点。在创建主机传输节点并随后将该节点添加到传输区域时，NSX-T 将在主机上安装一个主机交换机。对于主机所属的每个传输区域，将安装单独的主机交换机。主机交换机用于将虚拟机连接到 NSX-T 逻辑交换机以及创建 NSX-T 逻辑路由器上行链路和下行链路。

# 重要概念

在文档和用户界面中使用的常见 NSX-T 概念。

| | |
|---|---|
| **控制层面** | 根据管理层面中的配置计算运行时状态。控制层面传播数据层面元素报告的拓扑信息，并将无状态配置推送到转发引擎。 |
| **数据层面** | 根据控制层面填充的表执行无状态数据包转发或转换。数据层面向控制层面报告拓扑信息以及维护数据包级别统计信息。 |
| **外部网络** | 不是由 NSX-T 管理的物理网络或 VLAN。您可以通过 NSX Edge 将逻辑网络或覆盖网络链接到外部网络。例如，客户数据中心的物理网络或物理环境中的 VLAN。 |
| **架构节点** | 已在 NSX-T 管理层面中注册并安装了 NSX-T 模块的节点。要使管理程序主机或 NSX Edge 成为 NSX-T 覆盖网络的一部分，必须将该主机添加到 NSX-T 架构中。 |
| **架构配置文件** | 表示可以与 NSX Edge 群集关联的特定配置。例如，架构配置文件可能包含隧道属性以检测失效的对等项。 |
| **逻辑端口输出** | 到虚拟机或逻辑网络的入站网络流量称为输出，因为流量离开数据中心网络并进入虚拟空间。 |
| **逻辑端口输入** | 从虚拟机到数据中心网络的出站网络流量称为输入，因为流量进入物理网络。 |
| **逻辑路由器** | NSX-T 路由实体。 |
| **逻辑路由器端口** | 可以将逻辑交换机端口或物理网络的上行链路端口连接到的逻辑网络端口。 |
| **逻辑交换机** | 为虚拟机接口和网关接口提供虚拟第 2 层交换的 API 实体。逻辑交换机为租户网络管理员提供物理第 2 层交换机的逻辑等效项，从而允许他们将一组虚拟机连接到一个通用广播域。逻辑交换机是一个独立于物理管理程序基础架构的逻辑实体并跨很多管理程序，从而连接虚拟机而不考虑它们所在的物理位置。这样，就可以迁移虚拟机，而不要求租户网络管理员进行重新配置。<br><br>在多租户云中，很多逻辑交换机可能在同一管理程序硬件上并列存在，并且每个第 2 层分段与其他分段隔离。可以使用逻辑路由器连接逻辑交换机，逻辑路由器可以提供连接到外部物理网络的上行链路端口。 |
| **逻辑交换机端口** | 用于建立到虚拟机网络接口或逻辑路由器接口的连接的逻辑交换机连接点。逻辑交换机端口报告应用的交换配置文件、端口状态和链路状态。 |
| **管理层面** | 提供系统的单个 API 入口点，永久保留用户配置，处理用户查询以及在系统中的所有管理、控制和数据层面节点上执行操作任务。管理层面还负责查询、修改和永久保留用户配置。 |
| **NSX Controller 群集** | 部署为一组高可用性的虚拟设备，它们负责在整个 NSX-T 架构中以编程方式部署虚拟网络。 |
| **NSX Edge 群集** | 具有与高可用性监控中涉及的协议相同的设置的 NSX Edge 节点设备集合。 |

| | |
|---|---|
| **NSX Edge 节点** | 功能目标是提供计算能力以提供 IP 路由和 IP 服务功能的组件。 |
| **NSX-T 主机交换机或 KVM Open vSwitch** | 在管理程序上运行并提供物理流量转发的软件。主机交换机或 OVS 对租户网络管理员不可见，并提供每个逻辑交换机依赖的底层转发服务。要实现网络虚拟化，网络控制器必须为管理程序主机交换机配置网络流量表，它们构成租户管理员在创建和配置其逻辑交换机时定义的逻辑广播域。<br><br>每个逻辑广播域是使用隧道封装机制 Geneve 通过隧道传输虚拟机之间的流量以及虚拟机到逻辑路由器的流量实现的。网络控制器具有数据中心的全局视图，并确保在创建、移动或移除虚拟机时更新管理程序主机交换机流量表。 |
| **NSX Manager** | 托管 API 服务、管理层面和代理服务的节点。 |
| **Open vSwitch (OVS)** | 作为 XenServer、Xen、KVM 和其他基于 Linux 的管理程序中的管理程序主机交换机的开源软件交换机。NSX Edge 交换组件基于 OVS。 |
| **覆盖逻辑网络** | 使用“第 3 层中的第 2 层”隧道实现的逻辑网络，将虚拟机看到的拓扑从物理网络中解耦出来。 |
| **物理接口 (pNIC)** | 在其中安装管理程序的物理服务器上的网络接口。 |
| **第 0 层逻辑路由器** | 提供商逻辑路由器也称为物理网络的第 0 层逻辑路由器接口。第 0 层逻辑路由器是顶层路由器，可以实现为活动-活动或活动-备用服务路由器群集。该逻辑路由器运行 BGP 并作为物理路由器的对等项。在活动-备用模式下，该逻辑路由器还可以提供有状态服务。 |
| **第 1 层逻辑路由器** | 第 1 层逻辑路由器是第二层路由器，它连接到一个第 0 层逻辑路由器以建立北向连接，并连接到一个或多个覆盖网络以建立南向连接。第 1 层逻辑路由器可以是提供有状态服务的活动-备用服务路由器群集。 |
| **传输区域** | 定义逻辑交换机的最大范围的传输节点集合。传输区域表示一组以类似方式置备的管理程序以及连接这些管理程序上的虚拟机的逻辑交换机。NSX-T 可以将所需的支持软件包部署到主机中，因为它知道在逻辑交换机上启用了哪些功能。 |
| **虚拟机接口 (vNIC)** | 虚拟机上的网络接口，它在虚拟客户机操作系统和标准 vSwitch 或 vSphere Distributed Switch 之间提供连接。可以将 vNIC 连接到一个逻辑端口。您可以根据其唯一 ID (UUID) 识别 vNIC。 |
| **VTEP** | 虚拟隧道端点。管理程序主机可以通过隧道端点加入 NSX-T 覆盖网络。NSX-T 覆盖网络将帧封装到数据包中并通过底层传输网络传输数据包，从而在现有的第 3 层网络架构上部署第 2 层网络。底层传输网络可以是另一个第 2 层网络，也可以跨第 3 层边界。VTEP 是进行封装和解封的连接点。 |

# 创建逻辑交换机和配置虚拟机连接

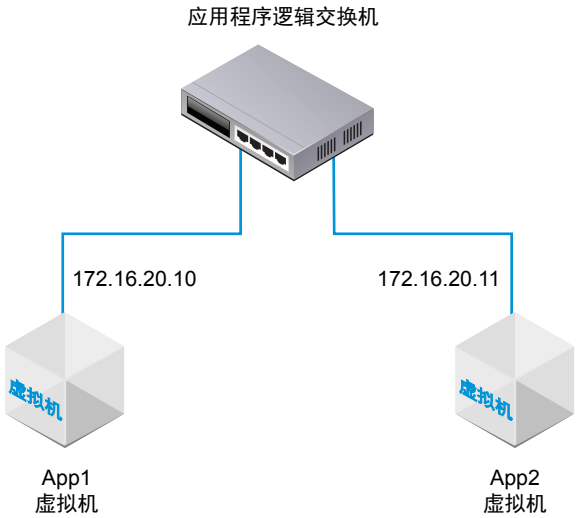
2

NSX-T 逻辑交换机在完全脱离底层硬件的虚拟环境中再现交换功能（广播、未知单播、多播 (BUM) 流量）。

逻辑交换机在提供可连接虚拟机的网络连接方式上类似于 VLAN。如果虚拟机连接到同一逻辑交换机，则虚拟机可以通过管理程序之间的隧道相互通信。每个逻辑交换机具有一个虚拟网络标识符 (Virtual Network Identifier, VNI)，与 VLAN ID 类似。与 VLAN 不同的是，VNI 远远超出 VLAN ID 的限制。

在添加逻辑交换机时，请务必规划要构建的拓扑。

**图 2-1. 逻辑交换机拓扑**




例如，拓扑显示连接到两个虚拟机的单个逻辑交换机。两个虚拟机可以位于不同主机群集或同一主机群集中的不同主机或同一主机上。由于示例中的虚拟机位于同一虚拟网络上，因此，在虚拟机上配置的基础 IP 地址必须位于同一子网中。

本章讨论了以下主题：

- 了解 BUM 帧复制模式
- 创建逻辑交换机
- 第 2 层桥接
- 为 NSX Edge 上行链路创建 VLAN 逻辑交换机
- 将虚拟机连接到逻辑交换机

# 了解 BUM 帧复制模式

每个主机传输节点是一个隧道端点。每个隧道端点具有一个 IP 地址。这些 IP 地址可以位于同一子网中，也可以位于不同的子网中，具体取决于传输节点的 IP 池或 DHCP 配置。

在不同主机上的两个虚拟机直接通信时，将在与两个管理程序关联的两个隧道端点 IP 地址之间传输单播封装流量，而无需进行泛洪。

不过，与任何第 2 层网络一样，虚拟机发出的流量有时需要进行泛洪，这意味着需要将其发送到属于同一逻辑交换机的所有其他虚拟机。第 2 层广播、未知单播和多播流量（BUM 流量）就属于这种情况。回想一下，单个 NSX-T 逻辑交换机可以跨多个管理程序。需要将给定管理程序上的虚拟机发出的 BUM 流量复制到远程管理程序，这些管理程序托管连接到同一逻辑交换机的其他虚拟机。要启用这种泛洪，NSX-T 支持两种不同的复制模式：

- 分层双层（有时称为 MTEP）
- 头（有时称为源）

以下示例说明了分层双层复制模式。假设主机 A 具有连接到虚拟网络标识符 (VNI) 5000、5001 和 5002 的虚拟机。请将 VNI 视为与 VLAN 类似，但每个逻辑交换机具有单个关联的 VNI。因此，有时可以将术语 VNI 和逻辑交换机换用。在我们说到主机位于 VNI 上时，我们的意思是主机的虚拟机连接到具有该 VNI 的逻辑交换机。

隧道端点表显示主机到 VNI 的连接。主机 A 检查 VNI 5000 的隧道端点表，并确定 VNI 5000 上的其他主机的隧道端点 IP 地址。

其中的一些 VNI 连接位于与主机 A 上的隧道端点相同的 IP 子网（也称为 IP 分段）上。对于其中的每个连接，主机 A 创建每个 BUM 帧的单独副本，并将该副本直接发送到每个主机。

其他主机的隧道端点位于不同的子网或 IP 分段上。对于具有多个隧道端点的每个分段，主机 A 将其中的一个端点提名为复制程序。

复制程序从主机 A 中接收 VNI 5000 的每个 BUM 帧的一个副本。该副本在封装标头中标记为本地复制。主机 A 不会将副本发送到与复制程序相同的 IP 分段中的其他主机。由复制程序负责为它了解的每个主机（位于 VNI 5000 上与复制程序主机相同的 IP 分段中）创建 BUM 帧副本。

将为 VNI 5001 和 5002 重复该过程。对于不同的 VNI，隧道端点列表和产生的复制程序可能是不同的。

头复制也称为头端复制，这种复制没有复制程序。主机 A 直接为 VNI 5000 上它了解的每个隧道端点创建每个 BUM 帧的副本并发送该副本。

如果所有主机隧道端点位于同一子网上，选择复制模式不会产生任何差异，因为这些行为是相同的。如果主机隧道端点位于不同的子网上，分层双层复制有助于在多个主机之间分摊负载。分层双层是默认模式。

# 创建逻辑交换机

逻辑交换机连接到网络中的一个或多个虚拟机。连接到逻辑交换机的虚拟机可以使用管理程序之间的隧道互相通信。

**前提条件**

- 确认配置了一个传输区域。请参阅《 NSX-T 安装指南》。
- 确认架构节点已成功连接到 NSX-T 管理层面代理 (MPA) 和 NSX-T 本地控制层面 (LCP)。

  在 `GET https://<nsx-mgr>/api/v1/transport-nodes/<transport-node-id>/state` API 调用中，`state` 必须为 `success`。请参阅《 NSX-T 安装指南》。
- 确认传输节点已添加到传输区域中。请参阅《 NSX-T 安装指南》。
- 确认管理程序已添加到 NSX-T 架构中，并且在这些管理程序上托管了虚拟机。
- 熟悉逻辑交换机拓扑和 BUM 帧复制概念。请参阅第 2 章，创建逻辑交换机和配置虚拟机连接和了解 BUM 帧复制模式。
- 确认 NSX Controller 群集处于稳定状态。

**步骤**

1. 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。
2. 选择**交换 > 交换机 (Switching > Switches)**。
3. 单击**添加 (Add)**。
4. 指定逻辑交换机的名称。
5. 为逻辑交换机选择一个传输区域。

   连接到位于同一传输区域中的逻辑交换机的虚拟机可以相互通信。
6. 为逻辑交换机择一种复制模式。

   覆盖网络逻辑交换机需要使用复制模式（分层双层或头），但基于 VLAN 的逻辑交换机不需要使用复制模式。

| 复制模式 | 说明 |
| --- | --- |
| 分层双层 | 复制程序是一个主机，它将 BUM 流量复制到同一 VNI 中的其他主机。<br>每个主机将每个 VNI 中的一个主机隧道端点提名为复制程序。将为每个 VNI 完成该操作。 |
| 头 | 主机创建每个 BUM 帧的副本，并将该副本发送到每个 VNI 中它了解的每个隧道端点。 |

7. （可选）单击**交换配置文件 (Switching Profiles)**选项卡，然后选择交换配置文件。
8. 单击**保存 (Save)**。

   在 NSX Manager UI 中，新逻辑交换机是一个可单击的链接。

**后续步骤**

将虚拟机连接到逻辑交换机。请参阅将虚拟机连接到逻辑交换机。

# 第 2 层桥接

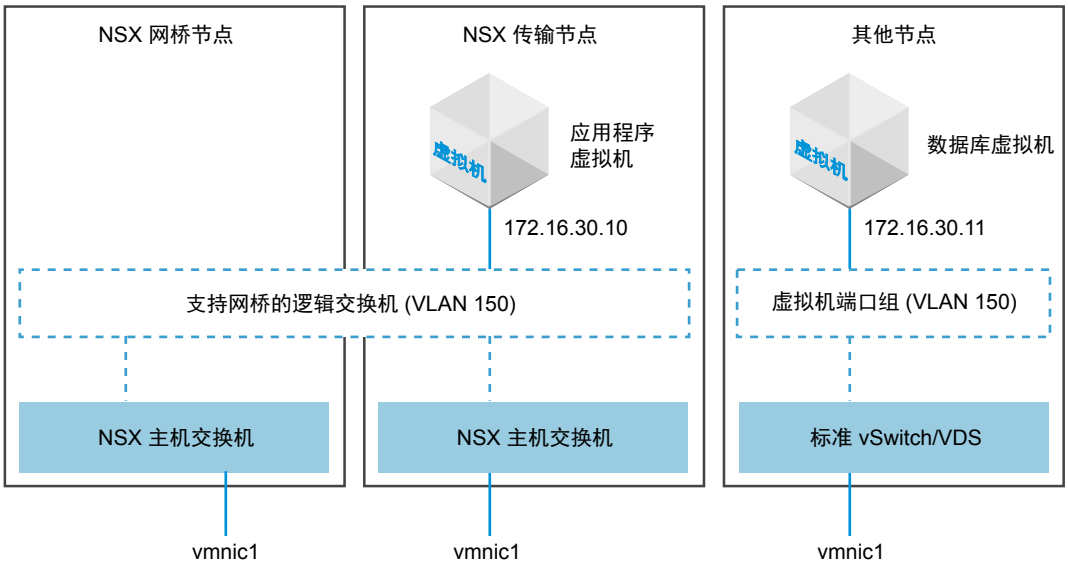
如果 NSX-T 逻辑交换机需要建立到支持 VLAN 的端口组的第 2 层连接，或者需要访问位于 NSX-T 部署外部的其他设备（如网关），您可以使用 NSX-T 第 2 层网桥。这在迁移情况下特别有用，此时，您需要在物理和虚拟工作负载之间拆分子网。

第 2 层桥接中涉及的 NSX-T 概念是网桥群集、网桥端点和网桥节点。网桥群集是高可用性 (High Availability, HA) 网桥节点集合。网桥节点是进行桥接的传输节点。用于桥接虚拟和物理部署的每个逻辑交换机具有关联的 VLAN ID。网桥端点确定网桥的物理属性，例如，网桥群集 ID 和关联的 VLAN ID。

在该 NSX-T 版本中，第 2 层桥接是由用作网桥节点的 ESXi 主机提供的。网桥节点是已添加到网桥群集的 ESXi 主机传输节点。

在以下示例中，两个 NSX-T 传输节点是同一覆盖网络传输区域的一部分。这样，就可以将其 NSX-T 主机交换机（有时称为 NSX-T vSwitch，如图中所示）连接到支持网桥的同一逻辑交换机。

**图 2-2. 网桥拓扑**




左侧的传输节点属于一个网桥群集，因此，它是一个网桥节点。

由于逻辑交换机连接到一个网桥群集，因此，它称为支持网桥的逻辑交换机。要能够支持网桥，逻辑交换机必须位于覆盖网络传输区域中，而不能位于 VLAN 传输区域中。

中间传输节点不是网桥群集的一部分。它是一个常规传输节点。它可以是 KVM 或 ESXi 主机。在该图中，该节点上称为“应用程序虚拟机”的虚拟机连接到支持网桥的逻辑交换机。

右侧的节点不是 NSX-T 覆盖网络的一部分。它可能是具有虚拟机的任何管理程序（如图中所示），也可能是物理网络节点。如果非 NSX-T 节点是 ESXi 主机，您可以使用标准 vSwitch 或 vSphere Distributed Switch 进行端口连接。一个要求是，与端口连接关联的 VLAN ID 必须与支持网桥的逻辑交换机上的 VLAN ID 相匹配。此外，还会在第 2 层上进行通信，因此，两个终端设备的 IP 地址必须位于同一子网中。

如上所述，网桥的用途是在两个虚拟机之间启用第 2 层通信。在两个虚拟机之间传输流量时，流量将通过网桥节点。

## 创建网桥群集

网桥群集是一组进行桥接并参与高可用性 (HA) 的传输节点。每次只有一个传输节点处于活动状态。具有多节点 NSX-T 网桥节点群集可以帮助确保至少一个 NSX-T 网桥节点始终可用。要创建支持网桥的逻辑交换机，您必须将其与一个网桥群集相关联。因此，即使您只有一个网桥节点，它也必须属于网桥群集才能使用。

在创建网桥群集后，以后可以编辑该群集以添加额外的网桥节点。

**前提条件**

- 创建至少一个 NSX-T 传输节点以用作网桥节点。
- 用作网桥节点的传输节点必须是 ESXi 主机。网桥节点不支持 KVM。
- 建议不要在网桥节点上托管任何虚拟机。
- 只能将传输节点添加到一个网桥群集中。您无法将同一传输节点添加到多个网桥群集中。

**步骤**

1. 在 NSX Manager UI 中，导航到**架构 > 配置 > 网桥 (Fabric > Configuration > Bridges)**。
2. 指定网桥群集的名称。
3. 为网桥群集选择一个传输区域。

   该传输区必须具有“覆盖网络”类型而不是 VLAN。
4. 从**可用 (Available)**列中，选择传输节点并单击右箭头以将其移到**已选择 (Selected)**列中。

**后续步骤**

现在，您可以将逻辑交换机与网桥群集相关联。

## 创建支持网桥的第 2 层逻辑交换机

在将虚拟机连接到 NSX-T 覆盖网络时，您可能希望这些虚拟机与位于 NSX-T 部署外部的其他设备或虚拟机之间具有第 2 层连接。在这种情况下，您可以使用支持网桥的逻辑交换机。

有关示例拓扑，请参阅图 2-2。

**前提条件**

- 至少将一个 ESXi 主机用作网桥节点。网桥节点是仅进行桥接的 ESXi 传输节点。必须将该传输节点添加到网桥群集中。请参阅创建网桥群集。
- 至少将一个 ESXi 或 KVM 主机用作常规传输节点。该节点托管的虚拟机需要与 NSX-T 部署外部的设备进行连接。
- 在 NSX-T 部署外部具有一个虚拟机或其他终端设备。该终端设备必须连接到与支持网桥的逻辑交换机的 VLAN ID 匹配的 VLAN 端口。
- 将覆盖网络传输区域中的一个逻辑交换机用作支持网桥的逻辑交换机。

**步骤**

1 从浏览器中，登录到 `https://<nsx-mgr>` 中的 NSX Manager。

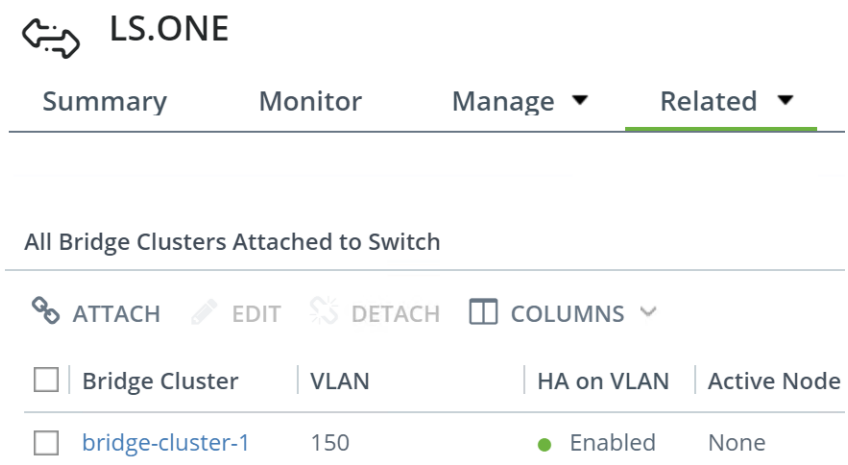
2 选择**交换 > 交换机 (Switching > Switches)**。

3 从交换机列表中，选择一个覆盖网络交换机（流量类型：覆盖网络）。

4 在交换机配置页面上，选择**相关设置 > 网桥群集 (Related > Bridge Clusters)**。

5 单击**连接 (ATTACH)**，选择一个网桥群集，然后输入一个 VLAN ID。

例如：




6 如果尚未将虚拟机连接到逻辑交换机，请连接虚拟机。

这些虚拟机必须位于与网桥群集相同的传输区域中的传输节点上。

您可以将 ping 命令从 NSX-T 内部虚拟机发送到 NSX-T 外部的节点以测试网桥是否正常工作。例如，在图 2-2 中，NSX-T 传输节点上的应用程序虚拟机应该能够 ping 通外部节点上的数据库虚拟机，反之亦然。

您可以导航到**交换 > 交换机 > 监控 (Switching > Switches > Monitor)**以监控网桥交换机上的流量。

您可以使用 `GET https://192.168.110.31/api/v1/bridge-endpoints/<endpoint-id>/statistics` API 调用查看网桥流量：

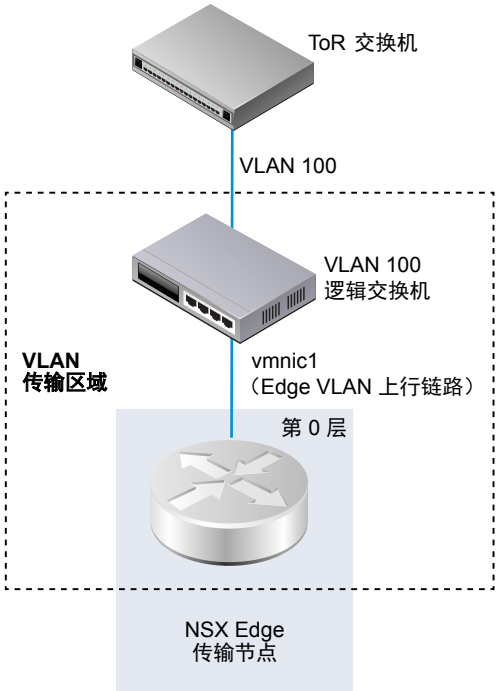

```
{
  "tx_packets": {
    "total": 134416,
    "dropped": 0,
    "multicast_broadcast": 0
  },
  "rx_bytes": {
    "total": 22164,
    "multicast_broadcast": 0
  },
  "tx_bytes": {
    "total": 8610134,
    "multicast_broadcast": 0
  },
  "rx_packets": {
```

```
    "total": 230,
    "dropped": 0,
    "multicast_broadcast": 0
  },
  "last_update_timestamp": 1454979822860,
  "endpoint_id": "ba5ba59d-22f1-4a02-b6a0-18ef0e37ef31"
}
```

# 为 NSX Edge 上行链路创建 VLAN 逻辑交换机

Edge 上行链路通过 VLAN 逻辑交换机连接到外部。

在创建 VLAN 逻辑交换机时，请务必记住要构建的特定拓扑。例如，以下简单拓扑显示 VLAN 传输区域中的单个 VLAN 逻辑交换机。该 VLAN 逻辑交换机具有 VLAN ID 100。这与用于 Edge 的 VLAN 上行链路的管理程序主机端口连接的 ToR 端口上的 VLAN ID 匹配。




**前提条件**

- 要创建 VLAN 逻辑交换机，必须先创建一个 VLAN 传输区域。
- 必须将一个 NSX-T vSwitch 添加到 NSX Edge 中。要在 Edge 上进行确认，请运行 `get host-switch` 命令。例如：

```
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
```

- 确认 NSX Controller 群集处于稳定状态。
- 确认架构节点已成功连接到 NSX-T 管理层面代理 (MPA) 和 NSX-T 本地控制层面 (LCP)。

  在 `GET https://<nsx-mgr>/api/v1/transport-nodes/<transport-node-id>/state` API 调用中，`state` 必须为 `success`。请参阅《 NSX-T 安装指南》。

**步骤**

1 从浏览器中，登录到 `https://<nsx-mgr>` 中的 NSX Manager。

2 选择**交换 > 交换机 (Switching > Switches)**。

3 单击**添加 (Add)**。

4 键入逻辑交换机的名称。

5 为逻辑交换机选择一个传输区域。

  在选择 VLAN 传输区域时，将显示 VLAN ID 字段。

6 键入 VLAN ID。

  如果没有到物理 ToR 的上行链路的 VLAN ID，请在 VLAN 字段中输入 0。

7 （可选）单击**交换配置文件 (Switching Profiles)**选项卡，然后选择交换配置文件。

**后续步骤**

添加逻辑路由器。

# 将虚拟机连接到逻辑交换机

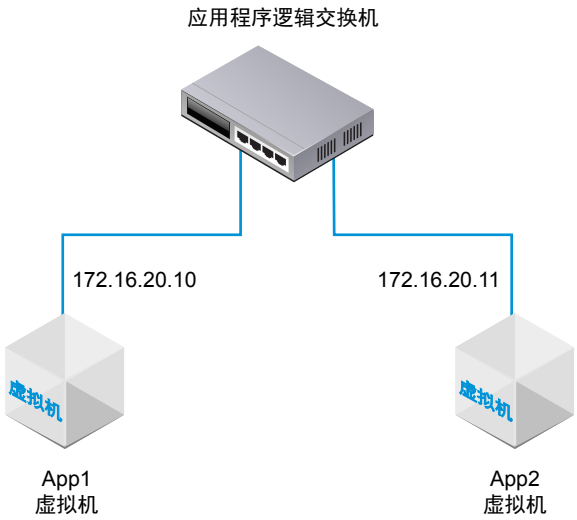
根据您的主机，将虚拟机连接到逻辑交换机的配置可能会有所不同。

可以连接到逻辑交换机的支持的主机是：在 vCenter Server 中管理的 ESXi 主机、单独的 ESXi 主机以及 KVM 主机。

## 将 vCenter Server 上托管的虚拟机连接到 NSX-T 逻辑交换机

如果在 vCenter Server 中管理某个 ESXi 主机，您可以通过基于 Web 的 vSphere Web Client 访问主机虚拟机。在这种情况下，您可以使用该过程将虚拟机连接到 NSX-T 逻辑交换机。

该过程中显示的示例说明了如何将名为 app-vm 的虚拟机连接到名为 app-switch 的逻辑交换机。




基于安装的 vSphere Client 应用程序不支持将虚拟机连接到 NSX-T 逻辑交换机。如果您没有（基于 Web 的）vSphere Web Client，请参阅将单独 ESXi 上托管的虚拟机连接到 NSX-T 逻辑交换机。

**前提条件**

- 必须在已添加到 NSX-T 结构层的管理程序上托管虚拟机。
- 结构层节点必须具有 NSX-T 管理层面 (MPA) 和 NSX-T 控制层面 (LCP) 连接。
- 必须将结构层节点添加到传输区域中。
- 必须创建一个逻辑交换机。

**步骤**

1 在 vSphere Web Client 中，编辑虚拟机设置并将虚拟机连接到 NSX-T 逻辑交换机。

例如：


2 单击**确定**。

在将虚拟机连接到逻辑交换机后，逻辑交换机端口将添加到逻辑交换机中。您可以在 NSX Manager 的**交换 > 端口**中查看逻辑交换机端口。

在 NSX-T API 中，您可以使用 `GET https://<nsx-mgr>/api/v1/fabric/virtual-machines` API 调用查看 NSX-T 连接的虚拟机。

在 NSX-T Manager UI 中的**交换 > 端口**下面，VIF 连接 ID 与在 API 调用中找到的 externalId 相匹配。找到与虚拟机的 externalId 匹配的 VIF 连接 ID，并确保管理和运行状态为“已连接/已连接”。

如果两个虚拟机连接到同一逻辑交换机并在同一子网中配置了 IP 地址，则它们应该可以 ping 通对方。

**后续步骤**

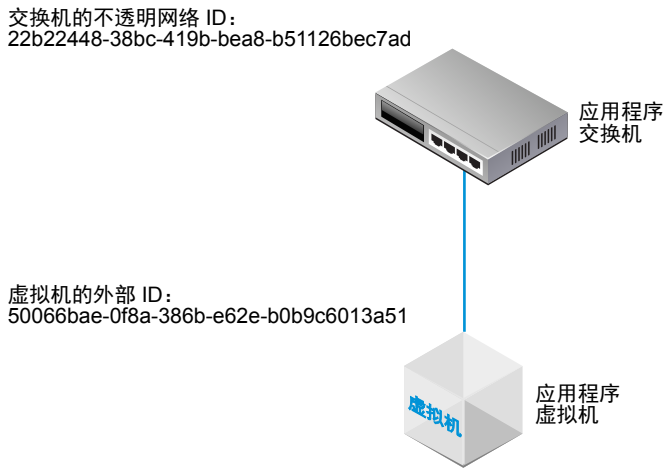
添加逻辑路由器。

您可以监控逻辑交换机端口上的活动以解决问题。请参阅监控逻辑交换机端口活动。

## 将单独 ESXi 上托管的虚拟机连接到 NSX-T 逻辑交换机

如果具有单独的 ESXi 主机，您无法通过基于 Web 的 vSphere Web Client 访问主机虚拟机。在这种情况下，您可以使用该过程将虚拟机连接到 NSX-T 逻辑交换机。

该过程中显示的示例说明了如何将名为 app-vm 的虚拟机连接到名为 app-switch 的逻辑交换机。




**前提条件**

- 必须在已添加到 NSX-T 架构的管理程序上托管虚拟机。
- 架构节点必须具有 NSX-T 管理层面 (MPA) 和 NSX-T 控制层面 (LCP) 连接。
- 必须将架构节点添加到传输区域中。
- 必须创建一个逻辑交换机。
- 您必须具有 NSX Manager API 的访问权限。
- 您必须具有虚拟机的 VMX 文件的写入访问权限。

### 步骤

1 通过使用（基于安装的）vSphere Client 应用程序或某种其他虚拟机管理工具，编辑虚拟机并添加一个 VMXNET 3 以太网适配器。

选择任何命名的网络。您将在后面的步骤中更改网络连接。


2 使用 NSX-T API 发出 `GET https://<nsx-mgr>/api/v1/fabric/virtual-machines/<VM-ID>` API 调用。

在结果中，找到虚拟机的 externalId。

例如：

```
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 关闭虚拟机电源并从主机中取消注册虚拟机。

您可以使用虚拟机管理工具或 ESXi CLI（如下所示）。

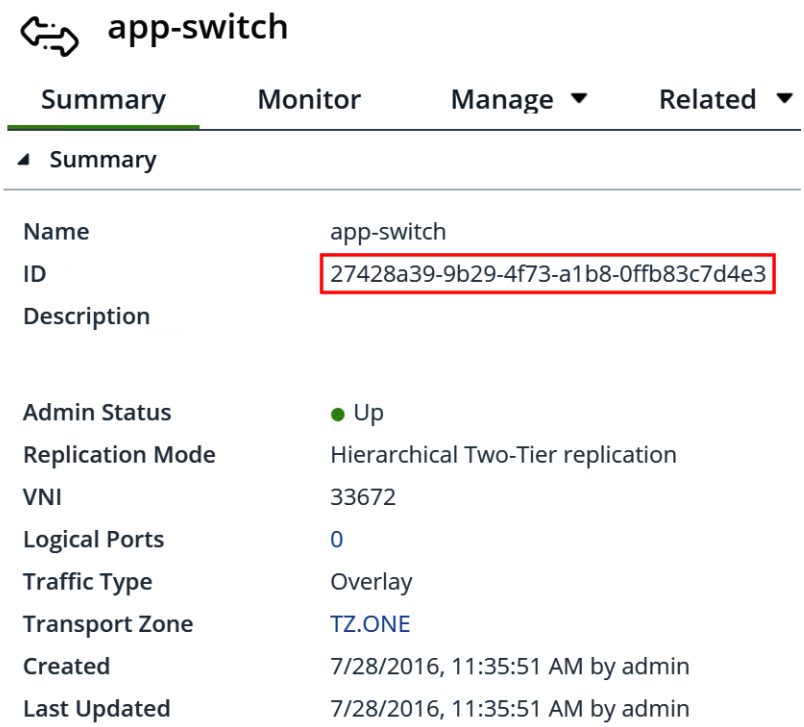
```
[user@host:~] vim-cmd /vmsvc/getallvms
Vmid    Name             File               Guest OS      Version   Annotation
5      app-vm    [ds2] app-vm/app-vm.vmx   ubuntuGuest     vmx-08
8      web-vm    [ds2] web-vm/web-vm.vmx   ubuntu64Guest   vmx-08

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

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

4 从 NSX Manager UI 中，获取逻辑交换机 ID。

例如：




5 修改虚拟机的 VMX 文件。

删除 **ethernet1.networkName = "<name>"** 字段并添加以下字段：

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

- ethernet1.startConnected = "TRUE"

例如：

**旧版本 (OLD)**

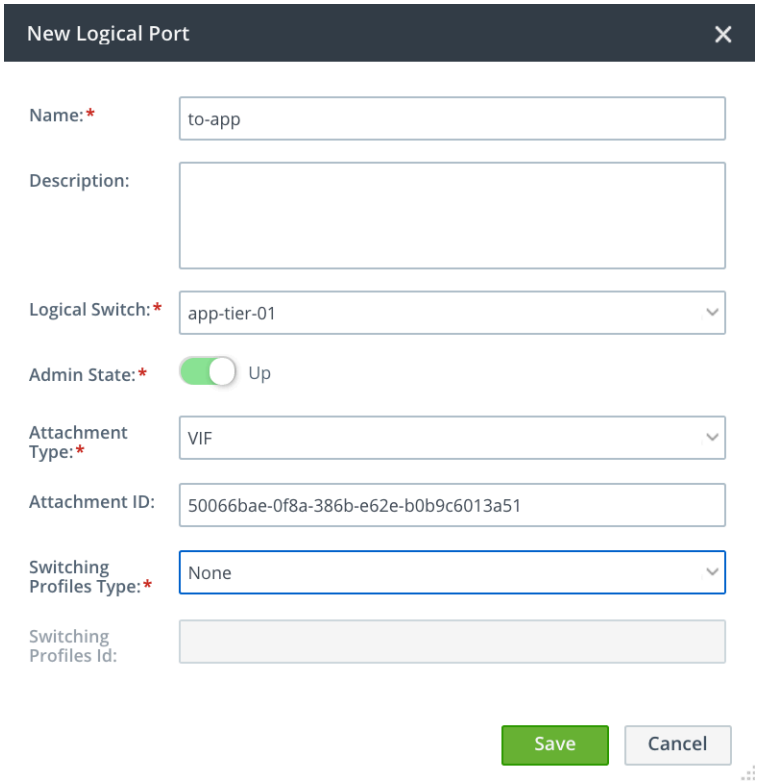
```
ethernet1.pciSlotNumber = "224"
ethernet1.virtualDev = "vmxnet3"
ethernet1.networkName = "VM Network"
ethernet1.addressType = "vpx"
ethernet1.generatedAddress = "00:50:56:86:7b:d7"
ethernet1.uptCompatibility = "true"
ethernet1.present = "TRUE"
```

**新版本 (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 在 NSX Manager UI 中，添加一个逻辑交换机端口，然后使用虚拟机的 externalId 进行 VIF 连接。

例如：




7 重新注册虚拟机，然后打开虚拟机电源。

您可以使用虚拟机管理工具或 ESXi CLI（如下所示）。

```
[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:
```

在 NSX Manager UI 中的**交换 > 端口 (Switching > Ports)**下面，找到与虚拟机的 externalId 匹配的 VIF 连接 ID，并确保管理和运行状态为“已连接/已连接”。

如果两个虚拟机连接到同一逻辑交换机并在同一子网中配置了 IP 地址，则它们应该可以 ping 通对方。

**后续步骤**

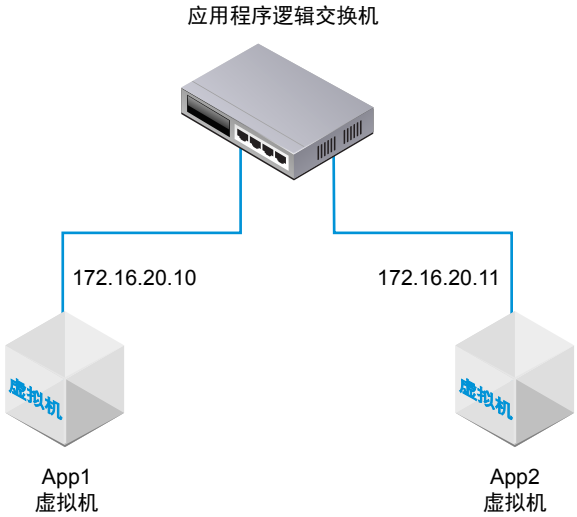
添加逻辑路由器。

您可以监控逻辑交换机端口上的活动以解决问题。请参阅监控逻辑交换机端口活动。

# 将 KVM 上托管的虚拟机连接到 NSX-T 逻辑交换机

如果具有 KVM 主机，您可以使用该过程将虚拟机连接到 NSX-T 逻辑交换机。

该过程中显示的示例说明了如何将名为 app-vm 的虚拟机连接到名为 app-switch 的逻辑交换机。




**前提条件**

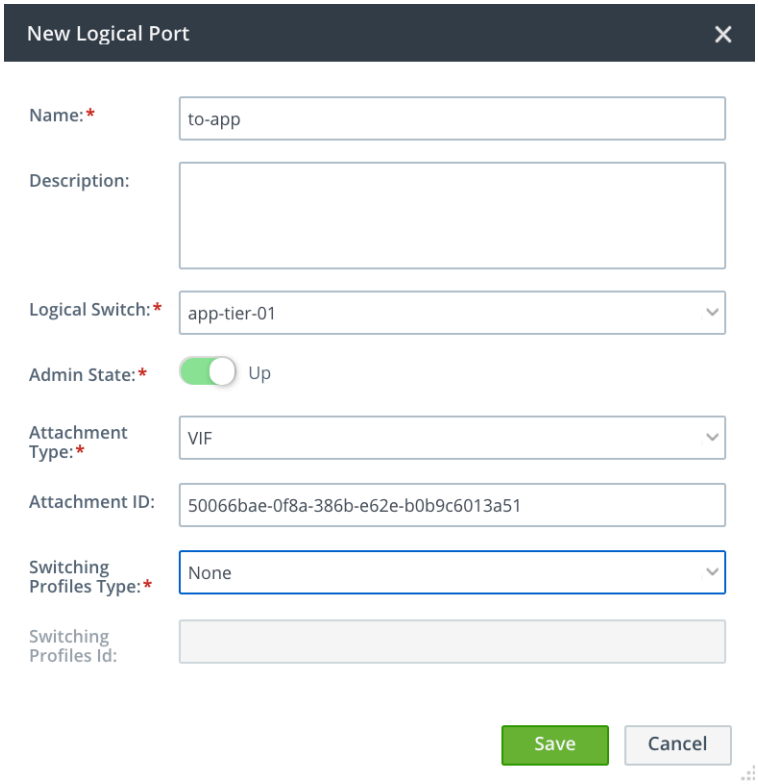
- 必须在已添加到 NSX-T 架构的管理程序上托管虚拟机。
- 架构节点必须具有 NSX-T 管理层面 (MPA) 和 NSX-T 控制层面 (LCP) 连接。
- 必须将架构节点添加到传输区域中。
- 必须创建一个逻辑交换机。

**步骤**

1 从 KVM CLI 中，运行 `virsh dumpxml <your vm> | grep interfaceid` 命令。

2 在 NSX Manager UI 中，添加一个逻辑交换机端口，然后使用虚拟机的接口 ID 进行 VIF 连接。

例如：




在 NSX Manager UI 中的**交换 > 端口 (Switching > Ports)**下面，找到 VIF 连接 ID 并确保管理和运行状态为“已连接/已连接”。

如果两个虚拟机连接到同一逻辑交换机并在同一子网中配置了 IP 地址，则它们应该可以 ping 通对方。

**后续步骤**

添加逻辑路由器。

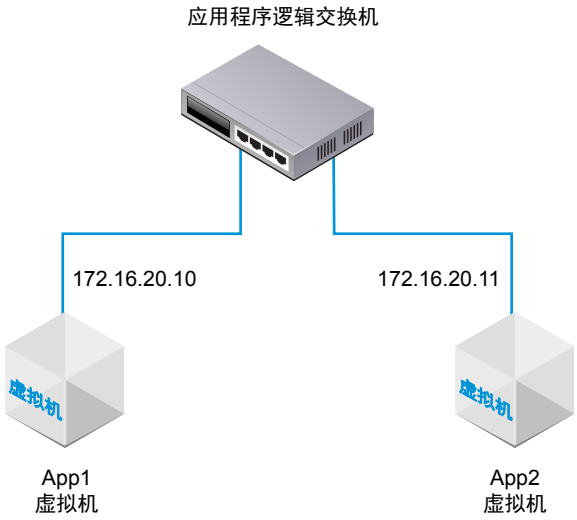
您可以监控逻辑交换机端口上的活动以解决问题。请参阅监控逻辑交换机端口活动。

# 测试第 2 层连接

在成功设置逻辑交换机并将虚拟机连接到逻辑交换机后，您可以测试连接的虚拟机的网络连接。

如果根据拓扑正确配置了您的网络环境，App2 虚拟机可以 ping 通 App1 虚拟机。

**图 2-3. 逻辑交换机拓扑**




**步骤**

1 使用 SSH 或虚拟机控制台登录到逻辑交换机连接的一个虚拟机。

例如，App2 虚拟机 172.16.20.11。

2 对连接到逻辑交换机的第二个虚拟机执行 ping 操作以测试连接。

```
$ 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 （可选）确定导致 ping 失败的问题。

a 验证虚拟机网络设置是否正确。

b 验证虚拟机网络适配器是否连接到正确的逻辑交换机。

c 验证逻辑交换机管理状态是否为“已连接”。

d 从 NSX Manager 中，选择**交换 > 交换机**。

e 单击逻辑交换机并记下 UUID 和 VNI 信息。

f 从 NSX Controller 中，运行以下命令以解决该问题。

| 命令 | 说明 |
| --- | --- |
| **get logical-switch <vni-or-uuid> arp-table** | 显示指定逻辑交换机的 ARP 表。<br>示例输出。<br>`nsx-controller1> get logical-switch 41866 arp-table`<br>`VNI      IP              MAC             Connection-ID`<br>`41866 172.16.20.11 00:50:56:b1:70:5e     295422` |
| **get logical-switch <vni-or-uuid> connection-table** | 显示指定逻辑交换机的连接。<br>示例输出。<br>`nsx-controller1> get logical-switch 41866 connection-table`<br>`Host-IP           Port   ID`<br>`192.168.110.37   36923 295420`<br>`192.168.210.53   37883 295421`<br>`192.168.210.54   57278 295422` |
| **get logical-switch <vni-or-uuid> mac-table** | 显示指定逻辑交换机的 MAC 表。<br>示例输出。<br>`nsx-controller1> get logical-switch 41866 mac-table`<br>`VNI     MAC               VTEP-IP         Connection-ID`<br>`41866 00:50:56:86:f2:b2 192.168.250.102   295421`<br>`41866 00:50:56:b1:70:5e 192.168.250.101   295422` |
| **get logical-switch <vni-or-uuid> stats** | 显示有关指定逻辑交换机的统计信息。<br>示例输出。<br>`nsx-controller1> get logical-switch 41866 stats`<br>`update.member 11`<br>`update.vtep 11`<br>`update.mac 4`<br>`update.mac.invalidate 0`<br>`update.arp 7`<br>`update.arp.duplicate 0`<br>`query.mac 2`<br>`query.mac.miss 0`<br>`query.arp 9`<br>`query.arp.miss 6` |
| **get logical-switch <vni-or-uuid> stats-sample** | 显示一段时间的所有逻辑交换机统计信息的摘要。<br>示例输出。<br>`nsx-controller1> get logical-switch 41866 stats-sample`<br>`21:00:00 21:10:00 21:20:00 21:30:00 21:40:00`<br>`update.member 0 0 0 0 0`<br>`update.vtep 0 0 0 0 0`<br>`update.mac 0 0 0 0 0`<br>`update.mac.invalidate 0 0 0 0 0`<br>`update.arp 0 0 0 0 0`<br>`update.arp.duplicate 0 0 0 0 0` |

| 命令 | 说明 |
| --- | --- |
| | query.mac 0 0 0 0 0<br>query.mac.miss 0 0 0 0 0<br>query.arp 0 0 0 0 0<br>query.arp.miss 0 0 0 0 0 |
| **get logical-switch <vni-or-uuid> vtep** | 显示与指定逻辑交换机相关的所有虚拟隧道端点。<br>示例输出。<br>nsx-controller1> get logical-switch 41866 vtep<br>VNI IP LABEL Segment MAC Connection-ID<br>41866 192.168.250.102 0x8801 192.168.250.0 00:50:56:65:f5:fc 295421<br>41866 192.168.250.100 0x1F801 192.168.250.0 02:50:56:00:00:00 295420<br>41866 192.168.250.101 0x16001 192.168.250.0 00:50:56:64:7c:28 295422 |

连接到逻辑交换机的第一个虚拟机可以将数据包发送到第二个虚拟机。

# 为逻辑交换机和逻辑端口配置交换配置文件

3

交换配置文件包括逻辑交换机和逻辑端口的第 2 层网络配置详细信息。NSX Manager 支持几种类型的交换配置文件，并为每种配置文件类型保留一个或多个系统定义的默认交换配置文件。

可以使用以下类型的交换配置文件。

- QoS（服务质量）
- 端口监控
- IP 发现
- SpoofGuard
- 交换机安全
- MAC 管理

注　您无法在 NSX Manager 中编辑或删除默认交换配置文件，但可以创建自定义交换配置文件。

每个默认或自定义交换配置文件具有唯一的保留标识符。您可以使用该标识符将交换配置文件与逻辑交换机或逻辑端口相关联。例如，默认 QoS 交换配置文件 ID 为 f313290b-eba8-4262-bd93-fab5026e9495。

可以将逻辑交换机或逻辑端口与每种类型的一个交换配置文件相关联。例如，您不能将两个不同的 QoS 交换配置文件与一个逻辑交换机或逻辑端口相关联。

如果在创建或更新逻辑交换机时未关联交换配置文件类型，则 NSX Manager 关联相应的默认系统定义交换配置文件。子逻辑端口从父逻辑交换机中继承默认系统定义的交换配置文件。

在创建或更新逻辑交换机或逻辑端口时，您可以选择关联默认或自定义交换配置文件。如果将交换配置文件与逻辑交换机关联或解除关联，将根据以下条件应用子逻辑端口的交换配置文件。

- 如果父逻辑交换机具有关联的配置文件，子逻辑端口将从父逻辑交换机中继承交换配置文件。
- 如果父逻辑交换机没有关联的交换配置文件，则为该逻辑交换机分配默认交换配置文件，并且该逻辑端口继承该默认交换配置文件。

- 如果明确将自定义配置文件与一个逻辑端口相关联，则该自定义配置文件覆盖现有的交换配置文件。

注 如果已将自定义交换配置文件与一个逻辑交换机相关联，但希望保留某个子逻辑端口的默认交换配置文件，您必须创建一个默认交换配置文件副本并将其与特定逻辑端口相关联。

如果将自定义交换配置文件与一个逻辑交换机或逻辑端口相关联，则无法删除该配置文件。您可以转到“摘要”视图的“分配给”部分并单击列出的逻辑交换机和逻辑端口，以确定任何逻辑交换机和逻辑端口是否与自定义交换配置文件相关联。

本章讨论了以下主题：



# 了解 QoS 交换配置文件

QoS 为需要高带宽的首选流量提供高质量和专用网络性能。QoS 机制确定分配足够带宽的优先顺序，控制延迟和抖动以及甚至在发生网络拥塞时减少首选数据包的数据丢失，从而实现该目的。该级别的网络服务是有效地使用现有的网络资源提供的。

对于该版本，支持调整和流量标记，即 CoS 和 DSCP。在由于拥塞而在逻辑交换机中缓冲流量时，第 2 层服务等级 (Class of Service, CoS) 允许您指定数据包的优先级。第 3 层差分服务代码点 (Differentiated Services Code Point, DSCP) 根据 DSCP 值检测数据包。CoS 始终应用于数据包，而不考虑受信任模式。

NSX-T 信任虚拟机应用的 DSCP 设置，或者在逻辑交换机级别修改和设置 DSCP 值。在每种情况下，DSCP 值将传播到 封装帧的外部 IP 标头。这样，外部物理网络就可以根据外部标头上的 DSCP 设置优先处理流量。在 DSCP 处于受信任模式时，将从内部标头中复制 DSCP 值。在处于不受信任模式时，不会保留内部标头的 DSCP 值。

注 DSCP 设置仅适用于隧道流量。这些设置不适用于同一管理程序中的流量。

您可以使用 QoS 交换配置文件配置平均输入和输出带宽值以设置传输限制速率。峰值带宽速率用于支持逻辑交换机允许的突发流量，以防止在北向网络链路上发生拥塞。不过，这些设置并不保证带宽，而是帮助限制使用网络带宽。

QoS 交换配置文件设置将应用于逻辑交换机，并且子逻辑交换机端口继承这些设置。

## 配置自定义 QoS 交换配置文件

您可以定义 DSCP 值并配置输入和输出设置以创建自定义 QoS 交换配置文件。

### 前提条件

- 熟悉 QoS 交换配置文件概念。请参阅了解 QoS 交换配置文件。
- 确定要优先处理的网络流量。

### 步骤

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**交换 (Switching) > 交换配置文件 (Switching Profiles)**。

3 单击**添加 (Add)**。

4 填写 QoS 交换配置文件详细信息。

| 选项 | 说明 |
| --- | --- |
| **名称和说明** | 指定自定义 QoS 交换配置文件的名称。<br>您可以选择描述在该配置文件中修改的设置。 |
| **类型** | 从下拉菜单中选择 **QoS**。 |
| **DSCP** | 从“模式”下拉菜单中选择**受信任 (Trusted)**或**不受信任 (Untrusted)**选项。<br>在选择“受信任”模式时，内部标头 DSCP 值将应用于 IP/IPv6 流量的外部 IP 标头。对于非 IP/IPv6 流量，外部 IP 标头使用默认值。在基于覆盖网络的逻辑端口上支持“受信任”模式。默认值为 0。<br>在基于覆盖网络和基于 VLAN 的逻辑端口上支持“不受信任”模式。对于基于覆盖网络的逻辑端口，出站 IP 标头的 DSCP 值设置为配置的值，而不考虑逻辑端口的内部数据包类型。对于基于 VLAN 的逻辑端口，IP/IPv6 数据包的 DSCP 值设置为配置的值。“不受信任”模式的 DSCP 值范围是 0 到 63 之间。<br>注 DSCP 设置仅适用于隧道流量。这些设置不适用于同一管理程序中的流量。 |
| **服务等级** | 配置流量优先级。<br>在基于 VLAN 的逻辑端口上支持 CoS。CoS 将网络中具有类似类型的流量划分到一起，并将每种类型的流量视为具有自己的服务优先级的等级。将减慢较低优先级的流量，或者在某些情况下，丢弃这些流量，以便为较高优先级的流量提供更好的吞吐量。也可以为具有零个数据包的 VLAN ID 配置 CoS。<br>CoS 值范围是 0 到 7，其中 0 是最佳效果服务。 |
| **输入** | 为从虚拟机到逻辑网络的出站网络流量设置自定义值。<br>您可以使用平均带宽以减少网络拥塞。峰值带宽速率用于支持突发流量，突发持续时间是在突发大小设置中设置的。您不能保证带宽。不过，您可以使用该设置限制网络带宽。默认值 0 禁用输入流量。<br>例如，在将逻辑交换机的平均带宽设置为 30 Mbps 时，该策略将限制带宽。您可以将突发流量限制为以 100 Mbps 速度传输 20 字节。 |
| **输入广播** | 为从虚拟机到逻辑网络的基于广播的出站网络流量设置自定义值。<br>默认值 0 禁用输入广播流量。<br>例如，在将逻辑交换机的平均带宽设置为 50 Kbps 时，该策略将限制带宽。您可以将突发流量限制为以 400 Kbps 速度传输 60 字节。 |
| **输出** | 为从逻辑网络到虚拟机的入站网络流量设置自定义值。<br>默认值 0 禁用输出流量。 |

如果未配置输入、输入广播和输出选项，则将默认值作为协议缓冲区。

5 单击**保存 (Save)**。

自定义 QoS 交换配置文件将显示为一个链接。

**后续步骤**

将该 QoS 自定义交换配置文件与一个逻辑交换机相关联，以便将交换配置文件中修改的参数应用于网络流量。请参阅将自定义配置文件与逻辑交换机相关联。

# 了解端口镜像交换配置文件

通过使用逻辑端口镜像，您可以复制和重定向流入或流出连接到虚拟机 VIF 端口的逻辑交换机端口的所有流量。镜像的流量在常规路由封装 (Generic Routing Encapsulation, GRE) 隧道中以封装形式发送到一个收集器，以便在通过网络传输到远程目标时保留所有原始数据包信息。

通常，端口镜像用于以下情况：

- 故障排除 - 分析流量以检测入侵，并调试和诊断网络上的错误。
- 合规性和监控 - 将所有监控的流量转发到网络设备以进行分析和修复。

与物理端口镜像相比，逻辑端口镜像确保捕获所有虚拟机网络流量。如果仅在物理网络中实现端口镜像，则无法镜像某些虚拟机网络流量。发生这种情况是因为，位于同一主机上的虚拟机之间的通信从不进入物理网络，因此，不会镜像这些通信。通过使用逻辑端口镜像，您可以继续镜像虚拟机流量，即使将该虚拟机迁移到其他主机。

对于 NSX-T 域中的虚拟机端口和物理应用程序的端口，端口镜像过程是类似的。您可以转发连接到逻辑网络的工作负载捕获的流量，并将该流量镜像到一个收集器。应该可以从托管虚拟机的客户机 IP 地址中访问该 IP 地址。该过程也适用于连接到网关节点的物理应用程序。

## 配置自定义端口镜像交换配置文件

您可以使用不同的目标和键值创建自定义端口镜像交换配置文件。

**前提条件**

- 熟悉端口镜像交换配置文件概念。请参阅了解端口镜像交换配置文件。
- 确定要将网络流量重定向到的目标逻辑端口 ID 的 IP 地址。

**步骤**

1. 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。
2. 从导航面板中选择**交换 (Switching) > 交换配置文件 (Switching Profiles)**。
3. 单击**添加 (Add)**。
4. 填写端口镜像交换配置文件详细信息。

| 选项 | 说明 |
| --- | --- |
| **名称和说明** | 指定自定义端口镜像交换配置文件的名称。<br>您可以选择描述自定义该配置文件时修改的设置。 |
| **类型** | 从下拉菜单中选择**端口镜像 (Port Mirroring)**。 |

| 选项 | 说明 |
| --- | --- |
| 方向 | 从下拉菜单中选择一个选项以将该源用于**输入 (Ingress)**、**输出 (Egress)**或**双向 (Bidirectional)**流量。<br>“输入”是从虚拟机到逻辑网络的出站网络流量。<br>“输出”是从逻辑网络到虚拟机的入站网络流量。<br>“双向”是从虚拟机到逻辑网络以及从逻辑网络到虚拟机的双向流量。这是默认选项。 |
| 数据包截断 | 可选。范围是 60-65535。 |
| 键 | 输入一个随机 32 位值以标识来自逻辑端口的镜像数据包。<br>该键值将复制到每个镜像数据包的 GRE 标头中的键字段。如果该键值设置为 0，则将默认定义复制到 GRE 标头中的键字段。<br>默认 32 位值由以下值组成。<br>■ 前 24 位是一个 VNI 值。VNI 是封装的帧的 IP 标头的一部分。<br>■ 第 25 位指示前 24 位是否为有效的 VNI 值。1 表示有效的值，0 表示无效的值。<br>■ 第 26 位指示镜像流量的方向。1 表示输入方向，0 表示输出方向。<br>■ 不使用剩余的 6 位。 |
| 目标 | 输入镜像会话的收集器的目标 ID。<br>目标 IP 地址 ID 只能是网络中的 IPv4 地址或 NSX-T 未管理的远程 IPv4 地址。您最多可以添加三个以逗号分隔的目标 IP 地址。 |

5 单击**保存 (Save)**。

自定义端口镜像交换配置文件将显示为一个链接。

### 后续步骤

验证自定义的端口镜像交换配置文件是否正常工作。请参阅验证自定义端口镜像交换配置文件。

## 验证自定义端口镜像交换配置文件

在开始使用自定义端口镜像交换配置文件之前，请验证自定义是否正常工作。

### 前提条件

- 确认配置了自定义端口镜像交换配置文件。请参见配置自定义端口镜像交换配置文件。
- 确认自定义端口镜像交换配置文件与一个逻辑交换机相关联。请参见将自定义配置文件与逻辑交换机相关联。

### 步骤

1 找到两个具有到配置了端口镜像的逻辑端口的 VIF 连接的虚拟机。

例如，虚拟机 1 10.70.1.1 和虚拟机 2 10.70.1.2 具有 VIF 连接，并且它们位于相同的逻辑网络中。

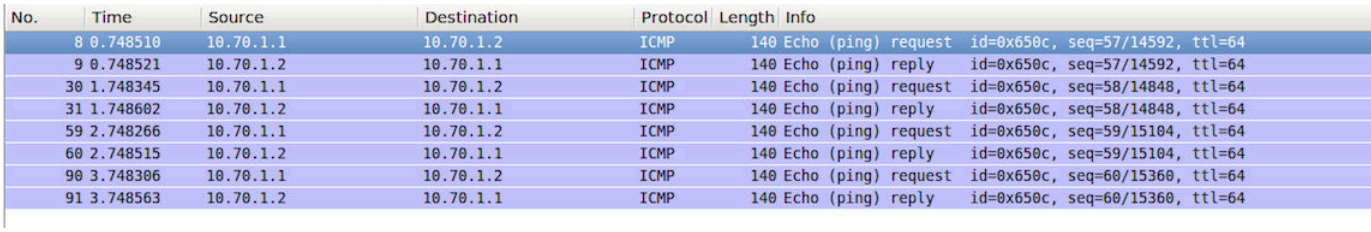
2 为某个目标 IP 地址运行 `tcpdump` 命令。

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

例如，目标 IP 地址为 10.24.123.196。

3 登录到第一个虚拟机并对第二个虚拟机执行 ping 操作，以验证在目标地址中是否收到相应的 ECHO 请求和回复。

例如，第一个虚拟机 10.70.1.1 对第二个虚拟机 10.70.1.2 执行 ping 操作以验证端口镜像。



| 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 |

#### 后续步骤

将该端口镜像自定义交换配置文件与一个逻辑交换机相关联，以便将交换配置文件中修改的参数应用于网络流量。请参见将自定义配置文件与逻辑交换机相关联。

# 了解 IP 发现交换配置文件

IP 发现使用 DHCP 或 ARP 侦听发现虚拟机 MAC 和 IP 地址。在发现 MAC 和 IP 地址后，将与 NSX Controller 共享这些条目以实现 ARP 抑制。ARP 抑制最大限度减少了连接到同一逻辑交换机的虚拟机中的 ARP 流量洪泛。

DHCP 侦听检查在虚拟机 DHCP 客户端和 DHCP 服务器之间传输的 DHCP 数据包以发现虚拟机 IP 和 MAC 地址。

ARP 侦听检查虚拟机的出站 ARP 和 GARP 以发现 IP 和 MAC 地址。

## 配置 IP 发现交换配置文件

您可以启用 ARP 侦听或 DHCP 侦听以创建自定义 IP 发现交换配置文件，以便发现 IP 和 MAC 地址以确保逻辑交换机的 IP 完整性。

#### 前提条件

熟悉 IP 发现交换配置文件概念。请参阅了解 IP 发现交换配置文件。

#### 步骤

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**交换 (Switching) > 交换配置文件 (Switching Profiles)**。

3 单击**添加 (Add)**。

4 填写 IP 发现交换配置文件详细信息。

| 选项 | 说明 |
|---|---|
| **名称和说明** | 指定自定义 IP 发现交换配置文件的名称。<br>您可以选择描述在该配置文件中启用的设置。 |
| **类型** | 从下拉菜单中选择 **IP 发现 (IP Discovering)**。 |

| 选项 | 说明 |
| --- | --- |
| **ARP 侦听** | 切换 **ARP 侦听 (ARP Snooping)**按钮以启用该功能。<br>ARP 侦听检查虚拟机出站 ARP 和 GARP 以发现虚拟机 MAC 和 IP 地址。如果虚拟机使用静态 IP 地址而不是 DHCP，ARP 侦听才适用。 |
| **DHCP 侦听** | 切换 **DHCP 侦听 (DHCP Snooping)**按钮以启用该功能。<br>DHCP 侦听检查在虚拟机 DHCP 客户端和 DHCP 服务器之间传输的 DHCP 数据包以发现虚拟机 MAC 和 IP 地址。 |

5 单击**保存 (Save)**。

自定义 IP 发现交换配置文件将显示为一个链接。

**后续步骤**

将该 IP 发现自定义交换配置文件与一个逻辑交换机相关联，以便将交换配置文件中修改的参数应用于网络流量。请参阅将自定义配置文件与逻辑交换机相关联。

# 了解 SpoofGuard

SpoofGuard 有助于防止一种称为“网络欺骗”或“网络钓鱼”的恶意攻击。SpoofGuard 策略阻止确定为欺骗的流量。

SpoofGuard 工具旨在防止您的环境中的虚拟机发送某种流量，该流量带有未授权终止流量的 IP 地址。如果虚拟机的 IP 地址与 SpoofGuard 上的相应逻辑端口和交换机地址绑定中的 IP 地址不匹配，将完全禁止虚拟机的 vNIC 访问网络。可以在端口或交换机级别配置 SpoofGuard。在您的环境中使用 SpoofGuard 可能有以下几个原因：

- 防止恶意虚拟机使用现有虚拟机的 IP 地址。
- 确保无法在没有干预的情况下更改虚拟机的 IP 地址 - 在某些环境中，在没有正确的更改控制检查的情况下，最好禁止虚拟机更改其 IP 地址。SpoofGuard 确保虚拟机所有者无法直接更改 IP 地址并继续工作而不会受到妨碍，从而简化了该过程。
- 保证不会无意（或有意）绕过分布式防火墙 (Distributed Firewall, DFW) 规则 - 对于将 IP 集作为源或目标创建的 DFW 规则，始终存在虚拟机可能在数据包标头中伪造其 IP 地址的可能性，从而绕过相关的规则。

NSX-T SpoofGuard 配置包括以下内容：

- MAC SpoofGuard - 验证数据包的 MAC 地址。
- IP SpoofGuard - 验证数据包的 IP 地址。
- 动态地址解析协议 (Dynamic Address Resolution Protocol, ARP) 检查（即，ARP）以及无故地址解析协议 (Gratuitous Address Resolution Protocol, GARP) SpoofGuard 和邻居发现 (Neighbor Discovery, ND) SpoofGuard 验证针对的都是 ARP/GARP/ND 负载中的 MAC 源、IP 源和 IP-MAC 源映射。

在端口级别，允许的 MAC/VLAN/IP 白名单是通过端口的地址绑定属性提供的。在虚拟机发送流量时，如果流量的 IP/MAC/VLAN 与端口的 IP/MAC/VLAN 属性不匹配，则会丢弃该流量。端口级别 SpoofGuard 处理流量验证，即，流量与 VIF 配置是否一致。

在交换机级别，允许的 MAC/VLAN/IP 白名单是通过交换机的地址绑定属性提供的。这通常是交换机的允许的 IP 范围/子网，交换机级别 SpoofGuard 处理流量授权。

端口级别和交换机级别 SpoofGuard 必须允许流量，然后才允许流量进入交换机。可以使用 SpoofGuard 交换机配置文件控制启用或禁用端口和交换机级别 SpoofGuard。

## 配置端口地址绑定

地址绑定指定逻辑端口的 IP 和 MAC 地址，并用于指定 SpoofGuard 中的端口白名单。

通过使用端口地址绑定，您可以指定逻辑端口的 IP 和 MAC 地址以及 VLAN（如果适用）。如果启用 SpoofGuard，它确保在数据路径中强制实施指定的地址绑定。除了 SpoofGuard 以外，端口地址绑定还用于 DFW 规则转换。

### 步骤

1 在 NSX Manager 中，导航到**交换 > 端口 (Switching > Ports)**。

2 单击要将地址绑定应用到的逻辑端口。

将显示逻辑端口摘要。

3 在“摘要”选项卡下面，展开**地址绑定 (Address Bindings)**。

4 单击**添加 (Add)**。

将显示“添加地址绑定”对话框。

5 指定要将地址绑定应用到的逻辑端口的 IP 和 MAC 地址。还可以选择指定 VLAN。

6 单击**保存 (Save)**。

### 后续步骤

在配置 SpoofGuard 交换配置文件时，请使用端口地址绑定。

## 配置交换机地址绑定

地址绑定允许将一定范围的 IP 和 MAC 地址以及 VLAN 绑定到交换机。

在 SpoofGuard 中，地址绑定提供了允许的 MAC/VLAN/IP 白名单。在启用相应 SpoofGuard 的情况下，它确保在数据路径中强制实施指定的地址绑定。

### 步骤

1 在 NSX Manager 中，导航到**交换 > 交换机 (Switching > Switches)**。

2 单击要将地址绑定应用到的逻辑交换机。

在右侧窗口中，将显示交换机摘要。

3 在“摘要”选项卡下面，展开**地址绑定 (Address Bindings)**。

4 单击**添加 (Add)**。

将显示“添加地址绑定”对话框。

5 在交换机地址绑定中输入交换机的 MAC 地址和 IP 范围以及 VLAN（如果适用）。

在指定 IP 范围/子网后，数据路径将在交换机上的所有端口中应用绑定。

6 单击**保存 (Save)**。

**后续步骤**

现在，您将配置 SpoofGuard 交换配置文件并将地址绑定添加到 SpoofGuard 白名单。

## 配置 SpoofGuard 交换配置文件

在配置 SpoofGuard 时，如果某个虚拟机的 IP 地址发生变化，可能会阻止来自该虚拟机的流量，直到将配置的相应端口/交换机地址绑定更新为新 IP 地址。

为包含客户机的端口组启用 SpoofGuard。如果为每个网络适配器启用 SpoofGuard，它将检查数据包以查找指定的 MAC 及其相应的 IP 地址。

**前提条件**

在配置 SpoofGuard 之前，请在每个逻辑交换机上添加地址绑定或交换机绑定。地址绑定允许将 IP 地址和 MAC 地址绑定到一个端口或交换机。配置端口地址绑定配置交换机地址绑定

**步骤**

1 在 NSX Manager 中，导航到**交换 > 交换配置文件 (Switching > Switching Profiles)**。

2 单击**添加 (Add)**。

将显示“新建交换配置文件”窗口。

3 命名配置文件并选择 **SpoofGuard** 以作为类型。也可以添加配置文件说明。

4 要启用端口级别 SpoofGuard，请选择**端口绑定 (port bindings)**；要启用交换机级别 SpoofGuard，请选择**交换机绑定 (switch bindings)**。

地址绑定是允许的端口和交换机 SpoofGuard 白名单。

5 单击**保存 (Save)**。

将使用 SpoofGuard 配置文件创建一个新的交换配置文件。

**后续步骤**

将 SpoofGuard 配置文件与逻辑交换机相关联。将自定义配置文件与逻辑交换机相关联

## 了解交换机安全交换配置文件

交换机安全通过以下方法提供无状态第 2 层和第 3 层安全性：检查逻辑交换机的输入流量，并将 IP 地址、MAC 地址和协议与一组允许的地址和协议进行匹配以丢弃从虚拟机发送的未授权的数据包。您可以使用交换机安全筛选掉来自网络中的虚拟机的恶意攻击以保护逻辑交换机完整性。

您可以配置网桥协议数据单元 (Bridge Protocol Data Unit, BPDU) 筛选器、DHCP 侦听、DHCP 服务器阻止以及速率限制选项以自定义逻辑交换机上的交换机安全交换配置文件。

# 配置自定义交换机安全交换配置文件

您可以使用允许的 BPDU 列表中的 MAC 目标地址创建自定义交换机安全交换配置文件并配置速率限制。

### 前提条件

熟悉交换机安全交换配置文件概念。请参阅了解交换机安全交换配置文件。

### 步骤

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**交换 (Switching) > 交换配置文件 (Switching Profiles)**。

3 单击**添加 (Add)**。

4 填写交换机安全配置文件详细信息。

| 选项 | 说明 |
| --- | --- |
| **名称和说明** | 指定自定义交换机安全配置文件的名称。<br>您可以选择描述在该配置文件中修改的设置。 |
| **类型** | 从下拉菜单中选择**交换机安全 (Switch Security)**。 |
| **BPDU 筛选器** | 切换 **BPDU 筛选器 (BPDU filter)**按钮以启用 BPDU 筛选。<br>如果启用了 BPDU 筛选器，将阻止到 BPDU 目标 MAC 地址的所有流量。如果启用，BPDU 筛选器还会在逻辑交换机端口上禁用 STP，因为这些端口应该不会加入 STP。 |
| **BPDU 筛选器允许列表** | 单击 BPDU 目标 MAC 地址列表中的目标 MAC 地址以允许将流量传输到允许的目标。 |
| **DHCP 筛选器** | 切换**服务器阻止 (Server Block)**按钮和**客户端阻止 (Client Block)**按钮以启用 DHCP 筛选。<br>“DHCP 服务器阻止”阻止从 DHCP 服务器到 DHCP 客户端的流量。请注意，它不会阻止从 DHCP 服务器到 DHCP 中继代理的流量。<br>“DHCP 客户端阻止”阻止 DHCP 请求以禁止虚拟机获取 DHCP IP 地址。 |
| **阻止非 IP 流量** | 切换**阻止非 IP 流量 (Block Non-IP Traffic)**按钮以仅允许 IPv4、IPv6、ARP、GARP 和 BPDU 流量。<br>将阻止其余非 IP 流量。允许的 IPv4、IPv6、ARP、GARP 和 BPDU 流量基于在地址绑定和 SpoofGuard 配置中设置的其他策略。<br>默认情况下，将禁用该选项以允许将非 IP 流量作为常规流量进行处理。 |
| **速率限制** | 为输入或输出广播和多播流量设置速率限制。<br>例如，配置速率限制以防止逻辑交换机或虚拟机受到广播流量风暴的影响。<br>为了避免任何连接问题，最小速率限制值必须大于或等于 10 pps。 |

5 单击**保存 (Save)**。

自定义交换机安全配置文件将显示为一个链接。

### 后续步骤

将该交换机安全自定义交换配置文件与一个逻辑交换机相关联，以便将交换配置文件中修改的参数应用于网络流量。请参阅将自定义配置文件与逻辑交换机相关联。

# 了解 MAC 管理交换配置文件

MAC 管理交换配置文件支持两种功能：MAC 发现和 MAC 地址更改。

MAC 发现提供到在一个 vNIC 后面配置多个 MAC 地址的部署的网络连接，例如，在嵌套管理程序部署中，ESXi 虚拟机在 ESXi 主机上运行，并且多个虚拟机在 ESXi 虚拟机中运行。如果未使用 MAC 发现，在 ESXi 虚拟机的 vNIC 连接到交换机端口时，其 MAC 地址是静态的。在 ESXi 虚拟机中运行的虚拟机没有网络连接，因为其数据包具有不同的源 MAC 地址。通过使用 MAC 发现，vSwitch 检查来自 vNIC 的每个数据包的源 MAC 地址，发现 MAC 地址并允许数据包通过。如果在特定时间段内未使用发现的 MAC 地址，则会将其移除。无法配置该到期属性。

MAC 发现还支持未知单播洪泛。通常，在端口收到的数据包具有未知目标 MAC 地址时，将丢弃该数据包。在启用未知单播洪泛的情况下，端口将未知单播流量洪泛到交换机上启用了 MAC 发现和单播洪泛的每个端口。默认情况下，将启用该属性，但只有在启用 MAC 发现时才启用。

MAC 管理交换配置文件还支持虚拟机更改其 MAC 地址的功能。连接到启用了 MAC 地址更改属性的端口的虚拟机可以运行管理命令以更改其 vNIC 的 MAC 地址，并且仍然在该 vNIC 上发送和接收流量。仅在 ESXi 上支持该功能，而在 KVM 上不支持。默认情况下，将禁用该属性。

如果启用 MAC 发现或 MAC 地址更改以提高安全性，还要配置 SpoofGuard。

有关创建 MAC 管理交换配置文件并将该配置文件与交换机或端口关联的详细信息，请参阅《《NSX-T API 指南》》。

注　在该版本中，只能通过 NSX API 使用 MAC 管理交换配置文件功能。无法从 NSX Manager UI 中使用该功能。

# 将自定义配置文件与逻辑交换机相关联

要将自定义交换配置文件应用于您的网络，您必须将其与一个逻辑交换机相关联。

如果将自定义交换配置文件与一个逻辑交换机相关联，它们将覆盖现有的默认交换配置文件。子逻辑交换机端口将继承自定义交换配置文件。

注　如果已将自定义交换配置文件与一个逻辑交换机相关联，但希望保留某个子逻辑交换机端口的默认交换配置文件，您必须创建一个默认交换配置文件副本并将其与特定逻辑交换机端口相关联。

**前提条件**

- 确认配置了一个逻辑交换机。请参阅创建逻辑交换机。
- 确认配置了一个自定义交换配置文件。请参阅第 3 章，为逻辑交换机和逻辑端口配置交换配置文件。

**步骤**

1. 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。
2. 从导航面板中选择**交换 (Switching) > 交换机 (Switches)**。
3. 双击逻辑交换机以应用自定义交换配置文件。

4 单击**管理 (Manage)**选项卡。

5 从下拉菜单中选择自定义交换配置文件类型。

- **QoS**
- **端口镜像 (Port Mirroring)**
- **IP 发现 (IP Discovering)**
- **SpoofGuard**
- **交换机安全 (Switch Security)**

6 单击**更改 (Change.)**。

7 从下拉菜单中选择以前创建的自定义交换配置文件。

8 单击**保存 (Save)**。

逻辑交换机现在与自定义交换配置文件相关联。

9 验证是否在**管理 (Manage)**选项卡下面显示具有修改的配置的新自定义交换配置文件。

10 （可选）单击**相关 (Related)**选项卡并从下拉菜单中选择**端口 (Ports)**，以验证是否将自定义交换配置文件应用于子逻辑端口。

**后续步骤**

如果不希望使用从逻辑交换机中继承的交换配置文件，您可以将自定义交换配置文件应用于子逻辑交换机端口。请参阅将自定义配置文件与逻辑交换机端口相关联。

# 将自定义配置文件与逻辑交换机端口相关联

逻辑交换机端口提供 VIF 的逻辑连接点、到路由器的修补连接或到外部网络的第 2 层网关连接。逻辑交换机端口还公开交换配置文件、端口统计信息计数器以及逻辑链路状态。

您可以将子逻辑交换机端口从逻辑交换机中继承的交换配置文件更改为不同的自定义交换配置文件。

**前提条件**

- 确认配置了一个逻辑交换机端口。请参阅将虚拟机连接到逻辑交换机。
- 确认配置了一个自定义交换配置文件。请参阅第 3 章，为逻辑交换机和逻辑端口配置交换配置文件。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**交换 (Switching) > 端口 (Port)**。

3 双击逻辑交换机端口以应用自定义交换配置文件。

4 单击**管理 (Manage)**选项卡。

5 从下拉菜单中选择自定义交换配置文件类型。

- **QoS**

- **端口镜像 (Port Mirroring)**
- **IP 发现 (IP Discovering)**
- **SpoofGuard**
- **交换机安全 (Switch Security)**

6 单击**更改 (Change)**。

7 从下拉菜单中选择以前创建的自定义交换配置文件。

8 单击**保存 (Save)**。

逻辑交换机端口现在与自定义交换配置文件相关联。

9 验证是否在**管理 (Manage)**选项卡下面显示具有修改的配置的新自定义交换配置文件。

**后续步骤**

您可以监控逻辑交换机端口上的活动以解决问题。请参阅监控逻辑交换机端口活动。

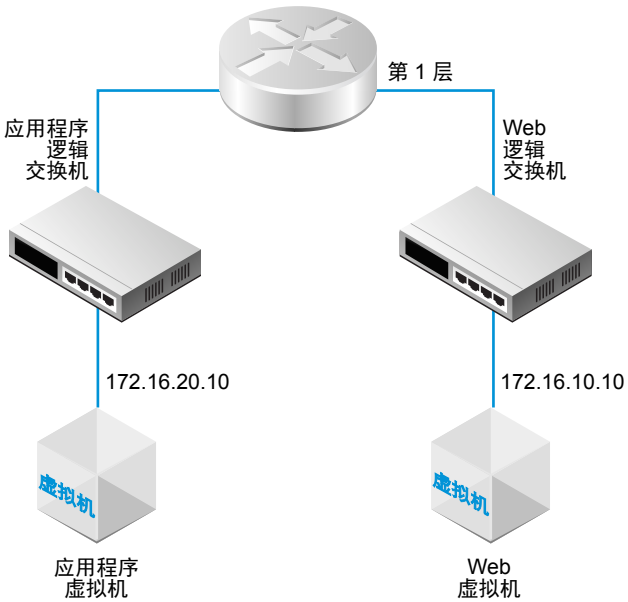
# 配置第 1 层逻辑路由器

4

NSX-T 逻辑路由器在完全脱离底层硬件的虚拟环境中再现路由功能。第 1 层逻辑路由器具有下行链路端口以连接到 NSX-T 逻辑交换机，并具有上行链路端口以连接到 NSX-T 第 0 层逻辑路由器。

在添加逻辑路由器时，请务必规划要构建的网络拓扑。

**图 4-1. 第 1 层逻辑路由器拓扑**




例如，该简单拓扑显示两个连接到第 1 层逻辑路由器的逻辑交换机。每个逻辑交换机连接了单个虚拟机。两个虚拟机可以位于不同主机群集或同一主机群集中的不同主机或同一主机上。如果逻辑路由器未隔离这些虚拟机，在这些虚拟机上配置的基础 IP 地址必须位于同一子网中。如果逻辑路由器隔离这些虚拟机，这些虚拟机上的 IP 地址必须位于不同的子网中。

本章讨论了以下主题：

- 创建第 1 层逻辑路由器
- 为第 1 层逻辑路由器添加下行链路端口
- 在第 1 层逻辑路由器上配置路由播发
- 配置第 1 层逻辑路由器静态路由

# 创建第 1 层逻辑路由器

必须将第 1 层逻辑路由器连接到第 0 层逻辑路由器以进行北向物理路由器访问。

### 前提条件

- 确认配置了逻辑交换机。请参阅创建逻辑交换机。
- 确认部署了一个 NSX Edge 群集以执行网络地址转换 (NAT) 配置。请参阅《 NSX-T 安装指南》。
- 熟悉第 1 层逻辑路由器拓扑。请参阅第 4 章，配置第 1 层逻辑路由器。

### 步骤

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 单击**添加 (Add)**，然后选择**第 1 层路由器 (Tier-1 Router)**。

4 指定逻辑路由器的名称。

5 （可选）选择一个第 0 层逻辑路由器以连接到该第 1 层逻辑路由器。

如果尚未配置任何第 0 层逻辑路由器，您可以将该字段暂时保留空白，以后再编辑路由器配置。

6 （可选）选择一个 Edge 群集以连接到该第 1 层逻辑路由器。

如果将第 1 层逻辑路由器用于 NAT 配置，则必须将其连接到一个 NSX Edge 群集。如果尚未配置任何 Edge 群集，您可以将该字段暂时保留空白，以后再编辑路由器配置。

7 单击**保存 (Save)**。

在 NSX Manager UI 中，新逻辑路由器是一个可单击的链接。

### 后续步骤

为第 1 层逻辑路由器创建下行链路端口。请参阅为第 1 层逻辑路由器添加下行链路端口。

# 为第 1 层逻辑路由器添加下行链路端口

在第 1 层逻辑路由器上创建下行链路端口时，该端口将作为同一子网中的虚拟机的默认网关。

### 前提条件

确认配置了一个第 1 层逻辑路由器。请参见创建第 1 层逻辑路由器。

### 步骤

1 单击第 1 层逻辑路由器链接以创建端口。

2 单击**配置 (Configuration)** 选项卡。

3 在“逻辑路由器端口”部分下面，单击**添加 (Add)**。

4 指定逻辑路由器端口的名称。

5 选择该连接是创建交换机端口还是更新现有的交换机端口。

如果该连接用于现有的交换机端口，请从下拉菜单中选择该端口。

6 以 CIDR 表示法输入路由器端口 IP 地址。

例如，IP 地址可以是 172.16.10.1/24。

也可以输入预配置的 DHCP 服务 IP 地址。

7 单击**保存 (Save)**。

8 （可选）重复步骤 1-7 以创建其他的第 1 层逻辑路由器端口。

9 验证第 1 层逻辑路由器是否可以路由东西向虚拟机流量。

在该示例中，第 1 层逻辑路由器具有两个下行链路端口，它们连接到两个逻辑交换机。每个逻辑交换机连接了一个虚拟机。这些虚拟机可以 ping 通对方。

```
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
```

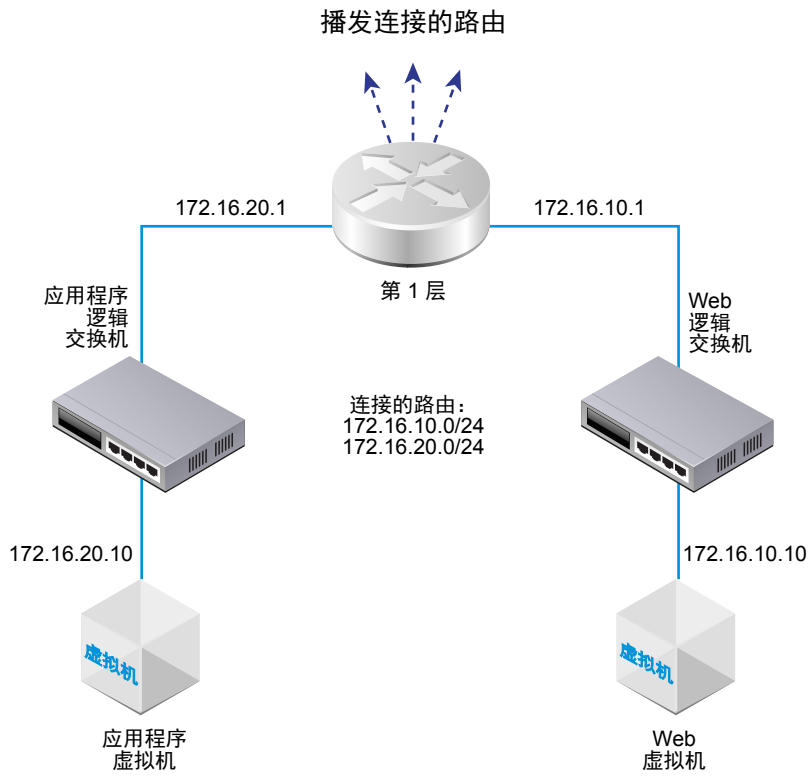
**后续步骤**

启用路由播发以在虚拟机和外部物理网络之间或连接到同一第 0 层逻辑路由器的不同第 1 层逻辑路由器之间提供南北向连接。请参见在第 1 层逻辑路由器上配置路由播发。

# 在第 1 层逻辑路由器上配置路由播发

要在连接到不同的第 1 层逻辑路由器的逻辑交换机连接的虚拟机之间提供第 3 层连接，必须允许将第 1 层路由播发到第 0 层。您不需要在第 1 层和第 0 层逻辑路由器之间配置路由协议或静态路由。在启用路由播发时，NSX-T 自动创建 NSX-T 静态路由。

例如，要通过其他对等路由器提供与虚拟机之间的连接，第 1 层逻辑路由器必须为连接的路由配置路由播发。如果不希望播发所有连接的路由，您可以指定要播发的路由。




**前提条件**

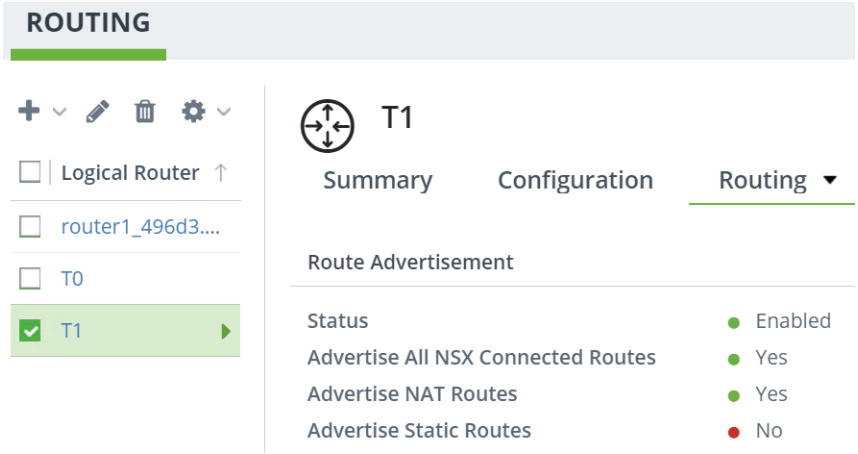
- 确认虚拟机已连接到逻辑交换机。请参见第 2 章，创建逻辑交换机和配置虚拟机连接。
- 确认配置了第 1 层逻辑路由器的下行链路端口。请参见为第 1 层逻辑路由器添加下行链路端口。

**步骤**

1. 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。
2. 选择**路由 (Routing)**。
3. 单击一个第 1 层逻辑路由器。
4. 从“路由”下拉菜单中选择**路由播发 (Route Advertisement)**。
5. 单击**编辑 (Edit)** 并确保“状态”按钮为“已启用”以启用路由播发。
6. 指定要播发的路由：所有路由或选定的路由。
   - 单击**编辑 (Edit)**，然后选择**播发所有 NSX 连接的路由 (Advertise All NSX Connected Routes)**。
   - 单击**添加 (Add)**，然后输入有关要播发的路由的信息。对于每个路由，您可以按 CIDR 格式输入名称和路由前缀。

7 单击**状态 (Status)** 切换按钮以启用路由播发。

例如：




8 单击**保存 (Save)**。

**后续步骤**

熟悉第 0 层逻辑路由器拓扑并创建第 0 层逻辑路由器。请参见第 5 章，配置第 0 层逻辑路由器。

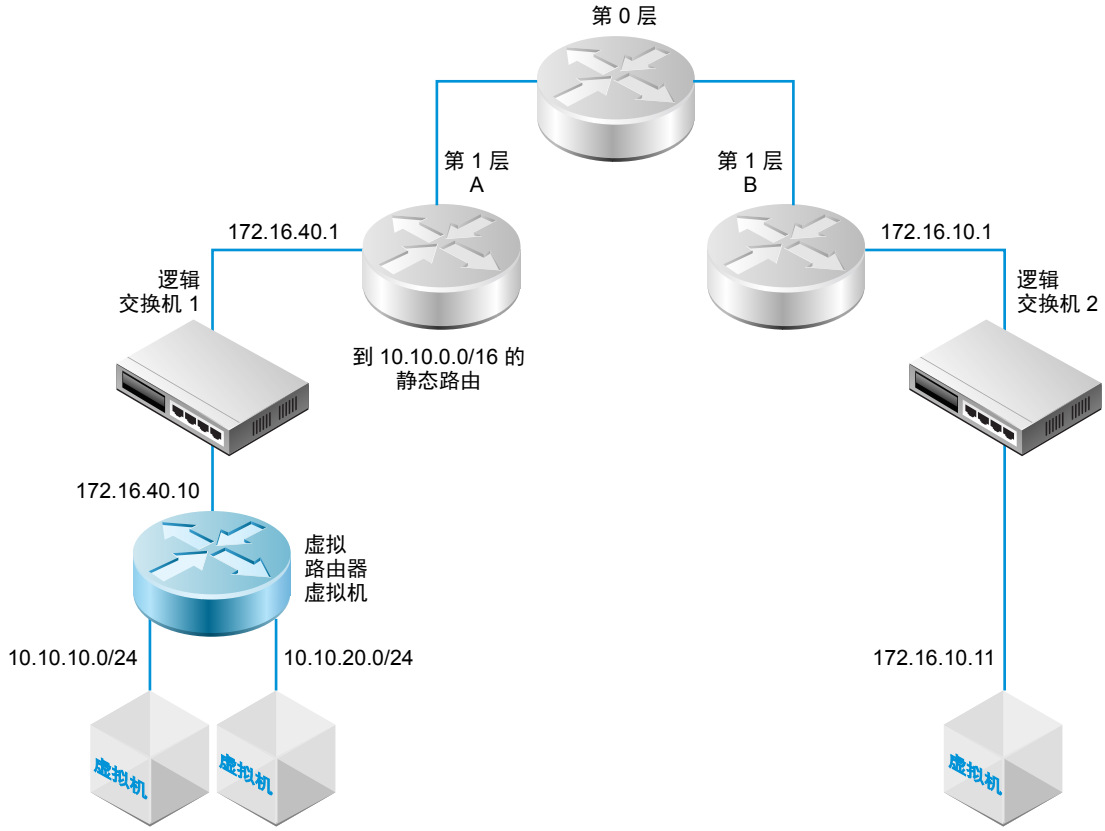
如果已将一个第 0 层逻辑路由器连接到第 1 层逻辑路由器，您可以验证该第 0 层路由器是否发现第 1 层路由器连接的路由。请参见验证 Tier-0 路由器是否发现来自 Tier-1 路由器的路由。

# 配置第 1 层逻辑路由器静态路由

您可以在第 1 层逻辑路由器上配置静态路由，以提供从 NSX-T 到一组可通过虚拟路由器访问的网络的连接。

例如，在下图中，第 1 层 A 逻辑路由器具有到 NSX-T 逻辑交换机的下行链路端口。该下行链路端口 (172.16.40.1) 为虚拟路由器虚拟机提供默认网关。虚拟路由器虚拟机和第 1 层 A 通过相同 NSX-T 逻辑交换机连接在一起。第 1 层逻辑路由器具有静态路由 10.10.0.0/16，它汇总了通过虚拟路由器访问的网络。第 1 层 A 配置了路由播发以将静态路由播发到第 1 层 B。

**图 4-2. 第 1 层逻辑路由器静态路由拓扑**




**前提条件**

确认配置了一个下行链路端口。请参阅为第 1 层逻辑路由器添加下行链路端口。

**步骤**

1. 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。
2. 从导航面板中选择**路由 (Routing)**。
3. 选择第 1 层逻辑路由器。
4. 单击**路由 (Routing)**选项卡，然后从下拉菜单中选择**静态路由 (Static Route)**。
5. 选择**添加 (Add)**。
6. 以 CIDR 格式输入一个网络地址。

   例如，10.10.10.0/16。
7. 单击**插入行 (Insert Row)**以添加一个下一跃点 IP 地址。

   例如，172.16.40.10。
8. 单击**保存 (Save)**。

   将在该行中显示新创建的静态路由网络地址。
9. 从第 1 层逻辑路由器中，选择**路由 > 路由播发 (Routing > Route Advertisement)**。

10 单击**编辑 (Edit)**，然后选择**播发静态路由 (Advertise Static Routes)**。

11 单击**保存 (Save)**。

将在 NSX-T 覆盖网络中传播静态路由。

# 配置第 0 层逻辑路由器

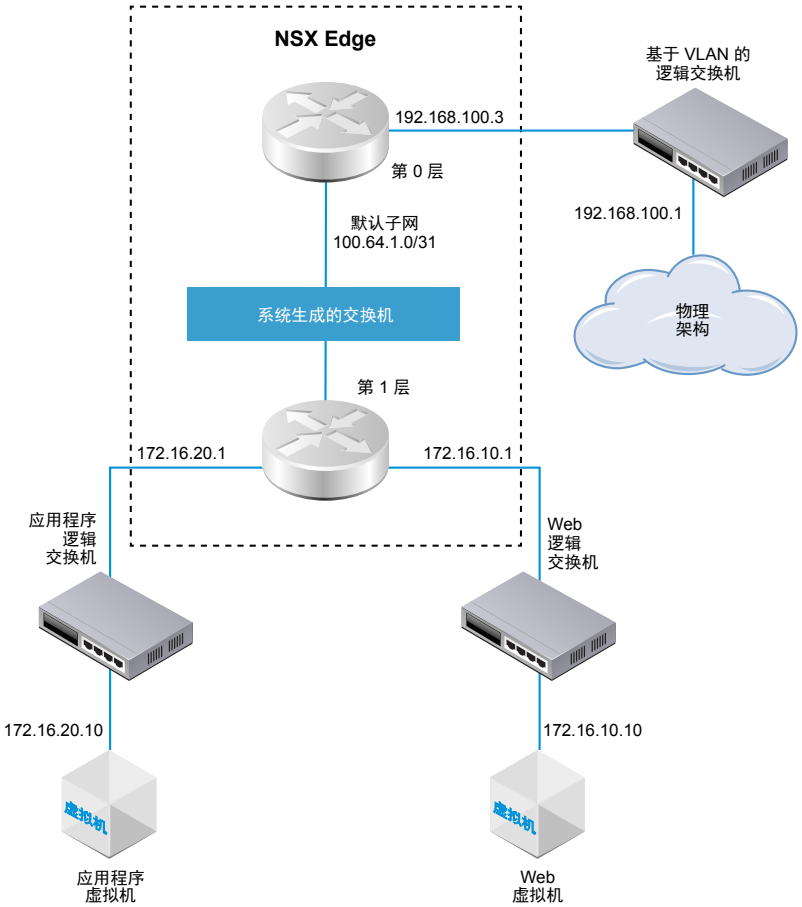
5

NSX-T 逻辑路由器在完全脱离底层硬件的虚拟环境中再现路由功能。第 0 层逻辑路由器在逻辑和物理网络之间提供打开和关闭网关服务。

NSX Edge 群集可以支持多个第 0 层逻辑路由器。第 0 层路由器支持 BGP 动态路由协议和 ECMP。

在添加第 0 层逻辑路由器时，请务必规划要构建的网络拓扑。

**图 5-1. 第 0 层逻辑路由器拓扑**



为了简单起见，示例拓扑显示单个第 1 层逻辑路由器，它连接到在单个 NSX Edge 节点上托管的单个第 0 层逻辑路由器。请记住，这不是建议的拓扑。理想情况下，您应该使用至少两个 NSX Edge 节点以充分利用逻辑路由器设计。

第 1 层逻辑路由器具有一个 Web 逻辑交换机和一个应用程序逻辑交换机，并且它们连接了相应的虚拟机。在将第 1 层路由器连接到第 0 层路由器时，将在第 1 层路由器和第 0 层路由器之间自动创建路由器-链路交换机。因此，该交换机标记为系统生成的交换机。

本章讨论了以下主题：

- 创建第 0 层逻辑路由器
- 连接第 0 层和第 1 层
- 将第 0 层逻辑路由器连接到 VLAN 逻辑交换机
- 配置静态路由
- BGP 配置选项
- 在第 0 层逻辑路由器上配置 BFD
- 在 Tier-0 逻辑路由器上启用路由重新分发
- 了解 ECMP 路由
- 创建 IP 前缀列表
- 创建路由映射

# 创建第 0 层逻辑路由器

第 0 层逻辑路由器具有下行链路端口以连接到 NSX-T 第 1 层逻辑路由器，并具有上行链路端口以连接到外部网络。

### 前提条件

- 确认安装了至少一个 NSX Edge。请参阅《 NSX-T 安装指南》。
- 确认 NSX Controller 群集处于稳定状态。
- 确认配置了一个 Edge 群集。请参见 NSX-T 安装指南。
- 熟悉第 0 层逻辑路由器的网络拓扑。请参见第 5 章，配置第 0 层逻辑路由器。

### 步骤

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 单击**添加**以创建一个第 0 层逻辑路由器。

4 从下拉菜单中选择**第 0 层路由器**。

5 指定第 0 层逻辑路由器的名称。

6 从下拉菜单中选择一个现有的 Edge 群集以支持该第 0 层逻辑路由器。

7 （可选）选择一种高可用性模式。

默认情况下，将使用活动-活动模式。在活动-活动模式下，将在所有成员之间进行流量负载平衡。在活动-备用模式下，将由选举的活动成员处理所有流量。如果活动成员发生故障，将选举新的成员以作为活动成员。

8 （可选）单击**高级**选项卡以输入一个子网以作为第 0 层内中转子网。

这是将第 0 层服务路由器连接到其分布式路由器的子网。如果将该字段保留空白，则使用默认 169.0.0.0/28 子网。

9 （可选）单击**高级**选项卡以输入一个子网以作为第 0 层到第 1 层的中转子网。

这是将第 0 层路由器连接到该第 0 层路由器连接的任何第 1 层路由器的子网。如果将该字段保留空白，则为这些第 0 层到第 1 层的连接分配的默认地址空间为 100.64.0.0/10。将在 100.64.0.0/10 地址空间中为每个第 0 层到第 1 层的对等连接提供一个 /31 子网。

10 单击**保存**。

新的第 0 层逻辑路由器将显示为一个链接。

11 （可选）单击第 0 层逻辑路由器链接以查看摘要。

**后续步骤**

将第 1 层逻辑路由器连接到该第 0 层逻辑路由器。

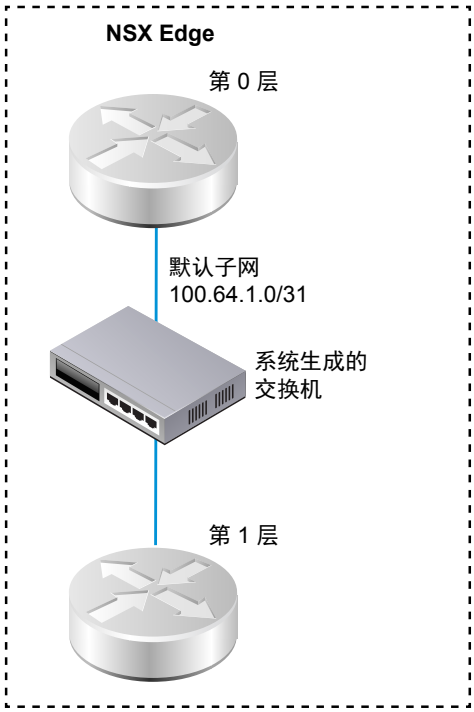
配置第 0 层逻辑路由器以将其连接到 VLAN 逻辑交换机，以便创建到外部网络的上行链路。请参见将第 0 层逻辑路由器连接到 VLAN 逻辑交换机。

# 连接第 0 层和第 1 层

您可以将第 0 层逻辑路由器连接到第 1 层逻辑路由器，以便第 1 层逻辑路由器具有北向和东西向网络连接。

在将第 1 层逻辑路由器连接到第 0 层逻辑路由器时，将在两个路由器之间创建路由器-链路交换机。该交换机在拓扑中标记为系统生成的交换机。为这些第 0 层到第 1 层的连接分配的默认地址空间为 100.64.0.0/10。将在 100.64.0.0/10 地址空间中为每个第 0 层到第 1 层的对等连接提供一个 /31 子网。您可以选择在第 0 层**摘要 > 高级**配置中配置地址空间。

下图显示了一个示例拓扑。




**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 选择第 1 层逻辑路由器。

4 从**摘要**选项卡中，单击**编辑**。

5 从下拉菜单中选择第 0 层逻辑路由器。

6 （可选）从下拉菜单中选择一个 Edge 群集。

如果要将第 1 层路由器用于服务（如 NAT），则需要使用 Edge 设备支持该路由器。如果未选择 Edge 群集，则第 1 层路由器无法执行 NAT。

7 指定成员和首选成员。

如果选择一个 Edge 群集并将成员和首选成员字段保留空白，NSX-T 将从指定的群集中设置支持 Edge 设备。

8 单击**保存**。

9 单击第 1 层路由器的**配置**选项卡，以验证是否创建了新的点对点链接端口 IP 地址。

例如，链接端口的 IP 地址可能是 100.64.1.1/31。

10 从导航面板中选择第 0 层逻辑路由器。

11 单击第 0 层路由器的**配置**选项卡，以验证是否创建了新的点对点链接端口 IP 地址。

例如，链接端口的 IP 地址可能是 100.64.1.1/31。

**后续步骤**

验证第 0 层路由器是否发现第 1 层路由器通告的路由。

## 验证 Tier-0 路由器是否发现来自 Tier-1 路由器的路由

在 Tier-1 逻辑路由器将路由通告到 Tier-0 逻辑路由器时，这些路由将在 Tier-0 路由器的路由表中列出为 NSX-T 静态路由。

**步骤**

1 在 NSX Edge 上，运行 `get logical-routers` 命令以查找 Tier-0 服务路由器的 VRF 编号。

```
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
```

2 运行 `vrf <number>` 命令以进入 Tier-0 服务路由器上下文。

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

3 在 Tier-0 服务路由器上，运行 `get route` 命令并确保在路由表中显示预期的路由。

请注意，NSX-T 静态路由 (ns) 是 Tier-0 路由器发现的，因为 Tier-1 路由器正在通告路由。

```
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
```

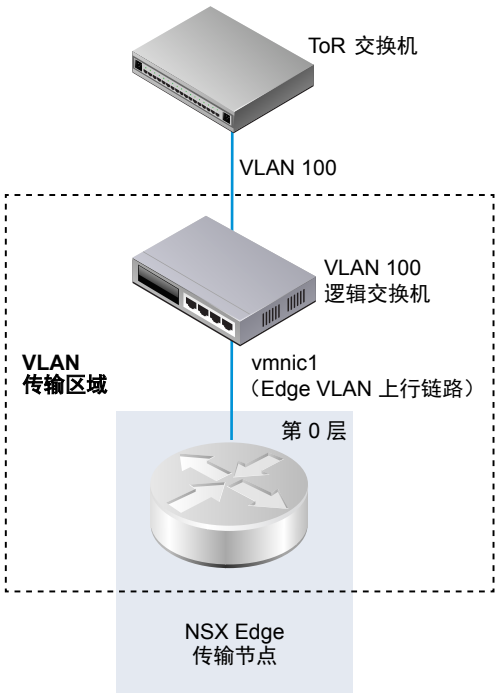
```
Total number of routes: 7

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
```

# 将第 0 层逻辑路由器连接到 VLAN 逻辑交换机

要创建 Edge 上行链路，请将第 0 层路由器连接到 VLAN 交换机。

以下简单拓扑显示 VLAN 传输区域中的 VLAN 逻辑交换机。VLAN 逻辑交换机具有一个 VLAN ID，它与 Edge 的 VLAN 上行链路的 ToR 端口上的 VLAN ID 相匹配。




### 前提条件

创建一个 VLAN 逻辑交换机。请参阅为 NSX Edge 上行链路创建 VLAN 逻辑交换机。

创建一个第 0 层路由器。

### 步骤

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 选择第 0 层逻辑路由器。

4 从**配置 (Configuration)**选项卡中，添加一个新的逻辑路由器端口。

5 键入该端口的名称，例如，uplink。

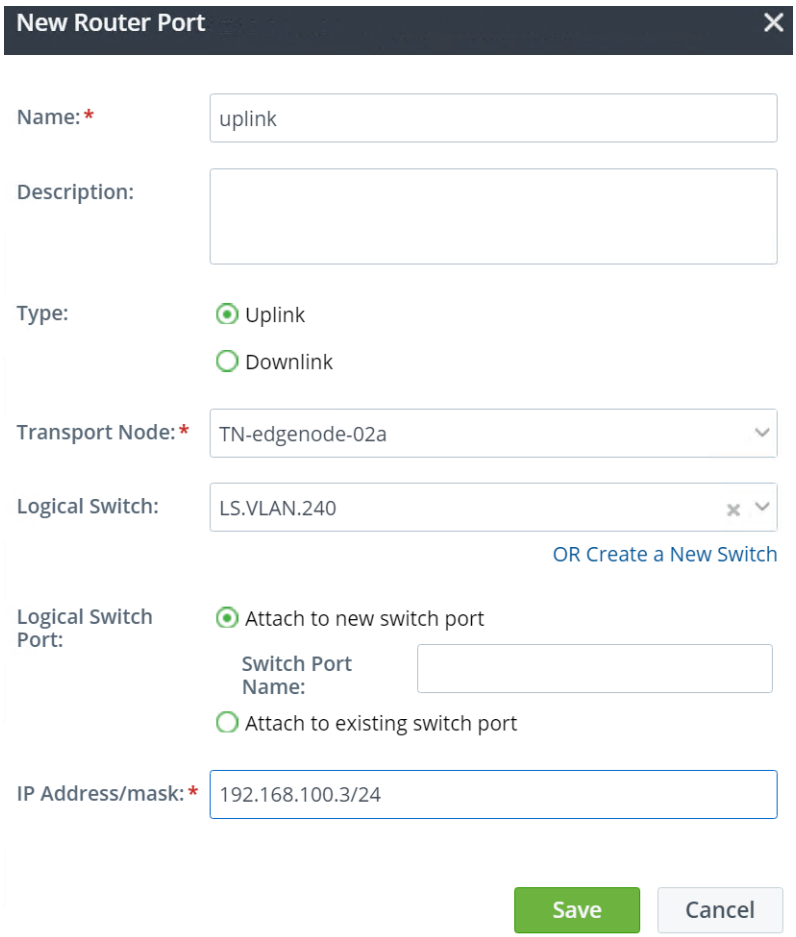
6 选择**上行链路 (Uplink)**类型。

7 选择一个 Edge 传输节点。

8 选择一个 VLAN 逻辑交换机。

9 以 CIDR 格式键入与 ToR 交换机上连接的端口位于同一子网中的 IP 地址。

例如：




将为第 0 层路由器添加一个新的上行链路端口。

**后续步骤**

配置 BGP 或静态路由。

## 验证 Tier-0 逻辑路由器和 TOR 连接

要使路由在 Tier-0 路由器的上行链路上正常工作，必须建立到架顶式设备的连接。

**前提条件**

- 确认 Tier-0 逻辑路由器连接到 VLAN 逻辑交换机。请参见将第 0 层逻辑路由器连接到 VLAN 逻辑交换机。

**步骤**

1 登录到 NSX Manager CLI。

2 在 NSX Edge 上，运行 `get logical-routers` 命令以查找 Tier-0 服务路由器的 VRF 编号。

```
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 运行 `vrf <number>` 命令以进入 Tier-0 服务路由器上下文。

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

4 在 Tier-0 服务路由器上，运行 `get route` 命令并确保在路由表中显示预期的路由。

请注意，到 TOR 的路由显示为 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 对 TOR 执行 ping 操作。

```
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
```

将在 Tier-0 逻辑路由器和物理路由器之间发送数据包以验证连接。

**后续步骤**

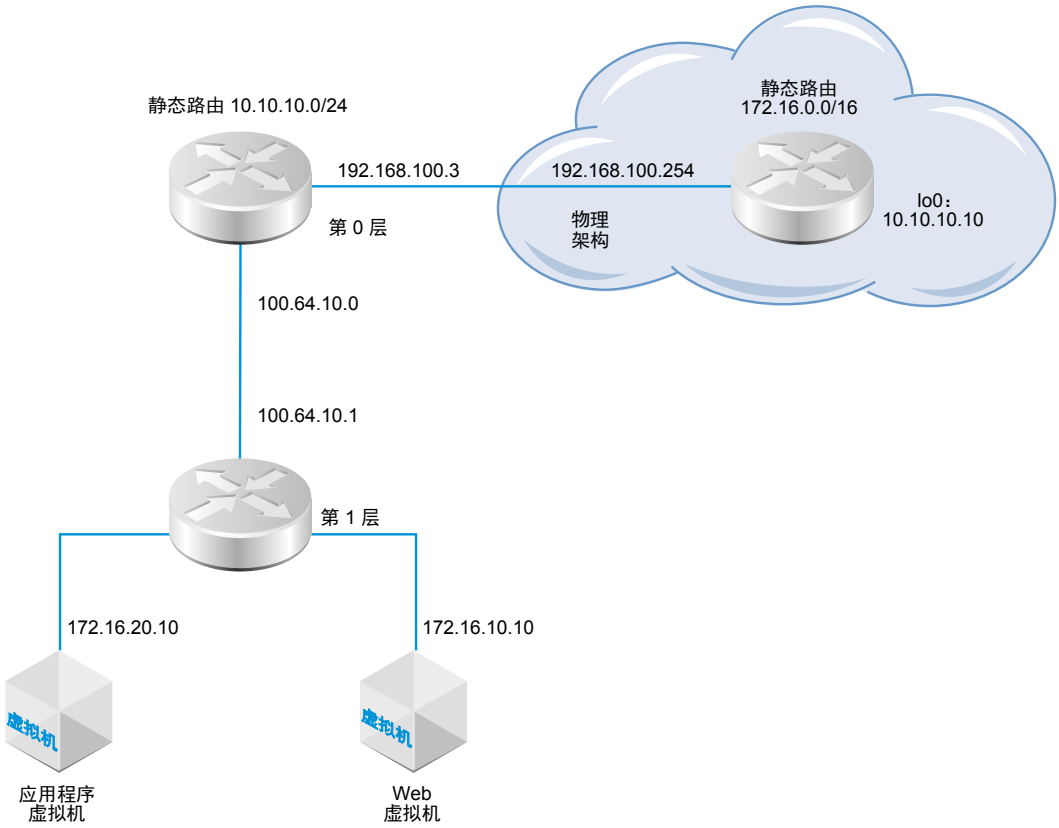
根据您的网络要求，您可以配置静态路由或 BGP。请参见配置静态路由或在第 0 层逻辑路由器上配置 BGP。

# 配置静态路由

您可以在第 0 层路由器上配置到外部网络的静态路由。在配置静态路由后，不需要将该路由从第 0 层播发到第 1 层，因为第 1 层路由器自动具有到它连接的第 0 层路由器的静态默认路由。

静态路由拓扑显示一个第 0 层逻辑路由器，它具有到物理架构中的 10.10.10.0/24 前缀的静态路由。出于测试目的，在外部路由器环回接口上配置了 10.10.10.10/32 地址。外部路由器具有到 172.16.0.0/16 前缀的静态路由以到达应用程序程序和 Web 虚拟机。

**图 5-2. 静态路由拓扑**




**前提条件**

- 确认连接了物理路由器和第 0 层逻辑路由器。请参阅验证 Tier-0 逻辑路由器和 TOR 连接。
- 确认配置了第 1 层路由器以播发连接的路由。请参阅创建第 1 层逻辑路由器。

**步骤**

1. 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。
2. 从导航面板中选择**路由 (Routing)**。
3. 选择第 0 层逻辑路由器。
4. 单击**路由 (Routing)**选项卡，然后从下拉菜单中选择**静态路由 (Static Route)**。
5. 选择**添加 (Add)**。
6. 以 CIDR 格式输入一个网络地址。

   例如，10.10.10.0/24。
7. 单击**插入行 (Insert Row)**以添加一个下一跃点 IP 地址。

   例如，192.168.100.254。
8. 单击**保存 (Save)**。

   将在该行中显示新创建的静态路由网络地址。

### 后续步骤

检查是否正确配置了静态路由。请参阅验证静态路由。

## 验证静态路由

可以使用 CLI 验证是否连接了静态路由。您还必须验证外部路由器是否可以 ping 通内部虚拟机，以及内部虚拟机是否可以 ping 通外部路由器。

### 前提条件

确认配置了一个静态路由。请参见配置静态路由。

### 步骤

1 登录到 NSX Manager CLI。

2 确认该静态路由。

a 获取服务路由器 UUID 信息。

```
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 从输出中找到 UUID 信息。

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

c 验证静态路由是否正常工作。

```
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 从外部路由器中，对内部虚拟机执行 ping 操作以确认可通过 NSX-T 覆盖网络访问这些虚拟机。

a 连接到外部路由器。

```
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 测试网络连接。

```
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  100.64.1.1 (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 从这些虚拟机中，对外部 IP 地址执行 ping 操作。

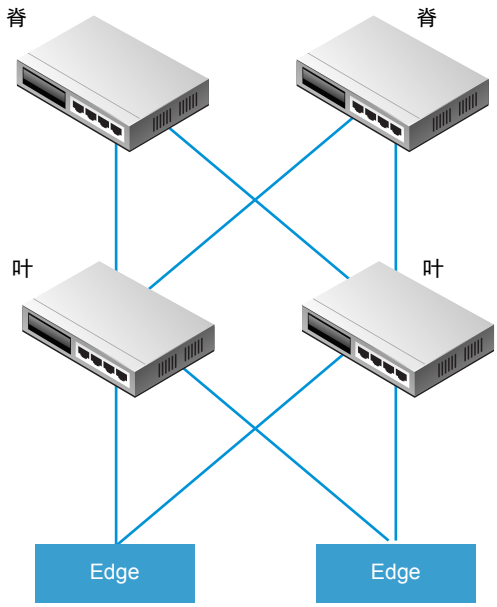
```
ping 10.10.10.10
```

```
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
```

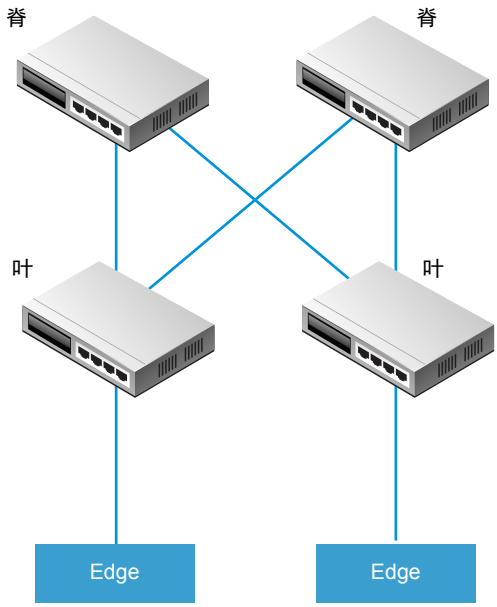
# BGP 配置选项

要充分利用 Tier-0 逻辑路由器，必须在 Tier-0 路由器和外部架顶式对等项之间使用 BGP 为拓扑配置冗余和对称性。这种设计有助于确保在链路和节点发生故障时保持连接。

共有两种配置模式：活动-活动和活动-备用。下图显示了两种对称配置选项。在每个拓扑中显示了两个 NSX Edge 节点。对于活动-活动配置，在创建 Tier-0 上行链路端口时，您可以将每个上行链路端口与最多 8 个 NSX Edge 传输节点相关联。每个 NSX Edge 节点可以具有两个上行链路。




选项 1




选项 2

对于选项 1，在配置物理叶节点路由器时，它们应该与 NSX Edge 之间具有 BGP 邻居关系。路由重新分发应包括相同的网络前缀并具有到所有 BGP 邻居的相等 BGP 衡量指标。在 Tier-0 逻辑路由器配置中，所有叶节点路由器应配置为 BGP 邻居。

在配置 Tier-0 路由器的 BGP 邻居时，如果未指定本地地址（源 IP 地址），则将 BGP 邻居配置发送到与 Tier-0 逻辑路由器上行链路关联的所有 NSX Edge 节点。如果配置了本地地址，则将配置发送到上行链路具有该 IP 地址的 NSX Edge 节点。

对于选项 1，如果上行链路位于 NSX Edge 节点上的同一子网中，则可以忽略本地地址。如果 NSX Edge 节点上的上行链路位于不同的子网中，则应该在 Tier-0 路由器的 BGP 邻居配置中指定本地地址以防止将配置发送到所有关联的 NSX Edge 节点。

对于选项 2，请确保 Tier-0 逻辑路由器配置包括 Tier-0 服务路由器的本地 IP 地址。叶节点路由器仅配置了它们作为 BGP 邻居直接连接到的 NSX Edge。

## 在第 0 层逻辑路由器上配置 BGP

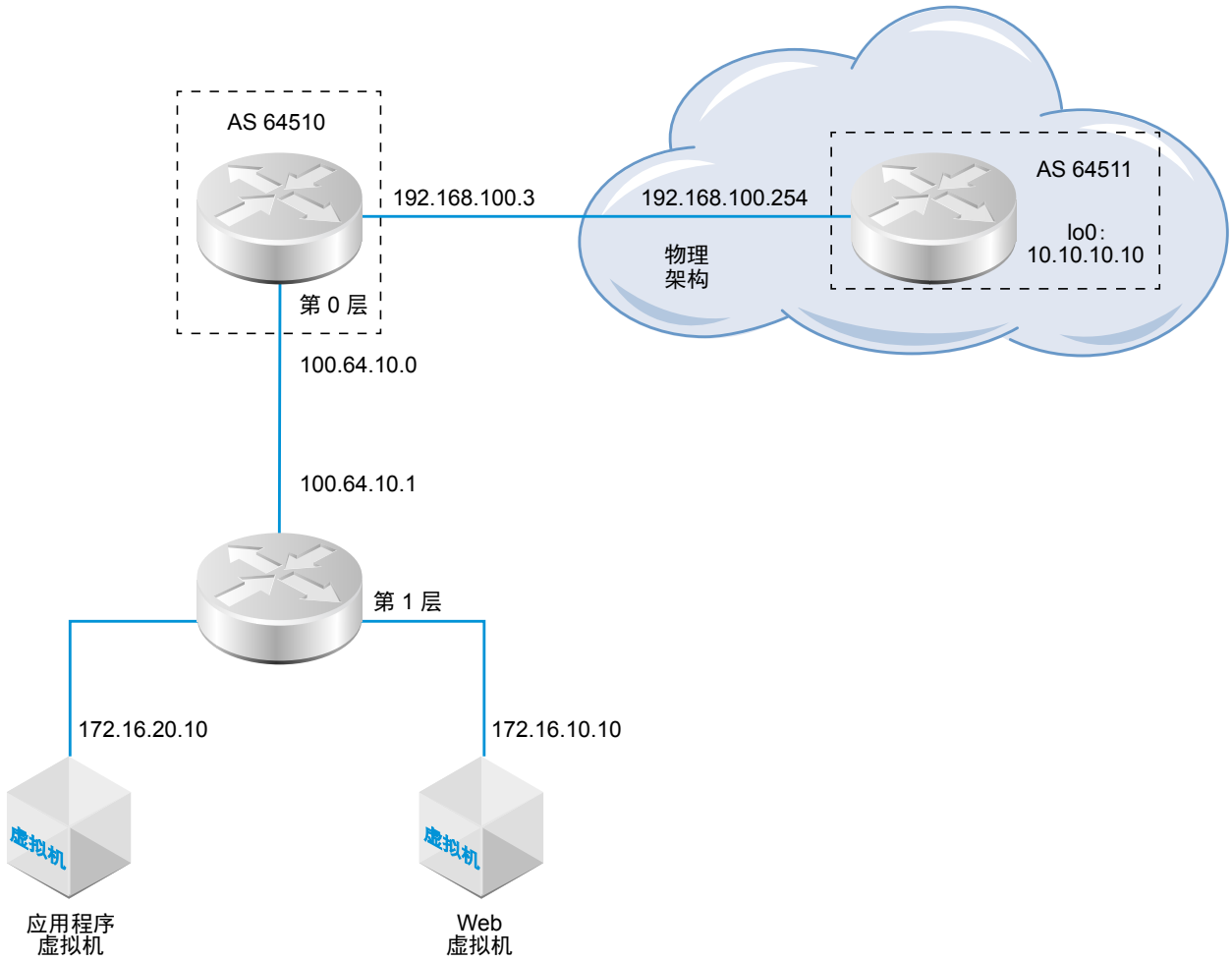
要在您的虚拟机和外界之间启用访问，您可以在第 0 层逻辑路由器和您的物理基础架构中的路由器之间配置外部 BGP (eBGP) 连接。

在配置 BGP 时，您必须为第 0 层逻辑路由器配置本地自治系统 (Autonomous System, AS) 编号。例如，以下拓扑显示本地 AS 编号为 64510。您还必须配置物理路由器的远程 AS 编号。在该示例中，远程 AS 编号为 64511。远程邻居 IP 地址为 192.168.100.254。该邻居必须位于与第 0 层逻辑路由器上的上行链路相同的 IP 子网中。不支持 BGP 多跃点。

出于测试目的，在外部路由器环回接口上配置了 10.10.10.10/32 地址。

**注** 系统会自动从在第 0 层逻辑路由器的上行链路上配置的 IP 地址中选择用于在 Edge 节点上形成 BGP 会话的路由器 ID。当路由器 ID 变化时，Edge 节点上的 BGP 会话可能会抖动。当删除为路由器 ID 自动选择的 IP 地址时，或者删除分配有此 IP 的逻辑路由器端口时，可能会发生这种情况。

**图 5-3. BGP 连接拓扑**




**前提条件**

- 确认配置了第 1 层路由器以播发连接的路由。请参阅在第 1 层逻辑路由器上配置路由播发。严格来说，这并不是 BGP 配置的必备条件，但如果您具有双层拓扑并打算将第 1 层网络重新分发到 BGP，则需要执行该步骤。
- 确认配置了一个第 0 层路由器。请参阅创建第 0 层逻辑路由器。
- 确保第 0 层逻辑路由器已从第 1 层逻辑路由器中发现路由。请参阅验证 Tier-0 路由器是否发现来自 Tier-1 路由器的路由。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 选择第 0 层逻辑路由器。

4 单击**路由 (Routing)**选项卡，然后从下拉菜单中选择 **BGP**。

5 单击**编辑 (Edit)**以配置本地 AS 编号，然后单击**保存 (Save)**。

例如，64.510。

6 单击**状态 (Status)**切换按钮以启用 BGP。

“状态”按钮必须显示为“已启用”。

7 （可选）配置路由聚合，启用正常重新启动以及启用 ECMP。

只有在与第 0 层路由器关联的 Edge 群集只有一个 Edge 节点时，才支持正常重新启动。

8 单击**保存 (Save)**。

9 在“邻居”部分下面，单击**添加 (Add)**以添加一个 BGP 邻居。

10 输入邻居 IP 地址。

例如，192.168.100.254。

11 （可选）从下拉菜单中选择一个本地地址。

12 输入远程 AS 编号。

例如，64.511。

13 （可选）配置定时器（保持活动状态时间和抑制时间）和密码。

14 （可选）添加一个地址系列并配置路由筛选和路由映射。

**后续步骤**

测试 BGP 是否正常工作。请参阅从 Tier-0 服务路由器中验证 BGP 连接。

## 从 Tier-0 服务路由器中验证 BGP 连接

可以使用 CLI 从 Tier-0 服务路由器中验证是否建立了到邻居的 BGP 连接。

**前提条件**

确认配置了 BGP。请参见在第 0 层逻辑路由器上配置 BGP。

**步骤**

1 登录到 NSX Manager CLI。

2 在 NSX Edge 上，运行 `get logical-routers` 命令以查找 Tier-0 服务路由器的 VRF 编号。

```
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 运行 `vrf <number>` 命令以进入 Tier-0 服务路由器上下文。

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

4 验证 BGP 状态是否为 `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
```

**后续步骤**

从外部路由器中检查 BGP 连接。请参见验证南北向连接和路由重新分发。

# 在第 0 层逻辑路由器上配置 BFD

BFD（双向转发检测）是一种可以检测转发路径故障的协议。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 选择第 0 层逻辑路由器。

4 单击**路由**选项卡，然后从下拉菜单中选择 **BFD**。

5 单击**编辑**以配置 BFD。

6 单击**状态**切换按钮以启用 BFD。

您可以选择更改全局 BFD 属性**接收间隔**、**发送间隔**和**声明失效间隔**。

7 （可选）在“静态路由下一跃点的 BFD 对等项”下面，单击**添加**以添加一个 BFD 对等项。

指定对等项 IP 地址并将管理状态设置为**已启用**。您可以选择覆盖全局 BFD 属性**接收间隔**、**发送间隔**和**声明失效间隔**。

# 在 Tier-0 逻辑路由器上启用路由重新分发

在启用路由重新分发时，Tier-0 逻辑路由器开始与其北向路由器共享指定的路由。

**前提条件**

- 确认连接了 Tier-0 和 Tier-1 逻辑路由器，以便通告 Tier-1 逻辑路由器网络以在 Tier-0 逻辑路由器上重新分发这些网络。请参见连接第 0 层和第 1 层。
- 如果要从路由重新分发中筛选特定的 IP 地址，请确认配置了路由映射。请参见创建路由映射。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 选择第 0 层逻辑路由器。

4 单击**路由**选项卡，然后从下拉菜单中选择**路由重新分发**。

5 单击**添加**以满足路由重新分发条件。

| 选项 | 说明 |
| --- | --- |
| **名称和说明** | 为路由重新分发指定一个名称。您可以选择提供相应的说明。<br>例如，名称为 advertise-to-bgp-neighbor。 |
| **源** | 选中要重新分发的源路由的复选框。<br>静态 - Tier-0 静态路由。<br>NSX 已连接 - Tier-1 连接的路由。<br>NSX 静态 - Tier-1 静态路由。将自动创建这些静态路由。<br>Tier-0 NAT - 在 Tier-0 逻辑路由器上配置 NAT 时生成的路由。<br>Tier-1 NAT - 在 Tier-1 逻辑路由器上配置 NAT 时生成的路由。 |
| **路由映射** | （可选）分配路由映射以从路由重新分发中筛选一组 IP 地址。 |

6 单击**保存**。

7 单击**状态**切换按钮以启用路由重新分发。

“状态”按钮将显示为“已启用”。

## 验证南北向连接和路由重新分发

可以使用 CLI 验证是否发现 BGP 路由。也可以从外部路由器中检查是否可以访问 NSX-T 连接的虚拟机。

### 前提条件

- 确认配置了 BGP。请参见在第 0 层逻辑路由器上配置 BGP。
- 确认将 NSX-T 静态路由设置为要进行重新分发。请参见在 Tier-0 逻辑路由器上启用路由重新分发。

### 步骤

1 登录到 NSX Manager CLI。

2 查看从外部 BGP 邻居中发现的路由。

```
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 从外部路由器中，检查是否发现了 BGP 路由，以及是否可以通过 NSX-T 覆盖网络访问虚拟机。

a 列出 BGP 路由。

```
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 从外部路由器中，对 NSX-T 连接的虚拟机执行 ping 操作。

```
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 检查通过 NSX-T 覆盖网络的路径。

```
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 从内部虚拟机中，对外部 IP 地址执行 ping 操作。

```
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
```

**后续步骤**

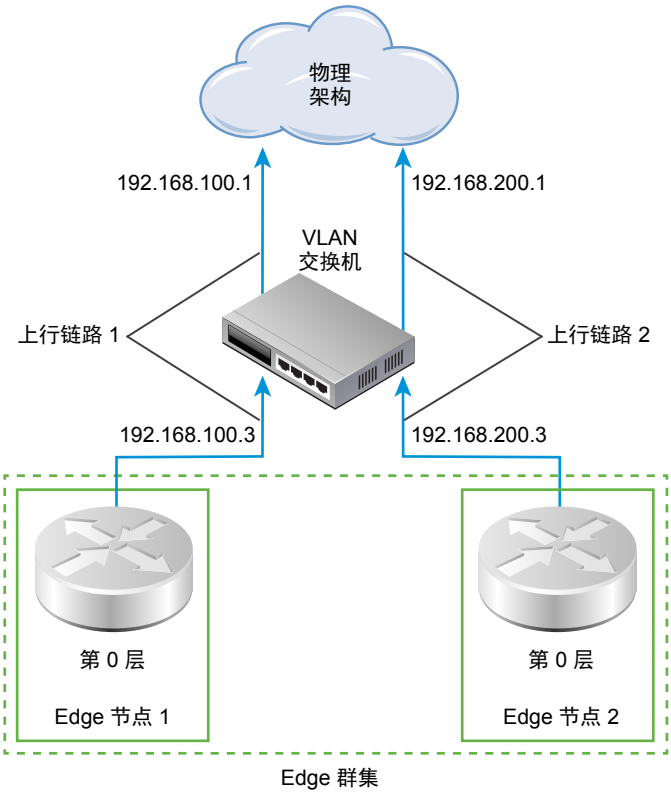
配置额外的路由功能，例如，ECMP。

# 了解 ECMP 路由

等价多路径 (Equal cost multi-path, ECMP) 路由协议将上行链路添加到第 0 层逻辑路由器，并为 Edge 群集中的每个 Edge 节点配置该上行链路以增加南北向通信带宽。ECMP 路由路径用于流量负载平衡并为发生故障的路径提供容错。

将自动创建从连接到逻辑交换机的虚拟机到在其中实例化第 0 层逻辑路由器的 Edge 节点的 ECMP 路径。最多支持 8 个 ECMP 路径。

**图 5-4. ECMP 路由拓扑**




例如，拓扑显示 Edge 群集中的两个第 0 层逻辑路由器。每个第 0 层逻辑路由器位于一个 Edge 节点中，并且这些节点是群集的一部分。上行链路端口 192.168.100.3 和 198.168.200.3 定义了传输节点如何连接到逻辑交换机以访问物理网络。如果启用了 ECMP 路由路径，这些路径将连接到逻辑交换机的虚拟机与 Edge 群集中的两个 Edge 节点相连。多个 ECMP 路由路径提高了网络吞吐量和弹性。

## 为第二个 Edge 节点添加上行链路端口

在启用 ECMP 之前，您必须配置一个上行链路以将第 0 层逻辑路由器连接到 VLAN 逻辑交换机。

**前提条件**

- 确认配置了一个传输区域和两个传输节点。请参阅《 NSX-T 安装指南》。
- 确认配置了两个 Edge 节点和一个 Edge 群集。请参阅《 NSX-T 安装指南》。
- 确认具有上行链路的 VLAN 逻辑交换机。请参阅为 NSX Edge 上行链路创建 VLAN 逻辑交换机。

- 确认配置了一个第 0 层逻辑路由器。请参阅创建第 0 层逻辑路由器。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 选择第 0 层逻辑路由器。

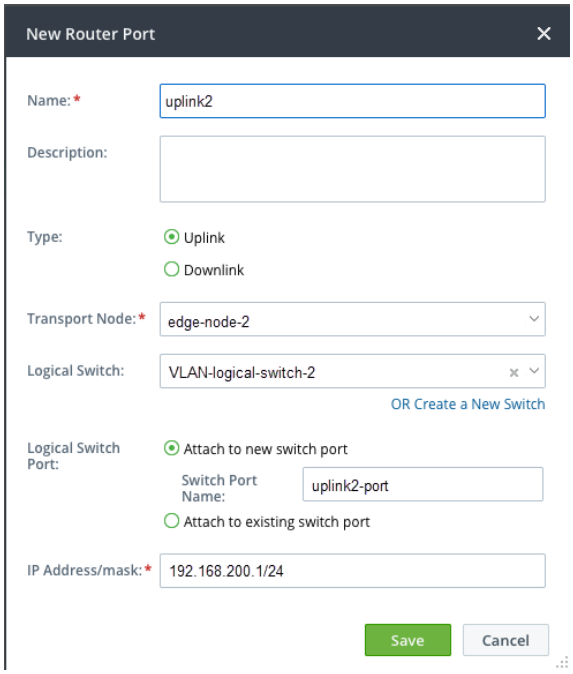
4 单击**配置 (Configuration)**选项卡以添加一个路由器端口。

5 单击**添加 (Add)**。

6 填写路由器端口详细信息。

| 选项 | 说明 |
| --- | --- |
| **名称** | 指定路由器端口的名称。 |
| **说明** | 提供额外的说明以指出该端口用于 ECMP 配置。 |
| **类型** | 接受默认类型**上行链路 (Uplink)**。 |
| **传输节点** | 从下拉菜单中分配主机传输节点。 |
| **逻辑交换机** | 从下拉菜单中分配 VLAN 逻辑交换机。 |
| **逻辑交换机端口** | 指定新交换机端口的名称。<br>也可以使用现有的交换机端口。 |
| **IP 地址/掩码** | 输入与 ToR 交换机上连接的端口位于同一子网中的 IP 地址。 |

示例路由器端口配置。




7 单击**保存 (Save)**。

将在第 0 层路由器和 VLAN 逻辑交换机中添加新的上行链路端口，并在两个 Edge 节点上配置第 0 层逻辑路由器。

**后续步骤**

为第二个邻居创建 BGP 连接并启用 ECMP 路由。请参阅添加第二个 BGP 邻居并启用 ECMP 路由。

## 添加第二个 BGP 邻居并启用 ECMP 路由

在启用 ECMP 路由之前，您必须添加一个 BGP 邻居并使用新添加的上行链路信息配置该邻居。

**前提条件**

确认第二个 Edge 节点配置了上行链路端口。请参阅为第二个 Edge 节点添加上行链路端口。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 选择第 0 层逻辑路由器。

4 单击**路由 (Routing)**选项卡，然后从下拉菜单中选择 **BGP**。

5 在“邻居”部分下面，单击**添加 (Add)**以添加一个 BGP 邻居。

6 输入邻居 IP 地址。

例如，192.168.200.254。

7 从下拉菜单中选择本地地址。

例如，上行链路 2 192.168.200.1。

8 输入远程 AS 编号。

例如，64.511。

9 单击**保存 (Save)**。

将显示新添加的 BGP 邻居。

10 单击“BGP 配置”部分旁边的**编辑 (Edit)**。

11 单击 **ECMP** 切换按钮以启用 ECMP。

“状态”按钮必须显示为“已启用”。

12 单击**保存 (Save)**。

多个 ECMP 路由路径将连接到逻辑交换机的虚拟机与 Edge 群集中的两个 Edge 节点相连。

**后续步骤**

测试 ECMP 路由连接是否正常工作。请参阅验证 ECMP 路由连接。

# 验证 ECMP 路由连接

可以使用 CLI 验证是否建立了到邻居的 ECMP 路由连接。

### 前提条件

确认配置了 ECMP 路由。请参见为第二个 Edge 节点添加上行链路端口和添加第二个 BGP 邻居并启用 ECMP 路由。

### 步骤

1 登录到 NSX Manager CLI。

2 获取分布式路由器 UUID 信息。

```
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 从输出中找到 UUID 信息。

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

4 键入 Tier-0 分布式路由器的 VRF。

```
vrf 5
```

5 验证 Tier-0 分布式路由器是否连接到 Edge 节点。

```
get forwarding
```

例如，edge-node-1 和 edge-node-2。

6 输入 **exit** 以退出 vrf 上下文。

7 打开 Tier-0 逻辑路由器的活动控制器。

8 验证是否连接了控制器节点上的 Tier-0 分布式路由器。

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

UUID 的路由类型应显示为 NSX_CONNECTED。

9 在两个 Edge 节点上启动 SSH 会话。

10 启动一个会话以捕获数据包。

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

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

11 导航到控制中心并双击 httpdata11.bat 和 httpdata12.bat 脚本。

将向两个 Web 虚拟机发送大量 HTTP 请求，并看到对流量进行哈希处理以发送到两个使用 Edge 节点的路径，这表明 ECMP 正常工作。

12 停止捕获会话。

```
del capture session 0
```

13 移除 bat 脚本。

# 创建 IP 前缀列表

IP 前缀列表包含一个或多个分配了访问权限以进行路由通告的 IP 地址。该列表中的 IP 地址是按顺序进行处理的。IP 前缀列表是通过输入或输出方向的 BGP 邻居筛选器或路由映射引用的。

例如，您可以将 IP 地址 192.168.100.3/27 添加到 IP 前缀列表中，并拒绝将路由重新分发到北向路由器。这意味着，除了 192.168.100.3/24 IP 地址以外，所有其他 IP 地址将共享该路由器。

也可以在 IP 地址后面附加小于或等于 (le) 或大于或等于 (ge) 修饰符以允许或限制路由重新分发。例如，192.168.100.3/27 ge 24 le 30 修饰符与长度大于或等于 24 位且小于或等于 30 位的子网掩码相匹配。

注 路由的默认操作为**拒绝**。在创建一个前缀列表以拒绝或允许特定的路由时，如果要允许所有其他路由，请务必创建一个具有空网络地址和**允许**操作的 IP 前缀。

### 前提条件

确认配置了一个第 0 层逻辑路由器。请参见创建第 0 层逻辑路由器。

### 步骤

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 选择第 0 层逻辑路由器。

4 单击**路由**选项卡，然后从下拉菜单中选择 **IP 前缀列表**。

5 选择**添加**。

6 指定 IP 前缀列表的名称。

7 单击**插入行**以按 CIDR 格式添加一个网络地址。

例如，192.168.100.3/27。

8 从下拉菜单中选择**拒绝**或**允许**。

可以根据您的要求允许或拒绝通告每个 IP 地址。

9 （可选）在 le 或 ge 修饰符中设置一定范围的 IP 地址位数。

例如，将 le 修饰符设置为 30 并将 ge 修饰符设置为 24。

10 单击**保存**。

将在该行中显示新创建的 IP 前缀列表。

# 创建路由映射

路由映射由一系列 IP 前缀列表、BGP 路径属性和关联的操作组成。路由器扫描该序列以查找匹配的 IP 地址。如果找到一个匹配的地址，路由器将执行操作，而不再扫描其他地址。

可以在 BGP 邻居级别和路由重新分发中引用路由映射。如果在路由映射中引用 IP 前缀列表并应用了路由映射允许或拒绝操作，在路由映射序列中指定的操作将覆盖 IP 前缀列表中指定的操作。

### 前提条件

确认配置了一个 IP 前缀列表。请参阅创建 IP 前缀列表。

### 步骤

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 选择第 0 层逻辑路由器。

4 选择**路由 (Routing) > 路由映射 (Route Maps)**。

5 单击**添加 (Add)**。

6 输入路由映射的名称和可选说明。

7 单击**添加 (Add)**以在路由映射中添加一个条目。

8 选择一个或多个 IP 前缀列表。

9 （可选）设置 BGP 属性。

| BGP 属性 | 说明 |
| --- | --- |
| AS 路径前置 | 在路径前面放置一个或多个 AS（自主系统）编号以使路径更长，因此，通常不是首选的路径。 |
| MED | 多出口区分符向外部对等项指示 AS 的首选路径。 |

| BGP 属性 | 说明 |
| --- | --- |
| 权重 | 设置权重以影响路径选择。范围是 0-65535。 |
| 团体 | 使用 aa:nn 格式指定团体，例如 300:500。或者，使用下拉菜单选择以下选项之一：<br>■ NO_EXPORT_SUBCONFED - 不播发到 EBGP 对等项。<br>■ NO_ADVERTISE - 不播发到任何对等项。<br>■ NO_EXPORT - 不播发到外部 BGP 联合。 |

10 在“操作”列中，选择**允许 (Permit)**或**拒绝 (Deny)**。

您可以允许或禁止 IP 前缀列表中的 IP 地址播发其地址。

11 单击**保存 (Save)**。

# 网络地址转换

6

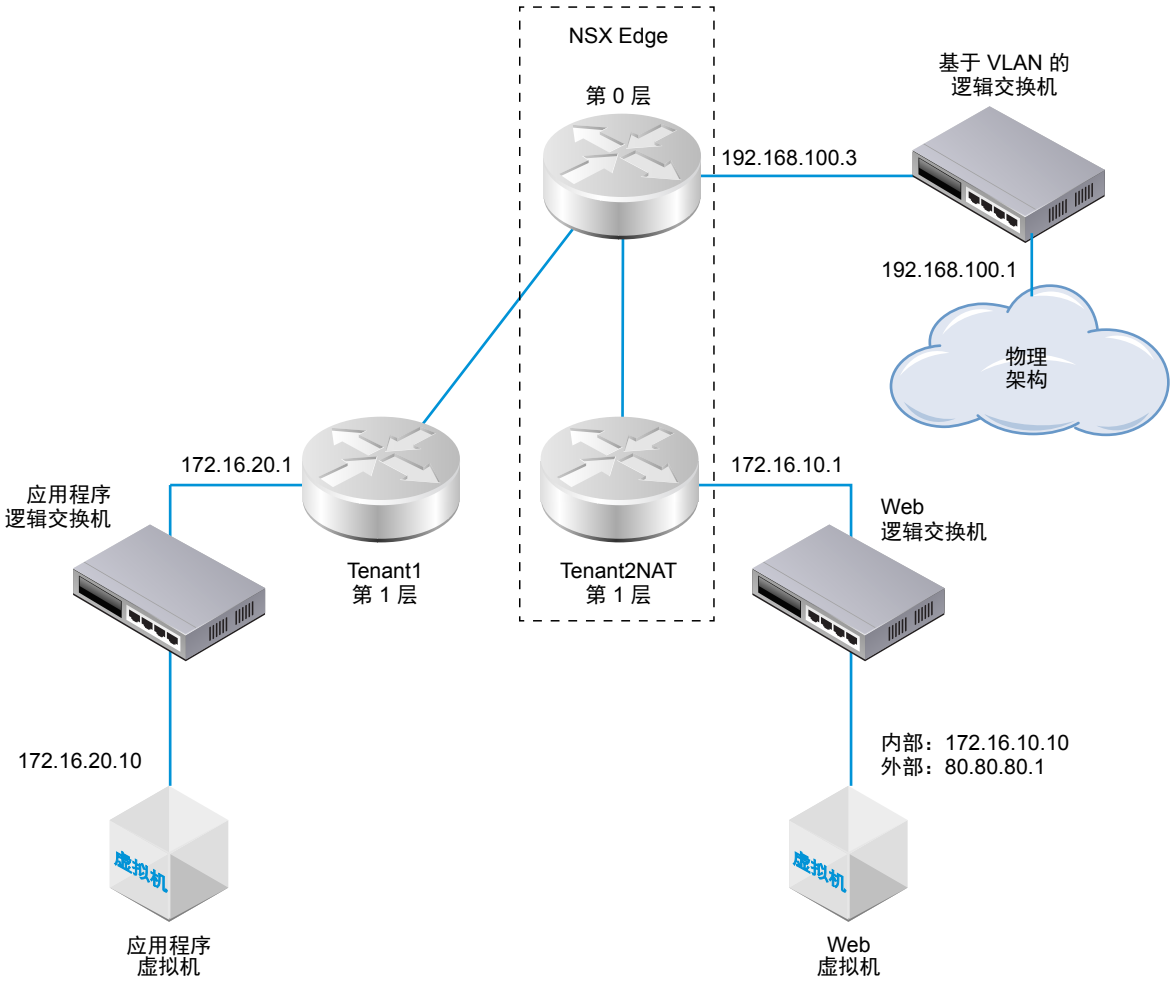
可以在 Tier-0 和 Tier-1 逻辑路由器上配置 NSX-T 中的网络地址转换 (NAT)。

例如，下图显示了两个在 Tenant2NAT 上配置了 NAT 的 Tier-1 逻辑路由器。Web 虚拟机简单配置为将 172.16.10.10 作为其 IP 地址，并将 172.16.10.1 作为其默认网关。

在 Tenant2NAT 逻辑路由器到 Tier-0 逻辑路由器的上行链路连接上强制实施了 NAT。

要启用 NAT 配置，Tenant2NAT 必须在 NSX Edge 群集上具有一个服务组件。因此，Tenant2NAT 显示在 NSX Edge 中。为了进行比较，可以将 Tenant1 放在 NSX Edge 外部，因为它不使用任何 Edge 服务。

图 6-1. NAT 拓扑




本章讨论了以下主题：

- Tier-1 NAT
- 第 0 层 NAT

# Tier-1 NAT

Tier-1 逻辑路由器支持源 NAT 和目标 NAT。

## 在第 1 层路由器上配置源 NAT

源 NAT (Source NAT, SNAT) 更改数据包 IP 标头中的源地址。它还可能会更改 TCP/UDP 标头中的源端口。典型的用途是将专用 (rfc1918) 地址/端口更改为离开您的网络的数据包的公共地址/端口。

在该示例中，从 Web 虚拟机中收到数据包时，Tenant2NAT 第 1 层路由器将数据包的源端口从 172.16.10.10 更改为 80.80.80.1。通过使用公共源地址，专用网络外部的目标可以路由回原始源。

**前提条件**

- 第 0 层路由器必须将一个上行链路连接到基于 VLAN 的逻辑交换机。请参阅将第 0 层逻辑路由器连接到 VLAN 逻辑交换机。
- 第 0 层路由器必须在到物理架构的上行链路上配置路由（静态或 BGP）和路由重新分发。请参阅配置静态路由、在第 0 层逻辑路由器上配置 BGP 和在 Tier-0 逻辑路由器上启用路由重新分发。
- 第 1 层路由器必须分别配置一个到第 0 层路由器的上行链路。Tenant2NAT 必须由一个 Edge 群集提供支持。请参阅连接第 0 层和第 1 层。
- 第 1 层路由器必须配置了下行链路端口和路由播发。请参阅为第 1 层逻辑路由器添加下行链路端口和在第 1 层逻辑路由器上配置路由播发。
- 虚拟机必须连接到正确的逻辑交换机。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 选择**路由 (Routing)**。

3 单击一个要在其中配置 NAT 的第 1 层逻辑路由器。

4 在“NAT”下面，单击**添加 (Add)**。

5 对于“操作”，请选择“SNAT”。

6 选择协议类型。

默认情况下，将选择**任何协议 (Any Protocol)**。

7 对于“源 IP”地址，请输入虚拟机的内部 IP 地址。

如果将“源 IP”保留空白，将转换路由器的下行链路端口上的所有源。在该示例中，源 IP 为 172.16.10.10。

8 对于“转换的 IP”地址，请输入虚拟机的外部 IP 地址。

请注意，不需要在虚拟机上配置外部/转换的 IP 地址。仅 NAT 路由器需要了解转换的 IP 地址。

在该示例中，转换的 IP 地址为 80.80.80.1。

9 对于“目标 IP”地址，您可以将其保留空白，或者输入一个 IP 地址。

如果将“目标 IP”保留空白，NAT 将应用于本地子网外部的所有目标。

10 启用规则。

11 （可选）启用日志记录。

将在“NAT”下面列出新规则。例如：


**后续步骤**

配置第 1 层路由器以播发 NAT 路由。

要将第 0 层路由器上游的 NAT 路由播发到物理架构，请配置第 0 层路由器以播发第 1 层 NAT 路由。

## 在第 1 层路由器上配置目标 NAT

目标 NAT 更改数据包 IP 标头中的目标地址。它还可能会更改 TCP/UDP 标头中的目标端口。它的典型用途是，将具有公共地址/端口目标的入站数据包重定向到您的网络中的专用 IP 地址/端口。

在该示例中，从应用程序虚拟机中收到数据包时，Tenant2NAT 第 1 层路由器将数据包的目标端口从 172.16.10.10 更改为 80.80.80.1。通过使用公共目标地址，可以从专用网络外部连接到专用网络中的目标。

**前提条件**

- 第 0 层路由器必须将一个上行链路连接到基于 VLAN 的逻辑交换机。请参阅将第 0 层逻辑路由器连接到 VLAN 逻辑交换机。
- 第 0 层路由器必须在到物理架构的上行链路上配置路由（静态或 BGP）和路由重新分发。请参阅配置静态路由、在第 0 层逻辑路由器上配置 BGP 和在 Tier-0 逻辑路由器上启用路由重新分发。
- 第 1 层路由器必须分别配置一个到第 0 层路由器的上行链路。Tenant2NAT 必须由一个 Edge 群集提供支持。请参阅连接第 0 层和第 1 层。
- 第 1 层路由器必须配置了下行链路端口和路由播发。请参阅为第 1 层逻辑路由器添加下行链路端口和在第 1 层逻辑路由器上配置路由播发。
- 虚拟机必须连接到正确的逻辑交换机。

**步骤**

1. 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。
2. 选择**路由 (Routing)**。
3. 单击一个要在其中配置 NAT 的第 1 层逻辑路由器。
4. 在“NAT”下面，单击**添加 (Add)**。
5. 对于“操作”，请选择“DNAT”。

6 选择协议类型。

默认情况下，将选择**任何协议 (Any Protocol)**。

7 对于“目标 IP”地址，请输入虚拟机的外部 IP 地址。

在该示例中，目标 IP 地址为 80.80.80.1。请注意，不需要在虚拟机上配置外部 IP 地址。仅 NAT 路由器需要了解外部 IP 地址。

8 对于“转换的 IP”地址，请输入虚拟机的内部 IP 地址。

必须在虚拟机上配置内部 IP 地址。

在该示例中，内部/转换的 IP 地址为 172.16.10.10。

9 对于“源 IP”地址，您可以将其保留空白，或者输入一个 IP 地址。

如果将“源 IP”保留空白，NAT 将应用于本地子网外部的所有源。

10 启用规则。

11 （可选）启用日志记录。

将在“NAT”下面列出新规则。例如：


**后续步骤**

配置第 1 层路由器以播发 NAT 路由。

要将第 0 层路由器上游的 NAT 路由播发到物理架构，请配置第 0 层路由器以播发第 1 层 NAT 路由。

## 将第 1 层 NAT 路由播发到上游第 0 层路由器

通过播发第 1 层 NAT 路由，可以使上游第 0 层路由器发现这些路由。

**步骤**

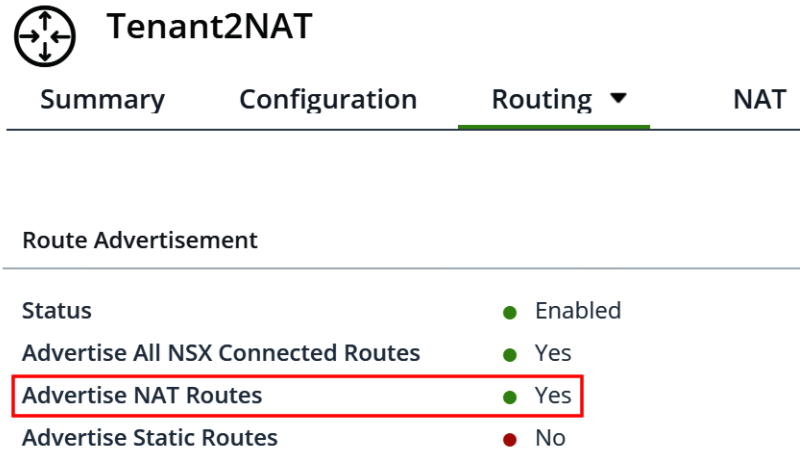
1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 选择**路由 (Routing)**。

3 单击一个已在其中配置 NAT 的第 1 层逻辑路由器。

4 从该第 1 层路由器中，选择**路由 > 路由播发 (Routing > Route Advertisement)**。

5 编辑路由播发规则以启用 NAT 路由播发。




**后续步骤**

将第 1 层 NAT 路由从第 0 层路由器播发到上游物理架构。

## 将 Tier-1 NAT 路由通告到物理架构

通过从 Tier-0 路由器中通告 Tier-1 NAT 路由，可以使上游物理架构发现这些路由。

**步骤**

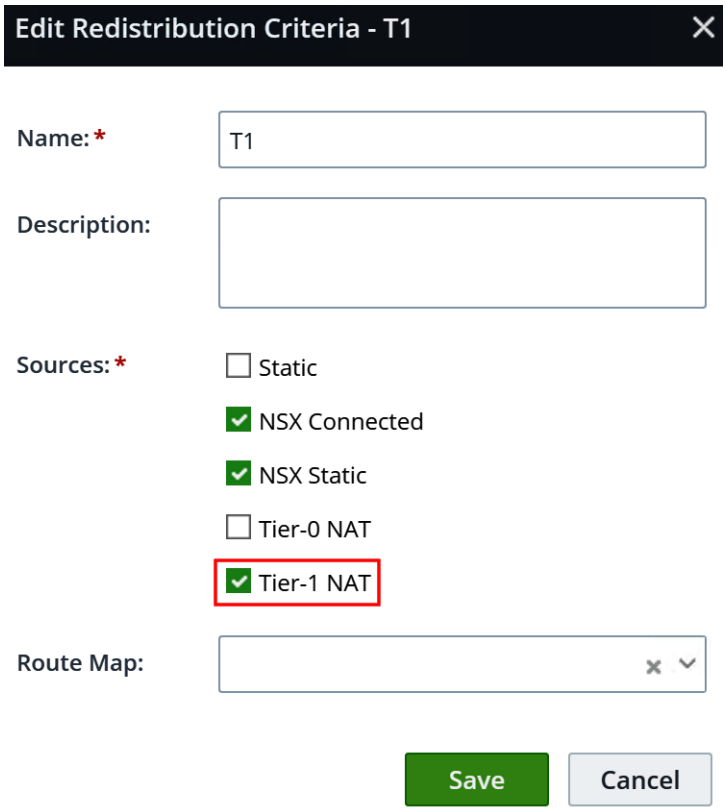
1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 选择**路由**。

3 单击一个连接到已在其中配置 NAT 的 Tier-1 路由器的 Tier-0 逻辑路由器。

4 从该 Tier-0 路由器中，选择**路由 > 路由重新分发**。

5 编辑路由通告规则以启用 Tier-1 NAT 路由通告。




**后续步骤**

验证 NAT 是否正常工作。

## 验证第 1 层 NAT

验证 SNAT 和 DNAT 规则是否正常工作。

**步骤**

1. 登录到 NSX Edge。
2. 运行 `get logical-routers` 以确定第 0 层服务路由器的 VRF 编号。
3. 运行 `vrf <number>` 命令以进入第 0 层服务路由器上下文。
4. 执行 show route 命令并确保显示第 1 层 NAT 地址。

```
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 如果 Web 虚拟机设置为提供网页，请确保您可以打开 http://80.80.80.1 中的网页。

6 确保物理架构中的第 0 层路由器的上游邻居可以 ping 通 80.80.80.1。

7 在 ping 仍在运行时，检查 DNAT 规则的“统计信息”列。

应该具有一个活动会话。

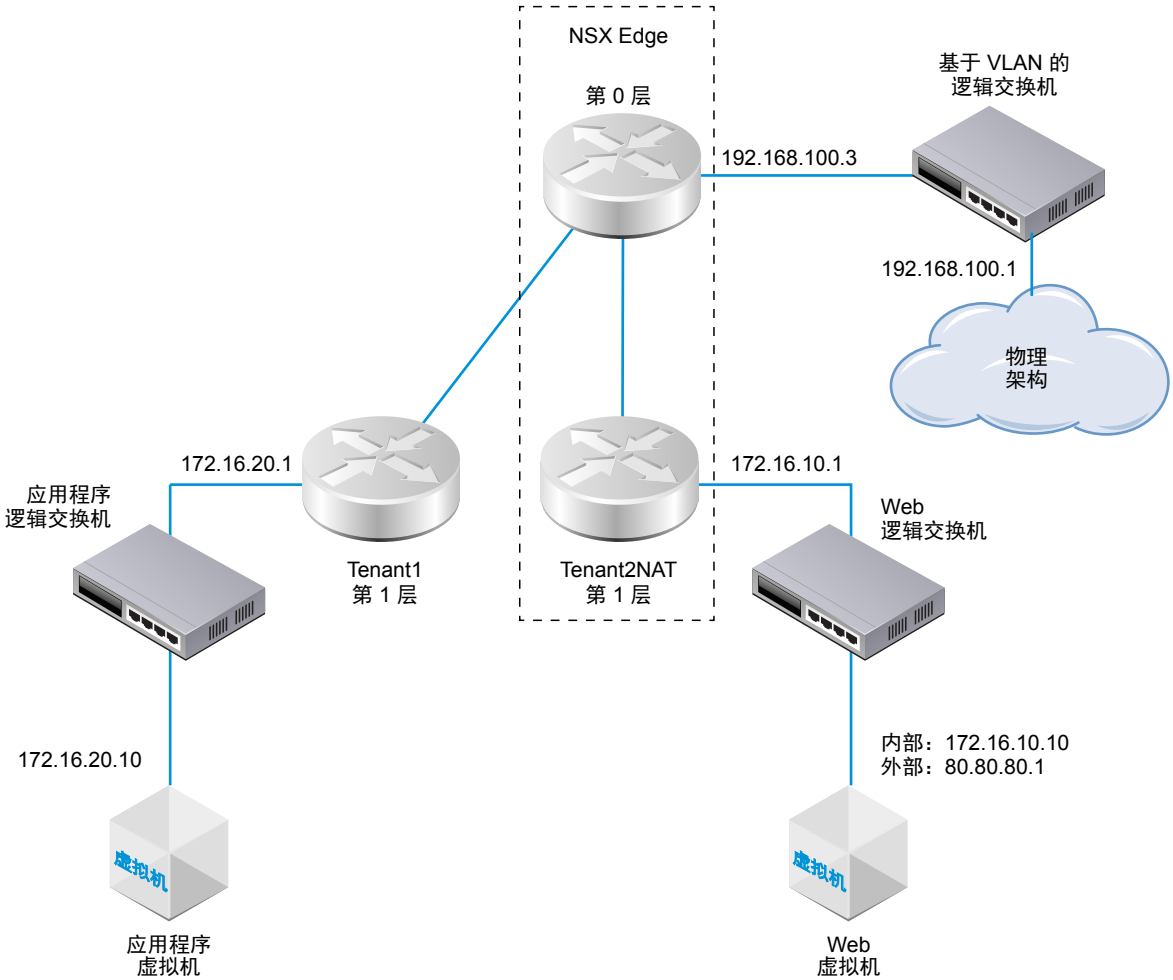
# 第 0 层 NAT

第 0 层逻辑路由器支持反射 NAT。

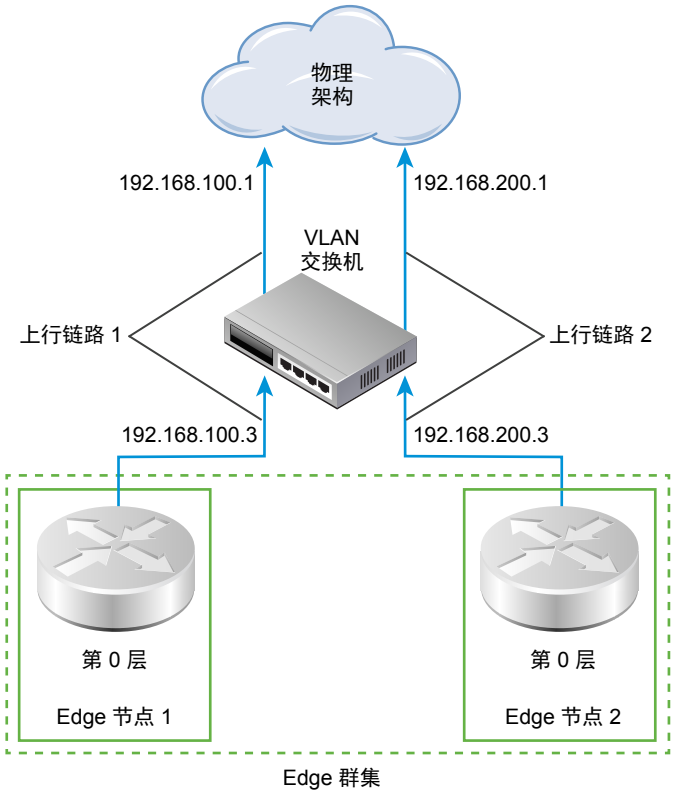
## 反射 NAT

如果第 0 层逻辑路由器在活动-活动 ECMP 模式下运行，您无法配置有状态 NAT，在该模式下，不对称的路径可能会导致出现问题。对于活动-活动 ECMP 路由器，您可以使用反射 NAT（有时称为无状态 NAT）。

在该示例中，从 Web 虚拟机中收到数据包时，Tenant2NAT 第 1 层路由器将数据包的源端口从 172.16.10.10 更改为 80.80.80.1。通过使用公共源地址，专用网络外部的目标可以路由回原始源。



不过，在涉及两个活动-活动第 0 层路由器时（如下所示），必须配置反射 NAT。




## 在第 0 层逻辑路由器上配置反射 NAT

如果第 0 层逻辑路由器在活动-活动 ECMP 模式下运行，您无法配置有状态 NAT，在该模式下，不对称的路径可能会导致出现问题。对于活动-活动 ECMP 路由器，您可以使用反射 NAT（有时称为无状态 NAT）。

### 前提条件

- 第 0 层路由器必须将两个上行链路连接到基于 VLAN 的逻辑交换机。请参阅将第 0 层逻辑路由器连接到 VLAN 逻辑交换机。
- 第 0 层路由器必须在到物理架构的上行链路上配置路由（静态或 BGP）和路由重新分发。请参阅配置静态路由、在第 0 层逻辑路由器上配置 BGP 和在 Tier-0 逻辑路由器上启用路由重新分发。
- 第 1 层路由器必须分别配置一个到第 0 层路由器的上行链路。Tenant2NAT 必须由一个 Edge 群集提供支持。请参阅连接第 0 层和第 1 层。
- 第 1 层路由器必须配置了下行链路端口和路由播发。请参阅为第 1 层逻辑路由器添加下行链路端口和在第 1 层逻辑路由器上配置路由播发。
- 虚拟机必须连接到正确的逻辑交换机。

### 步骤

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 选择**路由 (Routing)**。

3 单击一个要在其中配置反射 NAT 的第 0 层逻辑路由器。

4 在“NAT”下面，单击**添加 (Add)**。

5 对于“操作”，请选择“反射”。

6 对于“源 IP”地址，请输入虚拟机的外部 IP 地址。

在该示例中，源 IP 为 80.80.80.1。

7 对于“转换的 IP”地址，请输入虚拟机的内部 IP 地址。

在该示例中，转换的 IP 地址为 172.16.10.10。

8 对于“目标 IP”地址，您可以将其保留空白，或者输入一个 IP 地址。

如果将“目标 IP”保留空白，NAT 将应用于本地子网外部的所有目标。

9 启用规则。

10 （可选）启用日志记录。

将在“NAT”下面列出新规则。例如：


**后续步骤**

配置第 1 层路由器以播发 NAT 路由。

要将第 0 层路由器上游的 NAT 路由播发到物理架构，请配置第 0 层路由器以播发第 1 层 NAT 路由。

# 防火墙区域和防火墙规则

7

防火墙区域用于对一组防火墙规则进行分组。

防火墙区域由一个或多个单独的防火墙规则组成。每个单独的防火墙规则包含确定是应允许还是阻止数据包的说明；允许数据包使用哪些协议；允许数据包使用哪些端口，等等。区域用于多租户，例如，用于销售和工程部门的特定规则位于单独的区域中。

可以将一个区域定义为强制实施有状态或无状态规则。无状态规则被视为传统无状态 ACL。无状态区域不支持反射 ACL。建议不要在单个逻辑交换机端口上混用无状态和有状态规则，这可能会导致未定义的行为。

可以在区域中上下移动规则。对于尝试通过防火墙的任何流量，将按照区域中显示的顺序应用规则以处理数据包信息，从顶部开始并移到底部的默认规则。与数据包匹配的第一个规则将应用它配置的操作，并执行在该规则配置的选项中指定的任何处理，而忽略所有后续规则（即使后面的规则是更好的匹配项）。因此，您应该将具体的规则放在更常规的规则上面，以确保不会忽略这些规则。默认规则（位于规则表底部）是一个总括性规则；将为与任何其他规则不匹配的数据包强制实施默认规则。

本章讨论了以下主题：



## 添加防火墙规则区域

防火墙规则区域是单独编辑和保存的，用于将单独的防火墙配置应用于租户。

**步骤**

1 在导航面板中选择**防火墙 (Firewall)**。

确保您位于“常规”选项卡以添加 L3 规则。请单击“以太网”选项卡以添加 L2 规则。

2 要添加一个区域，请在第一列中单击滚轮 ( ) 图标或一个规则，然后选择**在上方添加区域 (Add Section Above)**或**在下方添加区域 (Add Section Below)**。

注 对于尝试通过防火墙的任何流量，将按照“规则”表中显示的顺序应用规则以处理数据包信息，从顶部开始并移到底部的默认规则。在某些情况下，两个或更多规则的优先级顺序在确定如何处理数据包方面可能是非常重要的。

3 输入区域名称和可选的说明。

4 选择**有状态 (Stateful)**为**否 (False)**或**是 (True)**。该选项仅适用于 L3。

无状态防火墙监控网络流量，并根据源和目标地址或其他静态值限制或阻止数据包。有状态防火墙可以监控从一端到另一端的流量流。在较高的流量负载情况下，无状态防火墙通常速度更快，性能更好。有状态防火墙在识别未授权和伪造的通信方面更好。在定义后，就不会在有状态和无状态之间进行切换。

5 选择要应用该区域的位置。

注 如果在某个区域中使用了**应用对象 (Applied To)**，它将覆盖该区域中的规则的所有**应用对象 (Applied To)**设置。

逻辑端口 - 显示所有逻辑端口

逻辑交换机 - 显示所有逻辑交换机

NS 组 - 显示所有 NS 组

6 单击可用的端口、交换机或组旁边的复选框，然后单击箭头。

项目将移到“已选择”列。

7 单击**保存 (Save)**以保存该区域。

将在**防火墙 (Firewall)**窗口中显示新添加的区域。

**后续步骤**

将防火墙规则添加到区域。

# 删除防火墙规则区域

在不再使用防火墙规则区域时，可以删除该区域。

在删除防火墙规则区域时，将删除该区域中的所有规则。不能在删除某个区域后将其重新添加到防火墙表中的其他位置。要执行该操作，必须删除该区域并发布配置。然后将已删除的区域添加到防火墙表，并重新发布配置。

**步骤**

1 在导航面板中选择**防火墙 (Firewall)**。

2 确保您位于“常规”选项卡以添加 L3 规则。

3 单击“以太网”选项卡以添加 L2 规则。

4 要删除某个区域，请在第一列中右键单击要删除的区域旁边的滚轮 。

5 单击**删除 (Delete)**以移除该区域。将立即删除该区域及其包含的所有规则。

# 启用和禁用区域规则

您可以在防火墙规则区域中启用或禁用所有规则。

**步骤**

1 在导航面板中选择**防火墙 (Firewall)**。

2 在第一列中单击滚轮图标，然后选择**禁用区域规则 (Disable Section Rules)**或**启用区域规则 (Enable Section Rules)**。

3 单击**保存 (Save)**。

# 禁用和启用区域日志

启用区域规则的日志将记录有关区域中的所有规则的数据包的信息。根据区域中的规则数，典型防火墙区域将生成大量日志信息，并且可能会影响性能。

日志存储在 vSphere ESXi 和 KVM 主机上的 /var/log/dfwpktlogs.log 文件中。

**步骤**

1 在导航面板中选择**防火墙 (Firewall)**。

2 在第一列中，单击滚轮 图标。选择**禁用区域规则的日志 (Disable Logs for Section Rules)**或**启用区域规则的日志 (Enable Logs for Section Rules)**。

3 单击**保存 (Save)**。

# 关于防火墙规则

NSX-T 使用防火墙规则指定流入和流出网络的流量处理。

防火墙提供了多组可配置的规则：第 3 层规则（ “常规”选项卡）和第 2 层规则（ “以太网”选项卡）。先处理第 2 层防火墙规则，然后再处理第 3 层规则。“配置”选项卡包含排除列表，其中包含要从防火墙强制实施中排除的逻辑交换机、逻辑端口和组。

防火墙规则是按以下方式强制实施的：

- 规则是按从上到下的顺序处理的。
- 根据规则表中的最上面规则检查每个数据包，然后向下移到表中的后续规则。
- 强制实施表中与流量参数匹配的第一个规则。

无法强制实施后续规则，因为随后将停止为该数据包搜索规则。由于这种行为，始终建议将最精细的策略放在规则表顶部。这将确保在较具体的规则之前强制实施这些规则。

默认规则（位于规则表底部）是一个总括性规则；将为与任何其他规则不匹配的数据包强制实施默认规则。在执行主机准备操作后，默认规则将设置为允许操作。这可确保虚拟机到虚拟机的通信在暂存或迁移阶段不会中断。最佳做法是将该默认规则更改为阻止操作，并通过积极控制模式强制实施访问控制（即，仅允许将防火墙规则中定义的流量传输到网络上）。

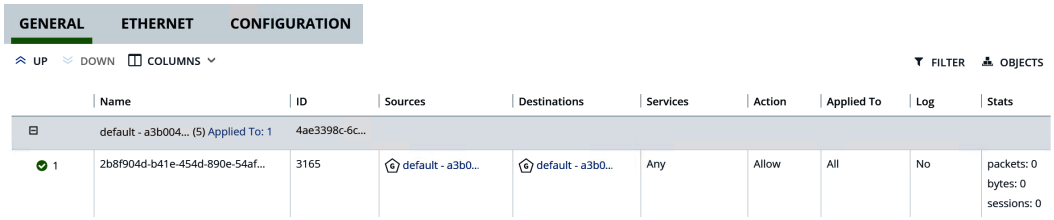
可以单击“列”旁边的下拉箭头，然后在“防火墙规则”窗口中选中要包含的列以访问防火墙规则选项。可以使用以下选项。

**表 7-1. “防火墙规则”屏幕中的列**

| 列名称 | 定义 |
| --- | --- |
| 名称 | 防火墙规则的名称。 |
| 源 | 规则源可以是 IP 或 MAC 地址或者 IP 地址以外的对象。如果未定义，源将与任何内容匹配。源或目标范围不支持 IPv6。 |
| ID | 系统为每个规则生成的唯一 ID。 |
| 方向 | 方向规则元素与数据包通过接口时的传输方向匹配。“输入”方向是指流量进入防火墙。“输出”方向是指流量离开防火墙。默认情况下，方向为输入输出（双向）。 |
| IP 协议 | 这仅适用于 L3 规则。支持 IPv4 和 IPv6。默认值为“二者”。 |
| 目标 | 受规则影响的连接的目标 IP 或 MAC 地址/网络掩码。如果未定义，目标将与任何内容匹配。源或目标范围不支持 IPv6。 |
| 服务 | 对于 L3，服务可以是预定义的端口协议组合。对于 L2，服务可以是以太网类型。对于 L2 和 L3，您可以手动定义新的服务或服务组。如果未指定，服务将与任何内容匹配。 |
| 操作(所需) | 规则应用的操作可以是允许、阻止或拒绝。 |
| 应用对象 | 定义该规则的适用范围。如果未定义，范围将是所有逻辑端口。如果在某个区域中添加了“应用对象”，它将覆盖规则。 |
| 日志 | 可以禁用或启用日志记录。日志存储在 ESX 和 KVM 主机上的 /var/log/dfwpktlogs.log 文件中。 |
| 统计信息 | 这是一个只读字段，其中显示了字节数、数据包计数和会话数。 |
| 备注 | 规则的备注。 |

下面是默认防火墙规则并显示了一部分列选项。

**图 7-1. “防火墙规则”窗口**




# 添加防火墙规则

防火墙是一个网络安全系统，它根据预定的防火墙规则监视和控制入站和出站的网络流量。

防火墙规则是在 NSX Manager 范围内添加的。然后，可以使用“应用对象”字段缩小要应用规则的范围。您可以在源级别和目标级别为每个规则添加多个对象，以帮助减少要添加的防火墙规则的总数。

注 默认情况下，规则与任何源、目标和服务规则元素的默认值匹配，从而与所有接口和流量方向匹配。如果要限制规则对特定接口或流量方向的影响，必须在规则中指定该限制。

**前提条件**

要使用一组地址，请先手动将每个虚拟机的 IP 和 MAC 地址与其逻辑交换机相关联。

**步骤**

1 在导航面板中选择**防火墙 (Firewall)**。

确保您位于“常规”选项卡以添加 L3 规则。请单击“以太网”选项卡以添加 L2 规则。

2 要添加一个规则，请在第一列中单击滚轮 (⚙) 图标，然后在列表底部选择**添加规则 (Add Rule)**。

将显示一个新行以定义防火墙规则。

注 对于尝试通过防火墙的任何流量，将按照“规则”表中显示的顺序应用规则以处理数据包信息，从顶部开始并移到底部的默认规则。在某些情况下，两个或更多规则的优先级顺序在确定如何处理数据包方面可能是非常重要的。

3 将在区域顶部添加一个新规则。如果要将规则添加到区域中的特定位置，请选择一个规则。在第一列中单击滚轮 (⚙) 图标，然后选择**在上方插入规则 (Insert Rule Above)**或**在下方插入规则 (Insert Rule Below)**。

将显示一个新行以定义防火墙规则。

4 在**名称 (Name)**列的右上角，单击铅笔图标。在“编辑名称”对话框中输入规则名称。

将显示一个具有指定名称的规则。

5 指向新规则的**源 (Sources)**单元格，单击铅笔图标，然后选择规则的源。如果未定义，源将与任何内容匹配。将显示**编辑源 (Edit Sources)**对话框。

注 在创建新的防火墙规则时，您可以拖放对象以用于“源”、“目标”、“服务”和“应用对象”字段，而不是每次都选择这些对象。这可以帮助加快规则创建过程，尤其是经常重复使用相同的对象时。

为此，请单击“防火墙规则”窗口左上角的**对象**，从列表中选择对象类型，然后将所需的对象拖放到右侧的字段，即防火墙规则中的“源”。

**表 7-2. “编辑源”窗口**

| 选项 | 说明 |
| --- | --- |
| IP 或 MAC 地址 | 在以逗号分隔的列表中输入多个 IP 或 MAC 地址。该列表最多可以包含 255 个字符。支持 IPv4 和 IPv6 格式。 |
| 对象 | 单击箭头并选择对象。<br>1 选择 IP 集、逻辑端口、逻辑交换机或 NS 组。<br>将显示选定容器的可用对象。<br>2 选择一个或多个对象，然后单击箭头。要选择所有可用的对象，请单击“可用”旁边的复选框，然后单击箭头。<br>3 这些对象将移到“已选择”列。<br>4 单击**确定 (OK)**。 |

6 指向新规则的**目标 (Destinations)**单元格。如果未定义，目标将与任何内容匹配。将显示**编辑目标 (Edit Destinations)**对话框。

**表 7-3. “编辑目标”窗口**

| 选项 | 说明 |
| --- | --- |
| IP 或 MAC 地址 | 您可以在以逗号分隔的列表中输入多个 IP 或 MAC 地址。该列表最多可以包含 255 个字符。支持 IPv4 和 IPv6 格式。 |
| 对象 | 单击箭头并选择对象。<br>1 您可以选择 IP 集、逻辑端口、逻辑交换机或 NS 组。<br>将显示选定容器的可用对象。<br>2 选择一个或多个对象，然后单击箭头。要选择所有可用的对象，请单击“可用”旁边的复选框，然后单击箭头。<br>3 这些对象将移到“已选择”列。<br>4 单击**确定 (OK)**。 |

7 指向新规则的**服务 (Service)**单元格。如果未定义，服务将与任何内容匹配。

将显示**编辑服务 (Edit Services)**对话框。该列表已显示很多预定义的服务，但您不限于选择这些服务。

8 要选择预定义的服务，请选择一个或多个可用的对象，然后单击箭头。单击**确定 (OK)**。

9 要定义新的服务，请单击**新建 (New)**。将显示“NS 服务”对话框。

| 选项 | 说明 |
| --- | --- |
| **名称** | 命名新服务。 |
| **说明** | 描述新服务。 |
| **服务类型** | ■ ALG<br>■ ICMP<br>■ IP<br>■ L4 端口集<br>■ IGMP |
| **协议** | 选择一个可用的协议。 |
| **源端口** | 输入源端口。 |

| 选项 | 说明 |
| --- | --- |
| **目标端口** | 选择目标端口。 |
| **将现有服务分组** | 单击单选按钮以添加现有的组服务。 |

10 指向**操作 (Action)**单元格，然后单击铅笔图标。该参数是必需的。将显示“编辑操作”对话框。

| 选项 | 说明 |
| --- | --- |
| **允许** | 允许具有指定的源、目标和协议的所有 L3 或 L2 流量通过当前防火墙上下文。与规则匹配并接受的数据包将通过系统，就好像没有防火墙一样。 |
| **丢弃** | 丢弃具有指定的源、目标和协议的数据包。丢弃数据包是一个静默操作，不会向源或目标系统发送通知。丢弃数据包将导致重试连接，直到达到重试阈值。 |
| **拒绝** | 拒绝具有指定的源、目标和协议的数据包。拒绝数据包是一种较友好的数据包阻止方式，因为将向发送方发送“无法到达目标 (destination unreachable)”消息。如果协议是 TCP，则会发送 TCP RST 消息。对于 UDP、ICMP 和其他 IP 连接，发送包含管理上被禁止的代码的 ICMP 消息。使用“拒绝”的一个好处是，在仅尝试一次后，就会向发送应用程序通知无法建立连接。 |

11 指向**应用对象 (Applied To)**单元格，然后单击铅笔图标。将显示“编辑应用对象”对话框。

从下拉列表中选择对象类型。单击**确定 (OK)**。

12 指向**日志 (Log)**单元格，然后单击铅笔图标。默认情况下，将禁用日志记录。选择**是 (Yes)**以启用日志记录，或者选择**否 (No)**以禁用日志记录。日志存储在 ESX 和 KVM 主机上的 /var/log/dfwpktlogs.log 文件中。也可以在此处添加备注。请注意，如果选择**是 (Yes)**，将记录与该规则匹配的所有会话的日志。启用日志记录功能可能会影响性能。

13 要使一个或多个规则生效，请单击**保存 (Save)**。

在单击**保存 (Save)**之前，您可以添加多个规则。

# 删除防火墙规则

防火墙是一个网络安全系统，它根据预定的防火墙规则监视和控制入站和出站的网络流量。可以添加和删除自定义规则。

**步骤**

1 在导航面板中选择**防火墙 (Firewall)**。

确保您位于“常规”选项卡以添加 L3 规则。请单击“以太网”选项卡以添加 L2 规则。

2 右键单击要移动的规则的编号。

将显示一个下拉列表。

3 选择**删除 (Delete)**。

将删除该防火墙规则。

4 单击**保存 (Save)**以使更改生效。

将删除该规则。

# 编辑默认分布式防火墙规则

您可以编辑应用于与任何用户定义的防火墙规则均不匹配的流量的默认防火墙设置。

默认防火墙设置应用于与任何用户定义的防火墙规则均不匹配的流量。分布式防火墙默认规则显示在集中式防火墙用户界面上。默认第 3 层规则显示在“常规”选项卡下面，而默认第 2 层规则显示在“以太网”选项卡下面。

默认分布式防火墙规则允许所有 L3 和 L2 流量通过基础架构中的所有准备的群集。默认规则始终位于规则表的底部，无法删除或添加。不过，您可以将规则的“操作”元素从“允许”更改为“丢弃”或“拒绝”（不建议），并指示是否应记录该规则的流量。

**步骤**

1 单击**防火墙 (Firewall)**。

将显示“常规防火墙”屏幕。

2 确保您位于**常规 (General)**选项卡中以编辑默认 L3 规则。请单击**以太网 (Ethernet)**选项卡以编辑 L2 规则。

3 在**操作 (Action)**列下面，展开该部分，然后选择以下选项之一：

- 允许 - 允许具有指定的源、目标和协议的所有 L3 或 L2 流量通过当前防火墙上下文。与规则匹配并接受的数据包将通过系统，就好像没有防火墙一样。
- 丢弃 - 丢弃具有指定的源、目标和协议的数据包。丢弃数据包是一个静默操作，不会向源或目标系统发送通知。丢弃数据包将导致重试连接，直到达到重试阈值。
- 拒绝 - 拒绝具有指定的源、目标和协议的数据包。拒绝数据包是一种较友好的数据包阻止方式，因为将向发送方发送“无法到达目标 (destination unreachable)”消息。如果协议是 TCP，则会发送 TCP RST 消息。对于 UDP、ICMP 和其他 IP 连接，发送包含管理上被禁止的代码的 ICMP 消息。使用“拒绝”的一个好处是，在仅尝试一次后，就会向发送应用程序通知无法建立连接。

注 不建议选择**拒绝 (Reject)**以作为默认规则的操作。

4 在**日志 (Log)**列下面，展开该部分，然后选择**是 (Yes)**以启用日志记录，或者选择**否 (No)**以禁用日志记录。也可以在此处添加备注。请注意，如果选择“是”，将记录与该规则匹配的所有会话的日志。启用日志记录功能可能会影响性能。

5 单击**保存 (Save)**并确认更改。

# 更改防火墙规则的顺序

规则是按从上到下的顺序处理的。您可以更改列表中的规则顺序。

对于尝试通过防火墙的任何流量，将按照“规则”表中显示的顺序应用规则以处理数据包信息，从顶部开始并移到底部的默认规则。在某些情况下，两个或更多规则的优先级顺序在确定流量流方面可能是非常重要的。

自定义规则可以在表中上下移动，而默认规则始终位于表底部且无法移动。

**步骤**

1 在导航面板中选择**防火墙 (Firewall)**。

确保您位于“常规”选项卡以添加 L3 规则。请单击“以太网”选项卡以添加 L2 规则。

2 右键单击要移动的规则的编号。

3 选择**上移 (Move Up)**或**下移 (Move Down)**。

规则将向上或向下移动一个位置。

# 筛选防火墙规则

您可以使用一些条件筛选规则集，从而轻松修改规则。

**步骤**

1 在“防火墙”窗口的左上角，单击**筛选器 (Filter)** FILTER。

将显示“筛选器”对话框。

2 要进行筛选，您可以根据以下条件进行搜索：

- 源 - 搜索入站防火墙规则。
- 目标 - 搜索出站防火墙规则。
- 应用对象 - 根据应用对象条件搜索规则。
- 服务 - 从该列表中，选择要允许或阻止的应用程序或服务。该列表已显示很多通用服务，但您不限于选择这些服务。请使用“服务”单元格添加尚未显示的任何其他服务或应用程序。

3 单击所需的搜索条件，它将显示在框的顶部。

4 选择搜索类型：

- IP 集 - 该选项列出规则源或目标的所有 IP 地址。
- 逻辑端口 - 筛选逻辑端口。
- 逻辑交换机 - 筛选逻辑交换机。
- NS 组 - 筛选 NS 组。

将在框中显示筛选结果。

# 从防火墙实施中排除对象

可以从防火墙规则中排除逻辑端口、逻辑交换机或 NS 组。

在创建一个具有防火墙规则的区域后，您可能希望将某个 NSX-T 设备端口从防火墙规则中排除。

**步骤**

1 在导航面板中选择**防火墙 (Firewall)**。选择**配置 (Configuration)**选项卡。

将显示排除列表屏幕。

2 在窗口右上角选择**对象 (Objects)**。

3 从下拉列表中，选择**逻辑端口 (Logical Ports)**、**逻辑交换机 (Logical Switch)**或 **NS 组 (NSGroup)**。

4 双击要从防火墙规则中排除的特定端口、交换机或组。要关闭“对象”对话框，请再次单击**对象 (Objects)**。

将在“排除列表”中填充要排除的对象的名称和类型。

5 要从排除列表中移除一个对象，请单击 **x**。

6 单击**保存 (Save)**。

# 8 配置组和服务

您可以配置组以划分对象。

您可以在防火墙规则中使用以下组：

- IP 集
- MAC 集
- 服务组
- NS 组；它可能包括 IP 集、MAC 集、逻辑端口、逻辑交换机以及其他 NS 组

此外，在创建传输节点时，您还可以创建 IP 池以分配 IP 地址。

本章讨论了以下主题：



## 创建 IP 集

IP 集是一组 IP 地址，可以用作防火墙规则中的源和目标。

IP 集可以包含各个 IP 地址、IP 范围和子网的组合。您可以指定 IPv4 和/或 IPv6 地址。IP 集可以是 NS 组的成员。

**注** 防火墙规则的源或目标范围不支持 IPv6。

**步骤**

1. 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。
2. 从导航面板中选择**清单 (Inventory) > 组 (Groups)**。
3. 选择主面板顶部的 **IP 集**。
4. 单击**添加**。

5 输入名称。

6 （可选）输入说明。

7 输入各个地址或地址范围。

8 单击**保存**。

# 创建 IP 池

在创建 L3 子网时，您可以使用 IP 池分配 IP 地址或子网。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**清单 (Inventory) > 组 (Groups)**。

3 选择主面板顶部的 **IP 池**。

4 单击**添加**。

5 输入名称。

6 （可选）输入说明。

7 单击**添加**。

8 输入 IP 范围。

将鼠标悬停在任何单元格的右上角，然后单击铅笔图标以编辑该单元格。

9 （可选）输入网关。

10 输入具有后缀的 CIDR IP 地址。

11 （可选）输入 DNS 服务器。

12 （可选）输入 DNS 后缀。

13 单击**保存**。

# 创建 MAC 集

MAC 集是一组 MAC 地址，可以用作第 2 层防火墙规则中的源和目标以及 NS 组的成员。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**清单 (Inventory) > 组 (Groups)**。

3 选择主面板顶部的 **MAC 集**。

4 单击**添加**。

5 输入名称。

6 （可选）输入说明。

7 输入 MAC 地址。

8 单击**保存**。

# 创建 NS 组

您可以配置一个 NS 组以包含 IP 集、MAC 集、逻辑端口、逻辑交换机和其他 NS 组的组合。您可以在防火墙规则（以及 `Applied To` 字段）中将 NS 组指定为源和目标。

NS 组具有以下特性：

- 您可以指定直接成员，它们可能是 IP 集、MAC 集、逻辑交换机、逻辑端口和 NS 组。
- 您可以为逻辑交换机和逻辑端口指定最多 5 个成员资格条件。对于每个条件，请指定标记和可选的范围。
- NS 组具有直接成员和有效成员。有效成员包括使用成员资格条件指定的成员，以及属于该 NS 组的成员的所有直接和有效成员。例如，假设 NSGroup-1 具有直接成员 LogicalSwitch-1。您添加 NSGroup-2 并将 NSGroup-1 和 LogicalSwitch-2 指定为成员。现在，NSGroup-2 具有直接成员 NSGroup-1 和 LogicalSwitch-2 以及有效成员 LogicalSwitch-1。接下来，您添加 NSGroup-3 并将 NSGroup-2 指定为成员。NSGroup-3 现在具有直接成员 NSGroup-2 以及有效成员 LogicalSwitch-1 和 LogicalSwitch-2。
- NS 组最多可以具有 500 个直接成员。
- NS 组中的建议有效成员数限制为 5000 个。超过该限制不会影响任何功能，但可能会对性能造成不利影响。在 NSX Manager 上，在 NS 组的有效成员数超过 5000 的 80% 时，将在日志文件中显示警告消息 `NS 组 xyz 即将超过最大成员限制。NS 组中的总数为... (NSGroup xyz is about to exceed the maximum member limit. Total number in NSGroup is ...)`；在该数字超过 5000 时，将显示警告消息 `NS 组 xyz 已达到最大数字限制。NS 组中的总数 = ... (NSGroup xyz has reached the maximum numbers limit. Total number in NSGroup = ...)`。在 NSX Controller 上，在 NS 组中的转换的 VIF/IP/MAC 数超过 5000 时，将在日志文件中显示警告消息`容器 xyz 已达到最大 IP/MAC/VIF 转换限制。容器中的当前转换计数 – IP:...，MAC:...，VIF:... (Container xyz has reached the maximum IP/MAC/VIF translations limit. Current translations count in Container – IPs:..., MACs:..., VIFs:...)`。NSX Manager 和 NSX Controller 每天检查两次 NS 组是否超过该限制（早晨 7 点和晚上 7 点）。

对于可作为成员添加到 NS 组的所有对象（即，逻辑交换机、逻辑端口、IP 集、MAC 集和 NS 组），您可以导航到任何对象的屏幕，然后选择**相关 (Related) > NS 组 (NSGroups)**以查看直接或间接将该对象作为成员的所有 NS 组。例如，在上面的示例中，在导航到 LogicalSwitch-1 屏幕后，选择**相关 (Related) > NS 组 (NSGroups)**将显示 NSGroup-1、NSGroup-2 和 NSGroup-3，因为所有三个组直接或间接将 LogicalSwitch-1 作为成员。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**清单 (Inventory) > 组 (Groups)**。

3 选择主面板顶部的 **NS 组 (NSGroups)**。

4 单击**添加 (Add)**。

5 输入 NS 组的名称。

6 （可选）输入说明。

7 （可选）单击**成员资格条件 (Membership Criteria)**以指定最多 5 个条件。

条件可以应用于逻辑交换机或逻辑端口。

条件可以指定标记值和/或范围值。

8 （可选）单击**成员 (Members)**以选择成员。

可用的类型是 **IP 集 (IP Set)**、**MAC 集 (MAC Set)**、**逻辑交换机 (Logical Switch)**、**逻辑端口 (Logical Port)**以及 **NS 组 (NSGroup)**。

9 单击**保存 (Save)**。

# 配置服务和服务组

您可以配置 NS 服务以指定用于匹配网络流量的参数，例如，端口和协议对。也可以使用 NS 服务在防火墙规则中允许或阻止某些类型的流量。

NS 服务可以具有以下类型：

- 以太网
- IP
- IGMP
- ICMP
- ALG
- L4 端口集

L4 端口集支持标识源端口和目标端口。您可以指定单个端口或一定范围的端口，最多为 15 个端口。

NS 服务也可以是一组其他 NS 服务。采用组形式的 NS 服务可以具有以下类型：

- 第 2 层
- 第 3 层和更高的层

在创建 NS 服务后，您无法更改类型。某些 NS 服务是预定义的。您无法修改或删除这些服务。

## 创建 NS 服务

您可以创建 NS 服务以指定网络匹配使用的特性，或者定义在防火墙规则中阻止或允许的流量类型。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**清单 (Inventory) > 服务 (Services)**。

3 单击**添加 (Add)**。

4 输入名称。

5 （可选）输入说明。

6 选择**指定协议 (Specify a protocol)**以配置单个服务，或者选择**将现有服务分组 (Group existing services)**以配置一组 NS 服务。

7 对于单个服务，请选择类型和协议。

可用的类型是**以太网 (Ether)**、**IP**、**IGMP**、**ICMP**、**ALG** 和 **L4 端口集 (L4 Port Set)**。

8 对于服务组，请为该组选择类型和成员。

可用的类型是**第 2 层 (Layer 2)**和**第 3 层和更高的层 (Layer 3 and above)**。

9 单击**保存 (Save)**。

# DHCP

9

通过使用 DHCP（动态主机配置协议），客户端可以从 DHCP 服务器中自动获取网络配置，例如，IP 地址、子网掩码、默认网关和 DNS 配置。

您可以创建 DHCP 服务器以处理 DHCP 请求，或者创建 DHCP 中继服务以将 DHCP 流量中继到外部 DHCP 服务器。

如果配置 DHCP 服务器，要提高安全性，请配置一个 DFW 规则以仅允许 UDP 端口 67 和 68 上来自有效 DHCP 服务器 IP 地址的流量通过。

**注** 将 `Logical Switch/Logical Port/NSGroup` 作为源、将 `Any` 作为目标并配置为丢弃端口 67 和 68 的 DHCP 数据包的 DFW 规则无法阻止 DHCP 流量。要阻止 DHCP 流量，请将 `Any` 配置为源和目标。

本章讨论了以下主题：



## 创建 DHCP 服务器配置文件

DHCP 服务器配置文件指定 NSX Edge 群集或 NSX Edge 群集成员。具有该配置文件的 DHCP 服务器处理来自连接到该配置文件中指定的 NSX Edge 节点的逻辑交换机上的虚拟机的 DHCP 请求。

**步骤**

1. 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。
2. 在导航面板中选择 **DHCP**。
3. 单击**服务器配置文件**，然后单击**添加**。
4. 输入名称和可选的说明。

5 从下拉菜单中选择一个 NSX Edge 群集。

6 （可选）选择该 NSX Edge 群集的成员。

您最多可以指定 2 个成员。

**后续步骤**

创建 DHCP 服务器。请参见创建 DHCP 服务器。

# 创建 DHCP 服务器

您可以创建 DHCP 服务器以处理来自连接到逻辑交换机的虚拟机的 DHCP 请求。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 在导航面板中选择 **DHCP**。

3 单击**服务器**，然后单击**添加**。

4 输入名称和可选的说明。

5 以 CIDR 格式输入 DHCP 服务器的 IP 地址及其子网掩码。

例如，输入 `192.168.1.2/24`。

6 从下拉菜单中选择一个 DHCP 配置文件。

7 （可选）输入常用的选项，例如，域名、默认网关、DNS 服务器和子网掩码。

8 （可选）输入无类静态路由选项。

9 （可选）输入其他选项。

10 单击**保存**。

11 选择新创建的 DHCP 服务器。

12 展开“IP 池”部分。

13 单击**添加**以添加 IP 范围、默认网关、租约期限、警告阈值、错误阈值、无类静态路由选项以及其他选项。

14 展开“静态绑定”部分。

15 单击**添加**以添加 MAC 地址和 IP 地址之间的静态绑定、默认网关、主机名、租约期限、无类静态路由选项以及其他选项。

**后续步骤**

将 DHCP 服务器连接到逻辑交换机。请参见将 DHCP 服务器连接到逻辑交换机。

# 将 DHCP 服务器连接到逻辑交换机

您必须将 DHCP 服务器连接到一个逻辑交换机，然后 DHCP 服务器才能处理连接到该交换机的虚拟机的 DHCP 请求。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**交换 (Switching) > 交换机 (Switches)**。

3 单击要将 DHCP 服务器连接到的逻辑交换机。

4 单击**操作 (Actions) > 连接 DHCP 服务器 (Attach DHCP Server)**。

# 将 DHCP 服务器与逻辑交换机断开连接

您可以将 DHCP 服务器与逻辑交换机断开连接以重新配置您的环境。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**交换 (Switching) > 交换机 (Switches)**。

3 单击要与 DHCP 服务器断开连接的逻辑交换机。

4 单击**操作 > 断开连接 DHCP 服务器**。

# 创建 DHCP 中继配置文件

DHCP 中继配置文件指定一个或多个外部 DHCP 服务器。在创建 DHCP 中继服务时，您必须指定一个 DHCP 中继配置文件。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 在导航面板中选择 **DHCP**。

3 单击**中继配置文件**，然后单击**添加**。

4 输入名称和可选的说明。

5 输入一个或多个外部 DHCP 服务器地址。

**后续步骤**

创建 DHCP 中继服务。请参见创建 DHCP 中继服务。

# 创建 DHCP 中继服务

您可以创建 DHCP 中继服务以中继未在 NSX-T 中创建的 DHCP 客户端和 DHCP 服务器之间的流量。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 在导航面板中选择 **DHCP**。

3 单击**中继服务**，然后单击**添加**。

4 输入名称和可选的说明。

5 从下拉菜单中选择一个 DHCP 中继配置文件。

**后续步骤**

将 DHCP 服务添加到逻辑路由器端口。请参见将 DHCP 服务添加到逻辑路由器端口。

# 将 DHCP 服务添加到逻辑路由器端口

在将 DHCP 中继服务添加到逻辑路由器端口时，连接到该端口的逻辑交换机上的虚拟机可以与中继服务中配置的 DHCP 服务器进行通信。

**前提条件**

- 确认您具有配置的 DHCP 中继服务。请参见创建 DHCP 中继服务。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**路由 (Routing)**。

3 选择连接到所需逻辑交换机的路由器，然后单击**配置**选项卡。

4 选择连接到所需逻辑交换机的路由器端口，然后单击**编辑**。

5 从 **DHCP 服务**下拉列表中选择一个 DHCP 中继服务，然后单击**保存**。

逻辑路由器端口将在 **DHCP 服务**列中显示该 DHCP 中继服务。

也可以在添加新的逻辑路由器端口时选择 DHCP 中继服务。

# 配置元数据代理

10

通过使用元数据代理服务器，虚拟机实例可以从 OpenStack Nova API 服务器中检索实例特定的元数据。

以下步骤说明了元数据代理的工作方式：

1 虚拟机将 HTTP GET 发送到 http://169.254.169.254:80 以请求某些元数据。

2 连接到与虚拟机相同的逻辑交换机的元数据代理服务器读取请求，对标头进行相应的更改，然后将请求转发到 Nova API 服务器。

3 Nova API 服务器从 Neutron 服务器中请求和接收有关虚拟机的信息。

4 Nova API 服务器查找元数据并将其发送到元数据代理服务器。

5 元数据代理服务器将元数据转发到虚拟机。

元数据代理服务器在一个 NSX Edge 节点上运行。为实现高可用性，您可以将元数据代理配置为在 NSX Edge 群集中的两个或更多 NSX Edge 节点上运行。

本章讨论了以下主题：



## 添加元数据代理服务器

通过使用元数据代理服务器，虚拟机可以从 OpenStack Nova API 服务器中检索元数据。

**前提条件**

确认您创建了一个 Edge 群集。有关详细信息，请参阅《 NSX-T 安装指南》。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 选择 **DHCP > 元数据代理 (DHCP > Metadata Proxies)**。

3 单击**添加 (Add)**。

4 输入元数据代理服务器的名称。

5 （可选）输入说明。

6 输入 Nova 服务器的 URL。

7 输入 `secret` 参数。

8 从下拉列表中选择一个 Edge 群集。

9 （可选）选择该 Edge 群集的成员。

例如：


**后续步骤**

将元数据代理服务器连接到逻辑交换机。

## 将元数据代理服务器连接到逻辑交换机

要向连接到逻辑交换机的虚拟机提供元数据代理服务，您必须将一个元数据代理服务器连接到该交换机。

**前提条件**

确认您创建了一个逻辑交换机。有关详细信息，请参阅创建逻辑交换机。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 选择 **DHCP > 元数据代理 (DHCP > Metadata Proxies)**。

3 选择一个元数据代理服务器。

4 选择**操作 (Actions) > 连接到逻辑交换机 (Attach to Logical Switch)**菜单选项。

5 从下拉列表中选择一个逻辑交换机。

也可以通过以下方法将元数据代理服务器连接到逻辑交换机：导航到**交换 > 交换机 (Switching > Switches)**，选择一个交换机，然后选择**操作 (Actions) > 连接元数据代理 (Attach Metadata Proxy)**菜单选项。

## 将元数据代理服务器与逻辑交换机断开连接

要停止为连接到逻辑交换机的虚拟机提供元数据代理服务或使用不同的元数据代理服务器，您可以将元数据代理服务器与逻辑交换机断开连接。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 选择 **DHCP > 元数据代理 (DHCP > Metadata Proxies)**。

3 选择一个元数据代理服务器。

4 选择**操作 (Actions) > 与逻辑交换机断开连接 (Detach from Logical Switch)**菜单选项。

5 从下拉列表中选择一个逻辑交换机。

也可以通过以下方法将元数据代理服务器与逻辑交换机断开连接：导航到**交换 > 交换机 (Switching > Switches)**，选择一个交换机，然后选择**操作 (Actions) > 断开连接元数据代理 (Detach Metadata Proxy)**菜单选项。

# 操作和管理

11

您可能需要更改安装的设备的配置，例如，添加许可证和证书以及更改密码。还应该执行一些日常维护任务，包括运行备份。此外，可以使用一些工具帮助您查找有关 NSX-T 基础架构和 NSX-T 创建的逻辑网络包含的设备的信息，其中包括远程系统日志记录、跟踪流和端口连接。

本章讨论了以下主题：



## 添加许可证密钥

您可以使用 NSX Manager UI 添加一个或多个许可证密钥。

可以使用以下非评估许可证类型：

- 标准
- 高级
- 企业

在安装 NSX Manager 时，预装的评估许可证将变为活动状态，有效期为 60 天。评估许可证提供了企业许可证的所有功能。您无法安装或取消分配评估许可证。

您可以安装一个或多个非评估许可证，但对于每种类型，您只能安装一个密钥。在安装标准、高级或企业许可证时，评估许可证不再可用。也可以取消分配非评估许可证。如果取消分配所有非评估许可证，将会恢复评估许可证。

如果具有多个相同许可证类型的密钥并且要合并这些密钥，您必须访问 https://my.vmware.com 并使用合并密钥功能。NSX Manager UI 不提供该功能。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**系统 (System) > 配置 (Configuration) > 许可证 (License)**。

3 单击**添加**以输入许可证密钥。

4 单击**保存**。

# 管理用户帐户

NSX-T 设备具有本地管理用户 admin。您无法创建或删除用户。

## 更改 admin 密码

您可以在任何 NSX-T 设备上更改 admin 用户的密码。

**步骤**

1 登录到 NSX Manager CLI。

2 运行 `set user` 命令。

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

密码必须满足以下密码复杂性要求：

- 长度至少为 8 个字符
- 至少一个大写字符
- 至少一个小写字符

- 至少一个数字字符
- 至少一个特殊字符

## 帐户锁定

在登录尝试连续失败 5 次后，将锁定管理员帐户 15 分钟。

对于 NSX Manager、NSX Controller 和 NSX Edge 节点，在登录尝试第 5 次连续失败后，将锁定管理员帐户 15 分钟。要重置锁定的帐户，请等待 15 分钟以再次登录。这种行为是有意设置的，因为这可通过观察登录失败消息从“密码不正确 (incorrect password)”变为“帐户已锁定 (account locked)”来防止攻击者了解某个帐户存在。

注 这适用于通过 SSH 或控制台的管理员登录。

# 设置证书

您可以在 NSX Manager 中生成一个证书签名请求 (Certificate Signing Request, CSR)，并将其发送到证书颁发机构 (Certificate Authority, CA) 以获取服务器证书。

还可以使用 CSR 生成自签名证书。如果当前具有证书或 CA 证书，您可以导入以使用该证书。也可以导入包含吊销的证书的证书吊销列表 (Certificate Revocation List, CRL)。

## 创建证书签名请求文件

证书签名请求 (CSR) 是包含特定信息的加密文本，例如，组织名称、公用名称、城市以及国家/地区。您可以将 CSR 文件发送到证书颁发机构 (CA) 以申请数字身份证书。

**前提条件**

- 收集填写 CSR 文件所需的信息。您必须知道服务器的 FQDN、组织单位、组织、城市、省/直辖市/自治区以及国家/地区。
- 确认具有公钥和私钥对。

**步骤**

1. 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。
2. 从导航面板中选择**系统 (System) > 设置 (Settings)**。
3. 单击**证书 (Certificates)**选项卡。
4. 从下拉菜单中选择 **CSR (CSRs)**。
5. 单击**生成 CSR (Generate CSR)**。

6 填写 CSR 文件详细信息。

| 选项 | 说明 |
| --- | --- |
| **名称** | 指定证书的名称。 |
| **公用名称** | 输入服务器的完全限定域名 (Fully Qualified Domain Name, FQDN)。<br>例如，test.vmware.com。 |
| **组织名称** | 输入具有相应后缀的组织名称。<br>例如，VMware Inc.。 |
| **组织单位** | 输入组织中处理该证书的部门。<br>例如，IT 部门。 |
| **城市** | 添加组织所在的城市。<br>例如，帕罗奥多。 |
| **省/直辖市/自治区** | 添加组织所在的省/直辖市/自治区。<br>例如，加利福尼亚。 |
| **国家/地区** | 添加组织所在的国家/地区。<br>例如，美国 (US)。 |
| **消息算法** | 为证书设置加密算法。<br>RSA 加密 - 用于数字签名和消息加密。因此，在创建加密令牌时比 DSA 慢，但分析和验证该令牌时较快。该加密的解密速度较慢，而加密速度较快。<br>DSA 加密 - 用于数字签名。因此，在创建加密令牌时比 RSA 快，但分析和验证该令牌时较慢。该加密的解密速度较快，而加密速度较慢。 |
| **密钥大小** | 设置加密算法的密钥位大小。<br>使用默认值 2048 就足够了，除非您明确需要使用不同的密钥大小。很多 CA 要求最小值为 2048。较大的密钥更安全，但对性能的影响更大。 |
| **说明** | 输入特定的详细信息以帮助您以后识别该证书。 |

7 单击**保存 (Save)**。

自定义 CSR 将显示为一个链接。

8 选择该 CSR，然后单击**操作 (Actions)**。

9 从下拉菜单中选择**下载 CSR PEM (Download CSR PEM)**。

您可以保存 CSR PEM 文件以进行存档和提交给 CA。

10 使用 CSR 文件内容按照 CA 注册过程向 CA 提交证书请求。

CA 根据 CSR 文件中的信息创建一个服务器证书，使用私钥对其进行签名，然后向您发送该证书。CA 还会向您发送根 CA 证书。

## 导入 CA 证书

您可以导入签名的 CA 证书以成为公司的临时 CA。在导入证书后，您有权对自己的证书进行签名。

**前提条件**

确认具有一个 CA 证书。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**系统 (System) > 设置 (Settings)**。

3 在“证书”选项卡上，单击**导入 (Import)**。

4 从下拉菜单中选择**导入 CA 证书 (Import CA Certificate)**并添加证书详细信息。

| 选项 | 说明 |
| --- | --- |
| **名称** | 指定 CA 证书的名称。 |
| **证书内容** | 浏览到计算机上的 CA 证书文件并添加该文件。 |
| **说明** | 输入该 CA 证书中包含的内容的摘要。 |

5 单击**保存 (Save)**。

您现在可以对自己的证书进行签名。

## 导入证书

您可以导入具有私钥的证书以创建自签名证书。

**前提条件**

确认具有一个证书。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**系统 (System) > 设置 (Settings)**。

3 在“证书”选项卡上，单击**导入 (Import)**。

4 从下拉菜单中选择**导入证书 (Import Certificate)**并添加证书详细信息。

| 选项 | 说明 |
| --- | --- |
| **名称** | 指定 CA 证书的名称。 |
| **证书内容** | 浏览到计算机上的证书文件并添加该文件。 |
| **私钥** | 浏览到计算机上的私钥文件并添加该文件。 |
| **密码** | 为该证书添加密码。 |
| **说明** | 输入该证书中包含的内容的摘要。 |

5 单击**保存 (Save)**。

您现在可以创建自己的自签名证书。

## 创建自签名证书

使用自签名证书可能不如使用受信任的证书安全。

在使用自签名证书时，客户端用户将收到一条警告消息，例如，无效的安全证书 (Invalid Security Certificate)。在首次连接到服务器时，客户端用户必须接受自签名证书才能继续。允许客户端用户选择该选项将提供比其他授权方法更低的安全性。

**前提条件**

确认具有一个 CSR。请参阅创建证书签名请求文件。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**系统 (System) > 设置 (Settings)**。

3 单击**证书 (Certificates)**选项卡。

4 从下拉菜单中选择 **CSR (CSRs)**。

5 选择现有的 CSR。

6 单击**操作 (Actions)**，然后从下拉菜单中选择 **CSR 的自签名证书 (Self Sign Certificate for CSR)**。

7 输入自签名证书的有效天数。

默认时间范围是 10 年。

8 单击**保存 (Save)**。

将在**证书 (Certificate)**列表中显示自签名证书，并将证书类型指定为自签名。

## 替换证书

如果您需要替换证书（例如，如果证书将要过期），您可以使用 API 请求替换现有的证书。

**前提条件**

确认在 NSX Manager 中具有一个证书。请参阅创建自签名证书和导入证书。

**步骤**

1 从导航面板中选择**系统 (System) > 设置 (Settings)**。

2 单击**证书 (Certificates)**选项卡，然后从下拉菜单中选择**证书 (Certificates)**。

3 单击要使用的证书的 ID，然后从弹出窗口中复制该证书 ID。

4 发送一个 `POST /api/v1/node/services/http?action=apply_certificate&certificate_id=<CertificateID>` API 请求以替换现有的证书。

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

该 API 请求重新启动 HTTP 服务，以便该服务开始使用新证书。在 POST 请求成功完成时，响应代码为 `200 Accepted`。

# 导入证书吊销列表

证书吊销列表 (Certificate Revocation List, CRL) 是订阅者及其证书状态列表。在潜在用户尝试访问服务器时，服务器将根据该特定用户的 CRL 条目拒绝访问。

该列表包含以下各项：

- 吊销的证书和吊销原因
- 证书颁发日期
- 颁发证书的实体
- 计划发行下一版本的日期

**前提条件**

确认具有一个 CRL。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**系统 (System) > 设置 (Settings)**。

3 单击**证书 (Certificates)**选项卡。

4 从下拉菜单中选择 **CRL (CRLs)**。

5 单击**导入 (Import)**并添加 CRL 详细信息。

| 选项 | 说明 |
| --- | --- |
| 名称 | 指定 CRL 的名称。 |
| 证书内容 | 复制 CRL 中的所有项目并将其粘贴到该部分中。<br>示例 CRL。<br>`-----BEGIN X509 CRL-----`<br>`MIIBODCB4zANBgkqhkiG9w0BAQQFADBgMQswCQYDVQQGEwJBVTEMMAoGA1UECBMD`<br>`UUxEMRkwFwYDVQQKExBNaW5jb20gUHR5LiBMdGQuMQswCQYDVQQLEwJDUzEbMBkG`<br>`A1UEAxMSU1NMZWF5IGRlbW8gc2VydmVyFw0wMTAxMTUxNjI2NTdaFw0wMTAyMTQx`<br>`NjI2NTdaMFIwEgIBARcNOTUxMDA5MjMzMjA1WjASAgEDFw05NTEyMDEwMTAwMDBa`<br>`MBMCAhI0Fw0wMTAxMTUxNjE5NDdaMBMCAhI1Fw0wMTAxMTUxNjIzNDZaMA0GCSqG`<br>`SIb3DQEBBAUAA0EAHPjQ3M93QOj8Ufi`<br>`+jZM7Y78TfAzG4jJn/E6MYBPFVQFYo/Gp`<br>`UZexfjSVo5CIyySOtYscz8oO7avwBxTiMpDEQg==`<br>`-----END X509 CRL--` |
| 说明 | 输入该 CRL 中包含的内容的摘要。 |

6 单击**保存 (Save)**。

导入的 CRL 将显示为一个链接。

# 配置设备

必须使用命令行或 API 完成某些系统配置任务。

有关完整的命令行界面信息，请参阅《NSX-T 命令行界面参考》。有关完整的 API 信息，请参阅《NSX-T API 指南》。

**表 11-1. 系统配置命令和 API 请求。**

| 任务 | 命令行（NSX Manager、NSX Controller、NSX Edge） | API 请求（仅限 NSX Manager） |
|---|---|---|
| 设置系统时区 | `set timezone <timezone>` | `PUT https://<nsx-mgr>/api/v1/node` |
| 设置 NTP 服务器 | `set ntp-server <ntp-server>` | `PUT https://<nsx-mgr>/api/v1/node/services/ntp` |
| 设置 DNS 服务器 | `set name-servers <dns-server>` | `PUT https://<nsx-mgr>/api/v1/node/network/name-servers` |
| 设置 DNS 搜索域 | `set search-domains <domain>` | `PUT https://<nsx-mgr>/api/v1/node/network/search-domains` |

# 管理标记

您可以在对象中添加标记以更轻松地搜索对象。在为对象指定标记时，您还可以指定范围。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 导航到一个对象类别。

例如，导航到**交换 (Switching) > 交换机 (Switches)**。

3 选择一个对象。

4 选择**操作 (Actions) > 管理标记 (Manage Tags)**菜单选项。

5 添加或删除标记。

| 选项 | 操作 |
|---|---|
| 添加标记 | 单击**添加 (Add)**以指定标记和可选的范围。 |
| 删除标记 | 选择一个现有的标记，然后单击**删除 (Delete)**。 |

逻辑端口最多可以具有 15 个标记。所有其他对象最多可以具有 10 个标记。

6 单击**保存 (Save)**。

# 搜索对象

您可以使用不同的条件搜索对象。

可以使用以下条件进行搜索：

- 资源类型
- 名称
- 描述
- 创建时间
- 修改时间
- 创建者
- 修改者
- 标记

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 单击主窗格右上角的放大镜图标。

3 输入对象或对象类型的正则表达式搜索模式。

默认情况下会定位搜索模式，即系统会假定字符串起始锚点为 `^`，字符串结束锚点为 `$`。请勿在搜索模式中使用这些锚点。例如，如果您希望搜索以“Logical”开头的资源，搜索模式可以是 `Logical.*`。如果您希望搜索以“Switch”结尾的资源，搜索模式可以是 `.*Switch`。

4 在显示结果的窗口中，单击窗口底部的**查看...结果 (View ... results)**链接可打开搜索窗格，您可以在其中细化搜索内容。

5 指定一个或多个条件以细化搜索内容。

# 查找远程服务器的 SSH 指纹

涉及与远程服务器之间复制文件的某些 API 请求要求在请求正文中提供远程服务器的 SSH 指纹。SSH 指纹是从远程服务器上的主机密钥中获取的。

要通过 SSH 进行连接，NSX Manager 和远程服务器必须具有相同的主机密钥类型。如果具有多个相同的主机密钥类型，将根据 NSX Manager 上的 HostKeyAlgorithm 配置确定首选的类型。

具有远程服务器的指纹可以帮助您确认连接到正确的服务器，从而防止受到中间人攻击。您可以询问远程服务器管理员他们是否可以提供服务器的 SSH 指纹。或者，您可以连接到远程服务器以查找指纹。通过控制台连接到服务器比通过网络更安全。

NSX Manager 设备基于 Ubuntu 14.04 并使用默认 HostKeyAlgorithm 顺序。默认情况下，该表按照从最优先到最不优先的顺序列出 NSX Manager 上存在的密钥。

表 11-2. 按优先顺序排列的 NSX Manager 主机密钥

| NSX Manager 上存在的主机密钥类型 | 该主机密钥类型的默认位置 |
| --- | --- |
| ECDSA（256 位） | /etc/ssh/ssh_host_ecdsa_key.pub |
| ED25519 | /etc/ssh/ssh_host_ed25519_key.pub |
| RSA | /etc/ssh/ssh_host_rsa_key.pub |
| DSA | /etc/ssh/ssh_host_dsa_key.pub |

**步骤**

1 登录到远程服务器的 CLI。

使用控制台登录比通过网络更安全。

2 列出 `/etc/ssh` 目录中的公钥文件。

```
$ 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
```

3 将可用的密钥与 HostKeyAlgorithm 顺序进行比较。

在该示例中，共有三个 SSH 密钥：DSA 和 RSA 以及 ED25519。ED25519 具有最高的优先顺序，因此，NSX Manager 在连接到远程服务器使用该密钥。

4 获取首选密钥的指纹。

```
$ 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)
```

该密钥的指纹是 d0:21:3e:ec:52:ff:19:a9:e7:71:b5:7f:63:23:57:f7。

注 您必须从备份和还原 API 请求的 SSH 指纹中移除冒号。

# 备份和还原 NSX Manager

如果 NSX Manager 虚拟设备变得无法运行，则可以通过备份还原该设备。NSX Manager 存储虚拟网络的所需状态。如果 NSX Manager 设备变得无法运行，并不会影响数据层面，但无法进行配置更改。

可以通过两种备份方法创建三种类型的备份：

**群集备份** 该备份包含虚拟网络的所需状态。

**节点备份** 该备份包含 NSX Manager 设备配置。

**清单备份** 该备份包含一组 ESX 和 KVM 主机以及 Edge。在还原操作期间，将使用该信息检测并修复管理层面的所需状态与这些主机之间的差异。

共有两种备份方法：

| | |
|---|---|
| **手动 NSX Manager 节点备份和群集备份** | 可以根据需要随时运行手动节点和群集备份。 |
| **自动 NSX Manager 节点备份、群集备份和清单备份** | 自动备份是根据您制订的计划运行的。强烈建议使用自动备份。请参阅计划自动备份。 |

要确保具有最新的备份，您应该配置自动备份。请务必定期运行群集和清单备份。

您可以将 NSX-T 配置还原回在任何群集备份中捕获的状态。在还原备份时，您必须还原到一个新 NSX Manager 设备，该设备运行与备份的设备相同的 NSX Manager 版本。

备份和还原要求管理程序、NSX Manager 设备和 NSX Controller 设备必须具有静态管理 IP 地址。不支持更改管理 IP 地址。不支持使用 DHCP 为 NSX Manager 和 NSX Controller 设备分配管理 IP 地址。只有在将 DHCP 服务器配置为始终为给定管理程序提供相同的 IP 地址时，才支持使用 DHCP 为管理程序分配管理 IP 地址。

# 备份 NSX Manager 配置

NSX Manager 配置备份包含 NSX Manager 节点备份、群集备份和清单备份。

**步骤**

1 配置备份位置

  备份保存到 NSX Manager 可访问的远程 SFTP 位置中。在进行备份之前，必须配置备份位置。

2 计划自动备份

  计划日常备份，以便还原无法运行的 NSX Manager 及其配置数据。默认情况下，将禁用自动备份。您可以计划在每周的特定日期或按指定的间隔执行自动备份。强烈建议使用计划的备份。

## 配置备份位置

备份保存到 NSX Manager 可访问的远程 SFTP 位置中。在进行备份之前，必须配置备份位置。

**步骤**

1 登录到 NSX Manager 虚拟设备。

2 单击**系统 > 实用程序 > 备份 (System > Utilities > Backup)**。

3 要提供备份位置的访问凭据，请单击页面右上角的**编辑 (Edit)**。

4 单击**自动备份 (Automatic Backup)**开关以启用自动备份。

5 输入 SFTP 服务器的 IP 地址或主机名。

6 根据需要编辑默认端口。

7 输入登录到 SFTP 服务器所需的用户名和密码。

8 在**目标目录 (Destination Directory)**字段中，输入存储备份的绝对目录路径。

9 输入用于加密备份数据的密码短语。

您需要使用该密码短语还原备份。如果忘记了备份密码短语，则无法还原任何备份。

10 输入存储备份的服务器的 SSH 指纹。请参阅查找远程服务器的 SSH 指纹。

11 单击**保存 (Save)**。

12 单击页面底部的**立即备份 (Backup Now)**，以确认可以将文件写入到 SFTP 服务器中。

**后续步骤**

计划自动备份。

## 计划自动备份

计划日常备份，以便还原无法运行的 NSX Manager 及其配置数据。默认情况下，将禁用自动备份。您可以计划在每周的特定日期或按指定的间隔执行自动备份。强烈建议使用计划的备份。

**前提条件**

- 确定相应的备份位置。选择一个可以防止单点故障的位置。例如，不要将备份放在与设备相同的文件存储中。该文件存储上的故障可能会影响设备及其备份。
- 找到存储备份的服务器的 SSH 指纹。请参阅查找远程服务器的 SSH 指纹。备份和还原 API 请求要求 SSH 指纹不包含冒号。

**步骤**

1 登录到 NSX Manager 虚拟设备。

2 单击**系统 > 实用程序 > 备份 (System > Utilities > Backup)**。

3 单击页面右上角的**编辑 (Edit)**。

4 单击**文件服务器 (File Server)**，然后验证是否启用了自动备份。

5 单击页面顶部的**计划 (Schedule)**。

6 对于节点/群集备份，请单击**每周 (Weekly)**并设置备份到 SFTP 服务器的日期和时间，或者单击**间隔 (Interval)**并设置备份时间。

7 默认情况下，清单备份设置为每隔 30 秒执行一次，并且应经常执行该备份。根据需要，接受或更改默认设置。

8 单击**保存 (Save)**。

---

**注** 第一个计划的每周备份在指定的工作日和时间执行。第一个计划的间隔备份在保存启用的自动备份的备份配置后立即执行。

---

NSX Manager 存储三个单独的备份文件：节点级别、群集级别和清单。备份文件保存到 SFTP 服务器上由备份配置指定的目录中。在该目录中，这些文件保存在以下目录中：

- /<user specified directory>/cluster-node-backups（群集和节点备份）
- /<user specified directory>/inventory-summary（清单备份）

# 还原 NSX Manager 配置

如果 NSX Manager 设备无法运行并创建了建议的备份，您可以还原 NSX Manager 设备。您需要使用在创建备份时指定的密码短语还原备份。

**步骤**

1 准备还原 NSX Manager 备份

在还原 NSX Manager 备份之前，您必须安装新的 NSX Manager 设备。必须使用与以前的 NSX Manager 相同的管理 IP 地址部署新的 NSX Manager。

2 还原群集备份

群集备份用于还原所需的网络状态。在还原节点备份之前，您必须还原群集备份。

3 还原 NSX Manager 节点备份

节点备份还原设备配置，以使 NSX Controller 群集能够连接到该设备。在还原节点备份之前，您必须还原群集备份。所选的节点备份文件应具有与群集备份文件相同的时间戳。

4 下载备份和还原帮助程序脚本

您必须从 NSX Manager 中下载备份和还原帮助程序脚本。

5 恢复在上次群集备份后进行的架构更改

备份和还原帮助程序脚本将还原备份后的所需状态与该脚本捕获的最新架构状态进行比较，并提供在还原备份后使架构状态与所需状态相匹配的说明。

6 还原 NSX Controller 群集

如果无法恢复 NSX Controller 群集，或者由于更改了群集成员资格而需要替换一个或多个控制器，则应该还原整个控制器群集。

## 准备还原 NSX Manager 备份

在还原 NSX Manager 备份之前，您必须安装新的 NSX Manager 设备。必须使用与以前的 NSX Manager 相同的管理 IP 地址部署新的 NSX Manager。

**前提条件**

- 确认您具有可还原的节点、群集和最新清单备份文件。
- 确认您具有节点和群集备份文件的密码短语。
- 确认您了解用于创建备份的 NSX Manager 版本，并具有相同版本的相应安装文件（OVA、OVF 或 QCOW2）。
- 确认您了解为用于创建节点备份的 NSX Manager 分配的 IP 地址。
- 确认没有人尝试对 NSX Manager 配置进行更改，直到还原过程完成。

**步骤**

1 如果旧 NSX Manager 设备仍在运行（例如，如果进行还原以回滚升级尝试），请将其关闭。

2 安装新的 NSX Manager 设备。

- 新 NSX Manager 设备的版本必须与用于创建备份的设备版本相同。
- 您必须为该设备配置用于创建节点备份的 NSX Manager 的 IP 地址。

有关这些步骤的信息和说明，请参阅《 NSX-T 安装指南》。

#### 后续步骤

还原群集备份。

## 还原群集备份

群集备份用于还原所需的网络状态。在还原节点备份之前，您必须还原群集备份。

#### 前提条件

- 找到存储备份的服务器的 SSH 指纹。请参阅查找远程服务器的 SSH 指纹。备份和还原 API 请求要求 SSH 指纹不包含冒号。

#### 步骤

1 在还原备份之前，检查 NSX Manager 的状态是否为 STABLE。

```
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"
  }
}
```

注 控制群集状态为 NO_CONTROLLERS，因为在还原节点备份后控制群集才会连接到 NSX Manager。

2 发送群集备份还原 API 请求 `POST /api/v1/cluster/backups?action=restore`，这会从远程位置中复制备份文件并在 NSX Manager 设备上还原该备份。请在该 API 请求中指定备份文件和位置信息。

**还原请求字段：**

| | |
|---|---|
| **passphrase** | 在创建备份时指定的密码短语。如果您不知道该密码，则无法还原该备份。 |
| **server** | 存储备份文件的远程服务器。 |
| **uri** | 远程服务器上的备份文件路径。 |
| **ssh_fingerprint** | 存储备份文件的远程服务器的 SSH 指纹。请参阅查找远程服务器的 SSH 指纹。 |

**还原请求字段：**

| | |
|---|---|
| **username** | 用于登录到远程服务器以复制备份文件的用户名。 |
| **password** | 用于登录到远程服务器以复制备份文件的密码。 |

示例群集备份还原请求：

```
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 等待系统再次变为稳定状态。

4 禁用自动备份。

a 登录到 NSX Manager 虚拟设备。

b 单击**系统 > 实用程序 > 备份 (System > Utilities > Backup)**。

c 单击页面右上角的**编辑 (Edit)**。

d 单击**自动备份 (Automatic Backup)**开关以禁用自动备份。

在完成备份和还原帮助程序脚本后，您可以再次启用自动备份。

**后续步骤**

重新启动所有 NSX Controller，以移除在还原节点备份之前以及 NSX Manager 和 NSX Controller 同步之前缓存的任何数据。请参阅重新引导 NSX Controller 群集成员。

## 还原 NSX Manager 节点备份

节点备份还原设备配置，以使 NSX Controller 群集能够连接到该设备。在还原节点备份之前，您必须还原群集备份。所选的节点备份文件应具有与群集备份文件相同的时间戳。

小心 在还原节点备份之前，您必须还原群集备份。在还原节点备份后，控制器现在可以与 NSX Manager 通信，并且它们将更新实现的网络状态，以便与在 NSX Manager 上配置的所需网络状态相匹配。如果尚未还原群集备份，则不会配置所需的网络状态并删除当前实现的网络状态。

### 前提条件

- 在 NSX Manager 上完成群集备份还原。请参阅还原群集备份。
- 确认您具有 NSX Manager 备份。请参阅备份 NSX Manager 配置。
- 找到存储备份的服务器的 SSH 指纹。请参阅查找远程服务器的 SSH 指纹。备份和还原 API 请求要求 SSH 指纹不包含冒号。

### 步骤

1 在还原备份之前，检查 NSX Manager 的状态是否为 STABLE。

```
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"
  }
}
```

注 控制群集状态为 NO_CONTROLLERS，因为在还原节点备份后控制群集才会连接到 NSX Manager。

2 发送节点备份还原 API 请求 `POST /api/v1/node/backups?action=restore`，这会从远程位置中复制备份文件并在 NSX Manager 设备上还原该备份。请在该 API 请求中指定备份文件和位置信息。

**还原请求字段：**

| | |
|---|---|
| **passphrase** | 在创建备份时指定的密码短语。如果您不知道该密码，则无法还原该备份。 |
| **server** | 存储备份文件的远程服务器。 |
| **uri** | 远程服务器上的备份文件路径。 |
| **ssh_fingerprint** | 存储备份文件的远程服务器的 SSH 指纹。请参阅查找远程服务器的 SSH 指纹。 |

**还原请求字段：**

| | |
|---|---|
| **username** | 用于登录到远程服务器以复制备份文件的用户名。 |
| **password** | 用于登录到远程服务器以复制备份文件的密码。 |

示例节点备份还原请求：

```
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"
            }
        }
    }
}
```

**后续步骤**

下载备份和还原帮助程序脚本。

## 下载备份和还原帮助程序脚本

您必须从 NSX Manager 中下载备份和还原帮助程序脚本。

**前提条件**

- 确认用于运行帮助程序脚本的计算机满足系统要求。帮助程序脚本需要使用 python 2 和 TLS 1.2，并且已在 Ubuntu 14.04 上进行了验证。

**步骤**

- 下载备份和还原帮助程序脚本。您可以从命令行或 API 中执行该操作。
  - 从命令行中：

运行 `copy file` 命令以将该脚本复制到远程服务器中。`url` 参数使用标准 URL 语法指定该脚本的目标，例如，*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/
```

- 从 API 中：

  发送以下 API 请求，然后将输出保存到名为 backup_restore_helper.py 的文件中。

```
GET https://nsx-manager-1/api/v1/node/file-store/backup_restore_helper.py/data
```

**后续步骤**

恢复在上次群集备份后进行的架构更改。

## 恢复在上次群集备份后进行的架构更改

备份和还原帮助程序脚本将还原备份后的所需状态与该脚本捕获的最新架构状态进行比较，并提供在还原备份后使架构状态与所需状态相匹配的说明。

**前提条件**

- 确认您下载了备份和还原帮助程序脚本。
- 确认您从 SFTP 服务器中下载了最新的清单备份。

**步骤**

1 登录到将备份和还原帮助程序脚本下载或复制到的计算机。

2 运行备份和还原帮助程序脚本，并使用 `-d` 选项指定要使用的检查点（清单）文件。

提供以下信息：

| | |
|---|---|
| **`-m`** | NSX Manager IP 地址 |
| **`-u`** | NSX Manager 用户名 |
| **`-p`** | NSX Manager 密码 |
| **`-d`** | 检查点（最新清单备份）文件名称 |

```
$ python backup_restore_helper.py -m 192.168.110.201 -u admin -p <password> -d
backups/backup_restore_checkpoint_20160318_013354.json
```

3 按照 `backup_restore_helper.py` 脚本输出中的说明更新架构状态，以便与所需状态相匹配。

## 还原 NSX Controller 群集

如果无法恢复 NSX Controller 群集，或者由于更改了群集成员资格而需要替换一个或多个控制器，则应该还原整个控制器群集。

在还原控制器群集之前，请先确定是否将控制群集成员资格从管理层面已知的成员资格更改为控制器本身已知的实际成员资格或相反。如果在备份后进行了更改，则成员资格可能会有所不同。

- 如果无法恢复整个群集，请参阅重新部署 NSX Controller 群集。
- 按照以下步骤确定是否更改了群集成员资格；如果已更改，请还原群集。

**前提条件**

- 确认您具有最新的群集级别备份。
- 执行群集级别还原。请参阅还原群集备份。

**步骤**

1 登录到 NSX Manager 的 CLI，然后运行 `get management-cluster status` 命令。

2 登录到 NSX Controller 的 CLI，然后运行 `get managers` 命令以确保在管理器中注册该控制器。

3 运行 `get control-cluster status` 命令。

4 要确定是否更改了成员资格，请将 `get management-cluster status` 命令输出中的 IP 地址与 `get control-cluster status` 命令输出进行比较。

如果 IP 地址集相同，则不需要执行任何操作。如果任何 IP 地址不相同，请继续执行其余步骤以还原整个控制器群集。

5 登录到 NSX Controller 的 CLI，然后运行 `get control-cluster status` 命令以确定哪个控制器是主控制器。

主控制器输出将显示 `is master: true`。

6 在某个非主控制器上运行 `stop service <controller>` 命令。

7 登录到主控制器，然后运行 `detach control-cluster <ip-address[:port]>` 命令以断开连接上一步中的非主控制器。

8 （可选）只有在 `get management-cluster status` 命令在 NSX Manager 上显示该控制器时，才应在 NSX Manager 上运行 `detach controller <uuid>` 命令以断开连接该控制器。

9 登录到 NSX Controller 的 CLI，然后运行 `deactivate control-cluster` 命令。

10 使用以下命令移除引导文件和 uuid 文件：`rm -r /opt/vmware/etc/bootstrap-config` 和 `rm -r /config/vmware/node-uuid`

11 对于其余非主控制器，执行步骤 6-10。

12 登录到主控制器的 CLI，然后运行 `stop service <controller>` 命令。

13 在 NSX Manager 上运行 `detach controller <uuid>` 命令以断开连接该控制器。

14 登录到主控制器的 CLI，然后运行 `deactivate control-cluster` 命令。

15 使用以下命令移除引导文件和 uuid 文件：`rm -r /opt/vmware/etc/bootstrap-config` 和 `rm -r /config/vmware/node-uuid`

16 从 NSX Manager 中运行 `get management-cluster status` 命令。如果在输出中仍显示控制器，请运行 `detach controller <uuid>` 命令以断开连接任何剩余的控制器。

**后续步骤**

按列出的顺序完成以下任务。

1 完成节点级别还原。请参阅还原 NSX Manager 节点备份。

2 将 NSX Controller 加入管理层面，如《NSX-T 安装指南》中所述。

3 重新部署 NSX Controller 群集，如《NSX-T 安装指南》中所述。

# 管理设备和设备群集

每个 NSX-T 安装仅需要使用和支持一个 NSX Manager 实例。NSX Controller 群集应具有三个成员。NSX Edge 群集应具有至少两个成员。

如果控制器或 Edge 群集中的设备无法运行，或者由于任何原因需要将其移除，您可以将其替换为新设备。

**重要** 如果对 NSX Controller 或 NSX Edge 群集成员资格进行任何更改，您必须随后创建群集备份以备份新配置。请参见备份和还原 NSX Manager。

# 管理 NSX Manager

您可以使用 CLI 命令检查 NSX Manager 的状态。如果 NSX Manager 无法运行并且无法恢复，您可以重新引导 NSX Manager 设备。

## 获取 NSX Manager 状态

您可以使用 CLI 命令获取 NSX Manager 状态。

**步骤**

1 登录到 NSX Manager 的 CLI。

2 运行 `get management-cluster status` 命令。例如，

```
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.
```

注 即使结果显示管理群集，也只能有一个 NSX Manager 实例。

## 重新引导 NSX Manager

您可以使用 CLI 命令重新引导 NSX Manager 以从严重错误中恢复。

### 步骤

1 登录到 NSX Manager 的 CLI。

2 运行 `reboot` 命令。例如，

```
nsx-manager> reboot
Are you sure you want to reboot (yes/no): y
```

## 管理 NSX Controller 群集

NSX Controller 群集应具有三个成员。在进行故障排除后，如果确定某个 NSX Controller 设备无法恢复，您可以在群集中添加一个设备以替换该设备，或者重新部署 NSX Controller 群集（如果需要）。

NSX Controller 群集必须占多数才能正常工作。如果三个成员中的两个成员处于联机状态，则群集仍占多数。您应该启动脱机 NSX Controller 以恢复三成员群集。如果无法启动该设备，您可以添加另一个 NSX Controller 设备以替换该设备并重新占多数。请参阅替换 NSX Controller 群集的成员。

如果三个成员中的一个成员处于联机状态，则群集占多数，而无法正常工作。如果启动脱机成员，群集将重新占多数。如果无法启动任一脱机成员，您可以重新部署 NSX Controller 群集。请参阅重新部署 NSX Controller 群集。

### 前提条件

通过故障排除确认无法恢复设备。例如，以下步骤可以恢复设备，而无需更换这些设备。

- 确认设备具有网络连接；如果没有，请解决该问题。
- 重新引导设备。

### 后续步骤

获取 NSX Controller 群集状态。请参阅获取 NSX Controller 群集状态。

## 获取 NSX Controller 群集状态

您可以从 NSX Manager 中确定 NSX Controller 群集的状态。也可以从命令行界面中检查每个 NSX Controller 的状态。

获取 NSX Controller 群集和群集成员的状态可以帮助您确定 NSX Controller 群集问题的来源。

**表 11-3. NSX Controller 群集状态**

| | 是否在 NSX Manager 中注册至少一个控制器? | NSX Controller 群集是否占多数? | 任何 NSX Controller 群集成员是否发生故障? |
|---|---|---|---|
| NO_CONTROLLERS | 否 | 不适用 | 不适用 |
| UNAVAILABLE | 未知 | 未知 | 未知 |
| STABLE | 是 | 是 | 否 |
| DEGRADED | 是 | 是 | 是 |
| UNSTABLE | 是 | 否 | 否 |

**步骤**

1 登录到 NSX Manager CLI。

2 运行 `get management-cluster status` 命令。

```
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 登录到 NSX Controller CLI。

4 运行 `get control-cluster status` 命令。

```
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
```

## 重新引导 NSX Controller 群集成员

如果需要重新引导多个 NSX Controller 群集成员，您必须每次重新引导一个成员。如果一个成员处于脱机状态，则三成员群集可以占多数。如果两个成员处于脱机状态，群集将不占多数，而无法正常工作。

**步骤**

1 登录到 NSX Manager 的 CLI。

2 获取管理和控制群集的状态。

```
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 登录到需要重新引导的 NSX Controller 的 CLI，然后将其重新引导。

```
nsx-controller-2> reboot
Are you sure you want to reboot (yes/no): y
```

4 再次获取管理和控制群集的状态。等到控制群集具有 STABLE 状态，然后再重新引导任何其他成员。

在该示例中，NSX Controller 192.168.110.53 正在重新引导，并且控制群集具有 DEGRADED 状态。这意味着群集占多数，但其中的一个成员已关闭。有关 NSX Controller 群集状态的详细信息，请参阅获取 NSX Controller 群集状态。

```
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
```

在 NSX Controller 群集处于 STABLE 状态后，就可以安全地重新引导任何其他成员。

```
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 如果需要了解有关各种 NSX Controller 设备状态的信息，您可以登录到 NSX Controller 并运行 `get control-cluster status` 命令。

```
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 如果需要，请重复这些步骤以重新引导其他 NSX Controller 设备。

## 替换 NSX Controller 群集的成员

NSX Controller 群集必须具有至少三个成员。如果 NSX Controller 设备无法运行，并且需要从群集中移除该设备，您必须先添加新的 NSX Controller 设备以创建四成员群集。在添加第四个成员后，您可以从群集中移除一个 NSX Controller 设备。

**前提条件**

- 通过故障排除确认无法恢复设备。例如，以下步骤可以恢复设备，而无需更换这些设备。
  - 确认设备具有网络连接；如果没有，请解决该问题。
  - 重新引导设备。
- 确认您了解要替换的 NSX Controller 版本，并具有相同版本的相应安装文件（OVA、OVF 或 QCOW2）。

**步骤**

1 安装并配置新的 NSX Controller。

有关这些步骤的信息和说明，请参阅《 NSX-T 安装指南》。

a 安装新的 NSX Controller 设备。

新 NSX Controller 的版本必须与要替换的 NSX Controller 相同。

b 将新的 NSX Controller 加入管理层面。

c 将新的 NSX Controller 加入控制群集。

2 关闭要从群集中移除的 NSX Controller。

3 登录到另一个 NSX Controller，然后检查要移除的 NSX Controller 是否具有 `not active` 状态。

```
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 将控制器与群集断开连接。

```
nsx-controller-1> detach control-cluster 192.168.110.52
Successfully detached node from the control cluster.
```

5 将控制器与管理层面断开连接。

```
nsx-manager-1> detach controller 863f9669-509f-4eba-b0ac-61a9702a242b
The detach operation completed successfully
```

6 验证控制器是否处于活动状态，以及控制群集是否处于稳定状态。

从 NSX Controller 中：

```
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
```

从 NSX Manager 中：

```
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
```

## 重新部署 NSX Controller 群集

如果替换一个控制器并未解决 NSX Controller 群集问题，或者无法恢复多个 NSX Controller 设备，您可以重新部署整个群集。NSX Manager 包含所需的所有配置状态，可以使用该管理器重新创建 NSX Controller 群集。

在还原 NSX Controller 群集期间，数据路径连接不会中断。

**前提条件**

- 通过故障排除确认无法恢复设备。例如，以下步骤可以恢复设备，而无需更换这些设备。
  - 确认设备具有网络连接；如果没有，请解决该问题。
  - 重新引导设备。
- 确认您了解要替换的 NSX Controller 版本，并具有相同版本的相应安装文件（OVA、OVF 或 QCOW2）。
- 确认您了解分配给 NSX Controller 设备的 IP 地址。

**步骤**

1 关闭 NSX Controller 群集中的所有控制器。

2 将控制器与 NSX Manager 断开连接。

a 登录到 NSX Manager CLI。

b 使用 `get management-cluster status` 命令获取控制器列表。

```
nsx-manager-1> get management-cluster status
Number of nodes in management cluster: 1
- 192.168.110.201  (UUID 422EC8D8-B43F-D206-5048-781A5AECDCC6) 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 使用 `detach controller` 命令断开连接控制器。

```
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 安装三个 NSX Controller 设备并创建新的 NSX Controller 群集。

有关这些步骤的信息和说明，请参阅《 NSX-T 安装指南》。

a 安装三个 NSX Controller 设备。

- 新 NSX Controller 设备的版本必须与要替换的 NSX Controller 设备相同。
- 为新控制器分配在旧控制器上使用的相同 IP 地址。

b 将 NSX Controller 设备加入管理层面。

c 在其中的一个 NSX Controller 设备上，初始化控制群集。

d 将两个其他控制器加入控制群集。

## 管理 NSX Edge 群集

您可以在以下情况下替换 NSX Edge：该设备无法运行或需要替换硬件。在安装新的 NSX Edge 并创建新的传输节点后，您可以修改 Edge 群集以将旧传输节点替换为新传输 节点。

注 移除第 1 层 Edge 群集将导致第 1 层分布式路由器 (DR) 实例暂时停止工作。

**步骤**

1 如果要替换的 NSX Edge 仍在运行，您可以将其置于维护模式以最大限度减少停机时间。如果在关联的逻辑路由器上启用了高可用性，进入维护模式将导致逻辑路由器使用不同的 Edge 群集成员。如果 NSX Edge 无法运行，您不需要执行该操作。

a 获取发生故障的结构层节点的结构层节点 ID。

```
https://192.168.110.201/api/v1/fabric/nodes
...
      "resource_type": "EdgeNode",
      "id": "a0f4fa74-e77c-11e5-8701-005056aeed61",
      "display_name": "edgenode-02a",

...
```

b 将发生故障的 NSX Edge 节点置于维护模式。

```
POST https://192.168.110.201/api/v1/fabric/nodes/a0f4fa74-e77c-11e5-8701-005056aeed61?
action=enter_maintenance_mode
```

2 安装新的 NSX Edge。

有关这些步骤的信息和说明，请参阅《 NSX-T 安装指南》。

3 使用 `join management-plane` 命令将新的 NSX Edge 加入管理层面。

有关这些步骤的信息和说明，请参阅《 NSX-T 安装指南》。

4 将 NSX Edge 配置为传输节点。

有关这些步骤的信息和说明，请参阅《 NSX-T 安装指南》。

您可以从 API 中获取发生故障的 NSX Edge 设备的传输节点配置，然后使用该信息创建新的传输节点。

a 获取新结构层节点的结构层节点 ID。

```
https://192.168.110.201/api/v1/fabric/nodes
...
      "resource_type": "EdgeNode",
      "id": "d61c8d86-f4b8-11e5-b1b2-005056ae3c10",
      "display_name": "edgenode-03a",

...
```

b 获取发生故障的传输节点的传输节点 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",
...
```

c 获取发生故障的传输节点的传输节点配置。

```
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 使用 `POST /api/v1/transport-nodes` 创建新的传输节点。

在请求正文中，为新传输节点提供以下信息：

- 新传输节点的 `description`（可选）
- 新传输节点的 `display_name`
- 用于创建新传输节点的结构层节点的 `node_id`

在请求正文中，从发生故障的传输节点中复制以下信息：

- `transport_zone_endpoints`
- `host_switches`
- `tags`（可选）

```
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 编辑 Edge 群集以将发生故障的传输节点替换为新传输节点。

a 获取新传输节点和发生故障的传输节点的 ID。id 字段包含传输节点 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",
```

b 获取 Edge 群集的 ID。id 字段包含 Edge 群集 ID。从 members 数组中获取 Edge 群集的成员。

```
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 编辑 Edge 群集以将发生故障的传输节点替换为新传输节点。member_index 必须与发生故障的传输节点的索引相匹配。

小心 如果 NSX Edge 仍在运行，这是一个破坏性操作。这会将所有逻辑路由器端口从发生故障的传输节点移动到新传输节点。

在该示例中，传输节点 TN-edgenode-01a (73cb00c9-70d0-4808-abfe-a12a43251133) 发生故障，并将其替换为 Edge 群集 Edge-Cluster-1 (9a302df7-0833-4237-af1f-4d826c25ad78) 中的传输节点 TN-edgenode-03a (890f0e3c-aa81-46aa-843b-8ac25fe30bd3)。

```
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 （可选）删除发生故障的传输节点和 NSX Edge 节点。

# 日志记录系统消息

除了在 ESXi 上运行的 NSX-T 组件以外，所有其他组件中的日志消息采用 RFC 5424 格式。您可以配置远程日志记录服务器以接收日志消息。

有关 RFC 5424 的详细信息，请参阅 https://tools.ietf.org/html/rfc5424。

RFC 5424 为日志消息定义以下格式：

```
<facility * 8 + severity> version UTC-TZ hostname APP-NAME procid MSGID [structured-data] msg
```

NSX Manager 中的示例日志消息：

```
<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 生成常规日志（设备 local6，它具有数值 22）和审核日志（设备 local7，它具有数值 23）。所有 API 调用将触发一个审核日志。

RFC 5424 定义以下严重性级别：

| 严重性值 | 说明 |
| --- | --- |
| 0 | 紧急：系统无法使用 |
| 1 | 警报：必须立即采取措施 |
| 2 | 严重：严重情况 |

| 严重性值 | 说明 |
| --- | --- |
| 3 | 错误：错误情况 |
| 4 | 警告：警告情况 |
| 5 | 通知：正常但重大情况 |
| 6 | 信息：信息性消息 |
| 7 | 调试：调试级别消息 |

具有“紧急”、“警报”、“严重”或“错误”严重性的所有日志在日志消息的结构化数据部分中包含唯一的错误代码。错误代码由一个字符串和一个十进制数字组成。该字符串表示特定的模块。

MSGID 字段指示日志消息类别。有关类别列表，请参阅日志消息类别。

## 配置远程日志记录

您可以配置 NSX-T 设备和管理程序以将日志消息发送到远程日志记录服务器。

在 NSX Manager、NSX Controller、NSX Edge 设备和管理程序上支持远程日志记录。

您可以根据以下条件筛选发送到日志记录服务器的日志消息：

- 级别：emerg、alert、crit、err、warning、notice、info、debug
- 设备：代码是在 RFC 5424 中定义的。设备 local7 用于审核消息，local6 用于非审核消息。
- 消息 ID 或类别：以下章节中列出了类别和示例：日志消息类别

有关相关的命令和请求的信息，请参阅《NSX-T 命令行参考》和《NSX-T API 指南》。

### 前提条件

- 配置一个远程日志记录服务器以从 NSX-T 设备中接收日志。
- 确定要发送到日志记录服务器的日志消息。

### 步骤

1 登录到要配置远程日志记录的 NSX-T 设备。

2 使用 `set logging-server` 命令和以下语法配置一个日志记录服务器。可以将多个设备或消息 ID 指定为逗号分隔列表（不含空格）。

```
set logging-server <hostname-or-ip-address[:port]> proto <proto> level <level> [facility <facility>]
[messageid <messageid>] [certificate <filename>]
```

您可以多次运行该命令以添加多个日志记录服务器配置。

```
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 （可选）使用 `get logging-server` 命令查看日志记录配置。

```
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
```

## 日志消息类别

日志消息属于某个类别。可以在 `set logging-server` 命令中使用这些类别以筛选发送到日志记录服务器的日志消息。

**表 11-4. 日志消息类别**

| RIZHIXIAOXILEIBIE | 示例 |
| --- | --- |
| FABRIC | 主机节点<br>主机准备<br>Edge 节点<br>传输区域<br>传输节点<br>上行链路配置文件<br>群集配置文件<br>Edge 群集<br>网桥群集和端点 |
| SWITCHING | 逻辑交换机<br>逻辑交换机端口<br>交换配置文件<br>交换机安全功能 |
| ROUTING | 逻辑路由器<br>逻辑路由器端口<br>静态路由<br>动态路由<br>NAT |
| FIREWALL | 防火墙规则<br>防火墙规则区域 |
| FIREWALL_PKTLOG | 防火墙连接日志<br>防火墙数据包日志 |
| GROUPING | IP 集<br>Mac 集<br>NS 组<br>NS 服务<br>NS 服务组<br>VNI 池<br>IP 池 |
| DHCP | DHCP 中继 |

**表 11-4. 日志消息类别（续）**

| RIZHIXIAOXILEIBIE | 示例 |
| --- | --- |
| SYSTEM | 设备管理（远程 syslog、ntp 等）<br>群集管理<br>信任管理<br>许可<br>用户和角色<br>任务管理<br>安装（NSX Manager、NSX Controller）<br>升级（NSX Manager、NSX Controller、NSX Edge 和主机软件包升级）<br>实现<br>标记 |
| MONITORING | SNMP<br>端口连接<br>跟踪流 |
| - | 所有其他日志消息。 |

# 配置 IPFIX

IPFIX（Internet 协议流量信息导出）是一个网络流量信息格式和导出标准。在启用 IPFIX 时，所有配置的主机传输节点使用端口 4739 将 IPFIX 消息发送到 IPFIX 收集器。

对于 ESXi，NSX-T 自动打开端口 4739。对于 KVM，如果未启用防火墙，则会打开端口 4739，但如果启用了防火墙，您必须确保打开该端口，因为 NSX-T 不会自动打开该端口。

**前提条件**

- 安装至少一个 IPFIX 收集器。
- 确认 IPFIX 收集器具有到管理程序的网络连接。
- 确认任何相关的防火墙（包括 ESXi 防火墙）允许 IPFIX 收集器端口上的流量通过。

**步骤**

1. 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。
2. 从导航面板中选择**工具 (Tools) > IPFIX**。
3. 如果尚未选择**收集器 (Collectors)**选项卡，请单击该选项卡。
4. 单击**配置收集器 (Configure Collectors)**。
5. 单击**添加 (Add)**并输入收集器 IP 地址和端口。

   您最多可以添加 8 个收集器。
6. （可选）在“收集选项”部分中，单击**编辑 (Edit)**以指定观察域 ID。

   观察域 ID 确定网络流量来自哪个观察域。默认值为 0，这表示没有特定的观察域。
7. 单击**交换机 IPFIX 配置文件 (Switch IPFIX Profiles)**选项卡。

8 单击**添加 (Add)**以添加一个配置文件。

| 设置 | 说明 |
| --- | --- |
| 活动超时(秒) | 在这段时间过后，即使收到与流量关联的更多数据包，流量也会超时。默认值为 300。 |
| 空闲超时(秒) | 在这段时间过后，如果没有收到与流量关联的更多数据包，流量将会超时（仅限 ESXi，KVM 根据活动超时确定所有流量是否超时）。默认值为 300。 |
| 最大流量数 | 在网桥上缓存的最大流量数（仅限 KVM，无法在 ESXi 上配置）。默认值为 16384。 |
| 采样概率(%) | 将采样的数据包比例（大致）。如果增加该设置，可能会影响管理程序和收集器的性能。如果所有管理程序将更多 IPFIX 数据包发送到收集器，收集器可能无法收集所有数据包。如果将概率设置为默认值 0.1%，则会将性能影响降到较低的程度。 |

9 单击**应用对象 (Applied To)**以将配置文件应用于一个或多个对象。

对象类型是逻辑端口和逻辑交换机。

ESXi 和 KVM 上的 IPFIX 使用不同的方法对隧道数据包进行采样。在 ESXi 上，隧道数据包将采样为两个记录：

- 具有一些内部数据包信息的外部数据包记录
  - SrcAddr、DstAddr、SrcPort、DstPort 和 Protocol 指的是外部数据包。
  - 包含一些企业条目以描述内部数据包。
- 内部数据包记录
  - SrcAddr、DstAddr、SrcPort、DstPort 和 Protocol 指的是内部数据包。

在 KVM 上，隧道数据包将采样为一个记录：

- 具有一些外部隧道信息的内部数据包记录
  - SrcAddr、DstAddr、SrcPort、DstPort 和 Protocol 指的是内部数据包。
  - 包含一些企业条目以描述外部数据包。

# 使用跟踪流跟踪数据包路径

在数据包从逻辑网络上的一个逻辑端口传输到同一网络上的另一个逻辑端口时，可以使用跟踪流检查数据包路径。跟踪流跟踪在逻辑端口中注入的数据包的传输节点级别路径。跟踪数据包通过逻辑交换机覆盖网络，但对于连接到逻辑交换机的接口不可见。也就是说，不会将数据包实际传送到测试数据包的预期接收方。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 导航到“跟踪流”屏幕。您可以使用两种方法。

- 从导航面板中选择**工具 (Tools) > 跟踪流 (Traceflow)**。
- 从导航面板中选择**交换 (Switching)**，单击**端口 (Ports)**选项卡，选择一个 VIF 连接的端口，然后单击**操作 (Actions) > 跟踪流 (Traceflow)**。

3 选择一种流量类型。

选项是“单播”、“多播”和“广播”。

4 根据流量类型，指定源和目标信息。

| 流量类型 | 指定源信息 | 指定目标信息 |
| --- | --- | --- |
| 单播 | 选择一个虚拟机和虚拟接口。<br>如果在虚拟机中安装了 VMtools，或者使用 OpenStack 插件部署了虚拟机，则会显示 IP 地址和 MAC 地址（在这种情况下，将使用地址绑定）。如果虚拟机具有多个 IP 地址，请从下拉菜单中选择一个地址。<br>如果未显示 IP 地址和 MAC 地址，请在文本框中输入 IP 地址和 MAC 地址。<br>这也适用于“多播”和“广播”。 | 从“类型”下拉菜单中选择“虚拟机名称”或“IP-MAC”。<br>■ 如果选择“虚拟机名称”，请选择一个虚拟机和虚拟接口。选择或输入一个 IP 地址和 MAC 地址。<br>■ 如果选择“IP-MAC”，请选择跟踪类型（第 2 层或第 3 层）。如果跟踪类型为第 2 层，请输入一个 IP 地址和 MAC 地址。如果跟踪类型为第 3 层，请输入一个 IP 地址。 |
| 多播 | 同上。 | 输入一个 IP 地址。它必须是 224.0.0.0-239.255.255.255 范围内的多播地址。 |
| 广播 | 同上。 | 输入一个子网前缀长度。 |

5 （可选）单击**高级 (Advanced)**以查看高级选项。

6 （可选）在左侧的列中，输入以下字段所需的值或输入：

| 选项 | 说明 |
| --- | --- |
| **帧大小** | 例如，128 |
| **TTL** | 例如，64 |
| **超时(毫秒)** | 例如，10000 |
| **以太网类型** | 例如，2048 |
| **负载类型** | 从下拉菜单中选择一个选项。 |
| **负载数据** | 根据选定的负载类型（Base64、十六进制、纯文本、二进制或十进制）设置格式的负载 |

7 （可选）在左侧的列中的“协议”下面，从“类型”下拉菜单中选择一种协议。

8 （可选）根据选择的协议，完成下表中的相关步骤。

| 协议 | 步骤 1 | 步骤 2 | 步骤 3 |
| --- | --- | --- | --- |
| TCP | 输入一个源端口。 | 输入一个目标端口。 | 从下拉菜单中选择所需的 TCP 标记。 |
| UDP | 输入一个源端口。 | 输入一个目标端口。 | 不适用 |
| ICMP | 输入一个 ICMP ID。 | 输入一个序列值。 | 不适用 |

9 单击**跟踪 (Trace)**。

将显示有关连接、组件和层的信息。如果选择单播和逻辑交换机作为目标，输出将包含一个表，其中列出了观察类型（已传送、已丢弃、已接收、已转发）、传输节点和组件以及拓扑图表。您可以为显示的观察应用一个筛选器（**全部 (All)**、**已传送 (Delivered)**、**已丢弃 (Dropped)**）。如果具有丢弃的观察，则默认应用**已丢弃 (Dropped)**筛选器。否则，将应用**全部 (All)**筛选器。

# 查看端口连接信息

您可以使用端口连接工具快速可视化两个虚拟机之间的连接和进行故障排除。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**工具 > 端口连接**。

3 从**源虚拟机**下拉菜单中选择一个虚拟机。

4 从**目标虚拟机**下拉菜单中选择一个虚拟机。

5 单击**查看**。

将显示端口连接拓扑的可视图表。您可以单击可视输出中的任何组件以显示有关该组件的详细信息。

# 监控逻辑交换机端口活动

例如，您可以监控逻辑端口活动以解决网络拥塞和数据包丢弃问题。

**前提条件**

确认配置了一个逻辑交换机端口。请参阅将虚拟机连接到逻辑交换机。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**交换 (Switching) > 端口 (Port)**。

3 双击要监控的逻辑交换机端口。

4 单击**监控 (Monitor)**选项卡。

5 选择**开始跟踪 (Begin Tracking)**。

将打开端口跟踪页面。

6 开始监控逻辑交换机端口上的活动。

您可以查看双向端口流量并确定丢弃的数据包。端口跟踪器页面还会列出与逻辑交换机端口关联的交换配置文件。

例如，如果注意到由于网络拥塞而丢弃的数据包，您可以为逻辑交换机端口配置 QoS 交换配置文件以防止首选数据包上的数据丢失。请参阅了解 QoS 交换配置文件。

# 监控端口镜像会话

您可以监控端口镜像会话以进行故障排除和用于其他用途。

该功能具有以下限制：

- 源镜像端口不能位于多个镜像会话中。

- 目标端口只能接收镜像流量。
- 通过使用 KVM，可以将多个网卡连接到同一 OVS 端口。镜像是在 OVS 上行链路端口上进行的，这意味着将镜像连接到 OVS 端口的所有 pNIC 上的流量。
- 镜像会话源和目标端口必须位于同一主机 vSwitch 上。因此，如果通过 vMotion 将具有源或目标端口的虚拟机移到另一个主机，则无法再镜像该端口上的流量。
- 在 ESXi 上，如果在上行链路上启用了镜像，则 VDL2 使用 Geneve 协议将原始生产 TCP 数据包封装为 UDP 数据包。支持 TSO（TCP 分段卸载）的物理网卡可以更改这些数据包，并使用 MUST_TSO 标记来标记这些数据包。在具有 VMXNET3 或 E1000 vNIC 的监控虚拟机上，该驱动程序将数据包视为常规 UDP 数据包，而无法处理 MUST_TSO 标记并丢弃这些数据包。

如果将大量流量镜像到一个监控虚拟机，则驱动程序的缓冲区环可能会变满并丢弃数据包。为了缓解该问题，您可以采取下面的一个或多个措施：

- 增加接收缓冲区环大小。
- 为虚拟机分配更多 CPU 资源。
- 使用数据层面开发工具包 (Data Plane Development Kit, DPDK) 提高数据包处理性能。

注 确保监控虚拟机的 MTU 设置以及管理程序的虚拟网卡设备的 MTU 设置（对于 KVM）足够大以处理数据包。这对于封装的数据包特别重要，因为封装增加了数据包大小。否则，可能会丢弃数据包。具有 VMXNET3 网卡的 ESXi 虚拟机不会出现该问题，但 ESXi 和 KVM 虚拟机上的其他类型的网卡可能会出现该问题。

注 在涉及 KVM 主机上的虚拟机的 L3 端口镜像会话中，您必须设置足够大的 MTU 以处理封装所需的额外字节。镜像流量经由 OVS 接口和 OVS 上行链路。您必须将 OVS 接口的 MTU 设置为比原始数据包大小（在封装和镜像之前）至少大 100 字节。如果您看到丢弃的数据包，请增加主机的虚拟网卡和 OVS 接口的 MTU 设置。可以使用以下命令设置 OVS 接口的 MTU：

```
ovs-vsctl -- set interface <ovs_Interface> mtu_request=<MTU>
```

注 在监控虚拟机的逻辑端口以及虚拟机所在的主机的上行链路端口时，根据主机是 ESXi 还是 KVM，将会看到不同的行为。对于 ESXi，将使用相同的 VLAN ID 标记逻辑端口镜像数据包和上行链路镜像数据包，它们在监控虚拟机中显示为相同的数据包。对于 KVM，不使用 VLAN ID 标记逻辑端口镜像数据包，但标记上行链路镜像数据包，它们在监控虚拟机中显示为不同的数据包。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**工具 (Tools) > 端口镜像会话 (Port Mirroring Session)**。

3 输入一个会话名称。

4 从下拉菜单中选择一个传输节点。

端口镜像会话必须位于同一传输节点上的网卡之间。

5 从下拉菜单中选择一个方向。

选项是**双向 (Bidirectional)**、**输入 (Ingress)**和**输出 (Egress)**。

6 （可选）选择一个数据包截断值。

7 单击**下一步 (Next)**。

8 选择源 PNIC。

9 （可选）切换**封装的数据包 (Encapsulated Packet)**开关以禁止捕获封装的流量。

默认情况下，将启用该开关。

10 选择源 VNIC。

11 选择一个目标。

您可以选择最多 3 个虚拟机和最多 3 个 VNIC。

12 单击**保存 (Save)**。

在保存端口镜像会话后，您无法更改源和目标。

# 监控结构层节点

您可以从 NSX Manager UI 中监控结构层节点，例如，主机、Edge、Edge 群集、网桥和传输节点。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**架构 (Fabric) > 节点 (Nodes)**。

3 选择以下选项卡之一。

- 主机
- Edge
- Edge 群集
- 网桥
- 传输节点

注 在“主机”屏幕上，如果主机的 MPA 连接状态为“关闭”或“未知”，请忽略 LCP 连接状态，因为该状态可能不准确。

# 收集支持包

您可以在注册的群集和架构节点上收集支持包，并将这些包下载到您的计算机或上载到文件服务器中。

如果您选择将包下载到您的计算机中，将获得一个存档文件，其中包含每个节点的清单文件和支持包。如果您选择将包上载到文件服务器中，则会将清单文件和各个包单独上载到文件服务器中。

**步骤**

1 从浏览器中，登录到 https://*nsx-manager-ip-address* 中的 NSX Manager。

2 从导航面板中选择**系统 (System) > 实用程序 (Utilities)**。

3 单击**支持包 (Support Bundle)**选项卡。

4 选择目标节点。

可用的节点类型是管理节点、控制器节点、Edge 和主机。

5 （可选）指定日志期限天数以排除早于指定天数的日志。

6 （可选）切换开关，选择包括或排除核心文件和审核日志。

核心文件和审核日志可能包含敏感信息，例如，密码或加密密钥。

7 （可选）选中相应的复选框以将包上载到文件服务器中。

8 单击**开始收集包 (Start Bundle Collection)**以开始收集支持包。

根据存在的日志文件数，每个节点可能需要几分钟的时间。

9 监控收集过程的状态。

状态字段显示完成支持包收集的节点的百分比。

10 如果未设置将包发送到文件服务器的选项，请单击**下载 (Download)**以下载包。