

NSX 故障排除指南

Update 8

修改日期：2020 年 2 月 21 日

VMware NSX Data Center for vSphere 6.3



vmware®

您可以从 VMware 网站下载最新的技术文档:

<https://docs.vmware.com/cn/>。

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

威睿信息技术(中国)有限公司
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

版权所有 © 2010 - 2020 VMware, Inc. 保留所有权利。 [版权和商标信息](#)

目录

1 NSX 故障排除指南	6
一般故障排除准则	6
使用 NSX 仪表板	7
NSX 命令行快速参考	10
NSX 主机运行状况检查	20
2 NSX 基础架构故障排除	21
主机准备	21
了解主机准备架构	26
主机准备的服务部署 workflow	31
第三方服务的服务部署 workflow	34
检查通信通道运行状况	36
安装状态为“未就绪”	38
服务没有响应	39
服务部署由于“无法访问 OVF/VIB”错误而失败	40
无法使用“解决”选项修复问题	41
关于 vSphere ESX Agent Manager (EAM)	42
解决 NSX Manager 问题	42
将 NSX Manager 连接到 vCenter Server	44
辅助 NSX Manager 停留在转换模式	47
配置 NSX SSO Lookup Service 失败	48
逻辑网络准备: VXLAN 传输	50
VXLAN VMkernel 网卡不同步	52
更改 VXLAN 绑定策略和 MTU 设置	53
逻辑交换机端口组不同步	55
3 NSX 路由故障排除	56
了解分布式逻辑路由器	57
简要 DLR 数据包流	58
DLR ARP 解析过程	59
了解 Edge 服务网关提供的路由	61
ECMP 数据包流	61
NSX 路由: 必备条件和注意事项	63
DLR 和 ESG UI	66
NSX 路由 UI	66
“NSX Edge” UI	67
新的 NSX Edge (DLR)	69

ESG 和 DLR 差异	71
典型的 ESG 和 DLR UI 操作	72
syslog 配置	72
静态路由	73
路由重新分发	74
NSX 路由故障排除	75
NSX 路由 CLI	75
路由简要概述	78
使用示例路由拓扑验证 DLR 状态	78
可视化 DLR 及其相关主机组件	86
分布式路由子系统架构	87
NSX 路由子系统组件	91
NSX 路由控制层面 CLI	93
NSX 路由子系统故障模式和影响	96
与路由有关的 NSX 日志	98
常见故障情况和修复	100
收集故障排除数据	101
4 NSX Edge 故障排除	105
Edge 防火墙数据包丢弃问题	109
Edge 路由连接问题	113
NSX Manager 与 Edge 通信问题	115
消息总线调试	116
Edge 诊断和恢复	118
5 防火墙故障排除	120
关于分布式防火墙	120
适用于 DFW 的 CLI 命令	121
对分布式防火墙进行故障排除	124
身份防火墙	129
6 负载均衡故障排除	133
场景：配置单路并联式负载均衡器	134
负载均衡器的故障排除流程图	139
使用 UI 验证负载均衡器配置和排除故障	139
使用 CLI 的负载均衡器故障排除	150
常见的负载均衡器问题	160
7 虚拟专用网络 (VPN) 故障排除	165
L2 VPN	165
L2 VPN 常见配置问题	165

用于缓解循环的 L2VPN 选项	167
使用 CLI 进行故障排除	170
SSL VPN	172
无法打开 SSL VPN Web 门户	172
SSL VPN-Plus: 安装故障	173
SSL VPN-Plus: 通信问题	175
SSL VPN-Plus: 身份验证问题	178
SSL VPN-Plus 客户端停止响应	178
基本日志分析	179
IPSEC VPN	180
成功协商 (阶段 1 和阶段 2)	180
阶段 1 策略不匹配	180
阶段 2 不匹配	182
PFS 不匹配	183
PSK 不匹配	184
成功协商的数据包捕获	184
8 NSX Controller 故障排除	190
了解控制器群集架构	190
NSX Controller 部署问题	193
磁盘延迟故障排除	197
查看磁盘延迟警报	197
磁盘延迟问题	198
NSX Controller 群集故障	200
方法 1: 删除已损坏的控制器并重新部署新的控制器	201
方法 2: 重新部署 NSX Controller 群集	204
幻影控制器	205
NSX Controller 已断开连接	206
控制层面代理 (netcpa) 问题	207
9 Guest Introspection 故障排除	211
Guest Introspection 架构	212
Guest Introspection 日志	213
ESX GI 模块 (MUX) 日志	213
GI 瘦代理日志	216
GI EPSecLib 和 SVM 日志	218
收集 Guest Introspection 环境和工作详细信息	220
Linux 或 Windows 上的瘦代理故障排除	221
ESX GI 模块 (MUX) 故障排除	224
EPSecLib 故障排除	225

NSX 故障排除指南

1

《NSX 故障排除指南》介绍了如何使用 NSX Manager 用户界面、vSphere Web Client 和其他 NSX 组件（根据需要）监控 VMware NSX[®] for vSphere[®] 系统和进行故障排除。

目标读者

本手册适用于要在 VMware vCenter 环境中使用 NSX 或解决任何 NSX 问题的用户。本手册的目标读者为熟悉虚拟机技术和虚拟数据中心操作且经验丰富的系统管理员。本手册假设您熟悉 VMware vSphere，包括 VMware ESXi、vCenter Server 和 vSphere Web Client。

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>。

本章讨论了以下主题：

- [一般故障排除准则](#)

一般故障排除准则

本主题介绍了在解决任何 NSX for vSphere 问题时遵循的一般准则。

- 1 转到 [使用 NSX 仪表板](#)，然后查看是否为组件显示任何错误或警告。
- 2 转到主 NSX Manager 的 **监控 (Monitor)** 选项卡，然后查看是否具有任何触发的系统事件。有关系统事件和警报的更多详细信息，请参阅 NSX 日志记录和系统事件。
- 3 使用 GET `api/2.0/services/systemalarms` API 查看有关 NSX 对象的警报。有关 API 的详细信息，请参阅 NSX API 指南。
- 4 按照 NSX 故障排除指南中所述解决该问题。
- 5 如果未解决您的问题，请下载技术支持日志并与 VMware 支持部门联系。请参见“[如何在 My VMware 中提出支持请求](#)”。有关如何下载日志的详细信息，请参阅 NSX 日志记录和系统事件。

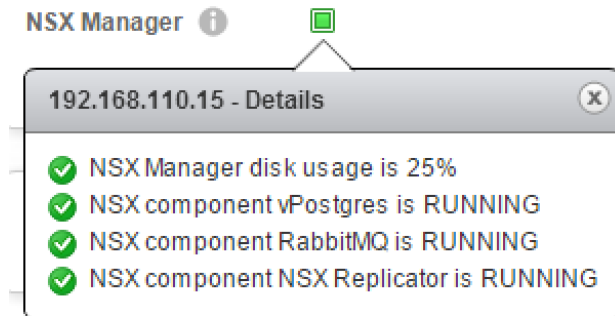
使用 NSX 仪表板

NSX 仪表板在一个中央视图中显示 NSX 组件的总体运行状况。NSX 仪表板显示各种 NSX 组件的状态以简化故障排除过程，如 NSX Manager、控制器、逻辑交换机、主机准备、服务部署、备份以及 Edge 通知。

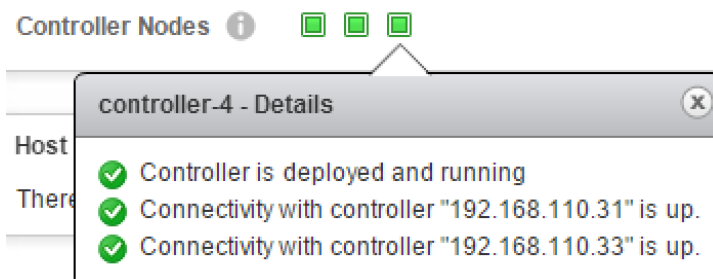
- 1 登录到 vSphere Web Client。
- 2 单击**网络和安全 (Networking & Security)**，然后单击**仪表板 (Dashboard)**。将显示“仪表板”页面。
- 3 在跨 vCenter NSX 环境中，选择具有主要角色或辅助角色的 NSX Manager。

仪表板提供以下信息：

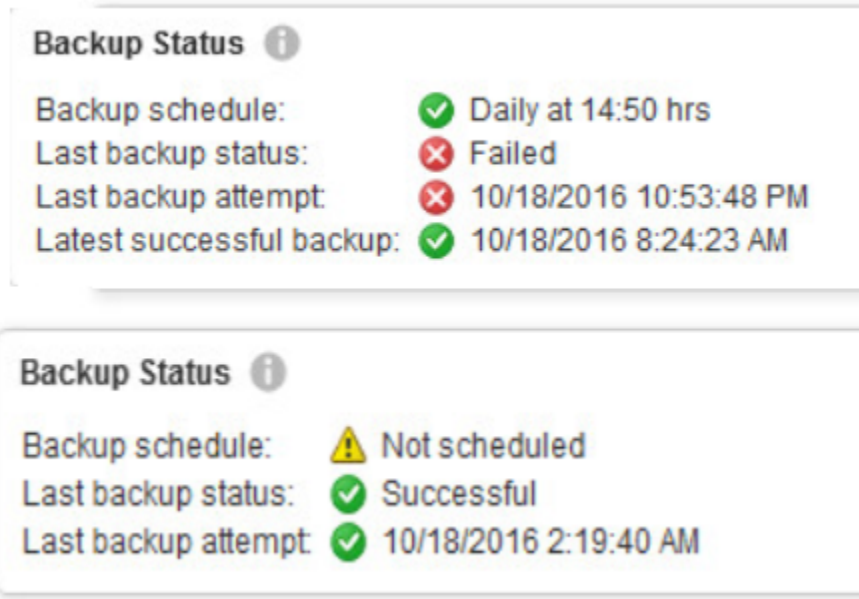
- NSX 基础架构 - 监控以下服务的 NSX Manager 组件状态：
 - 数据库服务 (vPostgres)。
 - 消息总线服务 (RabbitMQ)。
 - 复制程序服务 - 还会监控复制错误（如果已启用 跨 vCenter NSX）。
 - NSX Manager 磁盘使用率：
 - 黄色表示磁盘使用率超过 80%。
 - 红色表示磁盘使用率超过 90%。



- NSX 基础架构 - NSX Controller 状态：
 - 控制器节点状态（已启动/未启动/正在运行/正在部署/正在移除/失败/未知）。
 - 显示控制器对等连接状态。如果控制器已关闭并显示红色，对等控制器将显示为黄色。
 - 控制器虚拟机状态（已关闭电源/已删除）。
 - 控制器磁盘延迟警示。



- NSX Manager 备份状态：
 - 备份计划。
 - 上次备份状态（失败/成功/未调度以及日期和时间）。
 - 上次备份尝试（日期和时间以及详细信息）。
 - 上次成功备份（日期和时间以及详细信息）。



- NSX 基础架构 - 监控以下服务的主机状态：
 - 部署相关：
 - 具有安装失败状态的群集数。
 - 需要升级的群集数。
 - 正在进行安装的群集数。
 - 未准备的群集数。
 - 防火墙：
 - 禁用了防火墙的群集数。
 - 防火墙状态为黄色/红色的群集数：
 - 黄色表示在任何群集上禁用了分布式防火墙。
 - 红色表示无法在任何主机/群集上安装分布式防火墙。
 - VXLAN：
 - 未配置具有 VXLAN 的群集数。
 - VXLAN 状态为绿色/黄色/红色的群集数：
 - 绿色表示已成功配置该功能。

- 黄色表示在进行 **VXLAN** 配置时繁忙。
- 红色（错误）表示以下状态：**VTEP** 创建失败，**VTEP** 找不到 **IP** 地址，为 **VTEP** 分配了 *LinkLocal* **IP** 地址，等等。
- **NSX 基础架构 - 服务部署状态**
 - 部署失败 - 失败的部署的安装状态。
 - 服务状态 - 用于所有失败的服务。
- **NSX 基础架构 - NSX Edge 通知**

Edge 通知仪表板突出显示某些服务的活动警报。它监控下面列出的严重事件列表，并跟踪这些事件，直到解决了该问题。在报告恢复事件时，将自动解决警报，否则，将强制同步、重新部署或升级 Edge。

 - 负载均衡器（Edge 负载均衡器服务器状态）：
 - Edge 负载均衡器后端服务器已关闭。
 - Edge 负载均衡器后端服务器警告状态。
 - VPN（IPsec 隧道/IPsec 通道状态）：
 - Edge IPsec 通道已关闭。
 - Edge IPsec 隧道已关闭。
 - 设备（Edge 虚拟机、Edge 网关、Edge 文件系统、NSX Manager 以及 Edge 服务网关报告状态）：
 - Edge 服务网关缺少运行状况检查脉冲。
 - 已关闭 Edge 虚拟机电源。
 - Edge 虚拟机缺少运行状况检查脉冲。
 - NSX Edge 报告错误状态。
 - NSX Manager 报告 Edge 服务网关处于错误状态。
 - Edge 虚拟机在 VC 清单中不存在。

- 检测到 HA 脑裂。

注 在进行配置更新时，不会自动清除负载均衡器和 VPN 警报。在解决问题后，您必须通过 API 使用 `alarm-id` 命令手动清除警报。下面是可用于清除警报的 API 示例。有关详细信息，请参阅《NSX API 指南》。

```
GET https://<<NSX-IP>>/api/2.0/services/alarms/{source-Id}
POST https://<<NSX-IP>>/api/2.0/services/alarms?action=resolve

GET https://<<NSX-IP>>/api/2.0/services/systemalarms/<alarmId>
POST https://<<NSX-IP>>/api/2.0/services/systemalarms/<alarmId>?action=resolve
```

- NSX 服务 - 防火墙发布状态：
 - 防火墙发布状态为“失败”的主机数。在任何主机未成功应用发布的分布式防火墙配置时，状态为红色。
- NSX 服务 - 逻辑网络状态：
 - 具有错误或警告状态的逻辑交换机数。
 - 从 vCenter Server 中删除支持的分布式虚拟端口组时的标记。

NSX 命令行快速参考

您可以使用 NSX 命令行界面 (CLI) 解决问题。

表 1-1. 检查 ESXi 主机上的 NSX 安装 - 从 NSX Manager 中运行的命令

说明	NSX Manager 上的命令	备注
列出所有群集以获取群集 ID	<code>show cluster all</code>	查看所有群集信息
列出群集中的所有主机以获取主机 ID	<code>show cluster clusterID</code>	查看群集中的主机、主机 ID 和主机准备安装状态列表
列出主机上的所有虚拟机	<code>show host hostID</code>	查看特定主机信息、虚拟机、虚拟机 ID 和电源状态

表 1-2. 在主机上安装的 VIB 和模块名称（在命令中使用）

NSX 版本	ESXi 版本	VIB	模块
任何 6.3.x	5.5	esx-vxlan 和 esx-vsip	vdl2、vdrb、vsip、dvfilter-switch-security、bfd、traceflow
6.3.2 和更低版本	6.0 和更高版本	esx-vxlan 和 esx-vsip	vdl2、vdrb、vsip、dvfilter-switch-security、bfd、traceflow
6.3.3 和更高版本	6.0 和更高版本	esx-nsxv	nsx-vdl2、nsx-vdrb、nsx-vsip、nsx-dvfilter-switch-security、nsx-core、nsx-bfd、nsx-traceflow

表 1-3. 检查 ESXi 主机上的 NSX 安装 - 从主机中运行的命令

说明	主机上的命令	备注
包含的 VIB 取决于 NSX 和 ESXi 版本。 有关在您的安装上检查哪些模块的详细信息，请参见在主机上安装的 VIB 和模块名称表。	<code>esxcli software vib get -- vibname <name></code>	检查安装的版本/日期 <code>esxcli software vib list</code> 显示系统上的所有 VIB 的列表
列出当前在系统中加载的所有系统模块：	<code>esxcli system module list</code>	较旧的等效命令： <code>vmkload_mod -l grep -E vdl2 vdrb vsip dvfilter-switch-security</code>
包含的模块取决于 NSX 和 ESXi 版本。 有关在您的安装上检查哪些模块的详细信息，请参见在主机上安装的 VIB 和模块名称表。	<code>esxcli system module get -m <name></code>	为每个模块运行该命令
两个用户环境代理 (UWA)：控制层面代理、防火墙代理	<code>/etc/init.d/vShield-Stateful-Firewall status</code> <code>/etc/init.d/netcpad status</code>	
检查 UWA 连接：端口 1234 到控制器以及端口 5671 到 NSX Manager	<code>esxcli network ip connection list grep 1234</code> <code>esxcli network ip connection list grep 5671</code>	控制器 TCP 连接 消息总线 TCP 连接
检查 EAM 状态	vSphere Web Client，检查 管理 > vSphere ESX Agent Manager (Administration > vSphere ESX Agent Manager)	

表 1-4. 检查 ESXi 主机上的 NSX 安装 - 主机网络命令

说明	主机网络命令	备注
列出物理网卡/vmnic	<code>esxcli network nic list</code>	检查网卡类型、驱动程序类型、链路状态、MTU
物理网卡详细信息	<code>esxcli network nic get -n vmnic#</code>	检查驱动程序和固件版本以及其他详细信息

表 1-4. 检查 ESXi 主机上的 NSX 安装 - 主机网络命令（续）

说明	主机网络命令	备注
列出 vmk 网卡以及 IP 地址/MAC/MTU 等	<code>esxcli network ip interface ipv4 get</code>	确保正确实例化 VTEP
每个 vmk 网卡的详细信息，包括 vDS 信息	<code>esxcli network ip interface list</code>	确保正确实例化 VTEP
每个 vmk 网卡的详细信息，包括 VXLAN vmk 的 vDS 信息	<code>esxcli network ip interface list --netstack=vxlan</code>	确保正确实例化 VTEP
查找与该主机的 VTEP 关联的 VDS 名称	<code>esxcli network vswitch dvs vmware vxlan list</code>	确保正确实例化 VTEP
从 VXLAN 专用 TCP/IP 堆中执行 Ping 操作	<code>ping ++netstack=vxlan -I vmk1 x.x.x.x</code>	要解决 VTEP 通信问题：添加 <code>-d -s 1572</code> 选项以确保传输网络的 MTU 适用于 VXLAN
查看 VXLAN 专用 TCP/IP 堆的路由表	<code>esxcli network ip route ipv4 list -N vxlan</code>	解决 VTEP 通信问题
查看 VXLAN 专用 TCP/IP 堆的 ARP 表	<code>esxcli network ip neighbor list -N vxlan</code>	解决 VTEP 通信问题

表 1-5. 检查 ESXi 主机上的 NSX 安装 - 主机日志文件

说明	日志文件	备注
从 NSX Manager 中	<code>show manager log follow</code>	跟踪 NSX Manager 日志 适用于实时故障排除
主机的任何安装相关日志	<code>/var/log/esxupdate.log</code>	
与主机相关的问题	<code>/var/log/vmkernel.log</code>	
VMkernel 警告、消息、警示和可用性报告	<code>/var/log/vmksummary.log</code> <code>/var/log/vmkwarning.log</code>	
捕获模块加载故障	<code>/var/log/syslog</code>	IXGBE 驱动程序故障 NSX 模块相关性故障是重要指标
在 vCenter 上，ESX Agent Manager 负责进行更新	在 vCenter 日志 <code>eam.log</code> 中	

表 1-6. 检查逻辑交换 - 从 NSX Manager 中运行的命令

说明	NSX Manager 上的命令	备注
列出所有逻辑交换机	<code>show logical-switch list all</code>	列出在 API、传输区域和 <code>vdnscope</code> 中使用的所有逻辑交换机及其 UUID

表 1-7. 逻辑交换 - 从 NSX Controller 中运行的命令

说明	控制器上的命令	备注
查找作为 VNI 所有者的控制器	<code>show control-cluster logical-switches vni 5000</code>	记下输出中的控制器 IP 地址并通过 SSH 访问该地址
查找连接到该 VNI 的该控制器的所有主机	<code>show control-cluster logical-switch connection-table 5000</code>	输出中的源 IP 地址是主机的管理接口，而端口号是 TCP 连接的源端口

表 1-7. 逻辑交换 - 从 NSX Controller 中运行的命令（续）

说明	控制器上的命令	备注
查找注册以托管该 VNI 的 VTEP	<code>show control-cluster logical-switches vtep-table 5002</code>	
列出为该 VNI 上的虚拟机获悉的 MAC 地址	<code>show control-cluster logical-switches mac-table 5002</code>	指出 MAC 地址实际位于报告该地址的 VTEP 上
列出虚拟机 IP 更新填充的 ARP 缓存	<code>show control-cluster logical-switches arp-table 5002</code>	ARP 缓存在 180 秒后过期
对于特定的主机/控制器对，找出主机已加入的 VNI	<code>show control-cluster logical-switches joined-vnis <host_mgmt_ip></code>	

表 1-8. 逻辑交换 - 从主机中运行的命令

说明	主机上的命令	备注
检查主机 VXLAN 是否同步	<code>esxcli network vswitch dvs vmware vxlan get</code>	显示同步状态和用于封装的端口
查看连接的虚拟机以及用于数据路径捕获的本地交换机端口 ID	<code>net-stats -l</code>	提供了一种更好的方法以获取特定虚拟机的虚拟机交换机端口
验证是否加载了 VXLAN 内核模块 vdl2	<code>esxcli system module get -m vdl2</code>	显示指定的模块的完整详细信息 验证版本
验证是否安装了正确的 VXLAN VIB 版本 有关您的安装上检查哪些 VIB 的详细信息，请参见在主机上安装的 VIB 和模块名称表。	<code>esxcli software vib get --vibName esx-vxlan</code> 或 <code>esxcli software vib get --vibName esx-nsxv</code>	显示指定的 VIB 的完整详细信息 验证版本和日期
验证主机是否了解逻辑交换机中的其他主机	<code>esxcli network vswitch dvs vmware vxlan network vtep list --vxlan-id=5001 --vds-name=Compute_VDS</code>	显示该主机了解并托管 vtep 5001 的所有 VTEP 的列表
验证逻辑交换机的控制层面是否已启动并处于活动状态	<code>esxcli network vswitch dvs vmware vxlan network list --vds-name Compute_VDS</code>	确保控制器连接已启动并且端口/Mac 数与该主机上的 LS 中的虚拟机数相匹配。
验证主机是否获悉所有虚拟机的 MAC 地址	<code>esxcli network vswitch dvs vmware vxlan network mac list --vds-name Compute_VDS --vxlan-id=5000</code>	这会列出该主机上的 VNI 5000 虚拟机的所有 MAC
验证主机是否在本地缓存远程虚拟机的 ARP 条目	<code>esxcli network vswitch dvs vmware vxlan network arp list --vds-name Compute_VDS --vxlan-id=5000</code>	验证主机是否在本地缓存远程虚拟机的 ARP 条目

表 1-8. 逻辑交换 - 从主机中运行的命令（续）

说明	主机上的命令	备注
验证虚拟机是否连接到 LS 并映射到本地 VMKnic 还会显示虚拟机 dvPort 映射到的 vmknic ID	<code>esxcli network vswitch dvs vmware vxlan network port list --vds-name Compute_VDS --vxlan-id=5000</code>	只要 VNI 连接到路由器，就会始终列出 vdrport
查看 vmknic ID 以及它们映射到的交换机端口/上行链路	<code>esxcli network vswitch dvs vmware vxlan vmknic list --vds-name=DSwitch-Res01</code>	

表 1-9. 检查逻辑交换 - 日志文件

说明	日志文件	备注
主机始终连接到托管其 VNI 的控制器	<code>/etc/vmware/netcpa/config-by-vsm.xml</code>	该文件应始终列出环境中的所有控制器。 <code>config-by-vsm.xml</code> 文件是由 netcpa 进程创建的
<code>config-by-vsm.xml</code> 文件是 NSX Manager 使用 vsfwd 推送的 如果 <code>config-by-vsm.xml</code> 文件不正确，请查看 vsfwd 日志	<code>/var/log/vsfwd.log</code>	分析该文件以查找错误 要重新启动进程，请运行以下命令： <code>/etc/init.d/vShield-Stateful-Firewall stop start</code>
到控制器的连接是使用 netcpa 建立的	<code>/var/log/netcpa.log</code>	分析该文件以查找错误
逻辑交换模块日志位于 <code>vmkernel.log</code> 中	<code>/var/log/vmkernel.log</code>	在 <code>/var/log/vmkernel.log</code> 中检查“具有 VXLAN 前缀”的逻辑交换模块：

表 1-10. 检查逻辑路由 - 从 NSX Manager 中运行的命令

说明	NSX Manager 上的命令	备注
用于 ESG 的命令	<code>show edge</code>	用于 Edge 服务网关 (ESG) 的 CLI 命令以“show edge”开头
用于 DLR 控制虚拟机的命令	<code>show edge</code>	用于分布式逻辑路由器 (DLR) 控制虚拟机的 CLI 命令以“show edge”开头
用于 DLR 的命令	<code>show logical-router</code>	用于分布式逻辑路由器 (DLR) 的 CLI 命令以 <code>show logical-router</code> 开头
列出所有 Edge	<code>show edge all</code>	列出支持集中式 CLI 的所有 Edge
列出 Edge 的所有服务和部署详细信息	<code>show edge edgeID</code>	查看 Edge 服务网关信息
列出 Edge 的命令选项	<code>show edge edgeID ?</code>	查看详细信息，例如，版本、日志、NAT、路由表、防火墙、配置、接口和服务
查看路由详细信息	<code>show edge edgeID ip ?</code>	查看路由信息、BGP、OSPF 和其他详细信息
查看路由表	<code>show edge edgeID ip route</code>	查看 Edge 中的路由表
查看路由邻居	<code>show edge edgeID ip ospf neighbor</code>	查看路由邻居关系
查看逻辑路由器连接信息	<code>show logical-router host hostID connection</code>	验证连接的 LIF 数是否正确，成组策略是否正确以及是否使用相应的 vDS

表 1-10. 检查逻辑路由 - 从 NSX Manager 中运行的命令（续）

说明	NSX Manager 上的命令	备注
列出在主机上运行的所有逻辑路由器实例	<code>show logical-router host hostID dlr all</code>	验证 LIF 和路由数 在逻辑路由器的所有主机上，控制器 IP 应该相同 Control Plane Active 应该为 yes --brief 提供了精简响应
检查主机上的路由表	<code>show logical-router host hostID dlr dlrID route</code>	这是控制器推送到传输区域中的所有主机的路由表 在所有主机上，该表必须是相同的 如果在少数主机上缺少某些路由，请尝试从控制器中运行前面提到的 sync 命令 E 标记表示路由是通过 ECMP 获悉的
检查主机上的 DLR 的 LIF	<code>show logical-router host hostID dlr dlrID interface (all intName) verbose</code>	LIF 信息将从控制器推送到主机 可以使用该命令确保主机了解应了解的所有 LIF

表 1-11. 检查逻辑路由 - 从 NSX Controller 中运行的命令

说明	NSX Controller 上的命令	备注
查找所有逻辑路由器实例	<code>show control-cluster logical-routers instance all</code>	这会列出逻辑路由器实例以及传输区域中具有逻辑路由器实例的所有主机 此外，还会显示为该逻辑路由器提供服务的控制器
查看每个逻辑路由器的详细信息	<code>show control-cluster logical-routers instance 0x570d4555</code>	IP 列显示该 DLR 所在的所有主机的 vmk0 IP 地址
查看连接到逻辑路由器的所有接口	<code>show control-cluster logical-routers interface-summary 0x570d4555</code>	IP 列显示该 DLR 所在的所有主机的 vmk0 IP 地址
查看该逻辑路由器获悉的所有路由	<code>show control-cluster logical-routers routes 0x570d4555</code>	请注意，IP 列显示该 DLR 所在的所有主机的 vmk0 IP 地址
显示建立的所有网络连接，类似于 net stat 输出	<code>show network connections of-type tcp</code>	检查要排除故障的主机是否建立到控制器的 netcpa 连接
将接口从控制器同步到主机	<code>sync control-cluster logical-routers interface-to-host <logical-router-id> <host-ip></code>	如果新接口连接到逻辑路由器，但未同步到所有主机，这是非常有用的
将路由从控制器同步到主机	<code>sync control-cluster logical-routers route-to-host <logical-router-id> <host-ip></code>	如果在少数主机上缺少某些路由，但这些路由在大多数主机上可用，这是非常有用的

表 1-12. 检查逻辑路由 - 从 Edge 中运行的命令

说明	Edge 或逻辑路由器控制虚拟机上的命令	备注
查看配置	<code>show configuration <global bgp ospf ...></code>	
查看获悉的路由	<code>show ip route</code>	确保路由和转发表保持同步

表 1-12. 检查逻辑路由 - 从 Edge 中运行的命令（续）

说明	Edge 或逻辑路由器控制虚拟机上的命令	备注
查看转发表	<code>show ip forwarding</code>	确保路由和转发表保持同步
查看分布式逻辑路由器接口	<code>show interface</code>	<p>在输出中显示的第一个网卡是分布式逻辑路由器接口</p> <p>分布式逻辑路由器接口不是该虚拟机上的真正 vNIC</p> <p>连接到分布式逻辑路由器的所有子网具有 INTERNAL 类型</p>
查看其他接口（管理）	<code>show interface</code>	<p>管理/HA 接口是逻辑路由器控制虚拟机上的真正 vNIC</p> <p>如果启用 HA 而未指定 IP 地址，则使用 169.254.x.x/30</p> <p>如果为管理接口分配了 IP 地址，则会在此处显示该地址</p>
调试协议	<code>debug ip ospf</code> <code>debug ip bgp</code>	<p>在查看配置问题（例如，不匹配的 OSPF 区域、计时器和错误的 ASN）时，这是非常有用的</p> <p>注意：只能在 Edge 控制台上查看输出（而不能通过 SSH 会话）</p>
OSPF 命令	<code>show configuration ospf</code> <code>show ip ospf interface</code> <code>show ip ospf neighbor</code> <code>show ip route ospf</code> <code>show ip ospf database</code> <code>show tech-support</code> （并查找字符串“EXCEPTION”和“PROBLEM”）	
BGP 命令	<code>show configuration bgp</code> <code>show ip bgp neighbor</code> <code>show ip bgp</code> <code>show ip route bgp</code> <code>show ip forwarding</code> <code>show tech-support</code> （查找字符串“EXCEPTION”和“PROBLEM”）	

表 1-13. 检查逻辑路由 - 主机中的日志文件

说明	日志文件	备注
vsfwd 将分布式逻辑路由器实例信息推送到主机并保存为 XML 格式	/etc/vmware/netcpa/config-by-vsm.xml	如果在主机上缺少分布式逻辑路由器实例，请先查看该文件以确定是否列出该实例 如果未列出，请重新启动 vsfwd 此外，还可以使用该文件确保主机了解所有控制器
上述文件是 NSX Manager 使用 vsfwd 推送的 如果 config-by-vsm.xml 文件不正确，请查看 vsfwd 日志	/var/log/vsfwd.log	分析该文件以查找错误 要重新启动进程，请运行以下命令： <code>/etc/init.d/vShield-Stateful-Firewall stop start</code>
到控制器的连接是使用 netcpa 建立的	/var/log/netcpa.log	分析该文件以查找错误
逻辑交换模块日志位于 vmkernel.log 中	/var/log/vmkernel.log	在 /var/log/vmkernel.log 中检查“具有 vxlan 前缀”的逻辑交换模块：

表 1-14. 控制器调试 - 从 NSX Manager 中运行的命令

描述	命令（在 NSX Manager 上）	备注
列出所有控制器及其状态	show controller list all	显示所有控制器及其运行状态的列表

表 1-15. 控制器调试 - 从 NSX Controller 中运行的命令

说明	命令（在控制器上）	备注
检查控制器群集状态	show control-cluster status	应始终显示“Join complete”和“Connected to Cluster Majority”
检查抖动连接的统计信息和消息	show control-cluster core stats	丢弃的数据包计数器不应发生变化
查看与最初加入群集或重新启动后有关的节点活动	show control-cluster history	这对于解决群集加入问题非常有用
查看群集中的节点列表	show control-cluster startup-nodes	请注意，该列表不需要仅包含活动群集节点 它应该是具有当前部署的所有控制器的列表 在启动控制器以联系群集中的其他控制器时，可以使用该列表
显示建立的所有网络连接，类似于 net stat 输出	show network connections of-type tcp	检查要排除故障的主机是否建立到控制器的 netcpa 连接
重新启动控制器进程	restart controller	仅重新启动主控制器进程 强制重新连接到群集
重新引导控制器节点	restart system	重新引导控制器虚拟机

表 1-16. 控制器调试 - NSX Controller 上的日志文件

说明	日志文件	备注
查看控制器历史记录以及最近的加入、重新启动，等等	<code>show control-cluster history</code>	用于解决控制器问题的极佳工具，尤其是解决群集问题
检查速度较慢的磁盘	<code>show log cloudnet/cloudnet_java-zookeeper<timestamp>.log filtered-by fsync</code>	一种检查速度较慢的磁盘的可靠方法是，在 <code>cloudnet_java-zookeeper</code> 日志中查找“fsync”消息 如果同步所需的时间超过 1 秒，ZooKeeper 将输出该消息，这很好地指明了其他程序此时正在使用该磁盘
检查速度较慢/发生故障的磁盘	<code>show log syslog filtered-by collectd</code>	有关“collectd”的示例输出中的消息往往与速度较慢或发生故障的磁盘有关
检查磁盘空间使用率	<code>show log syslog filtered-by freespace:</code>	在空间使用率达到某个阈值时，名为“freespace”的后台作业定期从磁盘中清除旧日志和其他文件。在某些情况下，如果磁盘很小以及/或者填充速度很快，则会看到大量 <code>freespace</code> 消息。这可能表明磁盘已填满
查找当前的活动群集成员	<code>show log syslog filtered-by Active cluster members</code>	列出当前的活动群集成员的节点 ID。可能需要查看较早的 <code>syslog</code> ，因为并非始终输出该消息。
查看核心控制器日志	<code>show log cloudnet/cloudnet_java-zookeeper.20150703-165223.3702.1og</code>	可能具有多个 <code>zookeeper</code> 日志，请查看具有最新时间戳的文件 该文件包含有关选择的控制器群集主控制器的信息以及与控制器的分布式特性有关的其他信息
查看核心控制器日志	<code>show log cloudnet/cloudnet.nsx-controller.root.log.INFO.20150703-165223.3668</code>	主控制器工作日志，例如，LIF 创建时间、1234 上的连接侦听器、分片

表 1-17. 检查分布式防火墙 - 从 NSX Manager 中运行的命令

说明	NSX Manager 上的命令	备注
查看虚拟机信息	<code>show vm vmID</code>	DC、群集、主机、虚拟机名称、vNIC、安装的 dvfilter 等详细信息
查看特定的虚拟网卡信息	<code>show vnic icID</code>	VNIC 名称、mac 地址、端口组、应用的筛选器等详细信息
查看所有群集信息	<code>show dfw cluster all</code>	群集名称、群集 ID、数据中心名称、防火墙状态
查看特定的群集信息	<code>show dfw cluster clusterID</code>	主机名、主机 ID、安装状态
查看 dfw 相关主机信息	<code>show dfw host hostID</code>	虚拟机名称、虚拟机 ID、电源状态
查看 dvfilter 中的详细信息	<code>show dfw host hostID filter filterID <option></code>	列出每个 VNIC 的规则、统计信息、地址集等
查看虚拟机的 DFW 信息	<code>show dfw vm vmID</code>	查看虚拟机的名称、VNIC ID、筛选器等
查看 VNIC 详细信息	<code>show dfw vnic vnicID</code>	查看 VNIC 名称、ID、MAC 地址、端口组、筛选器

表 1-17. 检查分布式防火墙 - 从 NSX Manager 中运行的命令（续）

说明	NSX Manager 上的命令	备注
列出为每个 vNIC 安装的筛选器	<code>show dfw host hostID summarize-dvfilter</code>	查找感兴趣的虚拟机/vNIC，并获取名称字段以在后续命令中作为筛选器
查看特定筛选器/vNIC 的规则	<code>show dfw host hostID filter filterID rules</code> <code>show dfw vnic nicID</code>	
查看地址集的详细信息	<code>show dfw host hostID filter filterID addrsets</code>	这些规则仅显示地址集，可以使用该命令扩充地址集包含的内容
每个 vNIC 的 spoofguard 详细信息	<code>show dfw host hostID filter filterID spoofguard</code>	检查是否启用了 SpoofGuard 以及当前的 IP/MAC
查看流量记录详细信息	<code>show dfw host hostID filter filterID flows</code>	如果启用了流量监控，主机定期将流量信息发送到 NSX Manager 可以使用该命令查看每个 vNIC 的流量
查看 vNIC 的每个规则的统计信息	<code>show dfw host hostID filter filterID stats</code>	在查看是否命中规则时，这是非常有用的

表 1-18. 检查分布式防火墙 - 从主机中运行的命令

说明	主机上的命令	备注
列出在主机上下载的 VIB。 有关在您的安装上检查哪些 VIB 的详细信息，请参见在主机上安装的 VIB 和模块名称表。	<code>esxcli software vib list grep esx-vmip</code> 或 <code>esxcli software vib list grep esx-nsxv</code>	检查以确保下载了正确的 VIB 版本
有关当前加载的系统模块的详细信息 有关在您的安装上检查哪些模块的详细信息，请参见在主机上安装的 VIB 和模块名称表。	<code>esxcli system module get -m vmip</code> 或 <code>esxcli system module get -m nsx-vmip</code>	检查以确保安装/加载了模块
进程列表	<code>ps grep vsfwd</code>	查看是否使用多个线程运行 vsfwd 进程
守护程序命令	<code>/etc/init.d/vShield-Stateful-Firewall {start stop status restart}</code>	检查守护程序是否正在运行，并根据需要重新启动
查看网络连接	<code>esxcli network ip connection list grep 5671</code>	检查主机是否具有到 NSX Manager 的 TCP 连接

表 1-19. 检查分布式防火墙 - 主机上的日志文件

说明	日志	备注
进程日志	/var/log/vsfwd.log	vsfwd 守护程序日志，对 vsfwd 进程、NSX Manager 连接和 RabbitMQ 故障排除非常有用
数据包日志专用文件	/var/log/dfwpktlogs.log	数据包日志的专用日志文件
dvfilter 中的数据包捕获	pktpcap-uw --dvfilter nic-1413082-eth0-vmware-sfw.2 -- outfile test.pcap	

NSX 主机运行状况检查

从 NSX Manager 集中式 CLI 中，您可以检查每个 ESXi 主机的运行状态。

运行状态将报告为 critical、unhealthy 或 healthy。

例如：

```
nsxmgr> show host host-30 health-status
status: HEALTHY

nsxmgr> show host host-29 health-status
UNHEALTHY, Standard Switch vSwitch1 has no uplinks.
UNHEALTHY, Storage volume datastore1 has no enough free spaces: 19.% free.
status: UNHEALTHY

nsxmgr> show host host-28 health-status
CRITICAL, VXLAN VDS vds-site-a VNI 200000 multicast addr is not synchronized with VSM: 0.0.0.0.
CRITICAL, VXLAN VDS vds-site-a VNI 200003 multicast addr is not synchronized with VSM: 0.0.0.0.
CRITICAL, VXLAN VDS vds-site-a VNI 5000 multicast addr is not synchronized with VSM: 0.0.0.0.
Status: CRITICAL
```

也可以通过 NSX Manager API 调用 host-check 命令。

NSX 基础架构故障排除

2

NSX 准备是一个包含 4 个步骤的过程。

- 1 将 NSX Manager 连接到 vCenter Server。NSX Manager 和 vCenter Server 具有一对一关系。
 - a 在 vCenter Server 中注册。
- 2 部署 NSX Controller（仅处于单播或混合模式的逻辑交换、分布式路由或 VXLAN 需要。如果仅使用分布式防火墙 (Distributed Firewall, DFW)，则不需要使用控制器）。
- 3 主机准备：在群集中的所有主机上为 XLAN、DFW 和 DLR 安装 VIB。配置基于 Rabbit MQ 的消息传递基础架构。启用防火墙。通知控制器已为 NSX 准备好主机。
- 4 配置 IP 池设置并配置 VXLAN：在群集中的所有主机上创建 VTEP 端口组和 VMKNIC。在该步骤期间，您可以设置传输 VLAN ID、绑定策略和 MTU。

有关每个步骤的安装和配置的详细信息，请参阅《NSX 安装指南》和《NSX 管理指南》。

本章讨论了以下主题：

- [主机准备](#)
- [解决 NSX Manager 问题](#)
- [逻辑网络准备：VXLAN 传输](#)
- [逻辑交换机端口组不同步](#)

主机准备

vSphere ESX Agent Manager 将 vSphere 安装包 (VIB) 部署到 ESXi 主机上。

主机上的部署要求在主机、vCenter Server 和 NSX Manager 上配置 DNS。部署不需要重新引导 ESXi 主机，但任何 VIB 更新或移除需要重新引导 ESXi 主机。

VIB 是在 NSX Manager 上托管的，也可以作为 zip 文件提供。

可以从 <https://<NSX-Manager-IP>/bin/vdn/nwfabric.properties> 中访问该文件。可下载的 zip 文件因 NSX 和 ESXi 版本而异。例如，在 NSX 6.3.0 中，vSphere 6.0 主机使用 <https://<NSX-Manager-IP>/bin/vdn/vibs-6.3.0/6.0-buildNumber/vxlan.zip> 文件。

```
# 5.5 VDN EAM Info
VDN_VIB_PATH.1=/bin/vdn/vibs-6.3.0/5.5-4744075/vxlan.zip
VDN_VIB_VERSION.1=4744075
VDN_HOST_PRODUCT_LINE.1=embeddedEsx
VDN_HOST_VERSION.1=5.5.*

# 6.0 VDN EAM Info
VDN_VIB_PATH.2=/bin/vdn/vibs-6.3.0/6.0-4744062/vxlan.zip
VDN_VIB_VERSION.2=4744062
VDN_HOST_PRODUCT_LINE.2=embeddedEsx
VDN_HOST_VERSION.2=6.0.*

# 6.5 VDN EAM Info
VDN_VIB_PATH.3=/bin/vdn/vibs-6.3.0/6.5-4744074/vxlan.zip
VDN_VIB_VERSION.3=4744074
VDN_HOST_PRODUCT_LINE.3=embeddedEsx
VDN_HOST_VERSION.3=6.5.*

# Single Version associated with all the VIBs pointed by above VDN_VIB_PATH(s)
VDN_VIB_VERSION=6.3.0.4744320

# Legacy vib location. Used by code to discover available legacy vibs.
LEGACY_VDN_VIB_PATH_FS=/common/em/components/vdn/vibs/legacy/
LEGACY_VDN_VIB_PATH_WEB_ROOT=/bin/vdn/vibs/legacy/
```

在主机上安装的 VIB 取决于 NSX 和 ESXi 版本：

ESXi 版本	NSX 版本	安装的 VIB
5.5	任何 6.3.x	<ul style="list-style-type: none"> ■ esx-vsip ■ esx-vxlan
6.0 或更高版本	6.3.2 或更低版本	<ul style="list-style-type: none"> ■ esx-vsip ■ esx-vxlan
6.0 或更高版本	6.3.3 或更高版本	<ul style="list-style-type: none"> ■ esx-nsxv

您可以使用 `esxcli software vib list` 命令查看安装的 VIB。

```
[root@esx-01a:~] esxcli software vib list | grep -e vsip -e vxlan
esx-vsip                6.0.0-0.0.XXXXXXX    VMware  VMwareCertified
2016-04-20
esx-vxlan                6.0.0-0.0.XXXXXXX    VMware  VMwareCertified
2016-04-20
```

或

```
esxcli software vib list | grep nsxv
esx-nsxv                6.0.0-0.0.XXXXXXX          VMware  VMwareCertified
2017-08-11
```

主机准备期间的常见问题

在主机准备期间，可能遇到的典型问题如下所示：

- EAM 无法部署 VIB。
 - 可能是由于未在主机上正确配置 DNS。
 - 可能是由于防火墙阻止 ESXi、NSX Manager 和 vCenter Server 之间的所需端口。
 可以单击 **解决 (Resolve)** 选项以解决大多数问题。请参阅 [安装状态为“未就绪”](#)。
- 已安装以前的旧 VIB 版本。这需要用户干预以重新引导主机。
- NSX Manager 和 vCenter Server 遇到通信问题。“网络和安全”插件中的 **主机准备 (Host Preparation)** 选项卡未正确显示所有主机：
 - 检查 vCenter Server 是否可以枚举所有主机和群集。

如果无法使用 **解决 (Resolve)** 选项修复问题，请参阅 [无法使用“解决”选项修复问题](#)。

主机准备 (VIB) 故障排除

- 检查主机的通信通道运行状况。请参见 [检查通信通道运行状况](#)。
- 检查 vSphere ESX Agent Manager 以查找错误。

vCenter 主页 > 管理 > vCenter Server 扩展 > vSphere ESX Agent Manager (vCenter home > Administration > vCenter Server Extensions > vSphere ESX Agent Manager)。

在 vSphere ESX Agent Manager 上，检查带有“VCNS160”前缀的代理机构的状态。如果某个代理机构处于错误的状态，请选择该代理机构并查看其问题。

Agency	State	Status	Optimized Deployment
_VCNS_160_Management & Edge Cl...	Enabled	✓ Normal	✓
_VCNS_160_Compute Cluster A_VMwa...	Enabled	✗ Alert	✓

Issues for the selected agencies				
Trigger Time	Agency	Issue	Host	Agent VM
Thu Apr 28 12:03:12 GMT-0...	_VCNS_160_Compute Clu...	Agent VIB module is not installed	esx-01a.corp.local	

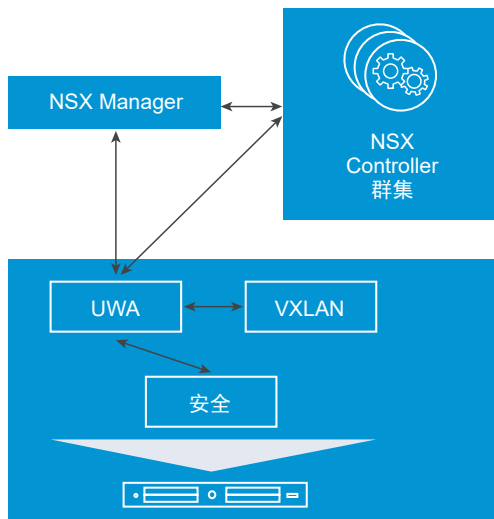
- 在出现问题的主机上，运行 `tail /var/log/esxupdate.log` 命令。

```
2016-04-28T19:02:52Z esxupdate: downloader: DEBUG: Downloading https://vcsa-0
o /tmp/tmpKT0wjN...
2016-04-28T19:03:12Z esxupdate: esxupdate: ERROR: An esxupdate error excepti
2016-04-28T19:03:12Z esxupdate: esxupdate: ERROR: Traceback (most recent call
2016-04-28T19:03:12Z esxupdate: esxupdate: ERROR: File "/usr/sbin/esxupdate
2016-04-28T19:03:12Z esxupdate: esxupdate: ERROR: cmd.Run()
2016-04-28T19:03:12Z esxupdate: esxupdate: ERROR: File "/build/mts/release/
site-packages/vmware/esx5update/Cmdline.py", line 106, in Run
2016-04-28T19:03:12Z esxupdate: esxupdate: ERROR: File "/build/mts/release/
site-packages/vmware/esximage/Transaction.py", line 73, in DownloadMetadatas
2016-04-28T19:03:12Z esxupdate: esxupdate: ERROR: MetadataDownloadError: ('ht
fd3f37ad4c', None, "('https://vcsa-01a.corp.local:443/eam/vib?id=facdb160-21f
rlopen error [Errno -3] Temporary failure in name resolution>")")
2016-04-28T19:03:12Z esxupdate: esxupdate: DEBUG: <<<
```

主机准备 (UWA) 故障排除

NSX Manager 在群集中的所有主机上配置两个用户环境代理:

- 消息总线 UWA (vsfwd)
- 控制层面 UWA (netcpa)



在极少数情况下，VIB 安装成功，但由于某种原因，一个或两个用户环境代理无法正常工作。这可能表现为：

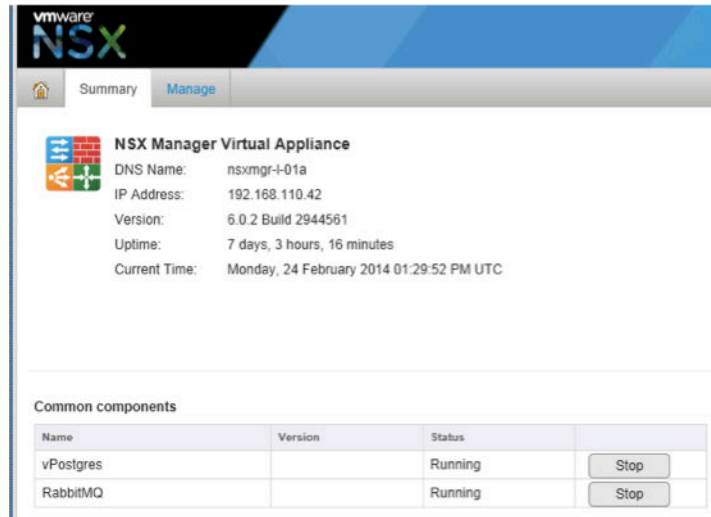
- 防火墙显示错误的状态。

Clusters & Hosts	Installation Status	Firewall
▶  dc-1	✔ 5.0 Uninstall	⚠ Error

- 管理程序和控制器之间的控制层面关闭。检查 **NSX Manager** 系统事件。请参阅 **NSX** 日志记录和系统事件。

Getting Started Summary Monitor Manage				
Audit Logs System Events Tasks				
Timestamp	Severity	Event Source	Code	Event Message
2/26/2014 10:56:38 AM	Critical	Host messaging infrastructure	391002	Messaging infrastructure down on host.
2/26/2014 10:51:56 AM	Critical	host-22	301502	Spoofguard configuration update number 139340752032...
2/26/2014 10:51:56 AM	Critical	host-20	301502	Spoofguard configuration update number 139340752032...

如果多个 ESXi 主机受到影响，请在 NSX Manager 设备 Web UI 摘要 (Summary) 选项卡下面检查消息总线服务的状态。如果已停止，请重新启动 RabbitMQ。



如果消息总线服务在 NSX Manager 上处于活动状态，请执行以下操作：

- 在 ESXi 主机上运行 `/etc/init.d/vShield-Stateful-Firewall status` 命令以检查主机上的消息总线用户环境代理状态。

```
[root@esx-01a:~] /etc/init.d/vShield-Stateful-Firewall status
vShield-Stateful-Firewall is running
```

- 检查主机上的消息总线用户环境代理日志 `/var/log/vsfwd.log`。
- 在 ESXi 主机上运行 `esxcfg-advcfg -l | grep Rmq` 命令以显示所有 Rmq 变量。应该有 16 个 Rmq 变量。

```
[root@esx-01a:~] esxcfg-advcfg -l | grep Rmq
/UserVars/RmqIpAddress [String] : Connection info for RMQ Broker
/UserVars/RmqUsername [String] : RMQ Broker Username
/UserVars/RmqPassword [String] : RMQ Broker Password
/UserVars/RmqVHost [String] : RMQ Broker VHost
/UserVars/RmqVsmRequestQueue [String] : RMQ Broker VSM Request Queue
/UserVars/RmqPort [String] : RMQ Broker Port
/UserVars/RmqVsmExchange [String] : RMQ Broker VSM Exchange
/UserVars/RmqClientPeerName [String] : RMQ Broker Client Peer Name
/UserVars/RmqHostId [String] : RMQ Broker Client HostId
/UserVars/RmqHostVer [String] : RMQ Broker Client HostVer
```

```

/UserVars/RmqClientId [String] : RMQ Broker Client Id
/UserVars/RmqClientToken [String] : RMQ Broker Client Token
/UserVars/RmqClientRequestQueue [String] : RMQ Broker Client Request Queue
/UserVars/RmqClientResponseQueue [String] : RMQ Broker Client Response Queue
/UserVars/RmqClientExchange [String] : RMQ Broker Client Exchange
/UserVars/RmqSslCertSha1ThumbprintBase64 [String] : RMQ Broker Server Certificate base64 Encoded
Sha1 Hash

```

- 在 ESXi 主机上运行 `esxcfg-advcfg -g /UserVars/RmqIpAddress` 命令。输出将显示 NSX Manager IP 地址。

```

[root@esx-01a:~] esxcfg-advcfg -g /UserVars/RmqIpAddress
Value of RmqIpAddress is 192.168.110.15

```

- 在 ESXi 主机上运行 `esxcli network ip connection list | grep 5671` 命令以查找活动消息总线连接。

```

[root@esx-01a:~] esxcli network ip connection list | grep 5671
tcp          0      0 192.168.110.51:29969      192.168.110.15:5671      ESTABLISHED
35505 newreno vsfwd
tcp          0      0 192.168.110.51:29968      192.168.110.15:5671      ESTABLISHED
35505 newreno vsfwd

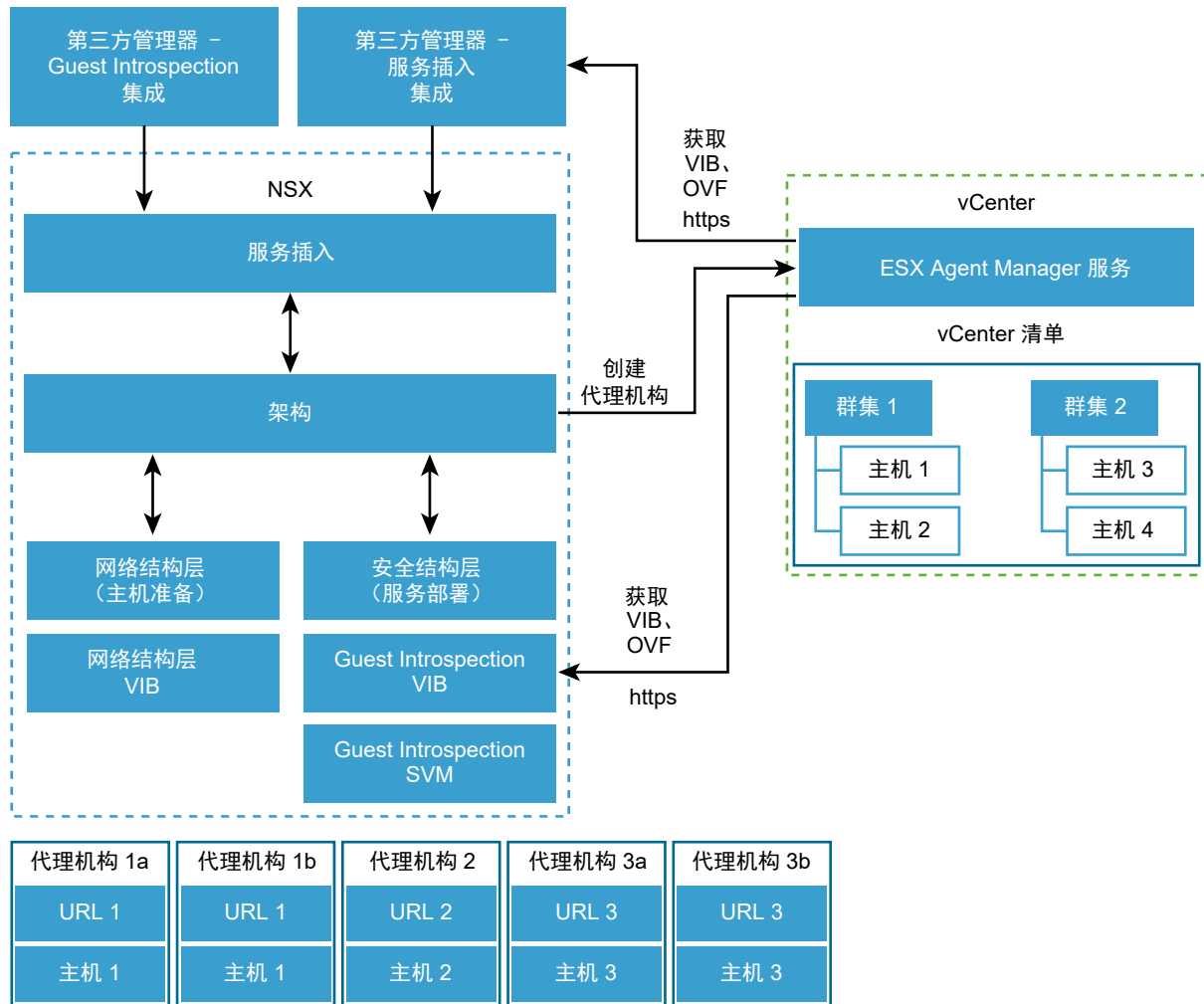
```

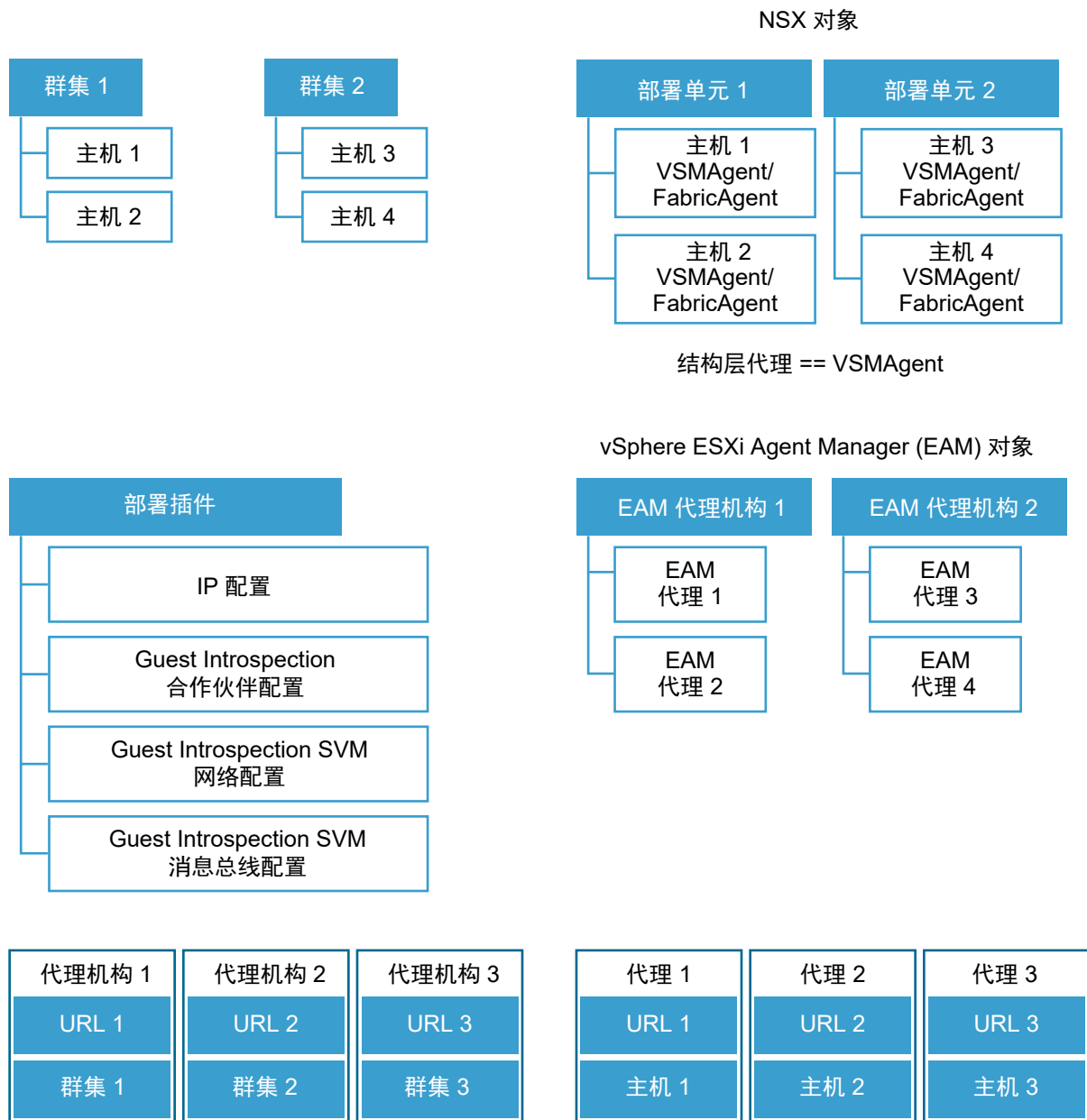
有关与控制层面代理相关的问题，请参阅[控制层面代理 \(netcpa\) 问题](#)。

了解主机准备架构

本主题介绍了基本主机准备架构。

- 要部署网络结构层，请转到[主机准备 \(Host Preparation\)](#)选项卡。
- 要部署安全结构层，请转到[服务部署 \(Service Deployment\)](#)选项卡。





以下术语可以帮助您了解主机准备架构：

结构层	结构层是 NSX Manager 中的软件层，它与 ESX Agent Manager 交互以在主机上安装网络和安全结构层服务。
网络结构层	网络结构层服务部署在群集上。网络结构层服务包括主机准备、VXLAN、分布式路由、分布式防火墙和消息总线。
安全结构层	安全结构层服务部署在群集上。安全结构层服务包括 Guest Introspection 和合作伙伴安全解决方案。
结构层代理	<p>结构层代理是 NSX Manager 数据库中的结构层服务和主机组合。将为部署网络或安全结构层服务的每个群集主机创建一个结构层代理。</p> <p>别名：VSM 代理</p>
部署单元	NSX Manager 数据库中的结构层服务和群集组合。必须创建部署单元以安装网络和安全服务。
ESX Agent Manager 代理	ESX Agent Manager 代理是 vCenter Server 数据库中的服务规范和主机组合。 ESX Agent Manager 代理映射到 NSX 结构层代理。
ESX Agent Manager 代理机构	<p>ESX Agent Manager 代理机构是 vCenter Server 数据库中的规范和群集组合。规范描述机构所管理的 ESX Agent Manager 代理、VIB、OVF 及其配置（如数据存储和网络设置）。</p> <p>NSX Manager 为每个准备的群集创建一个 ESX Agent Manager 代理机构。</p> <p>ESX Agent Manager 代理机构映射到 NSX 部署单元。部署单元的 NSX Manager 数据库和 ESX Agent Manager 代理机构的 vCenter ESX Agent Manager 数据库必须保持同步。在极少数情况下，如果两个数据库不同步，则 NSX 触发事件和警报以通知您这种情况。NSX Manager 在其数据库中为每个 ESX Agent Manager 代理机构创建一个部署单元。</p>

NSX Manager 为每个准备的群集创建一个 **ESX Agent Manager** 代理机构。**NSX Manager** 在其数据库中为每个 **ESX Agent Manager** 代理机构创建一个部署单元。一个 **ESX Agent Manager** 代理机构 = 一个部署单元。

您可以通过以下方法查看代理机构：

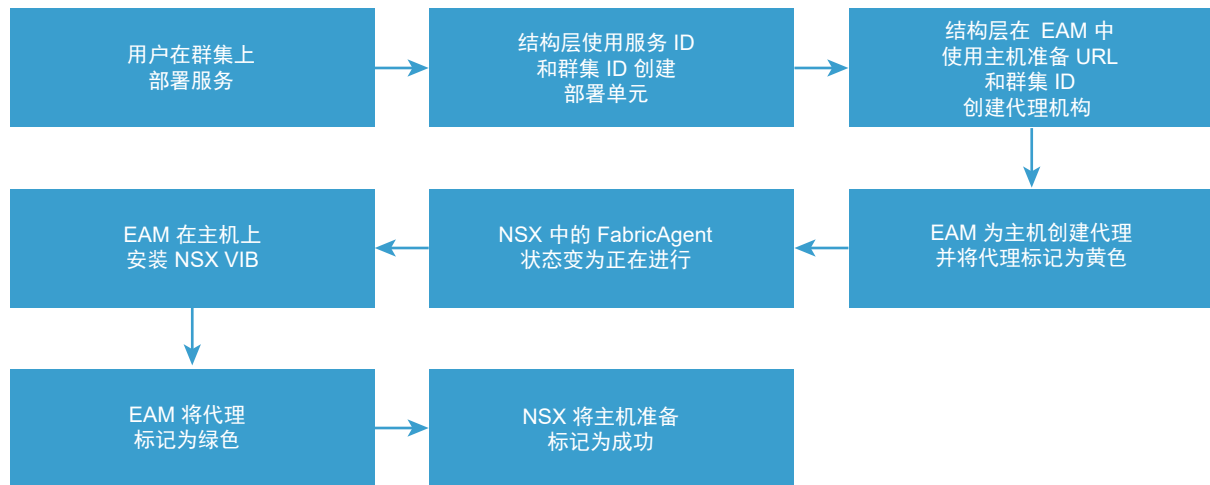
- 从 EAM MOB <https://<VC-hostname/IP>/eam/mob/> 中。
- 从 vSphere Web Client 中：
 - 转到 **vCenter Solutions Manager > vSphere ESX Agent Manager > 管理 (Manage)**。
 - 在 **ESX 代理机构 (ESX Agencies)** 下面，您可以看到这些代理机构（为主机准备的每个群集一个）。

部署单元的生命周期与代理机构的生命周期相关联：如果从 **ESX Agent Manager** 中移除一个代理机构，则会从 **NSX** 中移除相应的部署单元。

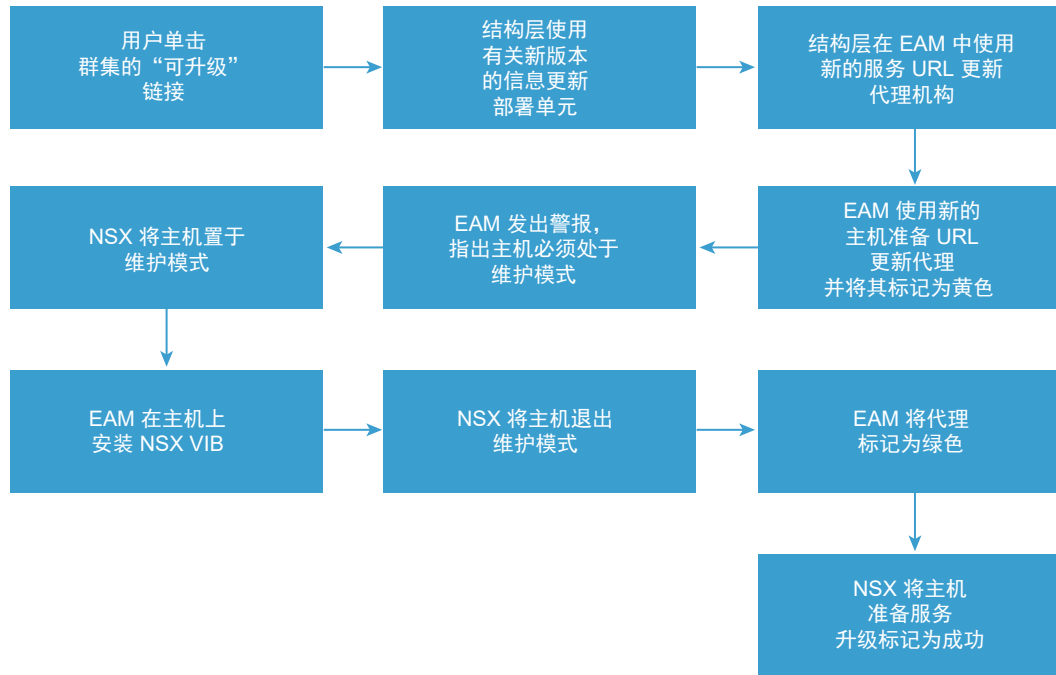
主机准备的服务部署 workflow

本主题显示了主机准备的服务部署 workflow（安装和升级）。

安装 workflow



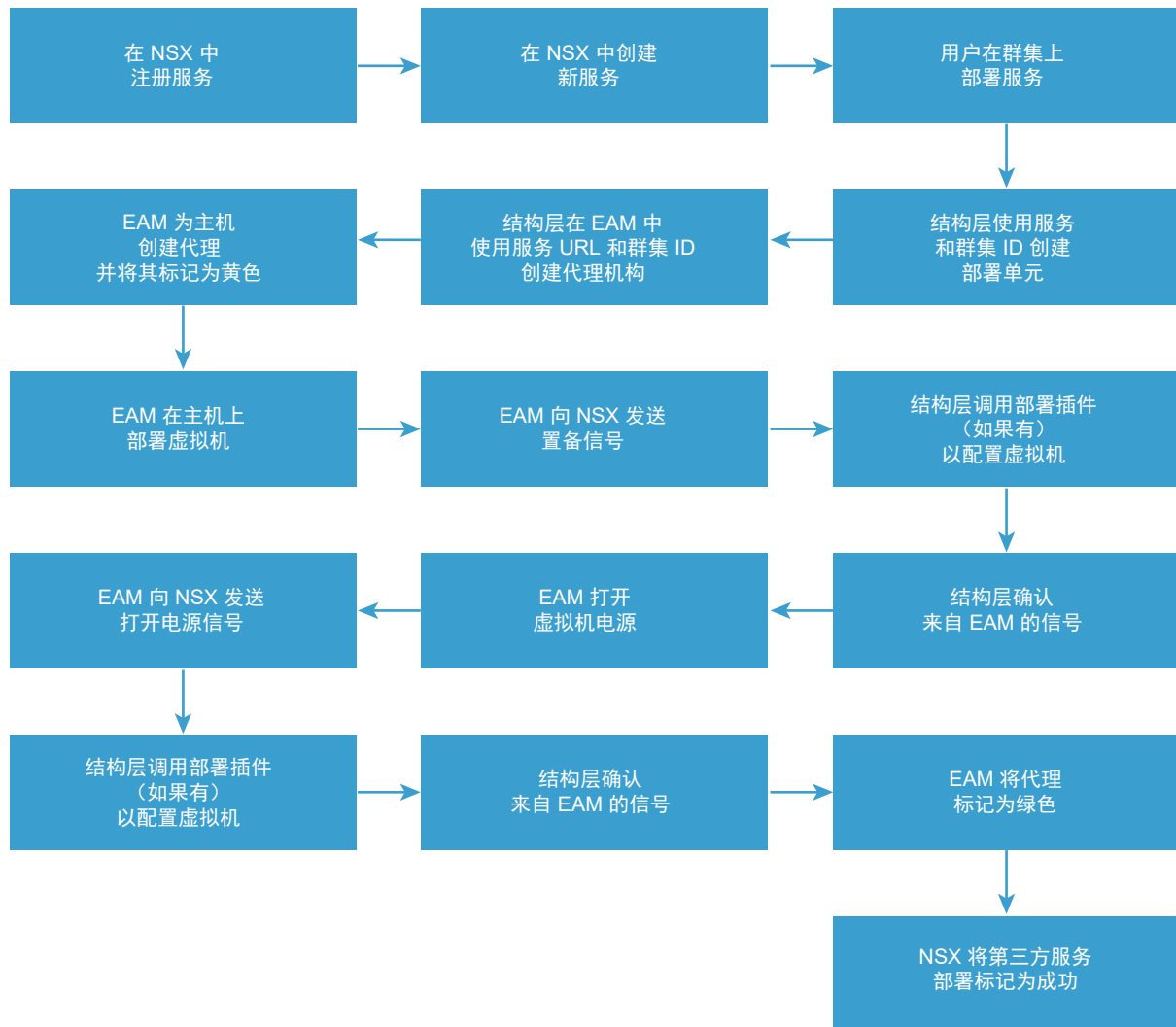
升级 workflow



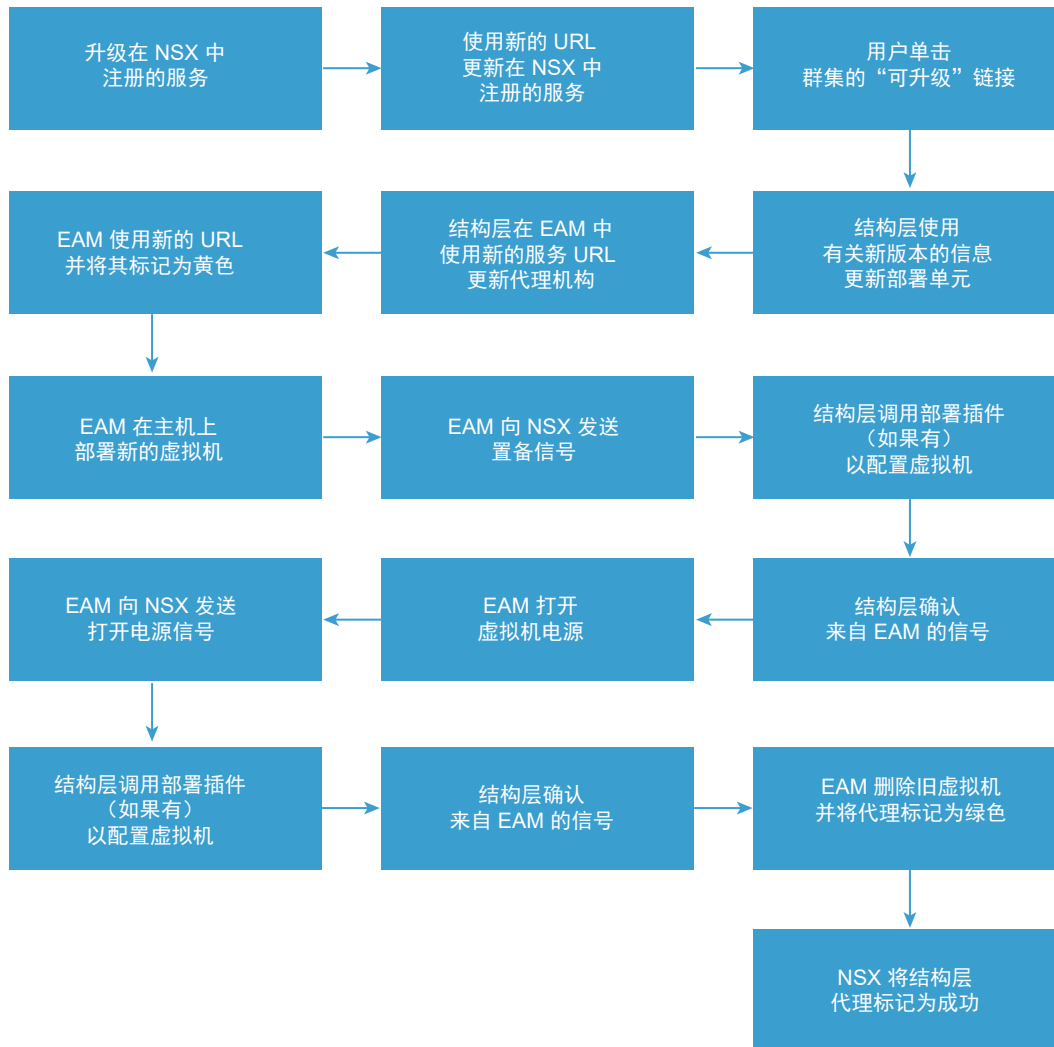
第三方服务的服务部署 workflow

本主题显示了第三方服务的服务部署 workflow（安装和升级）。

安装工作流



升级 workflow



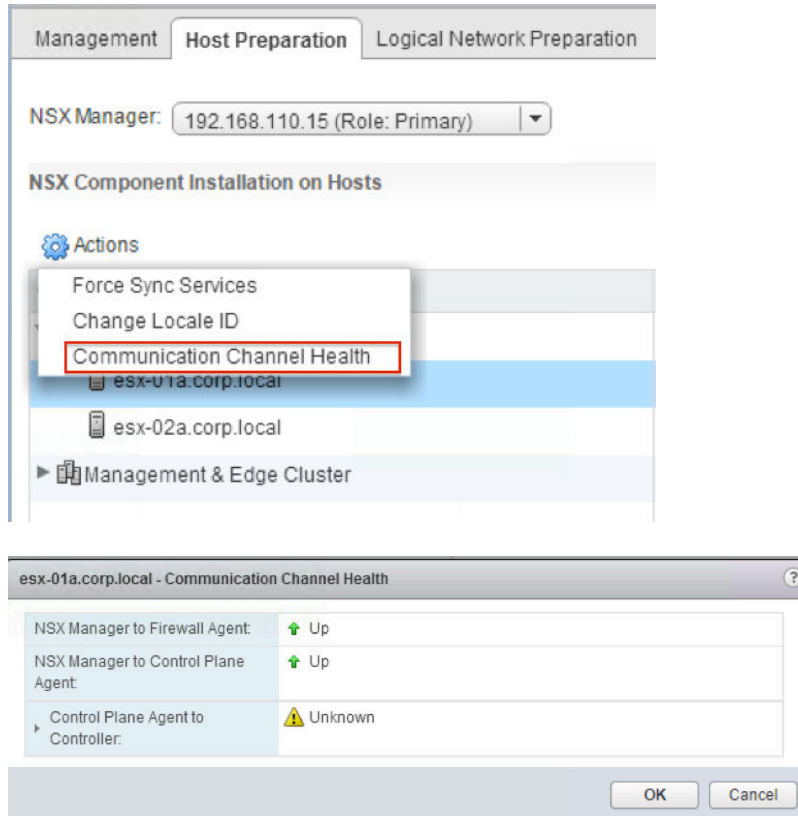
检查通信通道运行状况

从 vSphere Web Client 中，您可以检查各种组件之间的通信状态。

要检查 NSX Manager 和防火墙代理之间、NSX Manager 和控制层面代理之间以及控制层面代理和控制器之间的通信通道运行状况，请执行以下步骤：

- 1 在 vSphere Web Client 中，导航到**网络和安全 (Networking & Security) > 安装 (Installation) > 主机准备 (Host Preparation)**。
- 2 选择或展开一个群集，然后选择一个主机。单击**操作 (Actions)** (⚙️)，然后单击**通信通道运行状况 (Communication Channel Health)**。

随即显示通信通道运行状况信息。



如果主机的三个连接之一的状态发生变化，则会在 **NSX Manager** 日志中写入一条消息。在日志消息中，连接状态可能是“已连接”、“已关闭”或“不可用”（在 **vSphere Web Client** 中显示为“未知”）。如果状态由“已连接”变为“已关闭”或“不可用”，则会生成一条警告消息。例如：

```
2016-05-23 23:36:34.736 GMT+00:00 WARN TaskFrameworkExecutor-25 VdnInventoryFacadeImpl
$HostStatusChangedEventHandler:200 - Host Connection Status Changed: Event Code: 1941, Host:
esx-04a.corp.local (ID: host-46), NSX Manager - Firewall Agent: UP, NSX Manager - Control Plane
Agent: UP, Control Plane Agent - Controllers: DOWN.
```

如果状态由“已关闭”或“不可用”变为“已连接”，则会生成一条类似于警告消息的信息消息。例如：

```
2016-05-23 23:55:12.736 GMT+00:00 INFO TaskFrameworkExecutor-25 VdnInventoryFacadeImpl
$HostStatusChangedEventHandler:200 - Host Connection Status Changed: Event Code: 1938, Host:
esx-04a.corp.local (ID: host-46), NSX Manager - Firewall Agent: UP, NSX Manager - Control Plane
Agent: UP, Control Plane Agent - Controllers: UP.
```

如果控制层面通道发生通信故障，则会生成具有以下精细故障原因之一的系统事件：

- 1255601: 主机证书不完整
- 1255602: 控制器证书不完整
- 1255603: SSL 握手失败
- 1255604: 已拒绝连接
- 1255605: 保持活动状态超时

- 1255606: SSL 异常
- 1255607: 消息错误
- 1255620: 未知错误

此外，还会生成从 NSX Manager 到主机的检测信号消息。如果 NSX Manager 和 netcpa 之间的检测信号丢失，则会触发配置完全同步。

有关如何下载日志的详细信息，请参阅 NSX 管理指南。

安装状态为“未就绪”

在主机准备期间，您可能会注意到群集状态显示为未就绪。

问题

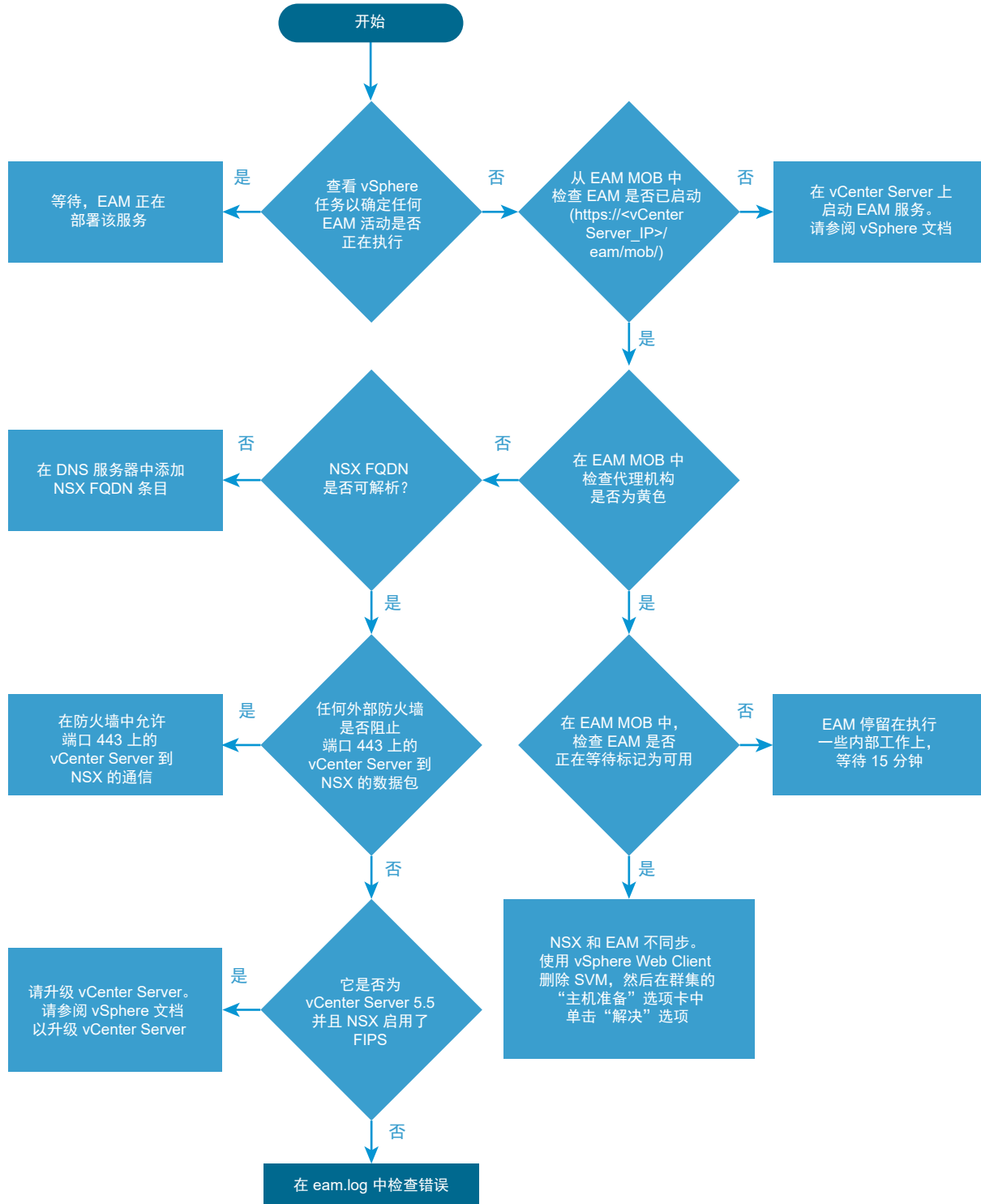
在主机准备 (Host Preparation)或服务部署 (Service Deployment)选项卡中，安装状态显示为未就绪。

解决方案

- 1 转到**网络和安全 (Networking & Security) > 安装 (Installation) > 主机准备 (Host Preparation)或服务部署 (Service Deployment)**选项卡。
- 2 在群集和主机上，单击未就绪。
将显示错误消息。
- 3 单击**解决 (Resolve)**选项。
要查看**解决 (Resolve)**选项解决的问题列表，请参阅 NSX 日志记录和系统事件。
- 4 如果仍显示未就绪，说明还没有解决错误，请参阅[无法使用“解决”选项修复问题](#)。

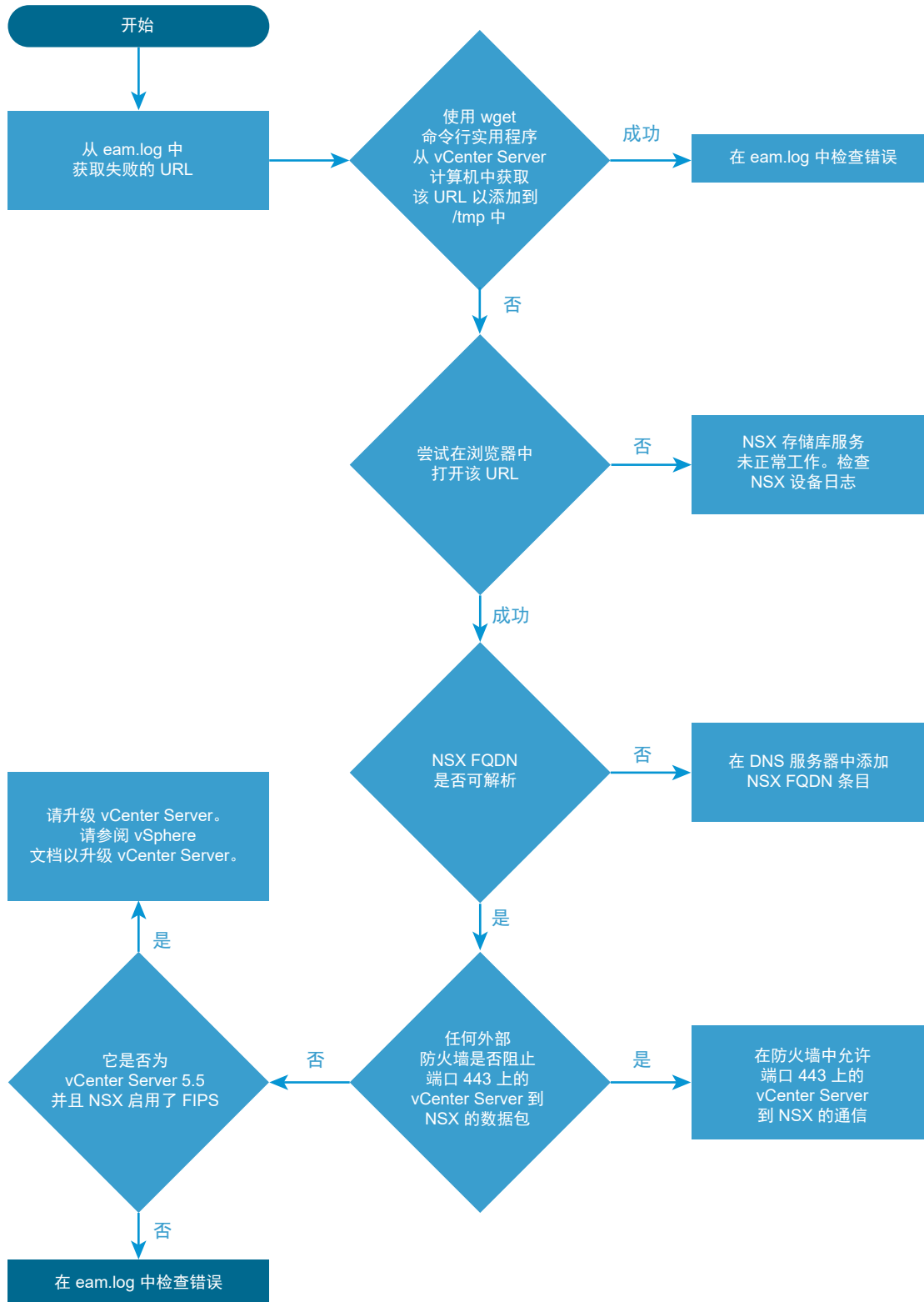
服务没有响应

该流程图简要说明了 NSX 主机准备过程，以及在服务长时间没有响应或长时间显示旋转图标时执行的操作。



服务部署由于“无法访问 OVF/VIB” 错误而失败

该流程图显示在服务部署由于无法访问 OVF/VIB (OVF/VIB not accessible) 错误而失败时执行的操作。



无法使用“解决”选项修复问题

在**网络和安全 (Networking & Security) > 安装 (Installation) > 主机准备 (Host Preparation)**或**服务部署 (Service Deployment)**选项卡中，群集和主机上的安装状态显示为未就绪。单击**解决 (Resolve)**选项无法修复该问题。

问题

- 单击未就绪链接将显示“未在主机上安装代理的 VIB 模块” (VIB module for agent is not installed on the host) 错误。
- ESXi 主机无法从 vCenter Server 中访问 VIB。
- 从 vShield Endpoint 更改为 NSX Manager 时，您可能会看到状态为失败。

解决方案

- 1 确认在 vCenter Server、ESXi 主机和 NSX Manager 上正确配置了 DNS。确保从 vCenter Server、ESXi 主机、NSX Manager 和 vSphere Update Manager 中进行的正向和反向 DNS 解析正常工作。
- 2 要确定问题是否与 DNS 相关，请查看 *esxupdate* 日志，并在 *esxupdate.log* 文件中查找“esxupdate: ERROR: MetadataDownloadError:IOError: <urlopen error [Errno -2] Name= or service not known”消息。
该消息表示 ESXi 主机无法访问 vCenter Server 的完全限定域名 (Fully Qualified Domain Name, FQDN)。有关详细信息，请参见[验证 VMware vCenter Server 受管 IP 地址 \(1008030\)](#)。
- 3 确认正确配置了网络时间协议 (Network Time Protocol, NTP)。VMware 建议配置 NTP。要确定 NTP 不同步问题是否影响您的环境，请检查 NSX Manager 6.2.4 和更高版本的支持包中的 */etc/ntp.drift* 文件。
- 4 确认防火墙未阻止 NSX for vSphere 6.x 所需的所有端口。有关相关的信息，请参阅：
 - [VMware NSX for vSphere 的网络端口要求 \(2079386\)](#)。
 - 访问 [VMware vCenter Server](#)、[VMware ESXi](#) 和 [ESX 主机](#) 以及其他网络组件所需的 TCP 和 UDP 端口 ([1012382](#))。

注 VMware vSphere 6.x 支持通过端口 443（而不是端口 80）下载 VIB。该端口是动态打开和关闭的。ESXi 主机和 vCenter Server 之间的中间设备必须允许使用该端口传输流量。

- 5 确认正确配置了 vCenter Server 受管 IP 地址。有关详细信息，请参见[验证 VMware vCenter Server 受管 IP 地址 \(1008030\)](#)。
- 6 确认 vSphere Update Manager 正常工作。从 vCenter Server 6.0U3 开始，NSX 安装和升级过程不再将 vSphere Update Manager 与 ESX Agent Manager 一起使用。VMware 强烈建议至少运行 vCenter Server 6.0U3 或更高版本。如果无法升级，请确保 vSphere Update Manager 服务正在运行。您可以按照[知识库文章 2053782](#) 配置 vSphere Update Manager 绕过选项。
- 7 如果在部署 vCenter Server 时指定非默认端口，请确保 ESXi 主机防火墙未阻止这些端口。
- 8 确认 vCenter Server *vpzd* 进程正在侦听 TCP 端口 8089。NSX Manager 仅支持默认端口 8089。

关于 vSphere ESX Agent Manager (EAM)

vSphere ESX Agent Manager 自动完成部署和管理 NSX 网络和安全服务的过程，同时扩展 ESXi 主机功能以提供 vSphere 解决方案所需的额外服务。

ESX Agent Manager 的日志和服务

ESX Agent Manager 日志包含在 vCenter 日志包中。

- Windows - C:\ProgramData\VMware\vCenterServer\logs\eam\eam.log
- VCSA - /var/log/vmware/vpx/eam.log
- ESXi - /var/log/esxupdate.log

监控 ESX Agent Manager

重要事项 确保在开始安装 NSX 之前将 *bypassVumEnabled* 标记更改为 **True**，并在安装后将其改回到 **False**。请参见 <https://kb.vmware.com/kb/2053782>。

要检查 ESX Agent Manager 状态，请执行以下操作：

- 1 转到 vSphere Web Client。
- 2 单击**管理 > vCenter Server 扩展 (Administration > vCenter Server Extensions)**，然后单击 vSphere ESX Agent Manager。
 - a 单击**管理 (Manage)**选项卡。

管理 (Manage)选项卡显示有关运行的代理的信息，列出任何孤立的 ESX 代理，并记录有关 ESX Agent Manager 管理的 ESX 代理的信息。

有关代理的详细信息，请参见 vSphere 文档。
 - b 单击**监控 (Monitor)**选项卡。

监控 (Monitor) > 事件 (Events)选项卡显示有关与 ESX Agent Manager 关联的事件的信息。

解决 NSX Manager 问题

验证每个故障排除步骤是否适用于您的环境。每个步骤提供了相应说明，以消除可能的根源并在必要时采取纠正措施。这些步骤按最适当的顺序进行排列，以查找问题并确定相应的解决方案。不要跳过某个步骤。

问题

- 安装 VMware NSX Manager 失败。
- 升级 VMware NSX Manager 失败。
- 登录到 VMware NSX Manager 失败。
- 访问 VMware NSX Manager 失败。

解决方案

1 请参阅当前版本的《NSX 发行说明》以查看是否在错误修复中解决了该问题。

2 确保在安装 VMware NSX Manager 时满足最低系统要求。

请参见 NSX 安装指南。

3 验证是否在 NSX Manager 中打开所需的所有端口。

请参见 NSX 安装指南。

4 安装问题：

- 如果配置 Lookup Service 或 vCenter Server 失败，请验证 NSX Manager 和 Lookup Service 设备上的时间是否同步。请在 NSX Manager 和 Lookup Service 上使用相同的 NTP 服务器配置。还要确保正确配置了 DNS。
- 验证是否正确安装了 OVA 文件。如果无法安装 NSX OVA 文件，vSphere Client 中的错误窗口将指出发生故障的位置。还要验证下载的 OVA/OVF 文件的 MD5 校验和。
- 验证 ESXi 主机上的时间是否与 NSX Manager 同步。
- VMware 建议您在安装 NSX Manager 后立即计划 NSX Manager 数据备份。

5 升级问题：

- 在升级之前，请参见“产品互操作性列表”页中的最新互操作性信息。
- VMware 建议在升级之前备份当前配置并下载技术支持日志。
- 在升级 NSX Manager 后，可能需要强制与 vCenter Server 重新进行同步。为此，请登录到 NSX Manager Web 界面 GUI。接下来，转到**管理 vCenter 注册 > NSX 管理服务 > 编辑 (Manage vCenter Registration > NSX Management Service > Edit)**，然后重新输入管理用户密码。

6 性能问题：

- 确保满足最低 vCPU 要求。
- 验证根 (/) 分区是否具有足够的空间。您可以登录到 ESXi 主机并键入 `df -h` 命令以验证这种情况。

例如：

```
[root@esx-01a:~] df -h
Filesystem      Size  Used Available Use% Mounted on
NFS             111.4G  80.8G   30.5G   73% /vmfs/volumes/ds-site-a-nfs01
vfat            249.7M 172.2M   77.5M   69% /vmfs/volumes/68cb5875-d887b9c6-a805-65901f83f3d4
vfat            249.7M 167.7M   82.0M   67% /vmfs/volumes/fe84b77a-b2a8860f-38cf-168d5dfe66a5
vfat            285.8M 206.3M   79.6M   72% /vmfs/volumes/54de790f-05f8a633-2ad8-00505603302a
```

- 使用 `esxtop` 命令检查哪些进程使用大量 CPU 和内存。
- 如果 NSX Manager 在日志中遇到任何内存不足错误，请验证 `/common/dumps/java.hprof` 文件是否存在。如果该文件存在，请创建该文件的副本，并在 NSX 技术支持日志包中包含该副本。

- 验证在环境中是否存在存储延迟问题。
- 尝试将 NSX Manager 迁移到另一个 ESXi 主机。

7 连接问题:

- 如果 NSX Manager 与 vCenter Server 或 ESXi 主机之间出现连接问题, 请登录到 NSX Manager CLI 控制台, 然后运行 `debug connection IP_of_ESXi_or_VC` 命令并检查输出。
- 确认已启动 Virtual Center Web 管理服务, 并且浏览器未处于错误状态。
- 如果未更新 NSX Manager Web 用户界面 (UI), 您可以尝试禁用并重新启用 Web 服务以解决该问题。请参见 <https://kb.vmware.com/kb/2126701>。
- 在 ESXi 主机上使用 `esxtop` 命令验证 NSX Manager 使用的端口组和上行链路网卡。有关详细信息, 请参见 <https://kb.vmware.com/kb/1003893>。
- 尝试将 NSX Manager 迁移到另一个 ESXi 主机。
- 从 vSphere Web Client 的**监控 (Monitor)**选项卡中检查 NSX Manager 虚拟机设备**任务和事件 (Tasks and Events)**选项卡。
- 如果 NSX Manager 与 vCenter Server 之间出现连接问题, 请尝试将 NSX Manager 迁移到运行 vCenter Server 虚拟机的相同 ESXi 主机以消除可能的底层物理网络问题。

请注意, 这仅适用于两个虚拟机位于同一 VLAN/端口组的情况。

将 NSX Manager 连接到 vCenter Server

通过使用 NSX Manager 和 vCenter Server 之间的连接, NSX Manager 可以使用 vSphere API 执行一些功能, 例如, 部署服务虚拟机, 准备主机以及创建逻辑交换机端口组。连接过程在 Web Client 服务器上为 NSX 安装 Web Client 插件。

要使连接正常工作, 您必须在 NSX Manager、vCenter Server 和 ESXi 主机上配置 DNS 和 NTP。如果按名称将 ESXi 主机添加到 vSphere 清单中, 请确保已在 NSX Manager 上配置 DNS 服务器并且名称解析正常工作。否则, NSX Manager 无法解析 IP 地址。必须指定 NTP 服务器, 以便 SSO 服务器时间和 NSX Manager 时间保持同步。在 NSX Manager 上, `/etc/ntp.drift` 中的偏移文件包含在 NSX Manager 的技术支持包中。

用于将 NSX Manager 连接到 vCenter Server 的帐户必须具有 vCenter “管理员” 角色。具有 “管理员” 角色允许 NSX Manager 在 Security Token Service 服务器中注册其自身。在使用特定用户帐户将 NSX Manager 连接到 vCenter 时, 还会在 NSX Manager 上为该用户创建一个 “企业管理员” 角色。

与将 NSX Manager 连接到 vCenter Server 有关的常见问题

- 未在 NSX Manager、vCenter Server 或 ESXi 主机上正确配置 DNS。
- 未在 NSX Manager、vCenter Server 或 ESXi 主机上正确配置 NTP。
- 使用没有 vCenter “管理员” 角色的用户帐户将 NSX Manager 连接到 vCenter。
- 在 NSX Manager 和 vCenter Server 之间出现网络连接问题。
- 用户使用在 NSX Manager 上没有角色的帐户登录到 vCenter。

您需要先通过用于将 NSX Manager 链接到 vCenter Server 的帐户登录到 vCenter。然后，您可以使用主页 > 网络和安全 > NSX Manager > {NSX Manager IP} > 管理 > 用户 (Home > Networking & Security > NSX Managers > {IP of NSX Manager} > Manage > Users) 在 NSX Manager 上创建其他具有角色的用户。

首次登录可能需要最多 4 分钟的时间，在此期间，vCenter 将加载并部署 NSX UI 包。

验证从 NSX Manager 到 vCenter Server 的连接

- 登录到 NSX Manager CLI 控制台。
- 要验证连接，请查看 ARP 和路由表。

```
nsxmgr# show arp
IP address      HW type  Flags   HW address    Mask     Device
192.168.110.31  0x1      0x2     00:50:56:ae:ab:01  *        mgmt
192.168.110.2   0x1      0x2     00:50:56:01:20:a5  *        mgmt
192.168.110.1   0x1      0x2     00:50:56:01:20:a5  *        mgmt
192.168.110.33  0x1      0x2     00:50:56:ae:4f:7c  *        mgmt
192.168.110.32  0x1      0x2     00:50:56:ae:50:bf  *        mgmt
192.168.110.10  0x1      0x2     00:50:56:03:19:4e  *        mgmt
192.168.110.51  0x1      0x2     00:50:56:03:30:2a  *        mgmt
192.168.110.22  0x1      0x2     00:50:56:01:21:f9  *        mgmt
192.168.110.55  0x1      0x2     00:50:56:01:23:21  *        mgmt
192.168.110.26  0x1      0x2     00:50:56:01:21:ef  *        mgmt
192.168.110.54  0x1      0x2     00:50:56:01:22:ef  *        mgmt
192.168.110.52  0x1      0x2     00:50:56:03:30:16  *        mgmt
```

```
nsxmgr# show ip route
Codes: K - kernel route, C - connected, S - static,
       > - selected route, * - FIB route

S>* 0.0.0.0/0 [1/0] via 192.168.110.1, mgmt
C>* 192.168.110.0/24 is directly connected, mgmt
```

- 在 NSX Manager 日志中查找错误，以找出未连接到 vCenter Server 的原因。用于查看日志的命令是 show log manager follow。

```
2014-02-26 12:53:23.815 GMT INFO VcEventsReaderThread DefaultRequestDirector:491 - I/O exception (org.apache.http.NoHttpResponseException: The target server failed to respond)
2014-02-26 12:53:23.815 GMT INFO VcEventsReaderThread DefaultRequestDirector:498 - Retrying request
2014-02-26 12:53:23.815 GMT WARN ViInventoryThread ViInventory:1482 - We received error from VC, probably lost connection
2014-02-26 12:53:23.817 GMT INFO VcEventsReaderThread VcEventsReader$VcEventsReaderThread:347 - Caught exception:com.vmware.vim.client.exception.ConnectionException: Connection to https://vc-1-01a.corp.local refused
2014-02-26 12:53:23.821 GMT DEBUG VcEventsReaderThread VcEventsReader$VcEventsReaderThread:348 - Caught exception during ping:com.vmware.vim.vimoml.client.exception.ConnectionException: org.apache.http.conn.HttpHostConnectException: Connection to https://vc-1-01a.corp.local refused
```

- 运行 debug connection IP_of_ESXi_or_VC 命令，然后检查输出。

在 NSX Manager 上执行数据包捕获以查看连接

使用 debug packet 命令: debug packet [capture|display] interface interface filter

NSX Manager 上的接口名称是 mgmt。

筛选器语法采用以下形式: “port_80_or_port_443”

该命令仅在特权模式下运行。要进入特权模式，请运行 **enable** 命令并提供管理员密码。

数据包捕获示例：

```
nsxmgr# en
nsxmgr# debug packet display interface mgmt port_80_or_port_443
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on mgmt, link-type EN10MB (Ethernet), capture size 262144 bytes
23:40:25.321085 IP 192.168.210.15.54688 > 192.168.210.22.443: Flags [P.], seq 2645022162:2645022199,
ack 2668322748, win 244, options [nop,nop,TS val 1447550948 ecr 365097421], length 37
...
```

在 NSX Manager 上验证网络配置

show running-config 命令显示管理接口、NTP 和默认路由设置的基本配置。

```
nsxmgr# show running-config
Building configuration...

Current configuration:
!
ntp server 192.168.110.1
!
ip name server 192.168.110.10
!
hostname nsxmgr
!
interface mgmt
 ip address 192.168.110.15/24
!
ip route 0.0.0.0/0 192.168.110.1
!
web-manager
```

NSX Manager 证书

NSX Manager 支持使用两种方法生成证书。

- NSX Manager 生成的 CSR：由于基本 CSR 而受到限制的功能
- KCS#12：建议将其用于生产环境

存在一个已知问题：在 CMS 无法执行 API 调用时，不显示任何提示。

在调用方不知道证书颁发者时，将会发生这种情况，因为这是不可信的根证书颁发机构，或者证书是自签名证书。要解决该问题，请使用浏览器导航到 NSX Manager IP 地址或主机名并接受证书。

辅助 NSX Manager 停留在转换模式

如果辅助 NSX Manager 停留在该问题中所述的转换模式，请使用下面所述的解决方案。如果辅助 NSX Manager 处于转换模式，在主 NSX Manager 上还原备份时，将出现该问题。

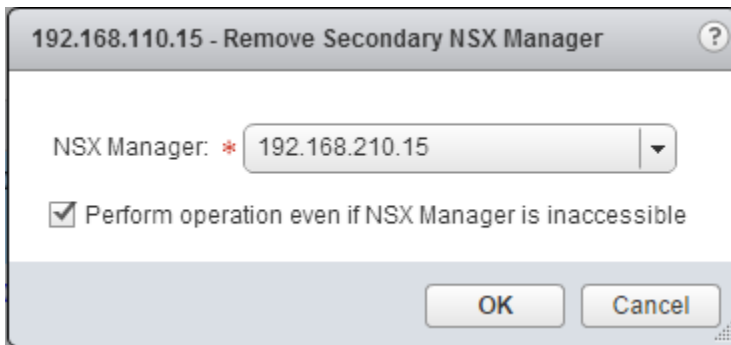
问题

- 1 您配置了主要和辅助 NSX Manager。
- 2 您创建了主 NSX Manager 的备份。
- 3 然后，您移除辅助 NSX Manager。辅助 NSX Manager 将处于转换模式。
- 4 现在，由于某些原因，您在主 NSX Manager 上还原备份。
- 5 在数据库中，转换 NSX Manager 将更新为**辅助 (Secondary)**，但它在 UI 上显示为**转换 (Transit)**，并且同步失败。
- 6 您可能无法移除辅助 NSX Manager 或将其再次升级为辅助管理器。
- 7 在升级转换 NSX Manager 时，将显示一条错误消息，指出具有 IP 地址/主机名的 NSX Manager 节点已存在。
- 8 在移除转换 NSX Manager 时，将显示一条错误消息，指出用户名或密码不正确。

解决方案

- 1 使用 vSphere Web Client 登录到与主 NSX Manager 链接的 vCenter。
- 2 导航到主页 (Home) > 网络和安全 (Networking & Security) > 安装 (Installation)，然后选择**管理 (Management)**选项卡。
- 3 选择要删除的辅助 NSX Manager，单击**操作 (Actions)**，然后单击**移除辅助 NSX Manager (Remove Secondary NSX Manager)**。

将显示确认对话框。



- 4 选中**即使 NSX Manager 无法访问仍执行操作 (Perform operation even if NSX Manager is inaccessible)**复选框。

- 5 单击**确定 (OK)**。

将从主数据库中删除辅助 NSX Manager。

- 6 再次添加辅助 NSX Manager。

后续步骤

有关添加辅助 NSX Manager 的详细信息，请参阅 NSX 安装指南。

配置 NSX SSO Lookup Service 失败

问题

- 将 NSX Manager 注册到 vCenter Server 失败
- 配置 SSO Lookup Service 失败
- 可能会出现以下错误：

```
nested exception is java.net.UnknownHostException: vc.local( vc.corp.local )
```

```
NSX Management Service operation failed.( Initialization of Admin Registration Service
Provider failed. Root Cause: Error occurred while registration of lookup service,
com.vmware.vim.sso.admin.exception.InternalError: General failure.
```

```
com.vmware.vshield.vsm.security.service.impl.SamlTokenSSOAuthenticator : SSO is not
configured or initialized properly so cannot authenticate user.
```

解决方案

1 连接问题：

- 如果 NSX Manager 与 vCenter Server 或 ESXi 主机之间出现连接问题，请登录到 NSX Manager CLI 控制台，然后运行 `debug connection IP_of_ESXi_or_VC` 命令并检查输出。
- 通过以下命令，使用 IP 地址和 FQDN 执行从 NSX Manager 到 vCenter Server 的 ping 操作，以检查 NSX Manager 中的路由（静态路由或默认路由）：

```
nsxmgr-l-01a# show ip route
```

代码：

K - 内核路由，

C - 已连接，

S - 静态

> - 选定路由，

* - FIB 路由

```
S>* 0.0.0.0/0 [1/0] via 192.168.110.2, mgmt
```

```
C>* 192.168.110.0/24 is directly connected, mgmt
```

2 DNS 问题

通过以下命令，使用 FQDN 执行从 NSX Manager 到 vCenter Server 的 ping 操作：

```
nsx-mgr> ping vc-l-01a.corp.local
```

将显示与以下示例类似的输出：


```

nsx-mgr> ping vc-l-01a.corp.local
PING vc-l-01a.corp.local (192.168.110.51): 56 data bytes
64 bytes from 192.168.110.51: icmp_seq=0 ttl=64 time=1.749 ms
64 bytes from 192.168.110.51: icmp_seq=1 ttl=64 time=2.111 ms
64 bytes from 192.168.110.51: icmp_seq=2 ttl=64 time=8.082 ms
64 bytes from 192.168.110.51: icmp_seq=3 ttl=64 time=2.010 ms
64 bytes from 192.168.110.51: icmp_seq=4 ttl=64 time=0.857 ms

```

如果此方法不起作用，请在 NSX Manager 中导航到**管理 > 网络 > DNS 服务器 (Manage > Network > DNS Servers)**，确保已正确配置 DNS。

3 防火墙问题

如果 NSX Manager 与 vCenter Server 之间存在防火墙，请确认防火墙在 TCP/443 上允许 SSL。此外，还请执行 ping 操作以检查连接。

4 确认已在 NSX Manager 中打开以下所需的端口。

表 2-1. NSX Manager 中打开的端口

端口	需要使用端口的情况
443/TCP	将 OVA 文件下载到 ESXI 主机上以进行部署 使用 REST API 使用 NSX Manager 用户界面
80/TCP	启动到 vSphere SDK 的连接 在 NSX Manager 与 NSX 主机模块之间进行消息传送
1234/TCP	在 NSX Controller 与 NSX Manager 之间进行通信
5671	Rabbit MQ（消息总线技术）
22/TCP	对 CLI 的控制台访问 (SSH) 注意：默认情况下，将关闭此端口

5 NTP 问题

确认已在 vCenter Server 与 NSX Manager 之间同步时间。为此，请在 NSX Manager 和 vCenter Server 上使用相同的 NTP 服务器配置。

要确定 NSX Manager 上的时间，请从 CLI 中运行以下命令：

```

nsxmgr-l-01a# show clock

Tue Nov 18 06:51:34 UTC 2014

```

要确定 vCenter Server 上的时间，请从 CLI 中运行以下命令：

```
vc-l-01a:~ # date
```

应当显示与以下内容类似的输出：

```
Tue Nov 18 06:51:31 UTC 2014
```

注意：配置时间设置后，请重新启动设备。

6 用户权限问题

确认用户具有 **admin** 特权。

要注册到 vCenter Server 或 SSO Lookup Service，您必须具有管理权限。

默认帐户是 `administrator user: administrator@vsphere.local`

7 通过输入凭据重新连接到 SSO。

逻辑网络准备：VXLAN 传输

NSX 为 VTEP VMkernel 网卡创建分布式虚拟端口组以准备为 VXLAN 选择的 vSphere Distributed Switch。

VTEP 的绑定策略、负载平衡方法、MTU 和 VLAN ID 是在 VXLAN 配置期间选择的。绑定和负载平衡方法必须与为 VXLAN 选择的 DVS 配置相匹配。

MTU 必须设置为至少 1600，并且不小于在 DVS 上已配置的大小。

创建的 VTEP 数取决于选择的成组策略和 DVS 配置。

VXLAN 准备期间的常见问题

由于以下几个原因，VXLAN 准备可能会失败：

- 为 VXLAN 选择的成组方法与 DVS 支持的方法不匹配。要查看支持的方法，请参见 VMware NSX for vSphere 网络虚拟化设计指南，网址为 <https://communities.vmware.com/docs/DOC-27683>。
- 为 VTEP 选择的 VLAN ID 不正确。
- 选择了 DHCP 以分配 VTEP IP 地址，但没有可用的 DHCP 服务器。
- 缺少 VMkernel 网卡。按照 [VXLAN VMkernel 网卡不同步](#) 中所述，解决该错误。
- VMkernel 网卡具有错误的 IP 地址。按照 <https://kb.vmware.com/kb/2137025> 中所述，解决该错误。
- 为 VTEP 选择的 MTU 设置不正确。您应该按照本主题后面所述调查 MTU 是否不匹配。
- 选择的 VXLAN 网关不正确。您应该按照本主题后面所述调查在配置 VXLAN 网关时是否出错。

重要的端口号

VXLAN UDP 端口用于 UDP 封装。在 NSX 6.2.3 之前，默认 VXLAN 端口号为 8472。在 NSX 6.2.3 中，新安装的默认 VXLAN 端口号更改为 4789。在使用硬件 VTEP 的 NSX 6.2 和更高版本的安装中，您必须使用 VXLAN 端口号 4789。有关更改 VXLAN 端口配置的信息，请参阅《NSX 管理指南》中的“更改 VXLAN 端口”。

在主机没有任何需要控制器连接的活动虚拟机时，控制层面状态显示为 *disabled*

可以使用 `show logical-switch` 命令查看主机上的 VXLAN 详细信息。有关详细信息，请参阅《NSX 命令行界面参考》。

如果未在主机中填充需要连接到控制器群集以转发表信息的任何虚拟机，则 `show logical-switch host hostID verbose` 命令将控制层面状态显示为 *disabled*。

```
Network count: 18
VXLAN network: 32003
Multicast IP: 0.0.0.0
Control plane: Disabled <<=====
MAC entry count: 0
ARP entry count: 0
Port count: 1
```

在配置 VXLAN 网关时出错

如果在网络和安全 > 安装 > 主机准备 > 配置 VXLAN (Networking & Security > Installation > Host Preparation > Configure VXLAN) 中使用静态 IP 池配置 VXLAN，并且配置无法在 VTEP 上设置 IP 池网关，主机群集的 VXLAN 配置将进入“错误 (红色)”状态。错误消息为“无法在主机上设置 VXLAN 网关” (VXLAN Gateway cannot be set on host)，错误状态为“VXLAN_GATEWAY_SETUP_FAILURE”。

在 REST API 调用 `GET https://<nsxmgr-ip>/api/2.0/nwfabric/status?resource=<cluster-moid>` 中，VXLAN 状态如下所示：

```
<nwFabricFeatureStatus>
<featureId>com.vmware.vshield.nsxmgr.vxlan</featureId>
  <featureVersion>5.5</featureVersion>
  <updateAvailable>false</updateAvailable>
  <status>RED</status>
  <message>VXLAN Gateway cannot be set on host</message>
  <installed>true</installed>
  <enabled>true</enabled>
  <errorStatus>VXLAN_GATEWAY_SETUP_FAILURE</errorStatus>
</nwFabricFeatureStatus>
```

解决办法：要解决该错误，可以使用两种方法。

- Option 1: 移除主机群集的 VXLAN 配置，确保正确配置并且可访问底层网关以修复 IP 池中的网关设置，然后为主机群集重新配置 VXLAN。
- Option 2: 执行以下步骤。
 - a 通过确保网关配置正确且可访问来修复 IP 池中的基础网关设置。
 - b 将主机置于维护模式，确保主机上没有任何活动的虚拟机流量。
 - c 从主机中删除 VXLAN VTEP。
 - d 将主机退出维护模式。如果将主机退出维护模式，则会在 NSX Manager 上触发创建 VXLAN VTEP 的过程。NSX Manager 将尝试在主机上重新创建所需的 VTEP。

调查 MTU 不匹配问题

- 运行以下命令以验证是否将 MTU 配置为 1600 或更高：

```
ping ++netstack=vxlan -d -s 1572 -I <vmkx hostname_or_IP>
```

其中 *vmkx* 是 VMkernel 端口的 ID，*hostname_or_IP* 是 VMkernel 端口的 IP 或主机名。

这样，您就可以检查所有上行链路的有效性。如果在多 VTEP 环境中工作，您可以从每个可能的 VTEP VMkernel 源/目标接口中运行 ping 命令以验证所有路径，从而验证所有上行链路。

- 检查物理基础架构。很多时候，可以更改物理基础架构配置以解决问题。
- 确定该问题是仅限于单个逻辑交换机，还是其他逻辑交换机也会受到影响。验证该问题是否影响所有逻辑交换机。

有关 MTU 检查的详细信息，请参见 NSX 升级指南中的“验证 NSX 工作状态”。

VXLAN VMkernel 网卡不同步

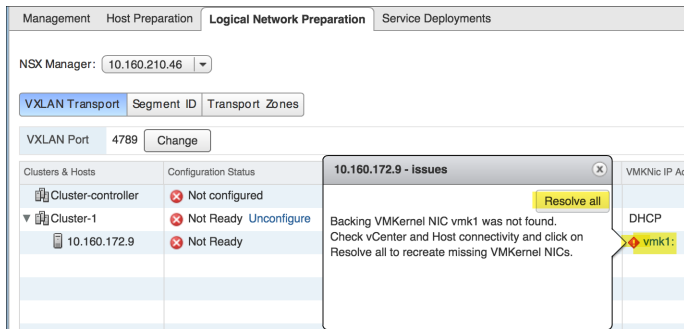
如果在主机上删除 VXLAN VMkernel 网卡，但在 NSX 中仍具有 VMkernel 网卡信息，则 NSX Manager 使用 **错误 (Error)** 图标指示删除的 VMkernel 网卡。

前提条件

在主机上删除了 VMkernel 网卡。

步骤

- 1 在 vSphere Web Client 中，导航到 **网络和安全 (Networking & Security) > 安装 (Installation) > 逻辑网络准备 (Logical Network Preparation)**。
- 2 在 **VXLAN 传输 (VXLAN Transport)** 选项卡上，展开“群集和主机”。



- 3 单击 **错误 (Error)** 图标以查看在主机上删除的 VMkernel 网卡的信息。
- 4 单击 **解决所有 (Resolve All)** 按钮以在主机上重新创建删除的 VMkernel 网卡。

结果

将在主机上重新创建删除的 VMkernel 网卡。

更改 VXLAN 绑定策略和 MTU 设置

可以在准备的主机和群集上更改 VXLAN 绑定策略和 MTU 设置，但仅在为 VXLAN 准备新的主机和群集时，才会应用这些更改。只能再次手动准备主机和群集以更改 VTEP VMkernel 的现有虚拟端口组。您可以使用 API 更改绑定策略和 MTU 设置。

问题

为 VTEP 选择的 MTU 设置不正确。

解决方案

- 1 使用 GET `https://<NSX-Manager-IP-Address>/api/2.0/vdn/switches` API 检索有关 VXLAN 准备的所有交换机的信息。

在 API 输出中，找到要修改的交换机并记下该名称。例如，*dvs-35*。

- 2 现在，使用前面记下的特定 vSphere Distributed Switch 进行查询。

例如，GET `https://<NSX-Manager-IP-Address>/api/2.0/vdn/switches/dvs-35` API。

将显示与以下示例类似的输出：

```
<vdsContext>
<switch>
  <objectId>dvs-35</objectId>
  <objectTypeName>VmwareDistributedVirtualSwitch</objectTypeName>
  <vsmUuid>423A993F-BEE6-1285-58F1-54E48D508D90</vsmUuid>
  <nodeId>916287b3-761d-430b-8ab2-83878dfe3e7f</nodeId>
  <revision>6</revision>
  <type>
    <typeName>VmwareDistributedVirtualSwitch</typeName>
  </type>
  < name>vds-site-a</name>
  <scope>
    <id>datacenter-21</id>
    <objectTypeName>Datacenter</objectTypeName>
    < name>Datacenter Site A</name>
  </scope>
  <clientHandle/>
  <extendedAttributes/>
  <isUniversal>false</isUniversal>
  <universalRevision>0</universalRevision>
</switch>
<mtu>1600</mtu>
<teaming>FAILOVER_ORDER</teaming>
<uplinkPortName>Uplink 4</uplinkPortName>
<promiscuousMode>false</promiscuousMode>
</vdsContext>
```

- 3 您可以使用 API 调用在 vSphere Distributed Switch 上修改参数，例如，绑定策略和/或 MTU。以下示例显示将 *dvs 35* 的绑定策略从 *FAILOVER_ORDER* 更改为 *LOADBALANCE_SRCMAC*，并将 MTU 从 *1600* 更改为 *9000*。

- 对于 NSX: PUT `https://<NSX-Manager-IP-Address>/api/2.0/vdn/switches`

将显示与以下示例类似的输出：

```
<vdsContext>
<switch>
  <objectId>dvs-35</objectId>
  <objectTypeName>VmwareDistributedVirtualSwitch</objectTypeName>
  <vsmUuid>423A993F-BEE6-1285-58F1-54E48D508D90</vsmUuid>
  <nodeId>916287b3-761d-430b-8ab2-83878dfe3e7f</nodeId>
  <revision>6</revision>
  <type>
    <typeName>VmwareDistributedVirtualSwitch</typeName>
  </type>
  <name>vds-site-a</name>
  <scope>
    <id>datacenter-21</id>
    <objectTypeName>Datacenter</objectTypeName>
    <name>Datacenter Site A</name>
  </scope>
  <clientHandle/>
  <extendedAttributes/>
  <isUniversal>false</isUniversal>
  <universalRevision>0</universalRevision>
</switch>
  <mtu>9000</mtu>
  <teaming>LOADBALANCE_SRCMAC</teaming>
  <uplinkPortName>Uplink 4</uplinkPortName>
  <promiscuousMode>false</promiscuousMode>
</vdsContext>
```

注 下面是 *<teaming>* 参数的有效绑定策略条目列表：

- FAILOVER_ORDER
- ETHER_CHANNEL
- LACP_ACTIVE
- LACP_PASSIVE
- LOADBALANCE_LOADBASED
- LOADBALANCE_SRCID
- LOADBALANCE_SRCMAC LACP_V2

- 4 使用 GET 命令验证使用的语法是否正确，以及是否将更改应用于使用的 vSphere Distributed Switch。例如，GET `https://<NSX-Manager-IP-Address>/api/2.0/vdn/switches/dvs-35`。
- 5 打开 vSphere Web Client 并确认反映了配置更改。

逻辑交换机端口组不同步

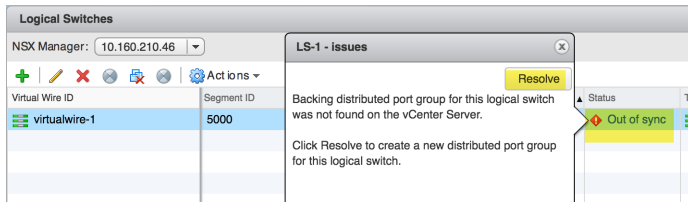
如果在 vCenter Server 上删除了逻辑交换机的备用分布式虚拟端口组 (Distributed Virtual Port Group, DVPG)，逻辑交换机 (Logical Switches) 页面的“状态”列将显示不同步 (Out of sync) 状态。

前提条件

在 vCenter Server 上删除了逻辑交换机的 DVPG。

步骤

- 1 在 vSphere Web Client 中，导航到主页 (Home) > 网络和安全 (Networking & Security) > 逻辑交换机 (Logical Switches)。



- 2 在“状态”列中，单击不同步 (Out of sync) 链接以查看这种不同步状态的详细原因。
- 3 单击解决 (Resolve) 按钮以解决该问题。

结果

这会调用 API 以重新创建备用 DVPG。

NSX 路由故障排除

3

NSX 具有两种类型的路由子系统，它们针对两种主要需求进行了优化。

NSX 路由子系统是：

- 逻辑空间中的路由；也称为“东-西”路由，这是由分布式逻辑路由器 (DLR) 提供的；
- 物理和逻辑空间之间的路由；也称为“北-南”路由，这是由 Edge 服务网关 (ESG) 提供的。

它们都提供了水平扩展选项。

您可以通过 DLR 横向扩展分布式东-西路由。

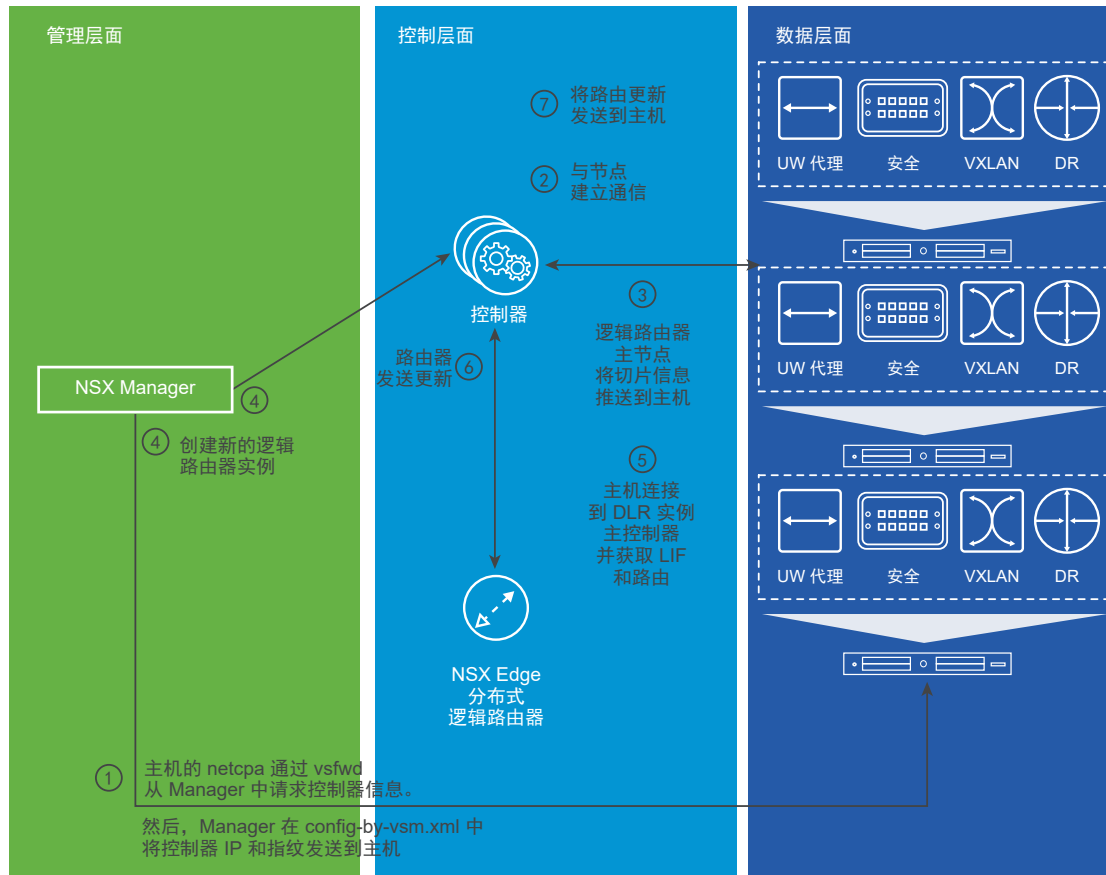
DLR 支持每次运行单个动态路由协议（OSPF 或 BGP）；而 ESG 支持同时运行两个路由协议。这样做的原因是，DLR 设计为“末端”路由器并具有单个输出路径，这意味着通常不需要使用更高级的路由配置。

DLR 和 ESG 均支持静态路由和动态路由组合。

DLR 和 ESG 均支持 ECMP 路由。

它们提供了 L3 域隔离，这意味着，每个分布式逻辑路由器或 Edge 服务网关实例具有自己的 L3 配置，类似于 L3VPN VRF。

图 3-1. 创建 DLR



本章讨论了以下主题：

- 了解分布式逻辑路由器
- 了解 **Edge** 服务网关提供的路由
- **ECMP** 数据包流
- **NSX** 路由：必备条件和注意事项
- **DLR** 和 **ESG** UI
- 新的 **NSX Edge (DLR)**
- 典型的 **ESG** 和 **DLR** UI 操作
- **NSX** 路由故障排除

了解分布式逻辑路由器

DLR 经过优化以在支持 **VXLAN** 或 **VLAN** 的端口组上的虚拟机之间的逻辑空间中转发。

DLR 具有以下属性：

- 高性能、低开销的第一跃点路由：

- 随主机数呈线性扩展
- 在上行链路上支持 8 向 ECMP
- 每个主机最多 1,000 个 DLR 实例
- 每个 DLR 上最多 999 个逻辑接口 (LIF) (8 个上行链路 + 991 个内部) + 1 个管理
- 在所有 DLR 实例中分布的每个主机最多 10,000 个 LIF (NSX Manager 未强制实施)

请注意以下问题：

- 无法将多个 DIR 连接到任何给定的 VLAN 或 VXLAN。
- 无法在每个 DLR 上运行多个路由协议。
- 如果使用 OSPF，则无法在多个 DLR 上行链路上运行 OSPF。
- 要在 VXLAN 和 VLAN 之间路由，传输区域必须跨单个 DVS。

概括来说，DLR 设计在以下方面与模块化路由器机箱类似：

- ESXi 主机类似于线路卡：
 - 它们具有连接了终端站（虚拟机）的端口。
 - 这是进行转发决策的位置。
- DLR 控制虚拟机类似于路由处理器引擎：
 - 它运行动态路由协议，以便与网络的其余部分交换路由信息。
 - 它根据接口配置、静态路由和动态路由信息计算“线路卡”的转发表。
 - 它将这些转发表写入到“线路卡”（通过控制器群集）以支持扩展和弹性。
- 将 ESXi 主机连接在一起的物理网络类似于背板：
 - 它在“线路卡”之间传送 VLAN 或 VXLAN 封装的数据。

简要 DLR 数据包流

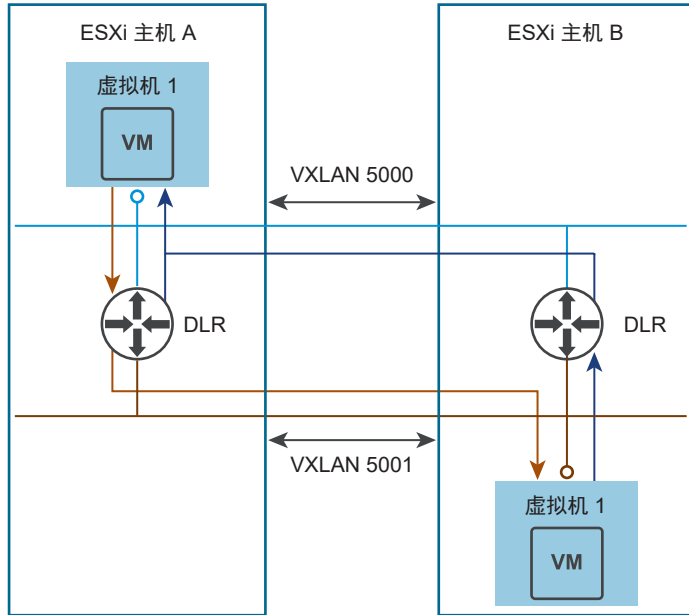
对于每个配置的 DLR 实例，每个 ESXi 主机具有自己的副本。每个 DLR 实例具有自己的一组独特的表，其中包含转发数据包所需的信息。将在该 DLR 实例所在的所有主机之间同步该信息。不同主机上的单个 DLR 实例具有完全相同的信息。

路由始终是由运行源虚拟机的相同主机上的 DLR 实例处理的。这意味着，在源和目标虚拟机位于不同的主机上时，在它们之间提供路由的 DLR 实例仅在一个方向（从源虚拟机到目标虚拟机）上看到数据包。仅目标虚拟机的主机上的相同 DLR 的相应实例看到返回流量。

如果源和目标虚拟机位于不同的主机上，在 DLR 完成路由后，DVS 负责通过 L2（VXLAN 或 VLAN）传送到最终目标；如果源和目标虚拟机位于相同的主机上，则 DVS 在本地进行传送。

图 3-2. 简要 DLR 数据包流 说明了在不同主机上运行并连接到两个不同逻辑交换机（VXLAN 5000 和 VXLAN 5001）的两个虚拟机（虚拟机 1 和虚拟机 2）之间的数据流。

图 3-2. 简要 DLR 数据包流



数据包流（跳过 ARP 解析）：

- 1 虚拟机 1 向虚拟机 2 发送一个数据包，它将发送到虚拟机 2 子网的虚拟机 1 网关（或默认位置）。该网关是 DLR 上的 VXLAN 5000 LIF。
- 2 ESXi 主机 A 上的 DVS 将数据包传送到该主机上的 DLR，将在其中执行查找并确定输出 LIF（此处为 VXLAN 5001 LIF）。
- 3 然后，从该目标 LIF 中发出数据包，这实际上将数据包返回到 DVS，但位于不同的逻辑交换机 (5001) 上。
- 4 接下来，DVS 执行 L2 传送以将该数据包传送到目标主机（ESXi 主机 B），DVS 在其中将数据包转发到虚拟机 2。

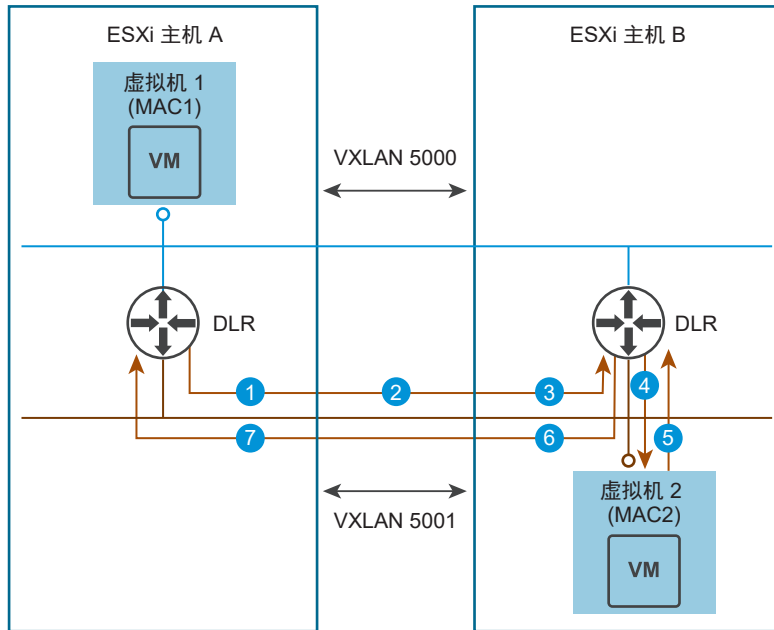
返回流量将采用相同的顺序，来自虚拟机 2 的流量转发到 ESXi 主机 B 上的 DLR 实例，然后通过 VXLAN 5000 上的 L2 进行传送。

DLR ARP 解析过程

在虚拟机 1 中的流量到达虚拟机 2 之前，DLR 需要获悉虚拟机 2 的 MAC 地址。在获悉虚拟机 2 的 MAC 地址后，DLR 可以为出站数据包创建正确的 L2 标头。

图 3-3. DLR ARP 过程 显示了 DLR 的 ARP 解析过程。

图 3-3. DLR ARP 过程



要获悉 MAC 地址，DLR 将执行以下步骤：

- 1 主机 A 上的 DLR 实例生成一个 ARP 请求数据包（源 MAC = vMAC，目标 MAC = 广播）。主机 A 上的 VXLAN 模块在输出 VXLAN 5001 上查找所有 VTEP，然后为每个 VTEP 发送该广播帧的一个副本。
- 2 在该帧通过 VXLAN 封装过程离开主机时，源 MAC 将从 vMAC 更改为 pMAC A，以便返回流量可以在主机 A 上找到源 DLR 实例。该帧现在为源 MAC = pMAC A，目标 MAC = 广播。
- 3 在主机 B 上收到并解封该帧时，将检查该帧并发现它来自于与 VXLAN 5001 上的本地 DLR 实例的 LIF 匹配的 IP 地址。这会将该帧标记为 **abrequest** 以执行代理 ARP 功能。目标 MAC 将从广播更改为 vMAC，以便该帧可以到达本地 DLR 实例。
- 4 主机 B 上的本地 DLR 实例收到 ARP 请求帧（源 MAC = pMAC A，目标 MAC = vMAC），然后查看自己的 LIF IP 地址以请求该信息。它保存源 MAC 并生成新的 ARP 请求数据包（源 MAC = vMAC，目标 MAC = 广播）。该帧将标记为“DVS 本地”，以防止它通过 dvUplink 发生洪泛。DVS 将该帧传送到虚拟机 2。
- 5 虚拟机 2 发送一个 ARP 回复（源 MAC = MAC2，目标 MAC = vMAC）。DVS 将其传送到本地 DLR 实例。
- 6 主机 B 上的 DLR 实例将目标 MAC 替换为在步骤 4 中保存的 pMAC A，然后将数据包发送回 DVS 以传送回主机 A。
- 7 在 ARP 回复到达主机 A 后，目标 MAC 将更改为 vMAC，并且 ARP 回复帧（源 MAC = MAC2，目标 MAC = vMAC）到达主机 A 上的 DLR 实例。

ARP 解析过程已完成，主机 A 上的 DLR 实例现在可以开始将流量发送到虚拟机 2。

DLR ARP 抑制

地址解析协议 (Address Resolution Protocol, ARP) 抑制是一种技术，用于降低连接到同一逻辑交换机的虚拟机之间的各个 VXLAN 分段中的 ARP 广播泛洪量。

在虚拟机 1 要了解虚拟机 2 的 MAC 地址时，它发送一个 ARP 请求。逻辑交换机截获该 ARP 请求；如果逻辑交换机已具有目标的 ARP 条目，则向该虚拟机发送 ARP 响应。

如果没有，则向 NSX Controller 发送一个 ARP 查询。如果控制器知道虚拟机 IP 到 MAC 绑定，则控制器回复该绑定并且逻辑交换机发送 ARP 响应。如果控制器没有 ARP 条目，则在逻辑交换机上重新广播该 ARP 请求。NSX Controller 通过侦测 ARP 请求/DHCP 数据包的交换机安全模块获悉 MAC 地址。

已扩展 ARP 抑制以包含分布式逻辑路由器 (DLR)。

- 来自分布式逻辑路由器的 ARP 请求与来自其他虚拟机的 ARP 请求的处理方式相同，并且可能会受到抑制。如果分布式逻辑路由器必须解析目标 IP 的 ARP 请求，逻辑交换机将抑制该 ARP 请求以防止在控制器已知 IP 到 MAC 绑定时发生泛洪。
- 在创建 LIF 时，分布式逻辑路由器在逻辑交换机中添加该 LIF IP 的 ARP 条目，因此，该 LIF IP 的 ARP 请求也会受到逻辑交换机抑制。

了解 Edge 服务网关提供的路由

NSX 路由的第二个子系统是由 Edge 服务网关提供的。

ESG 实际上是虚拟机中的路由器。其外形尺寸与设备类似并具有四个尺寸，并由 NSX Manager 管理其整个生命周期。ESG 的主要用途是作为外围路由器，它部署在多个 DLR 之间以及物理环境和虚拟网络之间。

ESG 具有以下属性：

- 每个 ESG 最多可以具有 10 个 vNIC 接口或 200 个中继子接口。
- 每个 ESG 支持 8 向 ECMP 以提供路径冗余和可扩展性。

ECMP 数据包流

假设部署了两个 ESG，以便在物理环境中提供具有 2 向 ECMP 上行链路的 DLR 实例。

图 3-4. 具有 ECMP 的简要 ESG 和 DLR 数据包流 显示了在两个 ESG 和物理基础架构之间启用等价多路径 (ECMP) 路由时的 ESG 和 DLR 数据包流。

因此，与具有单个 ESG 的部署相比，虚拟机 1 可以获得 2 倍的双向吞吐量。

VM1 连接到具有 VNI 5000 的逻辑交换机。

DLR 具有两个 LIF - VNI 5000 上的内部 LIF 以及 VNI 5001 上的上行链路 LIF。

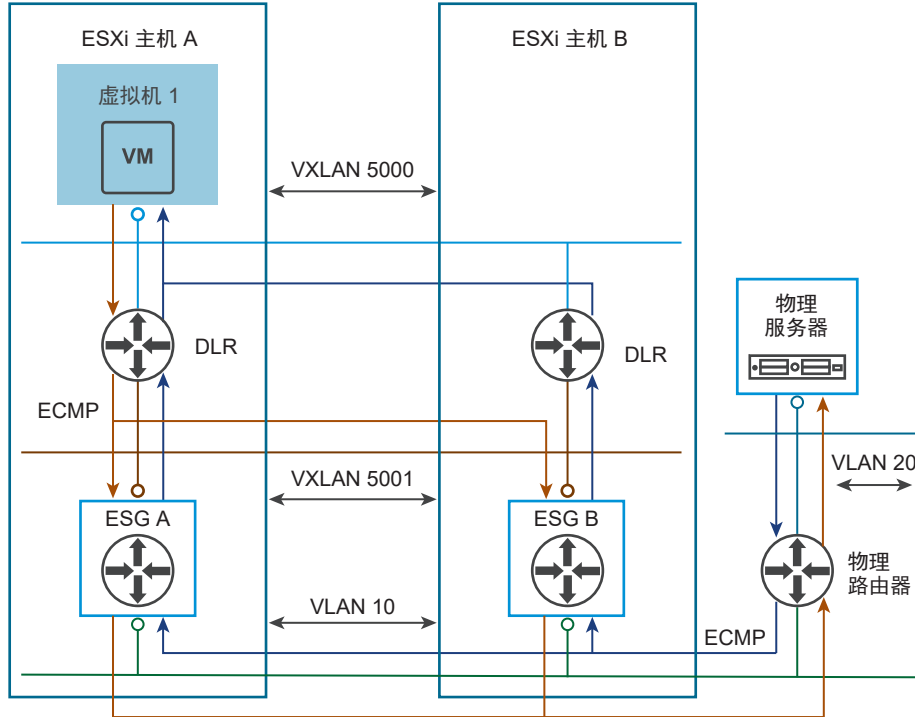
DLR 启用了 ECMP，并通过动态路由协议 (BGP 或 OSPF) 从一对 ESG (ESG A 和 ESG B) 中接收到 VLAN 20 的 IP 子网的等价路由。

两个 ESG 连接到与 VLAN 10 关联且支持 VLAN 的 dvPortgroup，还会在其中连接提供到 VLAN 20 的连接物理路由器。

ESG 通过动态路由协议从物理路由器中接收 VLAN 20 的外部路由。

进行交换的物理路由器从两个 ESG 中获悉与 VXLAN 5000 关联的 IP 子网，并对传输到该子网中的虚拟机的流量执行 ECMP 负载平衡。

图 3-4. 具有 ECMP 的简要 ESG 和 DLR 数据包流



DLR 可以接收最多 8 个等价路由并在这些路由之间平衡流量。图中的 ESG A 和 ESG B 提供了两个等价路由。

ESG 可以执行到物理网络的 ECMP 路由，假设存在多个物理路由器。为简单起见，该图显示单个物理路由器。

不需要在 ESG 上配置到 DLR 的 ECMP，因为所有 DLR LIF 位于 ESG 所在的同一主机“本地”。在 DLR 上配置多个上行链路接口并不会带来额外的好处。

在需要更多北-南带宽的情况下，可以将多个 ESG 放在不同的 ESXi 主机上以通过 8 个 ESG 纵向扩展到大约 80Gbps。

ECMP 数据包流（不包括 ARP 解析）：

- 1 虚拟机 1 将数据包发送到物理服务器，数据包将发送到 ESXi 主机 A 上的虚拟机 1 IP 网关（它是 DLR LIF）。
- 2 DLR 为物理服务器的 IP 执行路由查找，并发现它不是直接连接的，而是与从 ESG A 和 ESG B 中收到的两个 ECMP 路由相匹配。
- 3 DLR 计算 ECMP 哈希，确定下一跃点（可能是 ESG A 或 ESG B），然后将数据包从 VXLAN 5001 LIF 中发出。
- 4 DVS 将数据包传送到选定的 ESG。

- 5 ESG 执行路由查找，并发现可以通过 VLAN 10 上的物理路由器 IP 地址访问物理服务器的子网，它直接连接到 ESG 的某个接口。
- 6 数据包是通过 DVS 发出的，在使用 VLAN ID 10 的相应 801.Q 标记标记后，DVS 将数据包传送到物理网络。
- 7 数据包穿过物理交换基础架构以到达物理路由器，物理路由器执行查找以发现物理服务器直接连接到 VLAN 20 上的接口。
- 8 物理路由器将数据包发送到物理服务器。

在相反方向上：

- 1 物理服务器将数据包发送到虚拟机 1，并将物理路由器作为下一跃点。
- 2 物理路由器为虚拟机 1 的子网执行查找，并发现到该子网的两个等价路径，下一跃点分别为 ESG A 和 ESG B 的 VLAN 10 接口。
- 3 物理路由器选择其中的一个路径，然后将数据包发送到相应的 ESG。
- 4 物理网络将数据包传送到 ESG 所在的 ESXi 主机，然后将其传送到 DVS，DVS 解封数据包并在与 VLAN 10 关联的 dvPortgroup 上将其转发到 ESG。
- 5 ESG 执行路由查找，并发现可以通过与 VXLAN 5001 关联的接口访问虚拟机 1 的子网，下一跃点是 DLR 的上行链路接口 IP 地址。
- 6 ESG 将数据包发送到与 ESG 相同的主机上的 DLR 实例。
- 7 DLR 执行路由查找，并发现可以通过 VXLAN 5000 LIF 访问虚拟机 1。
- 8 DLR 将数据包从其 VXLAN 5000 LIF 发送到 DVS，DVS 执行最终传送。

NSX 路由：必备条件和注意事项

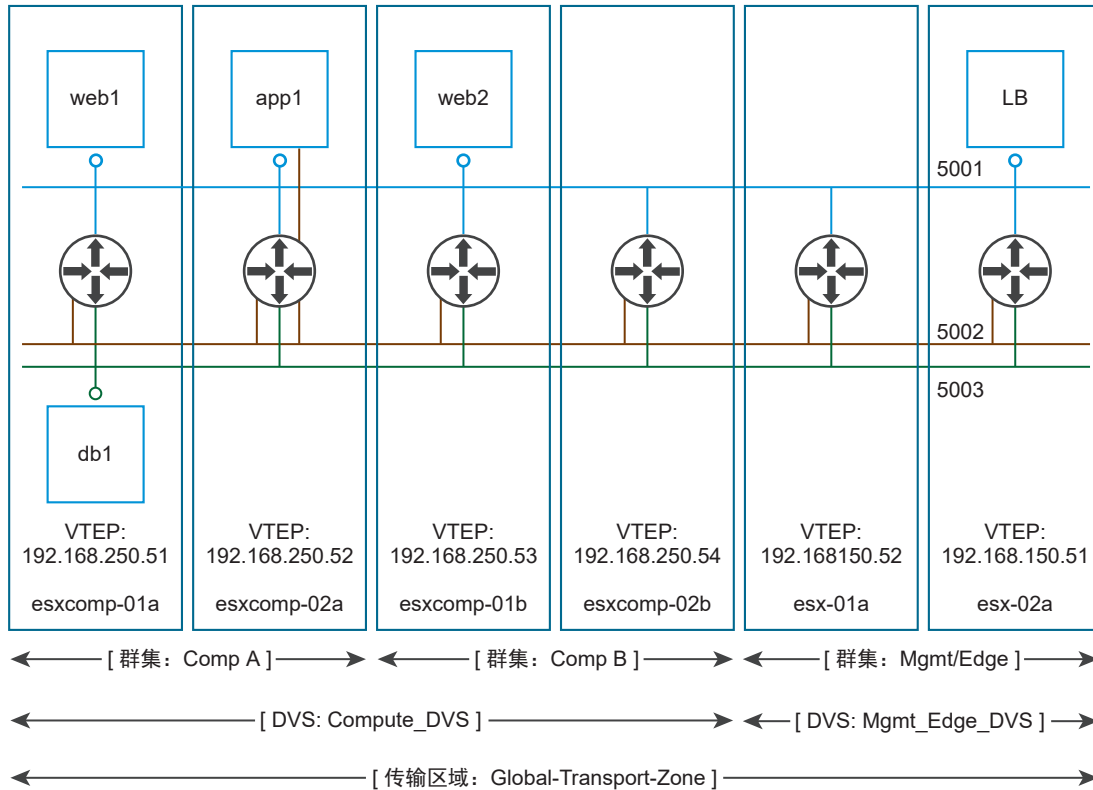
DLR 和 ESG 依靠 DVS 为 dvPortgroup 提供 L2 转发服务（基于 VXLAN 和 VLAN）以使端到端连接正常工作。

这意味着，必须配置连接到 DLR 或 ESG 的 L2 转发服务并且正常运行。在 NSX 安装过程中，“主机准备”和“逻辑网络准备”提供了这些服务。

在多群集 DVS 配置中创建传输区域时，请确保选定的 DVS 中的所有群集包含在传输区域中。这可确保 DLR 在提供了 DVS dvPortgroup 的所有群集上可用。

在传输区域与 DVS 边界对齐时，将正确创建 DLR 实例。

图 3-5. 传输区域与 DVS 边界正确对齐



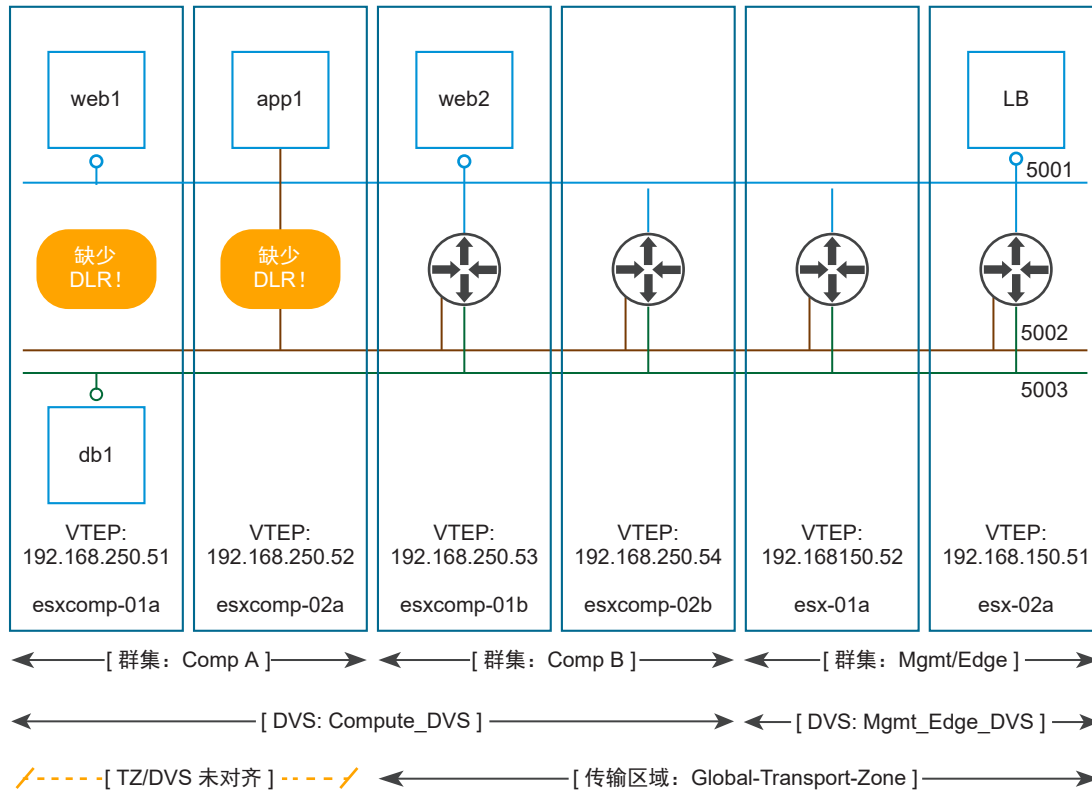
在传输区域与 DVS 边界不对应时，逻辑交换机（5001、5002 和 5003）的范围和这些逻辑交换机连接到的 DLR 实例将断开连接，从而导致 Comp A 群集中的虚拟机无法访问 DLR LIF。

在上图中，DVS “Compute_DVS” 包括两个群集：“Comp A” 和 “Comp B”。“Global-Transport-Zone” 包括 “Comp A” 和 “Comp B”。

这会导致逻辑交换机（5001、5002 和 5003）的范围与在所有群集中这些逻辑交换机所在的所有主机上创建的 DLR 实例正确对齐。

现在，让我们看一下另一种情况：未将传输区域配置为包含 “Comp A” 群集。

图 3-6. 传输区域与 DVS 边界不对齐



在这种情况下，在“Comp A”群集上运行的虚拟机具有所有逻辑交换机的完全访问权限。这是因为逻辑交换机是由主机上的 dvPortgroup 表示的，而 dvPortgroup 是 DVS 范围的结构。在我们的示例环境中，“Compute_DVS”包括“Comp A”和“Comp B”。

不过，创建的 DLR 实例与传输区域范围严格对齐，这意味着，不会在“Comp A”中的主机上创建任何 DLR 实例。

因此，“web1”虚拟机可以访问“web2”和“LB”虚拟机，因为它们位于相同的逻辑交换机上，但“app1”和“db1”虚拟机无法与任何设备进行通信。

DLR 依靠控制器群集才能正常工作，而 ESG 不依靠控制器群集。在创建或更改 DLR 配置之前，请确保控制器群集已启动并且可用。

如果要将 DLR 连接到 VLAN dvPortgroup，请确保配置了 DLR 的 ESXi 主机可以在 UDP/6999 上相互访问以使基于 DLR VLAN 的 ARP 代理正常工作。

注意事项：

- 给定的 DLR 实例无法连接到位于不同传输区域的逻辑交换机。这可确保所有逻辑交换机和 DLR 实例对齐。
- 如果 DLR 连接到跨多个 DVS 的逻辑交换机，则该 DLR 无法连接到支持 VLAN 的端口组。如上所述，这可确保 DLR 实例与主机中的逻辑交换机和 dvPortgroup 正确对齐。
- 在选择 DLR 控制虚拟机位置时，应使用 DRS 反关联性规则以避免将其放在与一个或多个上游 ESG 相同的主机上（如果它们位于同一群集中）。这可降低主机故障对 DLR 转发的影响。

- 只能在单个上行链路上启用 **OSPF**（但支持多个邻接）。另一方面，可以在多个上行链路接口上启用 **BGP**（如有必要）。

DLR 和 ESG UI

DLR 和 ESG UI 指示系统工作状态。

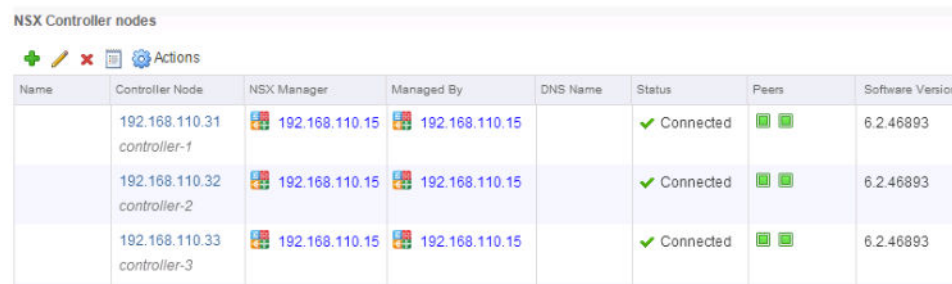
NSX 路由 UI

vSphere Web Client UI 提供了两个与 **NSX** 路由有关的主要部分。

这些部分包含 **L2** 和控制层面基础架构依赖关系和路由子系统配置。

NSX 分布式路由需要使用控制器群集提供的功能。下面的屏幕截图显示了处于正常状态的控制器群集。

NSX Controller nodes



Name	Controller Node	NSX Manager	Managed By	DNS Name	Status	Peers	Software Version
	192.168.110.31 controller-1	192.168.110.15	192.168.110.15		✓ Connected		6.2.46893
	192.168.110.32 controller-2	192.168.110.15	192.168.110.15		✓ Connected		6.2.46893
	192.168.110.33 controller-3	192.168.110.15	192.168.110.15		✓ Connected		6.2.46893

请注意以下事项：

- 部署了三个控制器。
- 所有控制器的“状态”为“已连接”。
- 所有控制器的软件版本相同。
- 每个控制器节点具有两个对等项。

分布式路由的主机内核模块是作为主机上的 **VXLAN** 配置的一部分安装和配置的。这意味着，分布式路由要求准备 **ESXi** 主机并在这些主机上配置 **VXLAN**。

Clusters & Hosts	Installation Status	Firewall	VXLAN
▶ Compute Cluster A	✓ 6.2.3.3771501	✓ Enabled	✓ Configured
▶ Management & Edge Cluster	✓ 6.2.3.3771501	✓ Enabled	✓ Configured

请注意以下事项：

- “安装状态”为绿色。
- “VXLAN”为“已配置”。

确保正确配置了 **VXLAN** 传输组件。

VLAN Transport		Segment ID	Transport Zones				
Clusters & Hosts	Configuration Status	Switch	VLAN	MTU	VMKNic IP Addressing	Teaming Policy	VTEP
▼ Compute Cluster A	Unconfigure	vds-site-a	0	1600	IP Pool	Fail Over	1
esx-02a.corp.local	Ready				vmk3: 192.168.130.51		
esx-01a.corp.local	Ready				vmk3: 192.168.130.52		
▼ Management & Edge	Unconfigure	vds-mgt-edge	0	1600	IP Pool	Fail Over	1
esxmgmt-02a.corp.l	Ready				vmk3: 192.168.120.52		
esxmgmt-01a.corp.l	Ready				vmk3: 192.168.120.51		

请注意以下事项：

- VTEP 传输 VLAN 的 VLAN ID 必须正确无误。请注意，在上面的屏幕截图中，它为“0”。在大多数实际部署中，情况并非如此。
- MTU 配置为 1600 或更大。确保 MTU 未设置为 9000，预计虚拟机上的 MTU 也设置为 9000。DVS 最大 MTU 为 9000；如果虚拟机也是 9000，则没有为 VXLAN 标头留出空间。
- VMKNic 必须具有正确的地址。确保它们未设置为 169.254.x.x 地址，以表明节点无法从 DHCP 中获取地址。
- 对于相同 DVS 的所有群集成员，绑定策略必须是一致的。
- VTEP 数必须与 dvUplink 数相同。确保列出了有效/预期的 IP 地址。

传输区域必须与 DVS 边界正确对齐，以避免出现在某些群集上缺少 DLR 的情况。

Name	NSX vSwitch	Status
Compute Cluster A	vds-site-a	Normal
Management & Edge ...	vds-mgt-edge	Normal

“NSX Edge” UI

NSX 路由子系统是在 UI 的“NSX Edge”部分中配置和管理的。

在选择 UI 的该部分时，将显示以下视图。

Home		NSX Manager: 192.168.110.15 (Role: Primary)						
Networking & Security		0 Installing 0 Failed						
NSX Home	Dashboard	Installation	Logical Switches	NSX Edges	Firewall	SpoolGuard		
Id	Name	Type	Version	Status	Tenant	Interfaces	Size	
edge-2	Local-Distributed-Router	Logical Router	6.2.3	Deployed	Default	4	Compact	
edge-3	Perimeter-Gateway-01	NSX Edge	6.2.3	Deployed	Default	2	Compact	
edge-4	OneArm-LoadBalancer-01	NSX Edge	6.2.3	Deployed	Default	1	Compact	
edge-5	Perimeter-Gateway-02	NSX Edge	6.2.3	Deployed	Default	2	Compact	
edge-6	OneArm-LoadBalancer-02	NSX Edge	6.2.3	Deployed	Default	1	Compact	
edge-9178...	Universal-Distributed-Router	Universal Distributed Router	6.2.3	Deployed	Default	4	Compact	

显示所有当前部署的 DLR 和 ESG，并分别显示以下信息：

- “Id” 显示 ESG 或 DLR Edge 设备 ID，可用于任何 API 调用以引用该 ESG 或 DLR。
- “租户” + “Id” 构成了 DLR 实例名称。可以在 NSX CLI 中看到和使用该名称。
- 对于 DLR，“大小”始终为“精简”；对于 ESG，它是操作员选择的大小。

除了表中的信息以外，还可以通过按钮或“操作”访问上下文菜单。

表 3-1. NSX Edge 上下文菜单

图标	操作
	“强制同步”操作清除 ESG 或 DLR 的控制虚拟机配置，重新引导，然后重新推送该配置。
	“重新部署”删除 ESG 或 DLR，然后使用相同的配置创建新的 ESG 或 DLR。将保留现有的 ID。
	“更改自动规则配置”适用于在 ESG 上启用服务时创建的 ESG 内置防火墙规则（例如，需要使用 TCP/179 的 BGP）。
	“下载技术支持日志”从 ESG 或 DLR 控制虚拟机中创建日志包。 对于 DLR，主机日志未包含在技术支持包中，需要单独进行收集。
	“更改设备大小”仅适用于 ESG。这会使用新设备执行“重新部署”（vNIC MAC 地址将发生变化）。
	通过使用“更改 CLI 凭据”，操作员可以强制更新 CLI 凭据。 如果在 5 次登录失败后在 ESG 或 DLR 控制虚拟机上锁定 CLI，这不会解除锁定。您需要等待 5 分钟，或“重新部署”ESG/DLR 以使用正确的凭据重新登录。
	“更改日志级别”更改发送到 ESG/DLR syslog 的详细信息级别。
	“配置高级调试”在启用核心转储的情况下重新部署 ESG 或 DLR，并连接额外的虚拟磁盘以存储核心转储文件。
	在创建而未部署 ESG 时，可以使用“部署”。 该选项仅执行部署步骤（部署 OVF，配置接口以及将配置推送到创建的设备）。
	如果 DLR/ESG 版本比 NSX Manager 早，则可以使用“升级版本”选项。
	“筛选器”可以按“名称”搜索 ESG/DLR。

新的 NSX Edge (DLR)

在操作员创建新的 DLR 时，将使用以下向导收集所需的信息。

New NSX Edge

1 Name and description
2 Settings
3 Configure deployment
4 Configure interfaces
5 Default gateway settings
6 Ready to complete

Name and description

Install Type: ☐ Edge Services Gateway
Provides common gateway services such as DHCP, Firewall, VPN, NAT, Routing and Load Balancing.
☒ Logical (Distributed) Router
Provides Distributed Routing and Bridging capabilities.
☐ Universal Logical (Distributed) Router
Provides Distributed Routing capabilities for Universal Logical Switches.

Name: * DLR-01
Hostname: dlr-01
Description:
Tenant: Tenant01

☒ Deploy Edge Appliance
Deploys NSX Edge Appliance to support Firewall and Dynamic routing.
☐ Enable High Availability
Enable HA, for enabling and configuring High Availability.

在“名称和描述”屏幕上，将收集以下信息：

- “名称”显示在“NSX Edge”UI 中。
- “主机名”用于设置 ESG 或 DLR 控制虚拟机的 DNS 名称，该名称显示在 SSH/控制台会话、syslog 消息以及 vCenter 的 ESG/DLR 虚拟机“摘要”页中的“DNS 名称”下面。
- “描述”位于 UI 中，它显示 NSX Edge 列表。
- “租户”用于组成 NSX CLI 使用的 DLR 实例名称。外部云管理平台也可能会使用该名称。

在“设置”屏幕上：

New NSX Edge

1 Name and description
2 Settings
3 Configure deployment
4 Configure interfaces
5 Default gateway settings
6 Ready to complete

Settings

CLI credentials will be set on the NSX Edge appliance(s). These credentials can be used to login to the read only command line interface of the appliance.

User Name: * admin
Password: *
Confirm password: *

☒ Enable SSH access
Edge Control Level Logging: EMERGENCY
Set the Edge Control Level Logging

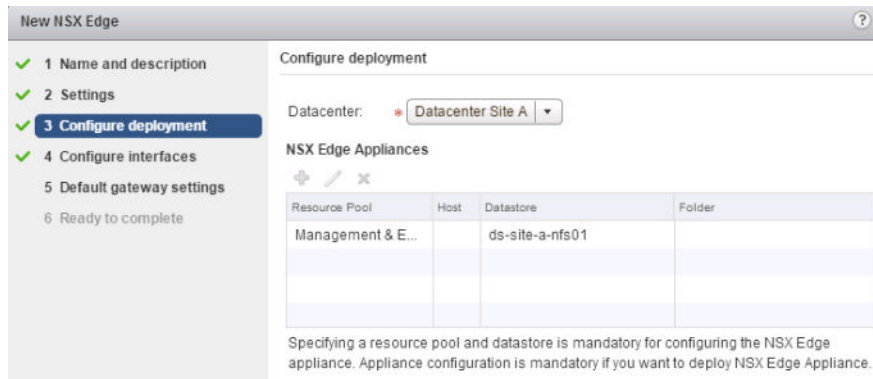
- “用户名”和“密码”设置 CLI/虚拟机控制台凭据以访问 DLR 控制虚拟机。NSX 在 ESG 或 DLR 控制虚拟机上不支持 AAA。该帐户具有 ESG/DLR 控制虚拟机的完全权限；但无法通过 CLI/虚拟机控制台更改 ESG/DLR 配置。

- “启用 SSH 访问” 允许启动 DLR 控制虚拟机上的 SSH 守护程序。
 - 需要调整控制虚拟机防火墙规则以允许 SSH 网络访问。
 - 操作员可以从控制虚拟机管理接口的子网上的主机中连接到 DLR 控制虚拟机，或者通过 OSPF/BGP “协议地址” 进行连接而没有此类限制（如果配置了协议地址）。

注 无法在 DLR 控制虚拟机和属于在该 DLR 的任何“内部”接口上配置的任何子网的任何 IP 地址之间建立网络连接。这是因为 DLR 控制虚拟机上的这些子网的输出接口指向伪接口“VDR”（它未连接到数据层面）。

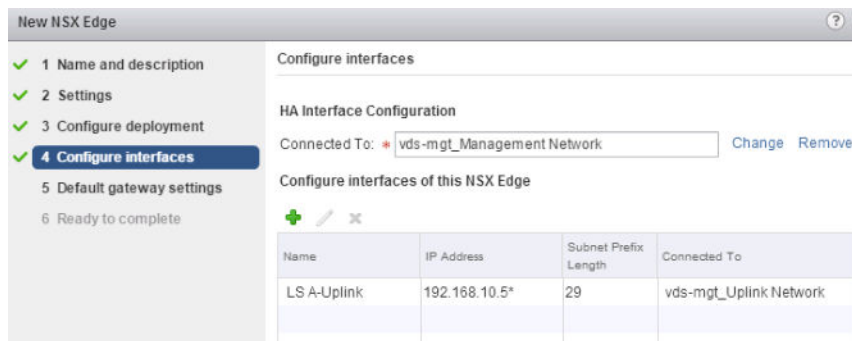
- “启用 HA” 将控制虚拟机部署为活动/备用 HA 对。
- “Edge 控制级别日志记录” 在 Edge 设备上设置 syslog 级别。

在“配置部署”屏幕上：



- “数据中心” 选择在其中部署控制虚拟机的 vCenter 数据中心。
- “NSX Edge 设备” 指的是 DLR 控制虚拟机，并允许定义恰好一个控制虚拟机（如图所示）。
 - 如果启用了“HA”，将在相同的群集、主机和数据存储上部署备用 Edge。将为活动和备用 DLR 控制虚拟机创建 DRS “单独虚拟机” 规则。

在“配置接口”屏幕上：



- “HA 接口”
 - 未创建为可路由的 DLR 逻辑接口。它仅是控制虚拟机上的 vNIC。
 - 该接口不需要使用 IP 地址，因为 NSX 通过 VMCI 管理 DLR 配置。

- 如果在“名称和描述”屏幕上选中 DLR “启用高可用性”，该接口将用于 HA 检测信号。
- “此 NSX Edge 的接口”指的是 DLR 逻辑接口 (LIF)
 - DLR 为“已连接到”的 dvPortgroup 或逻辑交换机上的虚拟机提供 L3 网关服务，这些虚拟机具有相应子网中的 IP 地址。
 - “上行链路”类型的 LIF 在控制虚拟机上创建为 vNIC，因此，最多支持 8 个 vNIC；最后两个可用的 vNIC 分配给 HA 接口并保留一个 vNIC。
 - 动态路由需要使用“上行链路”类型的 LIF 才能在 DLR 上正常工作。
 - “内部”类型的 LIF 在控制虚拟机上创建为伪 vNIC，最多可以具有 991 个伪 vNIC。

在“默认网关设置”屏幕上：

- 如果选定，“配置默认网关”在 DLR 上创建静态默认路由。如果在上一屏幕中创建了“上行链路”类型的 LIF，则可以使用该选项。
- 如果在上行链路上使用 ECMP，建议将该选项保持禁用状态，以防止在下一跃点失败时数据层面中断。

注 右上角的双右箭头可以“暂停”正在执行的向导，以便以后可以恢复该向导。

ESG 和 DLR 差异

与 DLR 相比，在部署 ESG 时，向导屏幕存在一些差异。

第一个差异是“配置部署”屏幕：

对于 ESG，可以在“配置部署”中选择 Edge 大小。如果 ESG 仅用于路由，“中型”是适用于大多数情况的典型大小。选择较大的大小并不会为 ESG 的路由进程提供更多 CPU 资源，并且不会导致更高的吞吐量。

也可以创建而不部署 ESG，这仍需配置 Edge 设备。

以后，可以通过 API 调用或使用“部署”UI 操作部署“未部署的”Edge。

如果选择 Edge HA，您必须创建至少一个“内部”接口，否则，HA 将静默失败，从而导致“脑裂”情况。

NSX UI 和 API 允许操作员移除最后一个“内部”接口，这会导致 HA 静默失败。

典型的 ESG 和 DLR UI 操作

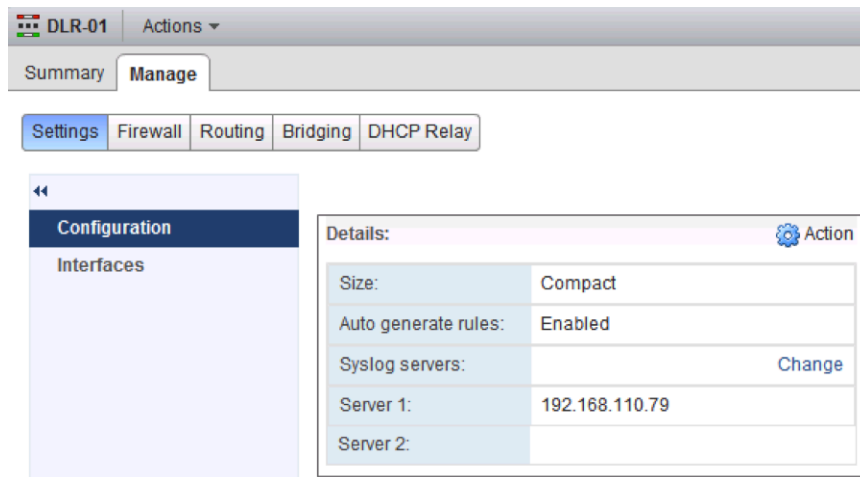
除了创建以外，在初始部署后通常还会执行一些配置操作。

其中包括：

- syslog 配置
- 静态路由管理
- 路由协议配置和路由重新分发

syslog 配置

配置 ESG 或 DLR 控制虚拟机以将日志条目发送到远程 syslog 服务器。



备注：

- 必须将 syslog 服务器配置为 IP 地址，因为 ESG/DLR 控制虚拟机没有配置 DNS 解析器。
 - 对于 ESG，可以“启用 DNS 服务”（DNS 代理），ESG 本身可以使用该服务解析 DNS 名称，但通常在具有较少依赖项的更可靠方法中将 syslog 服务器指定为 IP 地址。
- 无法在 UI 中指定 syslog 端口（它始终为 514），但可以指定协议 (UDP/TCP)。

- syslog 消息来自于 Edge 接口的 IP 地址，Edge 的转发表选择该接口以作为 syslog 服务器 IP 的输出。
- 对于 DLR，syslog 服务器的 IP 地址不能位于在 DLR 的任何“内部”接口上配置的任何子网上。这是因为 DLR 控制虚拟机上的这些子网的输出接口指向伪接口“VDR”（它未连接到数据层面）。

默认情况下，将禁用 ESG/DLR 路由引擎的日志记录。如果需要，请单击“动态路由配置”的“编辑”按钮以通过 UI 启用日志记录。

The screenshot displays the NSX Manager configuration page for a Distributed Logical Router (DLR-01). The 'Manage' tab is active, and the 'Routing' sub-tab is selected. On the left, a sidebar lists configuration options: Global Configuration, Static Routes, OSPF, BGP, and Route Redistribution. The main area shows the 'Routing Configuration' section with a 'Reset' button. Below this, the 'ECMP' setting is shown as 'Disabled' with a red power icon. The 'Default Gateway' section has 'Edit' and 'Delete' buttons. Below that, fields for 'Interface', 'Gateway IP', 'MTU', and 'Description' are visible. The 'Dynamic Routing Configuration' section has an 'Edit' button. Underneath, 'Router ID' is shown as an empty field. The 'OSPF', 'BGP', and 'Logging' settings are all 'Disabled' with red power icons. The 'Log Level' field is also visible.

您还必须配置路由器 ID，它通常是上行链路接口的 IP 地址。

静态路由

静态路由必须将下一跃点设置为与 DLR 的某个 LIF 或 ESG 的某个接口关联的子网上的 IP 地址。否则，配置将失败。

如果未选定，则自动将下一跃点与某个直接连接的子网匹配以设置“接口”。

Add Static Route

Network: * 10.10.10.0/24

Network should be entered in CIDR format
e.g. 192.169.1.0/24

Next Hop: * 192.168.10.1

Interface:

MTU: 1500

Description:

OK

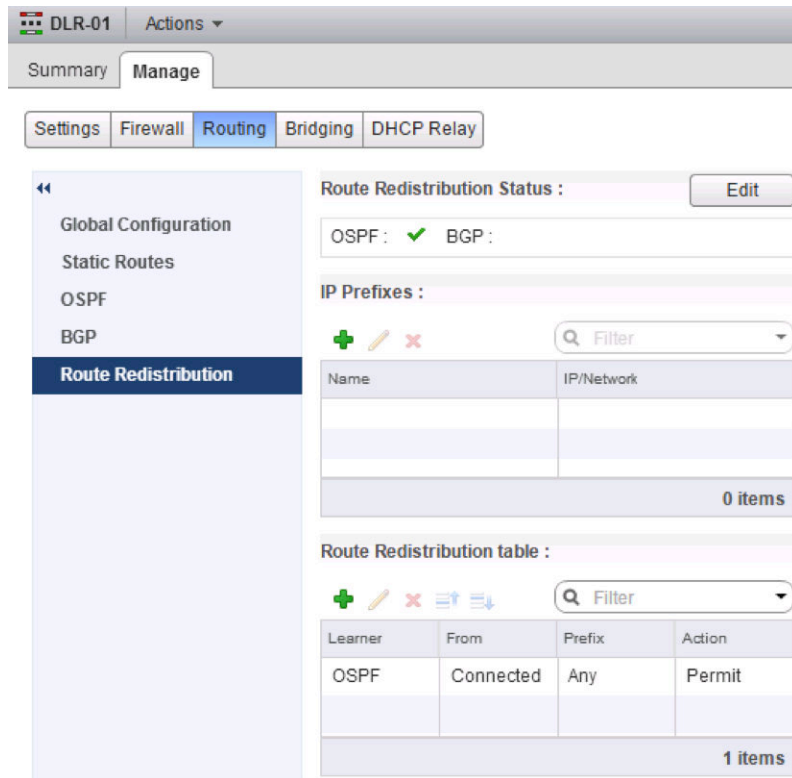
Cancel

路由重新分发

将条目添加到“路由重新分发表”并不会自动为选定的“学习者协议”启用重新分发。必须通过“路由重新分发状态”的“编辑”按钮明确完成该操作。

默认情况下，DLR 配置为将连接的路由重新分发到 OSPF，而 ESG 不是这样。

“路由重新分发表”是按从上到下的顺序处理的，并在首次匹配后停止处理。要从重新分发中排除某些前缀，请在顶部包含更具体的条目。



NSX 路由故障排除

NSX 提供了多种工具以确保路由正常工作。

NSX 路由 CLI

通过使用一组 CLI 命令，操作员可以检查 NSX 路由子系统的各个部分的运行状态。

由于 NSX 路由子系统的分布式特性，可以在各种 NSX 组件上访问一些 CLI。从 NSX 6.2 版开始，NSX 还具有一个集中式 CLI，可帮助缩短访问和登录到各种分布式组件所需的“行程时间”。它可以从一个位置中访问大多数信息：NSX Manager shell。

检查必备条件

每个 ESXi 主机必须满足两个主要的必备条件：

- 连接到 DLR 的任何逻辑交换机正常工作。
- 已成功为 VXLAN 准备 ESXi 主机。

逻辑交换机运行状况检查

NSX 路由与 NSX 逻辑交换配合使用。要验证连接到 DLR 的逻辑交换机是否正常工作，请执行以下操作：

- 查找连接到相关 DLR 的每个逻辑交换机的分段 ID (VXLAN VNI)，例如，5004..5007。

Logical Switches						
NSX Manager: 192.168.110.42						
Name	Status	Transport Zone	Segment ID	Control Plane Mode	Description	
LS A	✓ Normal	Global-Transport-Zone	5004	Unicast		
LS B	✓ Normal	Global-Transport-Zone	5005	Unicast		
LS C	✓ Normal	Global-Transport-Zone	5006	Unicast		
LS D	✓ Normal	Global-Transport-Zone	5007	Unicast		

- 在运行该 DLR 提供服务的虚拟机的 ESXi 主机上，检查连接到该 DLR 的逻辑交换机的 VXLAN 控制层面的状态。

```
# esxcli network vswitch dvs vmware vxlan network list --vds-name=Compute_VDS
```

VXLAN ID	Multicast IP	Control Plane	Controller Connection	Port Count
	MAC Entry Count	ARP Entry Count		
5004	N/A (headend replication)	Enabled (multicast proxy, ARP proxy)	192.168.110.201	
(up)	2	0		
5005	N/A (headend replication)	Enabled (multicast proxy, ARP proxy)	192.168.110.202	
(up)	1	0		
5006	N/A (headend replication)	Enabled (multicast proxy, ARP proxy)	192.168.110.203	
(up)	1	0		
5007	N/A (headend replication)	Enabled (multicast proxy, ARP proxy)	192.168.110.202	
(up)	1	0		

检查每个相关 VXLAN 的以下内容：

- 对于混合或单播模式下的逻辑交换机：
 - Control Plane 为 “Enabled”。
 - 列出了 “multicast proxy” 和 “ARP proxy”；即使禁用了 IP 发现，也会列出 “ARP proxy”。
 - 在 “Controller” 下面列出了有效的控制器 IP 地址，并且 “Connection” 为 “up”。
- “Port Count” 正确无误 - 至少为 1 个，即使在连接到相关逻辑交换机的该主机上没有虚拟机。该端口为 vdrPort，这是连接到 ESXi 主机上的 DLR 内核模块的特殊 dvPort。
- 运行以下命令以确保 vdrPort 连接到每个相关的 VXLAN。

```
~ # esxcli network vswitch dvs vmware vxlan network port list --vds-name=Compute_VDS --vxlan-id=5004
```

Switch Port ID	VDS Port ID	VMKNIC ID
50331656	53	0
50331650	vdrPort	0

```
~ # esxcli network vswitch dvs vmware vxlan network port list --vds-name=Compute_VDS --vxlan-id=5005
```

Switch Port ID	VDS Port ID	VMKNIC ID
50331650	vdrPort	0

- 在上面的示例中，VXLAN 5004 具有一个虚拟机和一个 DLR 连接，而 VXLAN 5005 仅具有一个 DLR 连接。
- 检查是否将相应的虚拟机正确连接到对应的 VXLAN，例如，VXLAN 5004 上的 web-sv-01a。

```
~ # esxcli network vswitch -l
```

DVS Name	Num Ports	Used Ports	Configured Ports	MTU	Uplinks
Compute_VDS	1536	10	512	1600	vmnic0

DVPort ID	In Use	Client
[..skipped..]		
53	1	web-sv-01a.eth0

VXLAN 准备检查

作为 ESXi 主机的 VXLAN 配置的一部分，还会安装和配置 DLR 内核模块，并将其连接到为 VXLAN 准备的 DVS 上的 dvPort。

- 1 运行 `show cluster all` 以获取群集 ID。
- 2 运行 `show cluster cluster-id` 以获取主机 ID。
- 3 运行 `show logical-router host hostID connection` 以获取状态信息。

```
nsxmgr-01a# show logical-router host <hostID> connection
```

Connection Information:

DvsName	VdrPort	NumLifs	VdrVmac
Compute_VDS	vdrPort	4	02:50:56:56:44:52

Teaming Policy: Default Teaming

Uplink : dvUplink1(50331650): 00:50:56:eb:41:d7(Team member)

Stats	Pkt Dropped	Pkt Replaced	Pkt Skipped
Input	0	0	1968734458
Output	303	7799	31891126

- 启用了 VXLAN 的 DVS 将创建一个 vdrPort，它由该 ESXi 主机上的所有 DLR 实例共享。
- “NumLifs” 指的是位于该主机上的所有 DLR 实例中的 LIF 总和。
- “VdrVmac” 是 DLR 在所有实例中的所有 LIF 上使用的 vMAC。该 MAC 在所有主机上是相同的。在 ESXi 主机外部的物理网络中传输的任何帧中，不会看到该内容。

- 对于启用了 VXLAN 的 DVS 的每个 dvUplink，具有一个匹配的 VTEP；但使用 LACP/以太网通道成组模式时除外，此时，仅创建一个 VTEP，而与 dvUplink 数无关。
 - 在离开主机时 DLR 路由的流量（源 MAC = vMAC）将源 MAC 更改为相应 dvUplink 的 pMAC。
 - 请注意，将使用原始虚拟机的源端口或源 MAC 确定 dvUplink（这是在 DVS 元数据中为每个数据包保留的）。
 - 如果在主机上具有多个 VTEP 并且某个 dvUplink 发生故障，与故障 dvUplink 关联的 VTEP 以及绑定到该 VTEP 的所有虚拟机将移动到剩下的某个 dvUplink。这样做是为了避免与将虚拟机移动到不同 VTEP 有关的控制层面更改发生洪泛。
- 每个“dvUplinkX”旁边的“()”中的数字是 dvPort 编号。这对于单个上行链路上的数据包捕获非常有用。
- 为每个“dvUplinkX”显示的 MAC 地址是与该 dvUplink 关联的“pMAC”。该 MAC 地址用于来自于 DLR 的流量，例如，DLR 生成的 ARP 查询以及在离开 ESXi 主机时 DLR 路由的任何数据包。可以在物理网络上看到该 MAC 地址：直接（如果 DLR LIF 具有 VLAN 类型）或从 VXLAN LIF 的 VXLAN 数据包中。
- Pkt Dropped/Replaced/Skipped 指的是与 DLR 内部实施详细信息有关的计数器，通常不用于故障排除或监控。

路由简要概述

为了有效地解决路由问题，了解路由的工作方式和查看相关的信息表是非常有用的。

- 1 收到一个数据包以发送到目标 IP 地址。
- 2 检查路由表并确定下一跃点的 IP 地址。
- 3 确定可访问该地址的网络接口。
- 4 获取该下一跃点的 MAC 地址（通过 ARP）。
- 5 生成一个 L2 帧。
- 6 将该帧从接口中发出。

因此，要进行路由，您需要使用：

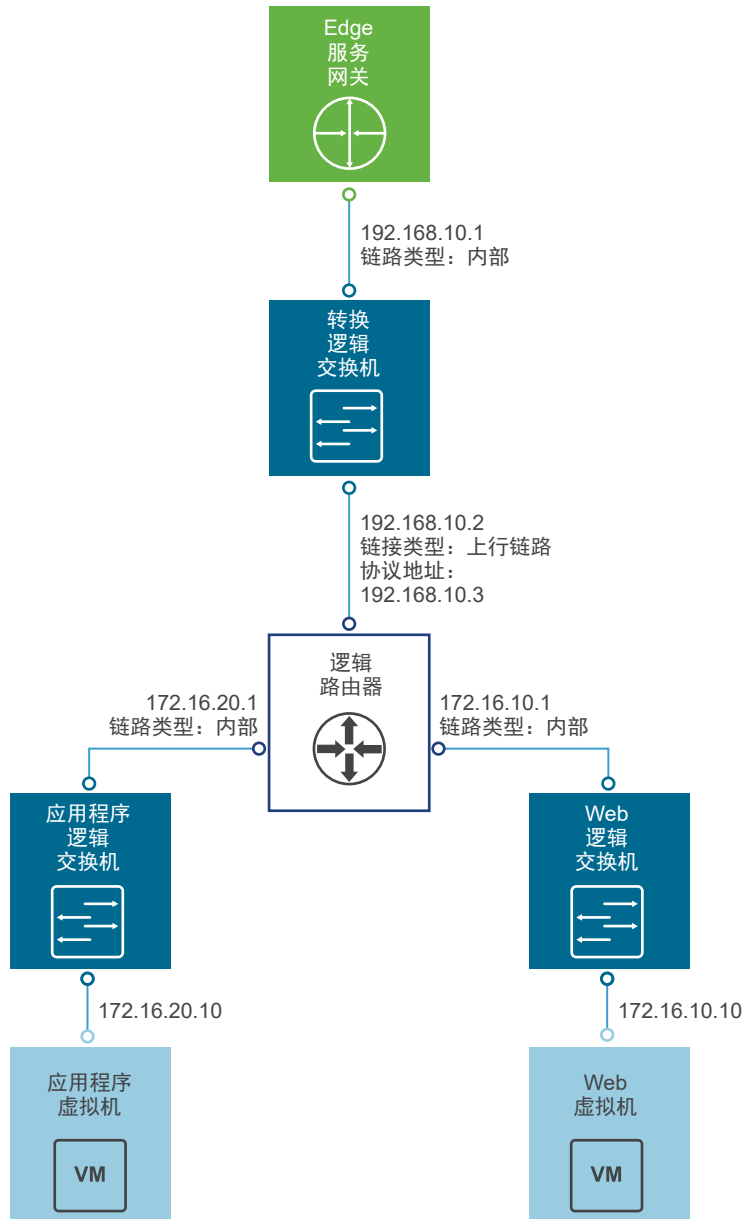
- 接口表（包含接口 IP 地址和子网掩码）
- 路由表
- ARP 表

使用示例路由拓扑验证 DLR 状态

本节讨论了如何验证 DLR 路由数据包所需的信息。

让我们使用示例路由拓扑，并创建一组逻辑交换机和 DLR 以在 NSX 中进行创建。

图 3-7. 示例路由拓扑



该图显示：

- 4 个逻辑交换机，每个交换机具有自己的子网
- 3 个虚拟机，每个逻辑交换机连接一个虚拟机
 - 每个虚拟机具有自己的 IP 地址和 IP 网关
 - 每个虚拟机具有 MAC 地址（显示了最后两个八位字节）
- 一个连接到 4 个逻辑交换机的 DLR；一个逻辑交换机用于“上行链路”，其余逻辑交换机是内部交换机
- 一个外部网关，它可能是 ESG 以作为 DLR 的上游网关

为上面的 DLR 显示了“即将完成”向导屏幕。

New NSX Edge

Ready to complete

Name and description
 Name: DLR1
 Install Type: Logical (Distributed) Router
 Tenant:
 HA: Disabled

Management Interface Configuration
 Connected To: Mgmt_Edge_VDS - Mgmt

IP Address	Subnet Prefix Length

NSX Edge Appliances

Resource Pool	Host	Datastore	Folder
Management and Edge Cluster		ds-site-a-nfs01	

Interfaces

Name	IP Address	Subnet Prefix Length	Connected To
LS A	172.16.10.1*	24	LS A
LS B	172.16.20.1*	24	LS B
LS C	172.16.30.1*	24	LS C
LS D	192.168.10.2*	29	LS D

Back Next Finish Cancel

在 DLR 部署完成后，可以使用 **ESXi CLI** 命令查看和验证涉及的主机上的相关 DLR 的分布状态。

确认 DLR 实例

首先要确认的是，是否创建了 DLR 实例以及其控制层面是否处于活动状态。

- 1 从 NSX Manager shell 中，运行 `show cluster all` 以获取群集 ID。
- 2 运行 `show cluster cluster-id` 以获取主机 ID。
- 3 运行 `show logical-router host hostID dlr all verbose` 以获取状态信息。

```
nsxmgr# show logical-router host host-id dlr all verbose
```

VDR Instance Information :

```
-----
Vdr Name:          default+edge-1
Vdr Id:            1460487509
Number of Lifs:    4
Number of Routes:  5
State:             Enabled
Controller IP:     192.168.110.201
Control Plane Active: Yes
Control Plane IP:  192.168.210.51
Edge Active:       No
```

请注意以下几点：

- 该命令显示位于给定 ESXi 主机上的所有 DLR 实例。

- “Vdr Name” 由“租户”和“Edge Id”组成。在该示例中，未指定“租户”，因此，使用“default”一词。“Edge Id”是“edge-1”，可以在 NSX UI 中看到该 ID。
 - 如果主机上具有多个 DLR 实例，一种查找正确实例的方法是查找在 UI “NSX Edge” 中显示的“Edge ID”。
- “Vdr Id” 对于进一步查找非常有用，包括日志。
- “Number of Lifs” 指的是位于该单个 DLR 实例上的 LIF。
- 此处，“Number of Routes” 为 5，它包含 4 个直接连接的路由（每个 LIF 一个）和一个默认路由。
- “State”、“Controller IP” 和 “Control Plane Active” 指的是 DLR 的控制层面状态，必须列出正确的控制器 IP 并且 Control Plane Active 为 Yes。请记住，DLR 功能需要使用正常工作的控制器；上面的输出显示正常 DLR 实例所需的设置。
- “控制层面 IP” 指的是 ESXi 主机用于与控制器通信的 IP 地址。该 IP 始终是与 ESXi 主机的管理 vmknics 关联的 IP 地址，在大多数情况下，该 IP 为 vmk0。
- “Edge Active” 显示该主机是否为运行该 DLR 实例的控制虚拟机的宿主以及是否处于活动状态。
 - 活动 DLR 控制虚拟机的位置决定了用于执行 NSX L2 桥接（如果启用）的 ESXi 主机。
- 还提供了上述命令的“brief”版本，以便生成压缩的输出以提供概要信息。请注意，此处以十六进制格式显示“Vdr Id”：

```
nsxmgr# show logical-router host host-id dlr all brief
```

```
VDR Instance Information :
```

```
State Legend: [A: Active], [D: Deleting], [X: Deleted], [I: Init]
```

```
State Legend: [SF-R: Soft Flush Route], [SF-L: Soft Flush LIF]
```

Vdr Name	Vdr Id	#Lifs	#Routes	State	Controller Ip	CP Ip
default+edge-1	0x570d4555	4	5	A	192.168.110.201	192.168.210.51

“Soft Flush” 状态指的是 LIF 生命周期的短暂过渡状态，通常在正常 DLR 中看不到该状态。

DLR 的逻辑接口

在确定已创建 DLR 后，请确保 DLR 的所有逻辑接口存在并具有正确的配置。

- 1 从 NSX Manager shell 中，运行 `show cluster all` 以获取群集 ID。
- 2 运行 `show cluster cluster-id` 以获取主机 ID。
- 3 运行 `show logical-router host hostID dlr all brief` 以获取 dlrID（Vdr 名称）。
- 4 运行 `show logical-router host hostID dlr dlrID interface all brief` 以获取所有接口的摘要状态信息。

- 5 运行 `show logical-router host hostID dlr dlrID interface (all | intName) verbose` 以获取所有接口或特定接口的状态信息。

```
nsxmgr# show logical-router host hostID dlr dlrID interface all verbose
```

VDR default+edge-1:1460487509 LIF Information :

```
Name:          570d45550000000a
Mode:          Routing, Distributed, Internal
Id:           Vxlan:5000
Ip(Mask):      172.16.10.1(255.255.255.0)
Connected Dvs: Compute_VDS
VXLAN Control Plane: Enabled
VXLAN Multicast IP: 0.0.0.1
State:         Enabled
Flags:         0x2388
DHCP Relay:    Not enabled
```

```
Name:          570d45550000000c
Mode:          Routing, Distributed, Internal
Id:           Vxlan:5002
Ip(Mask):      172.16.30.1(255.255.255.0)
Connected Dvs: Compute_VDS
VXLAN Control Plane: Enabled
VXLAN Multicast IP: 0.0.0.1
State:         Enabled
Flags:         0x2288
DHCP Relay:    Not enabled
```

```
Name:          570d45550000000b
Mode:          Routing, Distributed, Internal
Id:           Vxlan:5001
Ip(Mask):      172.16.20.1(255.255.255.0)
Connected Dvs: Compute_VDS
VXLAN Control Plane: Enabled
VXLAN Multicast IP: 0.0.0.1
State:         Enabled
Flags:         0x2388
DHCP Relay:    Not enabled
```

```
Name:          570d455500000002
Mode:          Routing, Distributed, Uplink
Id:           Vxlan:5003
Ip(Mask):      192.168.10.2(255.255.255.248)
Connected Dvs: Compute_VDS
VXLAN Control Plane: Enabled
VXLAN Multicast IP: 0.0.0.1
State:         Enabled
Flags:         0x2208
DHCP Relay:    Not enabled
```

请注意以下几点：

- LIF 的 “Name” 在主机上的所有 DLR 实例中是唯一的。它在主机和 DLR 的主控制器节点上是相同的。
- LIF 的 “Mode” 显示 LIF 是路由还是桥接，以及它是内部还是上行链路。
- “Id” 显示 LIF 类型和相应的服务 ID（VXLAN 和 VNI 或 VLAN 和 VID）。
- 将为 “Routing” LIF 显示 “Ip(Mask)”。
- 如果 LIF 在混合或单播模式下连接到 VXLAN，则 “VXLAN 控制层面” 为 “Enabled”。
- 对于 VXLAN 处于单播模式的 VXLAN LIF，“VXLAN Multicast IP” 显示为 “0.0.0.1”，否则，显示实际多播 IP 地址。
- 对于路由的 LIF，“State” 应该为 “Enabled”。对于桥接 LIF，在执行桥接的主机上为 “Enabled”，在所有其他主机上为 “Init”。
- “Flags” 是 LIF 状态的摘要表示形式，并显示 LIF 是：
 - 路由还是桥接
 - VLAN LIF 是否为 DI
 - 它是否启用了 DHCP 中继
 - 请注意 0x0100 标记，它是在 DLR 导致 VXLAN VNI 加入时设置的（与在该 VXLAN 上具有虚拟机的宿主相对）
 - 在 “brief” 模式下，将以更便于阅读的格式显示标记

```
nsxmgr# show logical-router host hostID dlr dlrID interface all brief
```

VDR default+edge-1 LIF Information :

State Legend: [A:Active], [d:Deleting], [X:Deleted], [I:Init],[SF-L:Soft Flush LIF]

Modes Legend: [B:Bridging],[E: Empty], [R:Routing],[S:Sedimented],[D:Distributed]

Modes Legend: [In:Internal],[Up:Uplink]

Lif Name	Id	Mode	State	Ip(Mask)
-----	--	-----	-----	-----
570d45550000000a	Vxlan:5001	R,D,In	A	172.16.10.1(255.255.255.0)
570d45550000000c	Vxlan:5003	R,D,In	A	172.16.30.1(255.255.255.0)
570d45550000000b	Vxlan:5002	R,D,In	A	172.16.20.1(255.255.255.0)
570d455500000002	Vxlan:5000	R,D,Up	A	192.168.10.5(255.255.255.248)

DLR 的路由

在确定 DLR 存在、正常工作并具有所有 LIF 后，接下来应检查路由表。

- 1 从 NSX Manager shell 中，运行 `show cluster all` 以获取群集 ID。
- 2 运行 `show cluster cluster-id` 以获取主机 ID。
- 3 运行 `show logical-router host hostID dlr all brief` 以获取 dlrID（Vdr 名称）。

4 运行 `show logical-router host hostID dlr dlrID route` 以获取所有接口的状态信息。

```
nsxmgr# show logical-router host hostID dlr dlrID route
```

VDR default+edge-1:1460487509 Route Table

Legend: [U: Up], [G: Gateway], [C: Connected], [I: Interface]

Legend: [H: Host], [F: Soft Flush] [!: Reject] [E: ECMP]

Destination	GenMask	Gateway	Flags	Ref	Origin	UpTime	Interface
0.0.0.0	0.0.0.0	192.168.10.1	UG	1	AUTO	10068944	570d455500000002
172.16.10.0	255.255.255.0	0.0.0.0	UCI	1	MANUAL	10068944	570d45550000000a
172.16.20.0	255.255.255.0	0.0.0.0	UCI	1	MANUAL	10068944	570d45550000000b
172.16.30.0	255.255.255.0	0.0.0.0	UCI	1	MANUAL	10068944	570d45550000000c
192.168.10.0	255.255.255.248	0.0.0.0	UCI	1	MANUAL	10068944	570d455500000002

请注意以下几点：

- “Interface” 显示为相应的 “Destination” 选择的输出 LIF。它设置为 DLR 的某个 LIF 的 “Lif Name”。
- 对于 ECMP 路由，存在多个具有相同 Destination、GenMask 和 Interface 但具有不同 Gateway 的路由。标记还包括 “E” 以反映这些路由的 ECMP 特性。

DLR 的 ARP 表

对于 DLR 转发的数据包，它必须能够解析下一跃点 IP 地址的 ARP 请求。该解析过程的结果存储在各个主机的 DLR 实例本地。

控制器在该过程中不起任何作用，不会使用控制器将生成的 ARP 条目分发到其他主机。

非活动缓存条目保留 600 秒，然后将其移除。有关 DLR ARP 解析过程的详细信息，请参见 [DLR ARP 解析过程](#)。

- 1 从 NSX Manager shell 中，运行 `show cluster all` 以获取群集 ID。
- 2 运行 `show cluster cluster-id` 以获取主机 ID。
- 3 运行 `show logical-router host hostID dlr all brief` 以获取 dlrID (Vdr 名称)。
- 4 运行 `show logical-router host hostID dlr dlrID arp` 以获取所有接口的状态信息。

```
nsxmgr# show logical-router host hostID dlr dlrID arp
```

VDR default+edge-1:1460487509 ARP Information :

Legend: [S: Static], [V: Valid], [P: Proxy], [I: Interface]

Legend: [N: Nascent], [L: Local], [D: Deleted]

Network	Mac	Flags	Expiry	SrcPort	Interface	Refcnt
172.16.10.1	02:50:56:56:44:52	VI	permanent	0	570d45550000000a	1
172.16.10.11	00:50:56:a6:7a:a2	VL	147	50331657	570d45550000000a	2
172.16.30.1	02:50:56:56:44:52	VI	permanent	0	570d45550000000c	1

172.16.30.11	00:50:56:a6:ba:09	V	583	50331650	570d45550000000c 2
172.16.20.11	00:50:56:a6:84:52	VL	568	50331658	570d45550000000b 2
172.16.20.1	02:50:56:56:44:52	VI	permanent	0	570d45550000000b 1
192.168.10.2	02:50:56:56:44:52	VI	permanent	0	570d455500000002 1
192.168.10.1	00:50:56:8e:ee:ce	V	147	50331650	570d455500000002 1

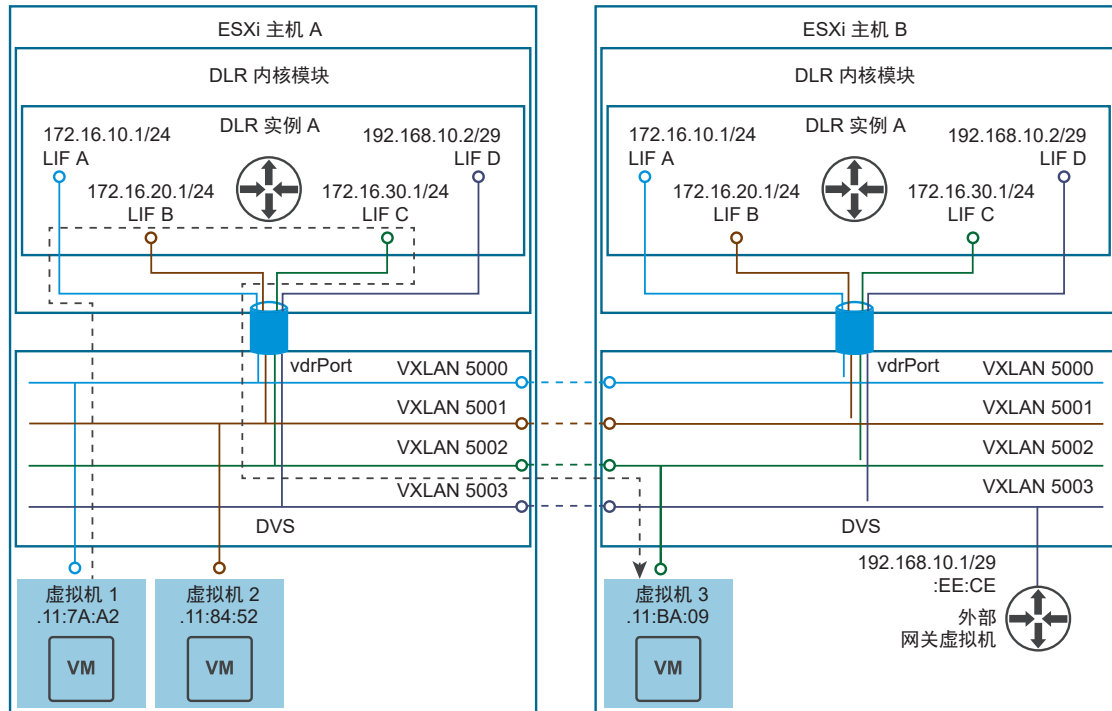
请注意以下事项：

- DLR 自己的 LIF 的所有 ARP 条目（“I” 标记）是相同的，并显示 [VXLAN 准备检查](#) 中讨论的相同 vMAC。
- 具有“L”标记的 ARP 条目对应于在运行 CLI 命令的主机上运行的虚拟机。
- “SrcPort”显示 ARP 条目来自的 dvPort ID。如果 ARP 条目来自于另一个主机，将显示 dvUplink 的 dvPort ID。该 dvPort ID 可以与 [VXLAN 准备检查](#) 中讨论的 dvUplink dvPort ID 交叉引用。
- 通常不会看到“Nascent”标记。在 DLR 等待 ARP 回复到达时，将设置该标记。设置了该标记的任何条目可能表明 ARP 解析出现问题。

可视化 DLR 及其相关主机组件

下图显示了两个主机（ESXi 主机 A 和 ESXi 主机 B），其中配置了示例“DLR 实例 A”并连接到四个 VXLAN LIF。

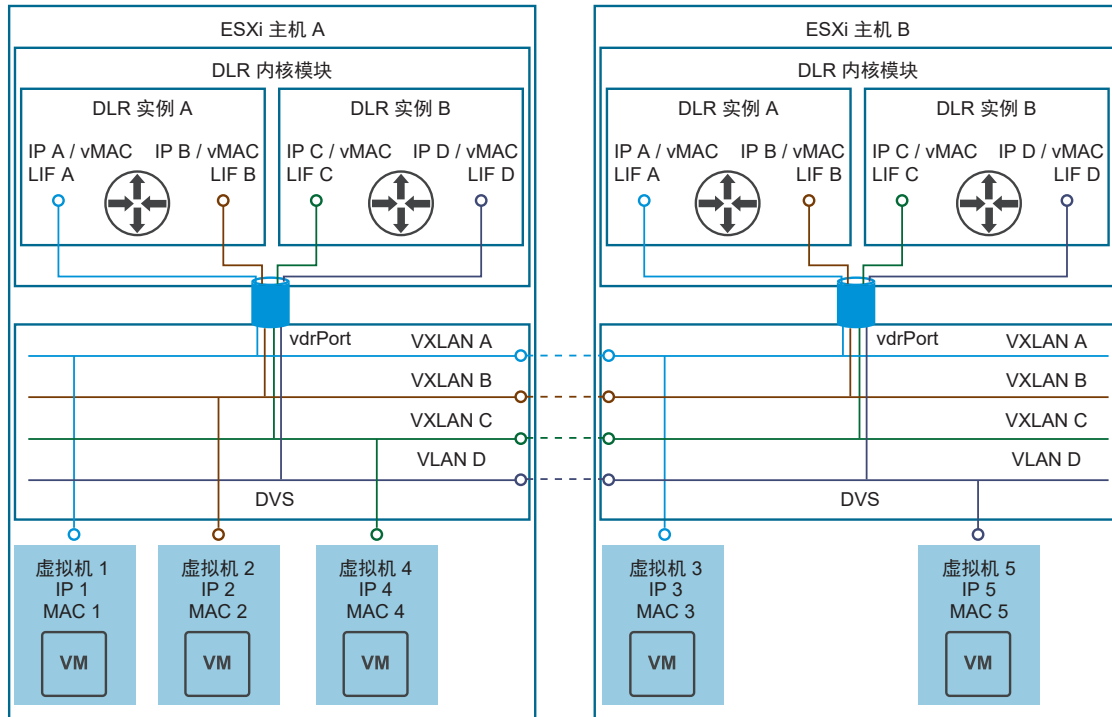
图 3-8. 两个具有单个 DLR 实例的主机



- 每个主机具有一个“L2 交换机” (DVS) 和一个“单臂路由器”（DLR 内核模块），该路由器通过“中继”接口 (vdrPort) 连接到该“交换机”。
 - 请注意，该“中继”接口可以传输 VLAN 和 VXLAN，但在通过 vdrPort 传输的数据包中不包含 801.Q 或 UDP/VXLAN 标头。相反，DVS 使用内部元数据标记方法将该信息传送到 DLR 内核模块。
- 在看到目标 MAC = VMAC 的帧时，DVS 知道应将其发送到 DLR 并将该帧转发到 vdrPort。
- 在数据包通过 vdrPort 到达 DLR 内核模块后，将检查其元数据以确定它们所属的 VXLAN VNI 或 VLAN ID。然后，使用该信息确定数据包所属的 DLR 实例的 LIF。
 - 该系统的不足之处是，无法将多个 DLR 实例连接到给定的 VLAN 或 VXLAN。

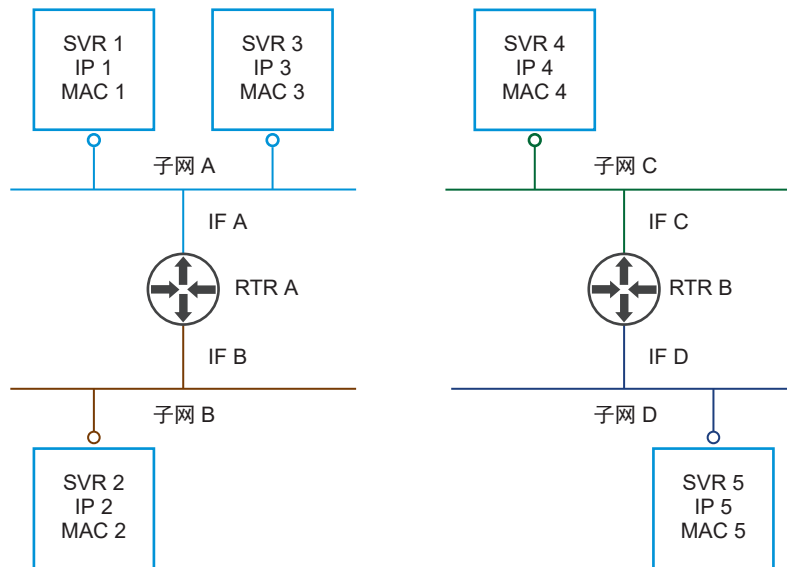
如果存在多个 DLR 实例，上图将如下所示：

图 3-9. 两个具有两个 DLR 实例的主机



这对应于具有两个独立路由域的网络拓扑，这两个域彼此完全隔离，并且可能具有重叠的 IP 地址。

图 3-10. 与两个主机和两个 DLR 实例对应的网络拓扑



分布式路由子系统架构

ESXi 主机上的 DLR 实例可以访问执行 L3 路由所需的所有信息。

- 直接连接网络（从接口的配置中了解）
- 每个子网的下一跃点（在路由表中查找）

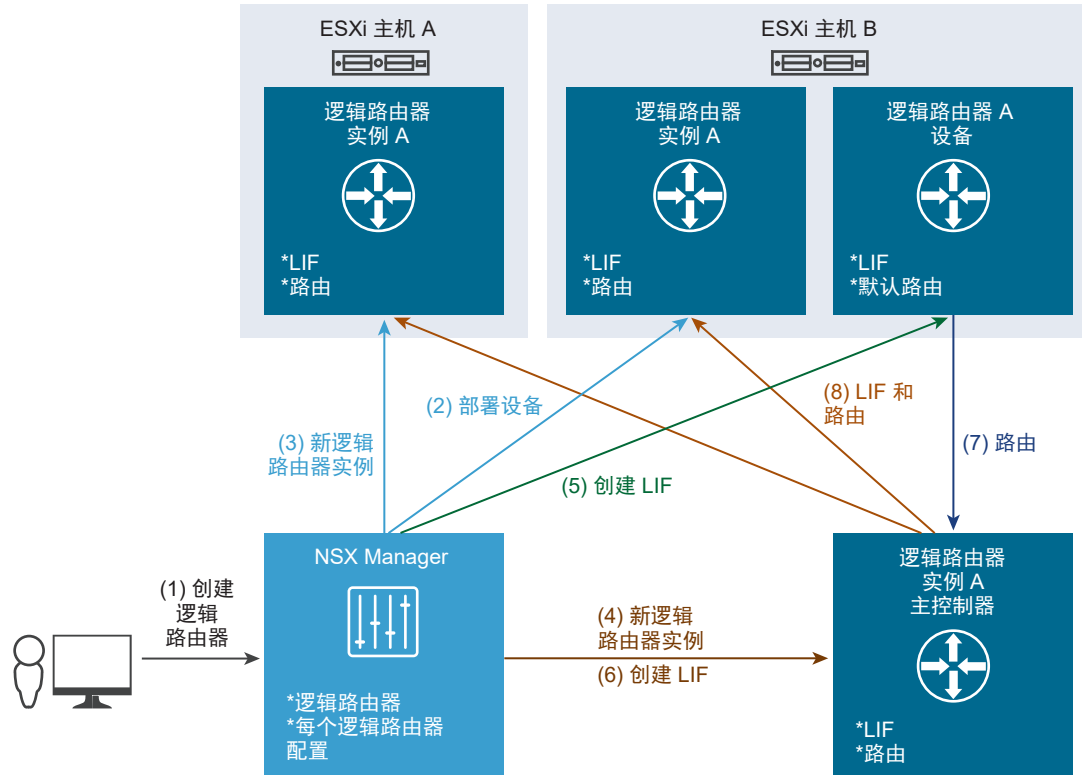
- 要插入到输出帧以到达下一跃点的 MAC 地址（ARP 表）

该信息将传送到在多个 ESXi 主机中分配的实例。

DLR 创建过程

下图简要说明了 NSX 在创建新的 DLR 时执行的过程。

图 3-11. DLR 创建过程



在使用“完成”按钮提交 UI 向导或进行 API 调用以部署新的 DLR 时，系统将执行以下步骤：

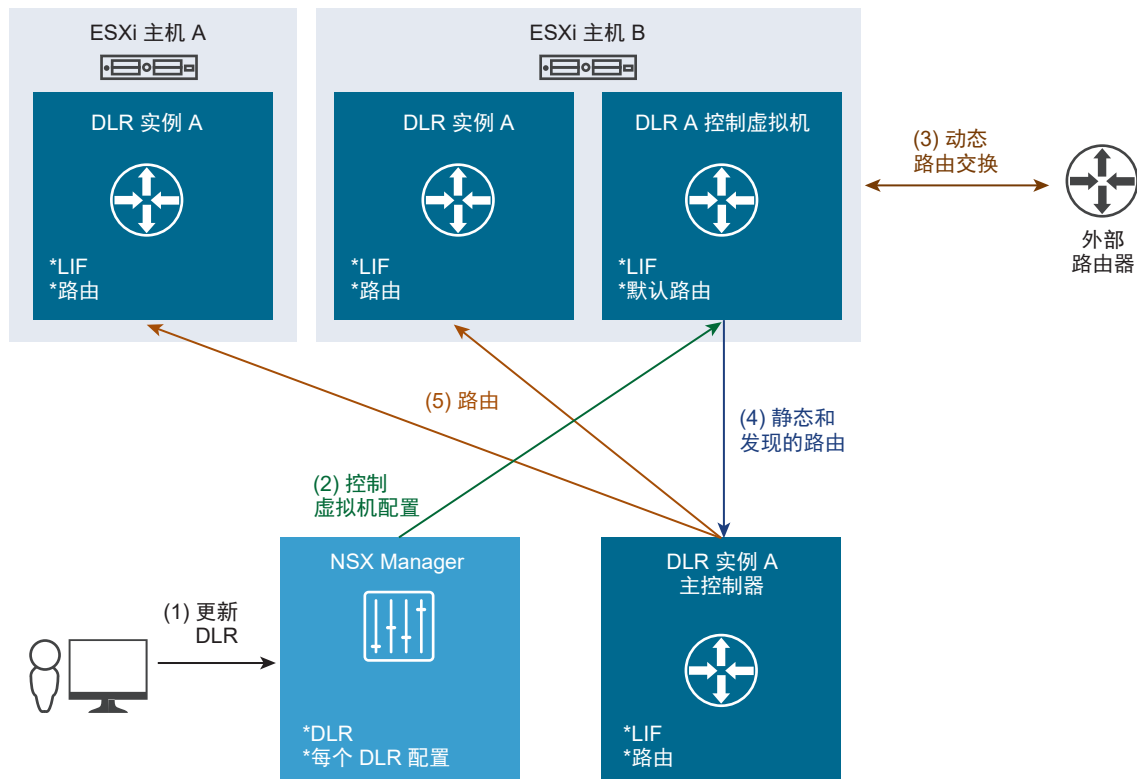
- 1 NSX Manager 收到 API 调用以部署新的 DLR（直接或从 UI 向导调用的 vSphere Web Client 中）。
- 2 NSX Manager 调用其链接的 vCenter Server 以部署一个或一对（如果请求 HA）DLR 控制虚拟机。
 - a 打开 DLR 控制虚拟机电源并连接回 NSX Manager 以准备接收配置。
 - b 如果部署了 HA 对，NSX Manager 配置反关联性规则以使 HA 对在不同的主机上运行。然后，DRS 采取措施以将它们分开。
- 3 NSX Manager 在主机上创建 DLR 实例：
 - a NSX Manager 查找要连接到新 DLR 的逻辑交换机，以确定它们属于哪个传输区域。
 - b 然后，它查找在该传输区域中配置的一组群集，并在这些群集中的每个主机上创建新的 DLR。
 - c 此时，主机仅知道新的 DLR ID，而没有任何相应的信息（LIF 或路由）。
- 4 NSX Manager 在控制器群集上创建新的 DLR 实例。
 - a 控制器群集将某个控制器节点分配为该 DLR 实例的主节点。

- 5 NSX Manager 将配置（包括 LIF）发送到 DLR 控制虚拟机。
 - a ESXi 主机（包括运行 DLR 控制虚拟机的宿主）从控制器群集中接收切片信息，确定负责新 DLR 实例的控制器节点，然后连接到该控制器节点（如果没有现有的连接）。
- 6 在 DLR 控制虚拟机上创建 LIF 后，NSX Manager 在控制器群集上创建新 DLR 的 LIF。
- 7 DLR 控制虚拟机连接到新 DLR 实例的控制器节点，然后将路由发送到该控制器节点：
 - a 首先，DLR 将其路由表转换为转发表（通过将前缀解析为 LIF）。
 - b 然后，DLR 将生成的表发送到该控制器节点。
- 8 通过在步骤 5.a 中建立的连接，控制器节点将 LIF 和路由推送到新 DLR 实例所在的其他主机。

将动态路由添加到 DLR 中

在通过“直接”API 调用（与使用 vSphere Web Client UI 相对）创建 DLR 时，可能会为其提供包含动态路由的完整配置 (1)。

图 3-12. DLR 上的动态路由



- 1 NSX Manager 收到 API 调用以更改现有 DLR 的配置，此处指的是添加动态路由。
- 2 NSX Manager 将新配置发送到 DLR 控制虚拟机。
- 3 DLR 控制虚拟机应用该配置并执行以下过程：建立路由邻接，交换路由信息，等等。
- 4 在交换路由后，DLR 控制虚拟机计算转发表，并将其发送到 DLR 的主控制器节点。
- 5 然后，DLR 的主控制器节点将更新的路由分发到 DLR 实例所在的 ESXi 主机。

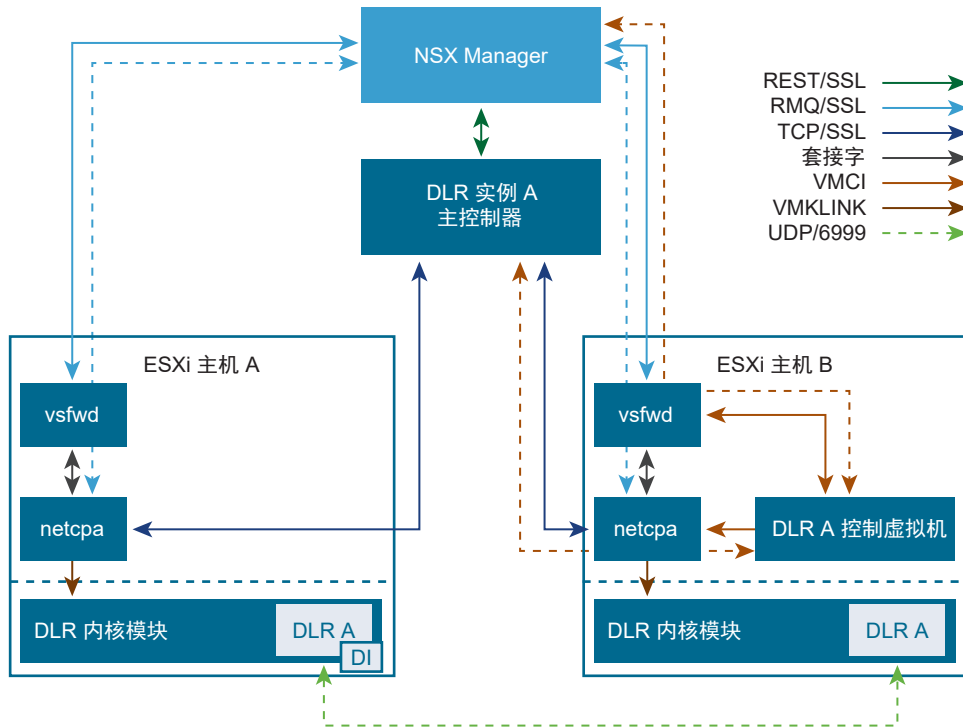
请注意，运行 DLR 控制虚拟机的 ESXi 主机上的 DLR 实例收到其 LIF，并且仅从 DLR 的主控制器节点中路由，而从不直接从 DLR 控制虚拟机或 NSX Manager 中路由。

DLR 控制和管理层面组件和通信

本节简要说明了 DLR 控制和管理层面的组件。

该图显示了这些组件以及它们之间的相应通信通道。

图 3-13. DLR 控制和管理层面组件



- **NSX Manager:**
 - 具有与控制器群集的直接通信
 - 具有到为 NSX 准备的每个主机上运行的消息总线客户端 (**vsfwd**) 进程的直接永久连接
- 对于每个 DLR 实例，将一个控制器节点（共有 3 个可用的节点）选择为主节点
 - 如果原始控制器节点发生故障，主节点功能可以移动到其他控制器节点
- 每个 ESXi 主机运行两个用户环境代理 (UWA): 消息总线客户端 (**vsfwd**) 和控制层面代理 (**netcpa**)
 - **netcpa** 需要使用 NSX Manager 中的信息才能正常工作（例如，在何处查找控制器以及如何从控制器中进行身份验证）；可以通过 **vsfwd** 提供的消息总线连接访问该信息
 - **netcpa** 还会与 DLR 内核模块通信，以使用从控制器中收到的相关信息对其进行编程
- 对于每个 DLR 实例，具有一个 DLR 控制虚拟机，它在某个 ESXi 主机上运行；DLR 控制虚拟机具有两个通信通道：
 - 通过 **vsfwd** 到 NSX Manager 的 VMCI 通道，用于配置控制虚拟机

- 通过 netcpa 到 DLR 主控制器的 VMCi 通道，用于将 DLR 的路由表发送到该控制器
- 如果 DLR 具有一个 VLAN LIF，则控制器将涉及的某个 ESXi 主机指定为指定实例 (DI)。其他 ESXi 主机上的 DLR 内核模块请求 DI 在关联的 VLAN 上执行代理 ARP 查询。

NSX 路由子系统组件

NSX 路由子系统是由多个组件实现的。

- NSX Manager
- 控制器群集
- ESXi 主机模块（内核和 UWA）
- DLR 控制虚拟机
- ESG

NSX Manager

NSX Manager 提供与 NSX 路由有关的以下功能：

- 作为集中式管理层面，从而为所有 NSX 管理操作提供统一的 API 访问点
- 在主机上安装分布式路由内核模块和用户环境代理，以做好准备以提供 NSX 功能
- 创建/破坏 DLR 和 DLR LIF
- 通过 vCenter 部署/删除 DLR 控制虚拟机和 ESG
- 通过 REST API 配置控制器群集，并通过消息总线配置主机：
 - 为主机控制层面代理提供控制器 IP 地址
 - 生成证书并将其分发到主机和控制器以保护控制层面通信安全
- 通过消息总线配置 ESG 和 DLR 控制虚拟机
 - 请注意，可以在未准备的主机上部署 ESG，在这种情况下，将使用 VIX 代替消息总线

控制器群集

NSX 分布式路由需要使用群集的控制器以扩展和提高可用性，从而提供以下功能：

- 支持 VXLAN 和分布式路由控制层面
- 提供 CLI 接口以获取统计信息和运行时状态
- 为每个 DLR 实例选择主控制器节点
 - 主节点从 DLR 控制虚拟机中接收路由信息，并将其分发到主机
 - 将 LIF 表发送到主机
 - 跟踪 DLR 控制虚拟机所在的主机
 - 为 VLAN LIF 选择指定的实例，并将该信息传送到主机；通过控制层面保持活动（超时为 30 秒，检测时间可以是 20-40 秒）监控 DI 主机；为主机发送更新（如果选定的 DI 主机消失）

ESXi 主机模块

NSX 路由直接利用两个用户环境代理 (UWA) 和路由内核模块，并且还依靠 VXLAN 内核模块建立 VXLAN 连接。

下面简要说明了其中的每个组件的功能：

- 控制层面代理 (netcpa) 是 TCP (SSL) 客户端，它使用控制层面协议与控制器通信。它可能会连接到多个控制器。netcpa 与消息总线客户端 (vsfwd) 通信，以便从 NSX Manager 中检索控制层面相关信息。
- netcpa 打包和部署：
 - 该代理打包为 VXLAN VIB (vSphere 安装包)
 - 在主机准备期间，NSX Manager 通过 EAM (ESX Agency Manager) 进行安装
 - 在 ESXi netcpa 上作为服务守护程序运行
 - 可以通过其启动脚本 /etc/init.d/netcpad 启动/停止/查询
 - 可以通过“网络和安全”用户界面中的“安装”->“主机准备”->“安装状态”在单个主机或整个群集上远程重新启动
- DLR 内核模块 (vdrb) 与 DVS 集成在一起以启用 L3 转发
 - 由 netcpa 配置
 - 作为 VXLAN VIB 部署的一部分进行安装
 - 通过名为“vdrPort”的特殊中继（支持 VLAN 和 VXLAN）连接到 DVS
 - 保留有关 DLR 实例的信息以及每个实例的：
 - LIF 和路由表
 - 主机本地 ARP 缓存
- netcpa、ESG 和 DLR 控制虚拟机使用消息总线客户端 (vsfwd) 与 NSX Manager 通信
 - vsfwd 通过 vpxa/hosd 从 vCenter 设置的 /UserVars/RmqIpAddress 中获取 NSX Manager 的 IP 地址，然后使用在其他 /UserVars/Rmq* 变量中存储的每个主机的凭据登录到消息总线服务器
- 在 ESXi 主机上运行的 netcpa 依靠 vsfwd 执行以下操作：
 - 从 NSX Manager 中获取主机的控制层面 SSL 私钥和证书。然后，将这些信息存储在 /etc/vmware/ssl/rui-for-netcpa.* 中
 - 从 NSX Manager 中获取控制器的 IP 地址和 SSL 指纹。然后，将这些信息存储在 /etc/vmware/netcpa/config-by-vsm.xml 中
 - 在主机上根据 NSX Manager 指令创建和删除 DLR 实例
- 打包和部署
 - 与 netcpa 相同，它是 VXLAN VIB 的一部分
 - 在 ESXi vsfwd 上作为服务守护程序运行
 - 可以通过其启动脚本 /etc/init.d/vShield-Stateful-Firewall 启动/停止/查询

- ESG 和 DLR 控制虚拟机使用到 vsfwd 的 VMCI 通道从 NSX Manager 中接收配置

DLR 控制虚拟机和 ESG

- DLR 控制虚拟机是其 DLR 实例的“路由处理器”
 - 具有每个 DLR LIF 的“占位符”或“真正 vNIC 接口”以及 IP 配置
 - 可以运行两个可用的动态路由协议（BGP 或 OSPF）之一以及/或者使用静态路由
 - 至少需要一个“上行链路”LIF 才能运行 OSPF 或 BGP
 - 通过直接连接的 (LIF) 子网、静态和动态路由计算转发表，然后通过到 netcpa 的 VMCI 链路将其发送到 DLR 实例的主控制器
 - 在活动/备用虚拟机对配置中支持 HA
- ESG 是虚拟机中的自包含路由器
 - 完全独立于 NSX DLR 路由子系统（无 NSX 控制层面集成）
 - 通常作为一个或多个 DLR 的上游网关
 - 支持多个同时运行的动态路由协议

NSX 路由控制层面 CLI

除了主机组件以外，NSX 路由还使用控制器群集和 DLR 控制虚拟机的服务，它们都是 DLR 控制层面信息来源，并具有自己的 CLI 以用于检查这些信息。

DLR 实例主控制器

每个 DLR 实例由某个控制器节点提供服务。可以使用以下 CLI 命令查看 DLR 实例的主控制器节点包含的信息：

```
nsx-controller # show control-cluster logical-routers instance 1460487509
```

LR-Id	LR-Name	Hosts[]	Edge-Connection	Service-Controller
1460487509	default+edge-1	192.168.210.57		192.168.110.201
		192.168.210.51		
		192.168.210.52		
		192.168.210.56		
		192.168.110.51		
		192.168.110.52		

```
nsx-controller # show control-cluster logical-routers interface-summary 1460487509
```

Interface	Type	Id	IP[]
570d455500000002	vxlان	5003	192.168.10.2/29
570d45550000000b	vxlان	5001	172.16.20.1/24
570d45550000000c	vxlان	5002	172.16.30.1/24
570d45550000000a	vxlان	5000	172.16.10.1/24

```
nsx-controller # show control-cluster logical-routers routes 1460487509
LR-Id      Destination      Next-Hop
1460487509  0.0.0.0/0         192.168.10.1
```

- “show control-cluster logical-routers” 命令的 “instance” 子命令显示连接到该 DLR 实例的该控制器的主机列表。在正常工作环境中，该列表包含所有群集中 DLR 所在的所有主机。
- “interface-summary” 显示控制器从 NSX Manager 中获悉的 LIF。该信息将发送到主机。
- “routes” 显示该 DLR 的控制虚拟机发送到该控制器的路由表。请注意，与 ESXi 主机上不同，该表不包含任何直接连接的子网，因为该信息是 LIF 配置提供的。

DLR 控制虚拟机

DLR 控制虚拟机具有 LIF 和路由/转发表。DLR 控制虚拟机的生命周期的主要输出是 DLR 路由表，这是 Interfaces 和 Routes 生成的。

```
edge-1-0> show ip route

Codes: 0 - OSPF derived, i - IS-IS derived, B - BGP derived,
C - connected, S - static, L1 - IS-IS level-1, L2 - IS-IS level-2,
IA - OSPF inter area, E1 - OSPF external type 1, E2 - OSPF external type 2

Total number of routes: 5

S      0.0.0.0/0      [1/1]      via 192.168.10.1
C      172.16.10.0/24 [0/0]      via 172.16.10.1
C      172.16.20.0/24 [0/0]      via 172.16.20.1
C      172.16.30.0/24 [0/0]      via 172.16.30.1
C      192.168.10.0/29 [0/0]      via 192.168.10.2

edge-1-0> show ip forwarding
Codes: C - connected, R - remote,
      > - selected route, * - FIB route
R>* 0.0.0.0/0 via 192.168.10.1, vNic_2
C>* 172.16.10.0/24 is directly connected, VDR
C>* 172.16.20.0/24 is directly connected, VDR
C>* 172.16.30.0/24 is directly connected, VDR
C>* 192.168.10.0/29 is directly connected, vNic_2
```

- 转发表的用途是显示选择哪个 DLR 接口以作为给定目标子网的输出。
 - 将为所有“内部”类型的 LIF 显示“VDR”接口。“VDR”接口是与 vNIC 不对应的伪接口。

DLR 控制虚拟机的接口可能如下所示：

```
edge-1-0> show interface
Interface VDR is up, line protocol is up
index 2 metric 1 mtu 1500 <UP,BROADCAST,RUNNING,NOARP>
HWaddr: be:3d:a1:52:90:f4
inet6 fe80::bc3d:a1ff:fe52:90f4/64
inet 172.16.10.1/24
```

```

inet 172.16.20.1/24
inet 172.16.30.1/24
proxy_arp: disabled
Auto-duplex (Full), Auto-speed (2460Mb/s)
  input packets 0, bytes 0, dropped 0, multicast packets 0
  input errors 0, length 0, overrun 0, CRC 0, frame 0, fifo 0, missed 0
  output packets 0, bytes 0, dropped 0
  output errors 0, aborted 0, carrier 0, fifo 0, heartbeat 0, window 0
  collisions 0

Interface vNic_0 is up, line protocol is up
index 3 metric 1 mtu 1500 <UP,BROADCAST,RUNNING,MULTICAST>
HWaddr: 00:50:56:8e:1c:fb
inet6 fe80::250:56ff:fe8e:1c:fb/64
inet 169.254.1.1/30
inet 10.10.10.1/24
proxy_arp: disabled
Auto-duplex (Full), Auto-speed (2460Mb/s)
  input packets 582249, bytes 37339072, dropped 49, multicast packets 0
  input errors 0, length 0, overrun 0, CRC 0, frame 0, fifo 0, missed 0
  output packets 4726382, bytes 461202852, dropped 0
  output errors 0, aborted 0, carrier 0, fifo 0, heartbeat 0, window 0
  collisions 0

Interface vNic_2 is up, line protocol is up
index 9 metric 1 mtu 1500 <UP,BROADCAST,RUNNING,MULTICAST>
HWaddr: 00:50:56:8e:ae:08
inet 192.168.10.2/29
inet6 fe80::250:56ff:fe8e:ae:08/64
proxy_arp: disabled
Auto-duplex (Full), Auto-speed (2460Mb/s)
  input packets 361446, bytes 30167226, dropped 0, multicast packets 361168
  input errors 0, length 0, overrun 0, CRC 0, frame 0, fifo 0, missed 0
  output packets 361413, bytes 30287912, dropped 0
  output errors 0, aborted 0, carrier 0, fifo 0, heartbeat 0, window 0
  collisions 0

```

感兴趣的注意事项：

- “VDI” 接口没有关联的虚拟机网卡 (vNIC)。这是单个“伪接口”，将为其配置 DLR 的所有“内部” LIF 的所有 IP 地址。
- 此示例中的 Nic_0 接口是 HA 接口。
 - 上面的输出是从启用了 HA 的部署 DLR 中提取的，并为 HA 接口分配一个 IP 地址。这显示为两个 IP 地址：169.254.1.1/30（为 HA 自动分配的）和 10.10.10.1/24（为 HA 接口手动分配的）。
 - 在 ESG 上，操作员可以手动将某个 vNIC 指定为 HA，或者保留默认设置以让系统自动从可用的“内部”接口中进行选择。具有“内部”类型是一项要求，否则，HA 将失败。
- vNic_2 接口具有“上行链路”类型；因此，它表示为“真实”vNIC。
 - 请注意，在该接口上看到的 IP 地址与 DLR 的 LIF 相同；但 DLR 控制虚拟机不会应答 LIF IP 地址（此处为 192.168.10.2/29）的 ARP 查询。可以为该 vNIC 的 MAC 地址应用一个 ARP 筛选器以进行应答。

- 在 DLR 上配置动态路由协议之前，上述观点是正确的，此时，将 IP 地址与 ARP 筛选器一起移除，并替换为在动态路由协议配置期间指定的“协议 IP”地址。
- 在 DLR 控制虚拟机上运行的动态路由协议使用该 vNIC 与其他路由器通信以发布和获悉路由。
- 在断开连接 Edge 并进行 HA 故障切换后，将从活动 Edge 路由信息库 (Routing Information Base, RIB)/转发信息库 (Forwarding Information Base, FIB) 中移除断开连接的 Edge 接口 IP 地址。但备用 Edge FIB 表或 `show ip forwarding` 命令仍显示该 IP，并且不会从 FIB 表中移除该 IP。这是预期的行为。

NSX 路由子系统故障模式和影响

本章介绍了可能会影响 NSX 路由子系统组件的典型故障场景，并简要说明了这些故障的影响。

NSX Manager

表 3-2. NSX Manager 故障模式和影响

故障模式	故障影响
到 NSX Manager 虚拟机的网络连接中断	<ul style="list-style-type: none"> ■ 所有 NSX Manager 功能完全中断，包括用于 NSX 路由/桥接的 CRUD ■ 不会丢失配置数据 ■ 数据层面或控制层面不会中断
NSX Manager 和 ESXi 主机之间的网络连接中断，或者 RabbitMQ 服务器发生故障	<ul style="list-style-type: none"> ■ 如果 DLR 控制虚拟机或 ESG 在受影响的主机上运行，这些主机上的 CRUD 操作将失败 ■ 在受影响的主机上创建和删除 DLR 实例失败 ■ 不会丢失配置数据 ■ 数据层面或控制层面不会中断 ■ 任何动态路由更新继续正常工作
NSX Manager 和控制器之间的网络连接中断	<ul style="list-style-type: none"> ■ NSX 分布式路由和桥接的创建、更新和删除操作失败 ■ 不会丢失配置数据 ■ 数据层面或控制层面不会中断
NSX Manager 虚拟机已破坏（数据存储故障）	<ul style="list-style-type: none"> ■ 所有 NSX Manager 功能完全中断，包括用于 NSX 路由/桥接的 CRUD ■ 如果 NSX Manager 还原为较旧的配置，一部分路由/桥接实例可能会变为孤立实例，从而需要手动进行清理和协调 ■ 数据层面或控制层面不会中断，除非需要进行协调

控制器群集

表 3-3. NSX Controller 故障模式和影响

故障模式	故障影响
控制器群集与 ESXi 主机之间的网络连接中断	<ul style="list-style-type: none"> ■ DLR 控制层面功能（创建、更新和删除路由，包括动态路由）完全中断 ■ DLR 管理层面功能（在主机上创建、更新和删除 LIF）中断 ■ 将影响 VXLAN 转发，这可能会导致端到端 (L2+L3) 转发过程也会失败 ■ 根据最后已知状态，数据层面继续正常工作
一个或两个控制器与 ESXi 主机之间的连接中断	<ul style="list-style-type: none"> ■ 如果受影响的控制器仍然可以访问群集中的其他控制器，该控制器控制的任何 DLR 实例将受到上面所述的相同影响。其他控制器不会自动接管
一个控制器与其他控制器之间的网络连接中断（或完全中断）	<ul style="list-style-type: none"> ■ 两个剩下的控制器接管隔离的控制器处理的 VXLAN 和 DLR ■ 受影响的控制器进入只读模式，丢弃到主机的会话并拒绝新的会话
控制器之间的连接中断	<ul style="list-style-type: none"> ■ 所有控制器将进入只读模式，关闭到主机的连接并拒绝新的连接 ■ DLR 的所有 LIF 和路由（包括动态路由）的创建、更新和删除操作失败 ■ NSX 路由配置 (LIF) 可能在 NSX Manager 和控制器群集之间不同步，从而需要手动干预以重新同步 ■ 主机将继续在最后已知控制层面状态下运行
一个控制器虚拟机丢失	<ul style="list-style-type: none"> ■ 控制器群集缺少冗余 ■ 管理/控制层面继续正常运行
两个控制器虚拟机丢失	<ul style="list-style-type: none"> ■ 其余控制器将进入只读模式；受到的影响与控制器之间的连接中断时相同（如上所述）。可能需要手动恢复群集

主机模块

netcpa 依靠主机 SSL 密钥和证书以及 SSL 指纹与控制器建立安全通信。这些信息是通过消息总线（由 vsfwd 提供）从 NSX Manager 中获取的。

如果证书交换过程失败，netcpa 将无法成功连接到控制器。

注意：本节不涉及内核模块故障，因为这种故障的影响非常严重 (PSOD) 并且很少会发生。

表 3-4. 主机模块故障模式和影响

故障模式	故障影响
vsfwd 使用用户名/密码身份验证访问消息总线服务器（可能会过期）	<ul style="list-style-type: none"> ■ 如果新准备的 ESXi 主机上的 vsfwd 在两小时内无法访问 NSX Manager，在安装期间提供的临时登录名/密码将过期，并且该主机上的消息总线无法运行
消息总线客户端 (vsfwd) 的故障影响取决于时间。	

表 3-4. 主机模块故障模式和影响（续）

故障模式	故障影响
如果它在 NSX 控制层面的其他部分进入稳定运行状态之前发生故障	<ul style="list-style-type: none"> ■ 主机上的分布式路由停止工作，因为主机无法与控制器通信 ■ 主机无法从 NSX Manager 中获悉 DLR 实例
如果它在主机进入稳定状态后发生故障	<ul style="list-style-type: none"> ■ 在主机上运行的 ESG 和 DLR 控制虚拟机无法接收配置更新 ■ 主机未获悉新的 DLR，并且无法删除现有的 DLR ■ 根据主机在发生故障时具有的配置，主机数据路径将继续运行

表 3-5. netcpa 故障模式和影响

故障模式	故障影响
控制层面代理 (netcpa) 的故障影响取决于时间。	
如果它在 NSX 数据路径内核模块进入稳定运行状态之前发生故障	<ul style="list-style-type: none"> ■ 主机上的分布式路由停止工作
如果它在主机进入稳定状态后发生故障	<ul style="list-style-type: none"> ■ 在主机上运行的 DLR 控制虚拟机无法将其转发表更新发送到控制器 ■ 分布式路由数据路径不会从控制器中收到任何 LIF 或路由更新，但根据故障前具有的状态继续运行

DLR 控制虚拟机

表 3-6. DLR 控制虚拟机故障模式和影响

故障模式	故障影响
DLR 控制虚拟机丢失或关闭电源	<ul style="list-style-type: none"> ■ 该 DLR 的 LIF 和路由的创建、更新和删除操作失败 ■ 不会将任何动态路由更新发送到主机（包括撤消通过现在断开的邻接收到的前缀）
DLR 控制虚拟机与 NSX Manager 和控制器之间的连接中断	<ul style="list-style-type: none"> ■ 影响与上面相同，所不同的是，如果 DLR 控制虚拟机及其路由邻接仍然启动，与以前获悉的前缀之间的流量将不会受到影响
DLR 控制虚拟机与 NSX Manager 之间的连接中断	<ul style="list-style-type: none"> ■ 该 DLR 的 LIF 和路由的 NSX Manager 创建、更新和删除操作失败，并且不会重试 ■ 动态路由更新继续进行传播
DLR 控制虚拟机与控制器之间的连接中断	<ul style="list-style-type: none"> ■ 该 DLR 的任何路由更改（静态或动态）不会传播到主机

与路由有关的 NSX 日志

最佳做法是，配置 NSX 的所有组件以将其日志发送到集中式收集器，以便在一个地方检查这些日志。

如有必要，您可以更改 NSX 组件的日志级别。有关详细信息，请参见 NSX 日志记录和系统事件中的“设置 NSX 组件的日志记录级别”主题。

NSX Manager 日志

- NSX Manager CLI 中的 `show log`
- 通过 NSX Manager UI 收集的技术支持日志包

NSX Manager Virtual Appliance Management



NSX Manager 日志包含与管理层面有关的信息，其中包括创建、读取、更新和删除 (CRUD) 操作。

控制器日志

控制器包含多个模块，很多模块具有自己的日志文件。可以使用 `show log <log file> [filtered-by <string>]` 命令访问控制器日志。与路由有关的日志文件如下所示：

- `cloudnet/cloudnet_java-vnet-controller.<start-time-stamp>.log`：该日志管理配置和内部 API 服务器。
- `cloudnet/cloudnet.nsx-controller.log`：这是控制器主进程日志。
- `cloudnet/cloudnet_cpp.log.nsx-controller.log`：该日志管理群集和引导。
- `cloudnet/cloudnet_cpp.log.ERROR`：如果出现任何错误，则包含该文件。

控制器日志非常详细，在大多数情况下，只有在请求 VMware 工程团队帮助解决更困难的问题时，才需要使用这些日志。

除了 `show log CLI` 以外，还可以使用 `watch log <logfile> [filtered-by <string>]` 命令在更新各个日志文件时实时观察这些文件。

这些日志包含在控制器支持包中，可以在 NSX UI 中选择一个控制器节点并单击 **下载技术支持日志 (Download tech support logs)** 图标以生成并下载该支持包。

ESXi 主机日志

在 ESXi 主机上运行的 NSX 组件写入几个日志文件：

- VMkernel 日志： `/var/log/vmkernel.log`
- 控制层面代理日志： `/var/log/netcpa.log`
- 消息总线客户端日志： `/var/log/vsfwd.log`

也可以将这些日志作为从 vCenter Server 中生成的虚拟机支持包的一部分进行收集。仅具有 `root` 特权的用户或用户组可以访问这些文件。

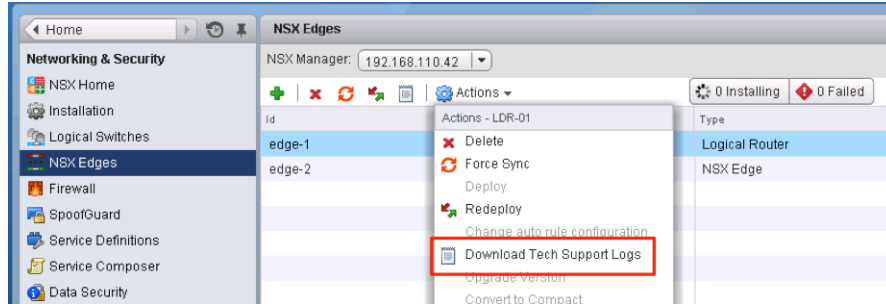
ESG/DLR 控制虚拟机日志

可以使用两种方法访问 ESG 和 DLR 控制虚拟机上的日志文件：使用 CLI 显示这些文件，或者使用 CLI 或 UI 下载技术支持包。

用于显示日志的 CLI 命令是 `show log [follow | reverse]`。

要下载技术支持包，请执行以下操作：

- 从 CLI 中，进入 **enable** 模式，然后运行 **export tech-support <[scp | ftp]> <URI>** 命令。
- 从 vSphere Web Client 中，在**操作 (Actions)**菜单中选择**下载技术支持日志 (Download Tech Support Logs)**选项。



其他有用的文件及其位置

虽然严格来说很多文件并不是日志，但它们可以帮助了解和解决 NSX 路由问题。

- 控制层面代理配置 `/etc/vmware/netcpa/config-by-vsm.xml` 包含有关以下组件的信息：
 - 控制器、IP 地址、TCP 端口、证书指纹、SSL 启用/禁用
 - 启用了 VXLAN 的 DVS 上的 dvUplink（绑定策略、名称、UUID）
 - 主机了解的 DLR 实例（DLR ID、名称）
- 控制层面代理配置 `/etc/vmware/netcpa/netcpa.xml` 包含各种 netcpa 配置选项，包括日志记录级别（默认为 **info**）。
- 控制层面证书文件：`/etc/vmware/ssl/rui-for-netcpa.*`
 - 两个文件：主机证书和主机私钥
 - 用于验证到控制器的主机连接

所有这些文件都是控制层面代理使用从 NSX Manager 收到的信息（通过 vsfwd 提供的消息总线连接）创建的。

常见故障情况和修复

最常见的故障情况分为两类。

它们是配置和控制层面问题。也可能是管理层面问题，但并不常见。

配置问题和修复

表 3-7. 常见配置问题和影响 中介绍了常见配置问题及其影响。

表 3-7. 常见配置问题和影响

问题	影响
动态路由的协议和转发 IP 地址是相反的	没有建立动态协议邻接
传输区域与 DVS 边界不对齐	分布式路由在一部分 ESXi 主机上无法正常工作（在传输区域中缺少这些主机）
动态路由协议配置不匹配（计时器、MTU、BGP ASN、密码、接口到 OSPF 区域的映射）	没有建立动态协议邻接
为 DLR HA 接口分配了 IP 地址并允许重新分发连接的路由	DLR 控制虚拟机可能会吸收 HA 接口子网的流量并产生流量黑洞

要解决这些问题，请查看配置并根据需要进行更正。

如果需要，请使用 `debug ip ospf` 或 `debug ip bgp` CLI 命令，并观察 DLR 控制虚拟机或 ESG 控制台（而不是通过 SSH 会话）上的日志以检测协议配置问题。

控制层面问题和修复

发现的控制层面问题通常是以下问题造成的：

- 主机控制层面代理 (netcpa) 无法通过 vsfwd 提供的消息总线通道连接到 NSX Manager
- 控制器群集在处理 DLR/VXLAN 实例的主角色时出现问题

通常，可以重新启动某个 NSX Controller（控制器的 CLI 上的 `restart controller`）以解决与处理主角色有关的控制器群集问题。

有关解决控制层面问题的详细信息，请参见 <http://kb.vmware.com/kb/2125767>。

收集故障排除数据

本节简要说明了通常用于排除 NSX 路由故障的 CLI 命令。

NSX Manager

从 NSX 6.2 开始，以前从 NSX Controller 和其他 NSX 组件中运行以解决 NSX 路由问题的命令现在从 NSX Manager 中直接运行。

- DLR 实例列表
- 每个 DLR 实例的 LIF 列表
- 每个 DLR 实例的路由列表
- 每个 DLR 桥接实例的 MAC 地址列表
- 接口
- 路由和转发表
- 动态路由协议（OSPF 或 BGP）状态
- NSX Manager 发送到 DLR 控制虚拟机或 ESG 的配置

DLR 控制虚拟机和 ESG

DLR 控制虚拟机和 ESG 提供在其接口上捕获数据包的功能。数据包捕获可以帮助解决路由协议问题。

- 1 运行 `show interfaces` 以列出接口名称。
- 2 运行 `debug packet [display | capture] interface <interface name>`。
 - 如果使用捕获，数据包将保存在 `.pcap` 文件中。
- 3 运行 `debug show files` 以列出保存的捕获文件。
- 4 运行 `debug copy [scp | ftp] ...` 以下载捕获包进行脱机分析。

```
d1r-01-0> debug packet capture interface vNic_2
tcpdump: listening on vNic_2, link-type EN10MB (Ethernet), capture size 65535 bytes
43 packets captured
48 packets received by filter
0 packets dropped by kernel
```

```
d1r-01-0> debug show files
total 4.0K
-rw----- 1 3.6K Mar 30 23:49 tcpdump_vNic_2.0
```

```
d1r-01-0> debug copy
  scp  use scp to copy
  ftp  use ftp to copy
```

```
d1r-01-0> debug copy scp
URL  user@<remote-host>:<path-to>
```

`debug packet` 命令在后台使用 `tcpdump` 并且可以接受筛选修饰符，这些修饰符的格式类似于 UNIX 上的 `tcpdump` 筛选修饰符。唯一需要注意的是，将筛选表达式中的任何空格替换为下划线（“_”）。

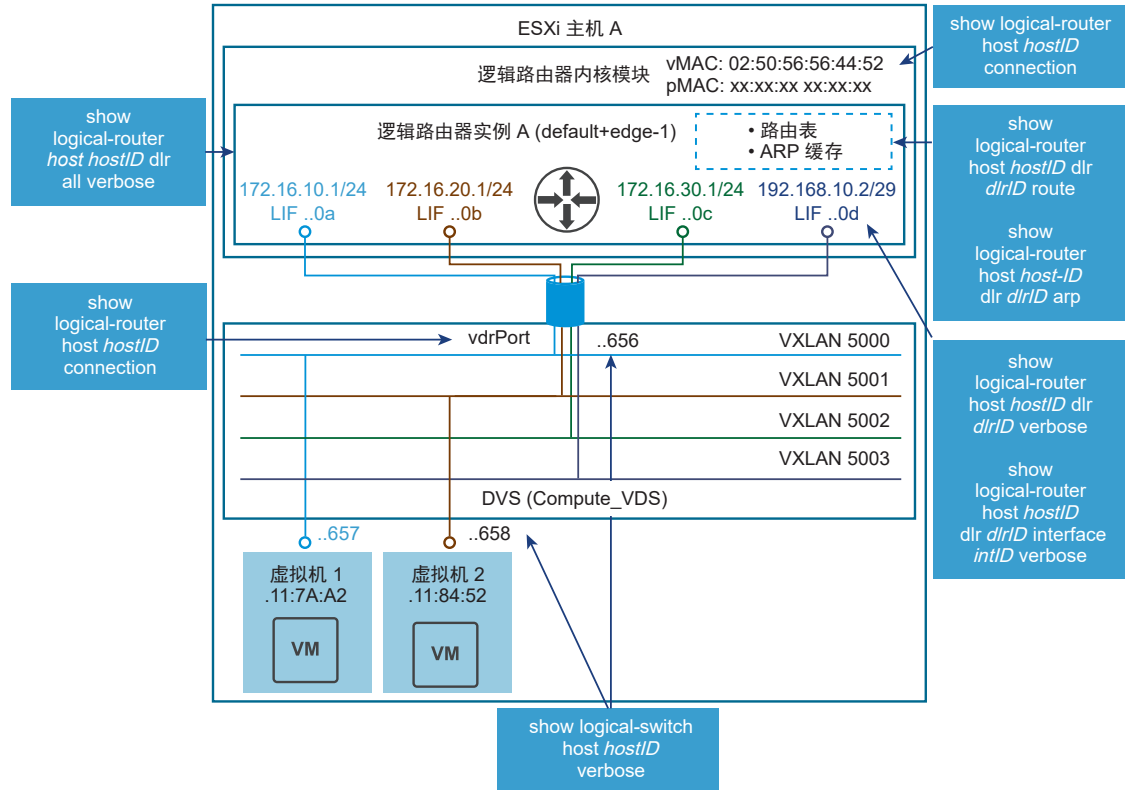
例如，以下命令显示通过 `vNic_0` 的所有流量（SSH 除外），以避免查看属于交互式会话本身的流量。

```
plr-02-0> debug packet display interface vNic_0 port_not_22
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on vNic_0, link-type EN10MB (Ethernet), capture size 65535 bytes
04:10:48.197768 IP 192.168.101.3.179 > 192.168.101.2.25698: Flags [P.], seq 4191398894:4191398913,
ack 2824012766, win 913, length 19: BGP, length: 19
04:10:48.199230 IP 192.168.101.2.25698 > 192.168.101.3.179: Flags [.] , ack 19, win 2623, length 0
04:10:48.299804 IP 192.168.101.2.25698 > 192.168.101.3.179: Flags [P.], seq 1:20, ack 19, win 2623,
length 19: BGP, length: 19
04:10:48.299849 IP 192.168.101.3.179 > 192.168.101.2.25698: Flags [.] , ack 20, win 913, length 0
04:10:49.205347 IP 192.168.101.3.179 > 192.168.101.2.25698: Flags [P.], seq 19:38, ack 20, win 913,
length 19: BGP, length: 19
```

ESXi 主机

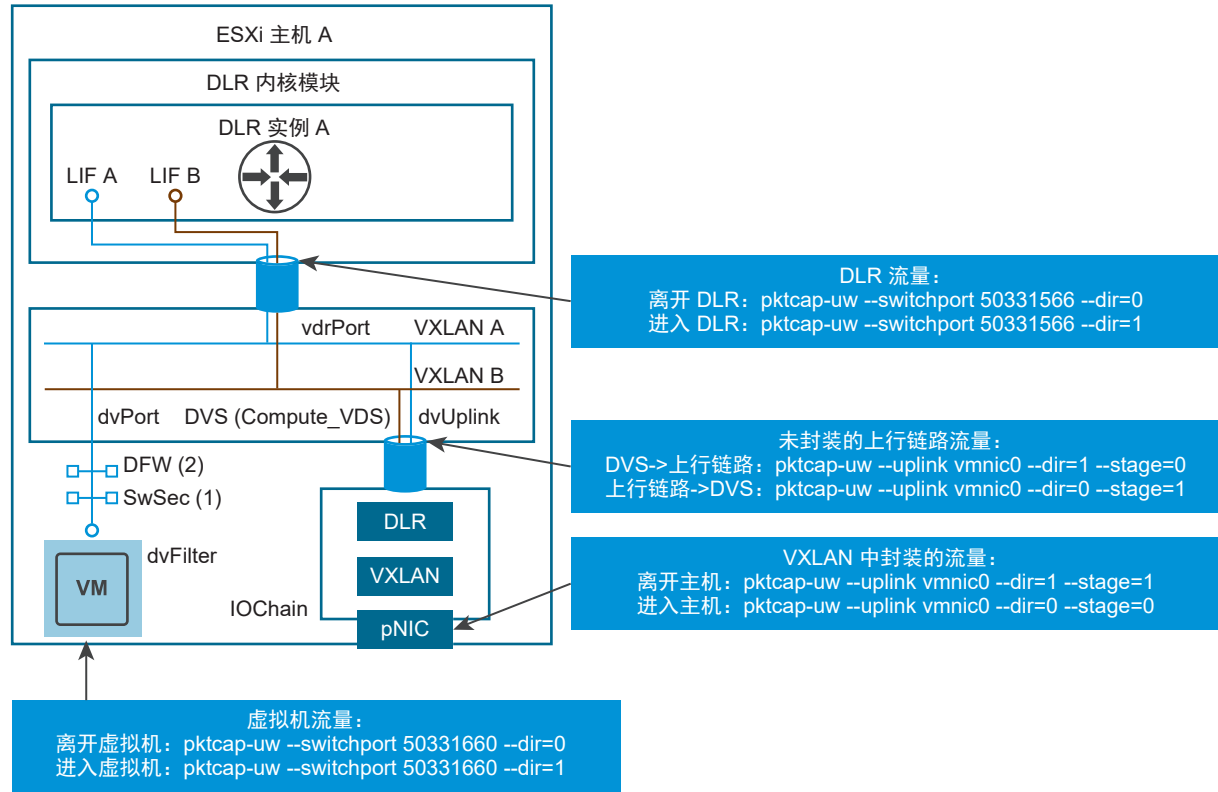
主机与 NSX 路由密切相关。图 3-14. 与解决 NSX 路由问题有关的主机组件 直观地显示了路由子系统涉及的组件以及用于显示相关信息的 NSX Manager CLI 命令：

图 3-14. 与解决 NSX 路由问题有关的主机组件



在数据路径中捕获的数据包可以帮助找出在数据包转发的各个阶段出现的问题。图 3-15. 捕获点和相关的 CLI 命令 涵盖了主要捕获点和使用的相应 CLI 命令。

图 3-15. 捕获点和相关的 CLI 命令



NSX Edge 故障排除

4

本主题提供了有助于了解 VMware NSX Edge 和进行故障排除的信息。

要解决 NSX Edge 设备问题，请验证下面的每个故障排除步骤是否适用于您的环境。每个步骤提供了相应说明或指向文档的链接，以消除可能的根源并在必要时采取纠正措施。这些步骤按最适当的顺序进行排列，以查找问题并确定相应的解决方案。不要跳过某个步骤。

请参阅当前版本的发行说明以查看是否解决了该问题。

确保在安装 VMware NSX Edge 时满足最低系统要求。请参见 NSX 安装指南。

安装和升级问题

- 验证遇到的问题是否与“Would Block”问题无关。有关详细信息，请参见 <https://kb.vmware.com/kb/2107951>。
- 如果升级或重新部署成功，但 Edge 接口没有连接，请验证后端 2 层交换机上的连接。请参见 <https://kb.vmware.com/kb/2135285>。
- 如果 Edge 部署或升级失败并出现以下错误：

```
/sbin/ifconfig vNic_1 up failed : SIOCSIFFLAGS: Invalid argument
```

或

- 如果部署或升级成功，但在 Edge 接口上没有连接：
- 运行 `show interface` 命令以及 Edge 支持日志将显示类似下面的条目：

```
vNic_0: <NO-CARRIER,BROADCAST,MULTICAST,UP> mtu 1500 qdisc mq state DOWN qlen 1000
link/ether 00:50:56:32:05:03 brd ff:ff:ff:ff:ff:ff
inet 21.12.227.244/23 scope global vNic_0
inet6 fe80::250:56ff:fe32:503/64 scope link tentative dadfailed
valid_lft forever preferred_lft forever
```

在这两种情况下，主机交换机未就绪或出现某些问题。要解决该问题，请调查主机交换机。

配置问题

- 收集 NSX Edge 诊断信息。请参见 <https://kb.vmware.com/kb/2079380>。

搜索字符串 `vse_die` 以筛选 NSX Edge 日志。包含该字符串的日志条目可能会提供有关配置错误的信息。

较高的 CPU 占用率

如果 NSX Edge 上的 CPU 占用率较高，请在 ESXi 主机上使用 `esxtop` 命令验证设备的性能。请参阅以下知识库文章：

- <https://kb.vmware.com/kb/1008205>
- <https://kb.vmware.com/kb/1008014>
- <https://kb.vmware.com/kb/1010071>
- <https://kb.vmware.com/kb/2096171>

另请参见 <https://communities.vmware.com/docs/DOC-9279>。

较高的 `ksoftirqd` 进程值表示传入数据包率较高。检查是否在数据路径上启用日志记录，例如，为防火墙规则启用。运行 `show log follow` 命令以确定是否记录了大量的日志命中数。

显示数据包丢弃统计信息

从 NSX for vSphere 6.2.3 开始，您可以使用 `show packet drops` 命令显示以下内容的数据包丢弃统计信息：

- 接口
- 驱动程序
- L2
- L3
- 防火墙

要运行该命令，请登录到 NSX Edge CLI 并进入基本模式。有关详细信息，请参见《NSX 命令行界面参考》。例如：

```
show packet drops
```

```
vShield Edge Packet Drop Stats:
```

```
Driver Errors
```

```
=====
```

	TX	TX	RX	RX	RX
Interface	Dropped	Error	Ring Full	Dropped	Error Out Of Buf
vNic_0	0	0	0	0	0
vNic_1	0	0	0	0	0
vNic_2	0	0	0	0	2

```

vNic_3    0      0      0      0      0      0
vNic_4    0      0      0      0      0      0
vNic_5    0      0      0      0      0      0

```

Interface Drops

```
=====
```

```
Interface RX Dropped TX Dropped
```

```

vNic_0          4      0
vNic_1        2710      0
vNic_2          0      0
vNic_3          2      0
vNic_4          2      0
vNic_5          2      0

```

L2 RX Errors

```
=====
```

```
Interface length crc frame fifo missed
```

```

vNic_0          0  0      0  0      0
vNic_1          0  0      0  0      0
vNic_2          0  0      0  0      0
vNic_3          0  0      0  0      0
vNic_4          0  0      0  0      0
vNic_5          0  0      0  0      0

```

L2 TX Errors

```
=====
```

```
Interface aborted fifo window heartbeat
```

```

vNic_0          0  0      0      0
vNic_1          0  0      0      0
vNic_2          0  0      0      0
vNic_3          0  0      0      0
vNic_4          0  0      0      0
vNic_5          0  0      0      0

```

L3 Errors

```
=====
```

IP:

```

ReasmFails : 0
InHdrErrors : 0
InDiscards : 0
FragFails : 0
InAddrErrors : 0
OutDiscards : 0
OutNoRoutes : 0
ReasmTimeout : 0

```

ICMP:

```

InTimeExcds : 0
InErrors : 227
OutTimeExcds : 0
OutDestUnreaches : 152
OutParmProbs : 0
InSrcQuenches : 0
InRedirects : 0
OutSrcQuenches : 0
InDestUnreaches : 151

```

```

OutErrors : 0
InParmProbs : 0

Firewall Drop Counters
=====

Ipv4 Rules
=====
Chain - INPUT
rid pkts bytes target prot opt in out source destination
0 119 30517 DROP all -- * * 0.0.0.0/0 0.0.0.0/0 state INVALID
0 0 0 DROP all -- * * 0.0.0.0/0 0.0.0.0/0
Chain - POSTROUTING
rid pkts bytes target prot opt in out source destination
0 101 4040 DROP all -- * * 0.0.0.0/0 0.0.0.0/0 state INVALID
0 0 0 DROP all -- * * 0.0.0.0/0 0.0.0.0/0

Ipv6 Rules
=====
Chain - INPUT
rid pkts bytes target prot opt in out source destination
0 0 0 DROP all * * ::/0 ::/0 state INVALID
0 0 0 DROP all * * ::/0 ::/0
Chain - POSTROUTING
rid pkts bytes target prot opt in out source destination
0 0 0 DROP all * * ::/0 ::/0 state INVALID
0 0 0 DROP all * * ::/0 ::/0

```

管理 NSX Edge 时的预期行为

- 在 vSphere Web Client 中，在 NSX Edge 上配置 L2 VPN 以及添加、移除或修改站点配置详细信息 (**Site Configuration Details**) 时，该操作将导致断开并重新连接所有现有的连接。这是预期的行为。
- NSX Edge 是一个虚拟机 (VM) 并包含一些存储在存储设备上的文件。关键文件包括配置文件、虚拟磁盘文件、NVRAM 设置文件、交换文件和日志文件。根据应用的虚拟机存储配置文件或手动放置方式，虚拟机配置文件、虚拟磁盘文件和交换文件可能放置在相同的位置，也可能放置在不同数据存储上的不同位置中。如果虚拟机文件位于不同的位置，NSX Manager 显示并使用包含 VMX 文件的数据存储进行虚拟机部署。在重新部署或升级操作期间，NSX Manager 在配置的数据存储或托管 VMX 文件的实时数据存储上部署 NSX Edge 虚拟机。*数据存储名称* 和 *数据存储 ID*（托管虚拟机的 VMX 文件）是作为 Appliance 参数的一部分返回的，并显示在 UI 上或作为 REST API 响应提供。您必须参阅 vCenter Server 以了解每个 NSX Manager 虚拟机文件的确切布局以及放置这些文件的一个或多个数据存储的详细信息。有关详细信息，请参阅以下文档：
 - *vSphere 虚拟机管理*。
 - *vSphere 资源管理*。
 - *vCenter Server 和主机管理*。

本章讨论了以下主题：

- [Edge 防火墙数据包丢弃问题](#)

- [Edge 路由连接问题](#)
- [NSX Manager 与 Edge 通信问题](#)
- [消息总线调试](#)
- [Edge 诊断和恢复](#)

Edge 防火墙数据包丢弃问题

显示防火墙数据包丢弃统计信息

从 NSX for vSphere 6.2.3 开始，您可以使用 `show packet drops` 命令显示防火墙的数据包丢弃统计信息。

要运行该命令，请登录到 **NSX Edge CLI** 并进入基本模式。有关详细信息，请参见《**NSX 命令行界面参考**》。例如：

```
show packet drops

vShield Edge Packet Drop Stats:

Firewall Drop Counters
=====

Ipv4 Rules
=====
Chain - INPUT
rid pkts bytes target prot opt in out source destination state
0 119 30517 DROP all -- * * 0.0.0.0/0 0.0.0.0/0 state INVALID
0 0 0 DROP all -- * * 0.0.0.0/0 0.0.0.0/0
Chain - POSTROUTING
rid pkts bytes target prot opt in out source destination state
0 101 4040 DROP all -- * * 0.0.0.0/0 0.0.0.0/0 state INVALID
0 0 0 DROP all -- * * 0.0.0.0/0 0.0.0.0/0

Ipv6 Rules
=====
Chain - INPUT
rid pkts bytes target prot opt in out source destination state
0 0 0 DROP all * * ::/0 ::/0 state INVALID
0 0 0 DROP all * * ::/0 ::/0
Chain - POSTROUTING
rid pkts bytes target prot opt in out source destination state
0 0 0 DROP all * * ::/0 ::/0 state INVALID
0 0 0 DROP all * * ::/0 ::/0
```

Edge 数据包防火墙问题

要运行一个命令，请登录到 NSX Edge CLI 并进入基本模式。有关详细信息，请参见《NSX 命令行界面参考》。

- 1 使用 `show firewall` 命令检查防火墙规则表。`usr_rules` 表显示配置的规则。

```
nsxedge> show firewall
Chain PREROUTING (policy ACCEPT 3146M packets, 4098G bytes)
rid  pkts bytes target    prot opt in     out     source      destination

Chain INPUT (policy ACCEPT 0 packets, 0 bytes)
rid  pkts bytes target    prot opt in     out     source      destination
0    78903 16M ACCEPT    all  --  lo      *        0.0.0.0/0    0.0.0.0/0
0      0 0 DROP      all  --  *       *        0.0.0.0/0    0.0.0.0/0
state INVALID
0    140K 9558K block_in  all  --  *       *        0.0.0.0/0    0.0.0.0/0
0    23789 1184K ACCEPT    all  --  *       *        0.0.0.0/0    0.0.0.0/0
state RELATED,ESTABLISHED
0    116K 8374K usr_rules all  --  *       *        0.0.0.0/0    0.0.0.0/0
0      0 0 DROP      all  --  *       *        0.0.0.0/0    0.0.0.0/0

Chain FORWARD (policy ACCEPT 3146M packets, 4098G bytes)
rid  pkts bytes target    prot opt in     out     source      destination

Chain OUTPUT (policy ACCEPT 173K packets, 22M bytes)
rid  pkts bytes target    prot opt in     out     source      destination

Chain POSTROUTING (policy ACCEPT 0 packets, 0 bytes)
rid  pkts bytes target    prot opt in     out     source      destination
0    78903 16M ACCEPT    all  --  *       lo       0.0.0.0/0    0.0.0.0/0
0    679K 41M DROP      all  --  *       *        0.0.0.0/0    0.0.0.0/0
state INVALID
0    3146M 4098G block_out all  --  *       *        0.0.0.0/0    0.0.0.0/0
0      0 0 ACCEPT    all  --  *       *        0.0.0.0/0    0.0.0.0/0
PHYSDEV match --physdev-in tap0 --physdev-out vNic_+
0      0 0 ACCEPT    all  --  *       *        0.0.0.0/0    0.0.0.0/0
PHYSDEV match --physdev-in vNic_+ --physdev-out tap0
0      0 0 ACCEPT    all  --  *       *        0.0.0.0/0    0.0.0.0/0
PHYSDEV match --physdev-in na+ --physdev-out vNic_+
0      0 0 ACCEPT    all  --  *       *        0.0.0.0/0    0.0.0.0/0
PHYSDEV match --physdev-in vNic_+ --physdev-out na+
0    3145M 4098G ACCEPT    all  --  *       *        0.0.0.0/0    0.0.0.0/0
state RELATED,ESTABLISHED
0    221K 13M usr_rules all  --  *       *        0.0.0.0/0    0.0.0.0/0
0      0 0 DROP      all  --  *       *        0.0.0.0/0    0.0.0.0/0

Chain block_in (1 references)
rid  pkts bytes target    prot opt in     out     source      destination

Chain block_out (1 references)
rid  pkts bytes target    prot opt in     out     source      destination

Chain usr_rules (2 references)
rid  pkts bytes target    prot opt in     out     source      destination
```

```

131074 70104 5086K ACCEPT    all  --  *      *      0.0.0.0/0      0.0.0.0/0
match-set 0_131074-os-v4-1 src
131075 116K 8370K ACCEPT    all  --  *      *      0.0.0.0/0      0.0.0.0/0
match-set 1_131075-ov-v4-1 dst
131073 151K 7844K ACCEPT    all  --  *      *      0.0.0.0/0      0.0.0.0/0

```

在 `show firewall` 命令的 `POST_ROUTING` 部分中检查 `DROP invalid` 规则的递增值。典型的原因包括：

- 不对称路由问题
- 基于 **TCP** 的应用程序处于不活动状态的时间超过一小时。如果出现不活动超时问题，并且应用程序处于空闲状态的时间太长，请使用 **REST API** 增加不活动超时设置。请参见 <https://kb.vmware.com/kb/2101275>

2 收集 `show ipset` 命令输出。

```

nsxedge> show ipset
Name: 0_131074-os-v4-1
Type: bitmap:if (Interface Match)
Revision: 3
Header: range 0-64000
Size in memory: 8116
References: 1
Number of entries: 1
Members:
vse (vShield Edge Device)

Name: 0_131074-os-v6-1
Type: bitmap:if (Interface Match)
Revision: 3
Header: range 0-64000
Size in memory: 8116
References: 1
Number of entries: 1
Members:
vse (vShield Edge Device)

Name: 1_131075-ov-v4-1
Type: hash:oservice (Match un-translated Ports)
Revision: 2
Header: family inet hashsize 64 maxelem 65536
Size in memory: 704
References: 1
Number of entries: 2
Members:
Proto=6, DestPort=179, SrcPort=Any      (encoded: 0.6.0.179,0.6.0.0/16)
Proto=89, DestPort=Any, SrcPort=Any      (encoded: 0.89.0.0/16,0.89.0.0/16)

Name: 1_131075-ov-v6-1
Type: hash:oservice (Match un-translated Ports)
Revision: 2
Header: family inet hashsize 64 maxelem 65536

```

```

Size in memory: 704
References: 1
Number of entries: 2
Members:
Proto=89, DestPort=Any, SrcPort=Any    (encoded: 0.89.0.0/16,0.89.0.0/16)
Proto=6, DestPort=179, SrcPort=Any    (encoded: 0.6.0.179,0.6.0.0/16)

```

- 3 使用 REST API 或 Edge 用户界面在特定防火墙规则上启用日志记录，然后使用 `show log follow` 命令监视日志。

如果看不到日志，请使用以下 REST API 在 DROP Invalid 规则上启用日志记录。

```

URL : https://NSX_Manager_IP/api/4.0/edges/{edgeId}/firewall/config/global

PUT Method
Input representation
<globalConfig>    <!-- Optional -->
<tcpPickOngoingConnections>false</tcpPickOngoingConnections>    <!-- Optional. Defaults to false -->
>
<tcpAllowOutOfWindowPackets>false</tcpAllowOutOfWindowPackets>    <!-- Optional. Defaults to false -->
<tcpSendResetForClosedVsePorts>true</tcpSendResetForClosedVsePorts>    <!-- Optional. Defaults to true -->
<dropInvalidTraffic>true</dropInvalidTraffic>    <!-- Optional. Defaults to true -->
<logInvalidTraffic>true</logInvalidTraffic>    <!-- Optional. Defaults to false -->
<tcpTimeoutOpen>30</tcpTimeoutOpen>    <!-- Optional. Defaults to 30 -->
<tcpTimeoutEstablished>3600</tcpTimeoutEstablished>    <!-- Optional. Defaults to 3600 -->
<tcpTimeoutClose>30</tcpTimeoutClose>    <!-- Optional. Defaults to 30 -->
<udpTimeout>60</udpTimeout>    <!-- Optional. Defaults to 60 -->
<icmpTimeout>10</icmpTimeout>    <!-- Optional. Defaults to 10 -->
<icmp6Timeout>10</icmp6Timeout>    <!-- Optional. Defaults to 10 -->
<ipGenericTimeout>120</ipGenericTimeout>    <!-- Optional. Defaults to 120 -->
</globalConfig>
Output representation
No payload

```

使用 `show log follow` 命令查找类似下面的日志：

```

2016-04-18T20:53:31+00:00 edge-0 kernel: nf_ct_tcp: invalid TCP flag combination IN= OUT=
SRC=172.16.1.4 DST=192.168.1.4 LEN=40 TOS=0x00 PREC=0x00 TTL=64 ID=43343 PROTO=TCP
SPT=5050 DPT=80 SEQ=0 ACK=1572141176 WINDOW=512 RES=0x00 URG PSH FIN URG=0
2016-04-18T20:53:31+00:00 edge-0 kernel: INVALID IN= OUT=vNic_1 SRC=172.16.1.4
DST=192.168.1.4 LEN=40 TOS=0x00 PREC=0x00 TTL=63 ID=43343 PROTO=TCP SPT=5050 DPT=80
WINDOW=512 RES=0x00 URG PSH FIN URG=0

```

- 4 使用 `show flowtable rule_id` 命令在 Edge 防火墙状态表中查找匹配的连接。

```

nsxedge> show flowtable
1: tcp 6 21554 ESTABLISHED src=192.168.110.10 dst=192.168.5.3 sport=25981
d port=22 pkts=52 bytes=5432 src=192.168.5.3 dst=192.168.110.10 sport=22 dport=259

```



```
81 pkts=44 bytes=7201 [ASSURED] mark=0 rid=131073 use=1
2: tcp 6 21595 ESTABLISHED src=127.0.0.1 dst=127.0.0.1 sport=53194
dport=10 001 pkts=33334 bytes=11284650 src=127.0.0.1 dst=127.0.0.1 sport=10001 dport=5319
4 pkts=33324 bytes=1394146 [ASSURED] mark=0 rid=0 use=1
```

使用 `show flowstats` 命令将活动连接数与最大允许连接数进行比较：

```
nsxedge> show flowstats
Total Flow Capacity: 65536
Current Statistics :
cpu=0 searched=3280373 found=3034890571 new=52678 invalid=659946 ignore=77605
delete=52667 delete_list=49778 insert=49789 insert_failed=0 drop=0 early_drop=0
error=0 search_restart=0
```

- 5 使用 `show log follow` 命令检查 Edge 日志并查找任何 ALG 丢弃问题。搜索类似于 `tftp_alg`、`msrpc_alg` 或 `oracle_tns` 的字符串。有关其他信息，请参见：

- <https://kb.vmware.com/kb/2126674>
- <https://kb.vmware.com/kb/2137751>

Edge 路由连接问题

- 1 使用 `ping <destination_IP_address>` 命令从客户端中启动控制的流量。
- 2 在两个接口上同时捕获流量，将输出写入到一个文件中，然后使用 SCP 导出该文件。

例如：

使用以下命令在输入接口上捕获流量：

```
debug packet display interface vNic_0 - n_src_host_1.1.1.1
```

使用以下命令在输出接口上捕获流量：

```
debug packet display interface vNic_1 - n_src_host_1.1.1.1
```

对于同时的数据包捕获，请在 ESXi 中使用 ESXi 数据包捕获实用程序 `pktcap-uw` 工具。请参见 <https://kb.vmware.com/kb/2051814>。

如果数据包丢弃一贯出现，请查找与以下内容相关的配置错误：

- IP 地址和路由
- 防火墙规则或 NAT 规则
- 不对称路由
- RP 筛选器检查
 - a 使用 `show interface` 命令检查接口 IP/子网。

- b 如果在数据层面中缺少路由，请运行以下命令：
 - `show ip route`
 - `show ip route static`
 - `show ip route bgp`
 - `show ip route ospf`
- c 运行 `show ip forwarding` 命令以在路由表中查找所需的路由。
- d 如果具有多个路径，请运行 `show rpfilter` 命令。

```
nsxedge> show rpfilter
net.ipv4.conf.VDR.rp_filter = 0
net.ipv4.conf.all.rp_filter = 0
net.ipv4.conf.br-sub.rp_filter = 1
net.ipv4.conf.default.rp_filter = 1
net.ipv4.conf.lo.rp_filter = 0
net.ipv4.conf.vNic_0.rp_filter = 1
net.ipv4.conf.vNic_1.rp_filter = 1
net.ipv4.conf.vNic_2.rp_filter = 1
net.ipv4.conf.vNic_3.rp_filter = 1
net.ipv4.conf.vNic_4.rp_filter = 1
net.ipv4.conf.vNic_5.rp_filter = 1
net.ipv4.conf.vNic_6.rp_filter = 1
net.ipv4.conf.vNic_7.rp_filter = 1
net.ipv4.conf.vNic_8.rp_filter = 1
net.ipv4.conf.vNic_9.rp_filter = 1

nsxedge> show rpfstats
RPF drop packet count: 484
```

要查找 RPF 统计信息，请运行 `show rpfstats` 命令。

```
nsxedge> show rpfstats
RPF drop packet count: 484
```

如果数据包丢弃是随机出现的，请检查资源限制：

- a 对于 CPU 或内存使用率，请运行以下命令：
 - `show system cpu`
 - `show system memory`
 - `show system storage`
 - `show process monitor`
 - `top`

对于 ESXi，请运行 `esxtop n` 命令。

```
PCPU USED(%): 2.5 5.0 3.7 77 AVG: 22
PCPU UTIL(%): 0.5 2.7 3.3 92 AVG: 24
```

ID	GID	NAME	NWLD	%USED	%RUN	%SYS	%WAIT
98255269	98255269	esxtop.11224149	1	67.04	69.86	0.00	6.26
2	2	system	139	3.03	4.61	0.00	12053.58
86329	86329	app-01a	6	0.69	0.57	0.00	466.09
78730	78730	db-01a	6	0.48	0.67	0.00	441.44
90486	90486	app-02a	6	0.38	0.32	0.00	463.42

%VMWAIT	%RDY	%IDLE	%OVRLP	%CSTP	%MLMTD	%SWPWT
11.01	-	0.39	0.00	0.09	0.00	0.00
600.00	53.81	0.10	93.13	0.00	0.00	0.00
13900.00	-	28.68	0.00	2.69	0.00	0.00
600.00	53.81	0.10	93.13	0.00	0.00	0.00
600.00	0.00	0.19	151.92	0.00	0.00	0.00

NSX Manager 与 Edge 通信问题

NSX Manager 通过 VIX 或消息总线与 NSX Edge 通信。NSX Manager 在部署 Edge 时选择 VIX 或消息总线，并且从不发生变化。

注 VIX 在 NSX 6.3.0 和更高版本中不受支持。

VIX

- 如果未准备 ESXi 主机，则将 VIX 用于 NSX Edge。
- NSX Manager 先从 vCenter Server 中获取主机凭据以连接到 ESXi 主机。
- NSX Manager 使用 Edge 凭据登录到 Edge 设备。
- Edge 上的 vmttoolsd 进程处理 VIX 通信。

发生 VIX 故障的原因如下所示：

- NSX Manager 无法与 vCenter Server 通信。
- NSX Manager 无法与 ESXi 主机通信。
- 存在 NSX Manager 内部问题。
- 存在 Edge 内部问题。

VIX 调试

在 NSX Manager 日志中查找 VIX 错误 `VIX_E_<error>` 以缩小原因范围。查找类似下面的错误：

```
Vix Command 1126400 failed, reason com.vmware.vshield.edge.exception.VixException: vShield
Edge:10013:Error code 'VIX_E_FILE_NOT_FOUND' was returned by VIX API.:null

Health check failed for edge edge-13 VM vm-5025 reason:
com.vmware.vshield.edge.exception.VixException: vShield Edge:10013:Error code
'VIX_E_VM_NOT_RUNNING' was returned by VIX API.:null
```

通常，如果很多 Edge 同时发生相同的故障，则问题不是出在 Edge 上。

消息总线调试

在准备 ESXi 主机时，将使用消息总线进行 NSX Edge 通信。

在遇到问题时，NSX Manager 日志可能包含类似下面的条目：

```
GMT ERROR taskScheduler-6 PublishTask:963 - Failed to configure VSE-vm index 0, vm-id vm-117,
edge edge-5. Error: RPC request timed out
```

在以下情况下，将出现该问题：

- Edge 处于错误状态
- 消息总线连接中断

要在 Edge 上诊断该问题，请执行以下操作：

- 要检查 rmq 连接，请运行以下命令：

```
nsxedge> show messagebus messages
-----
Message bus is enabled
cmd conn state : listening
init_req       : 1
init_resp      : 1
init_req_err    : 0
...
```

- 要检查 vmci 连接，请运行以下命令：

```
nsxedge> show messagebus forwarder
-----
Forwarder Command Channel
vmci_conn       : up
app_client_conn : up
vmci_rx         : 3649
vmci_tx         : 3648
vmci_rx_err     : 0
```

```

vmci_tx_err      : 0
vmci_closed_by_peer: 8
vmci_tx_no_socket : 0
app_rx           : 3648
app_tx           : 3649
app_rx_err       : 0
app_tx_err       : 0
app_conn_req     : 1
app_closed_by_peer : 0
app_tx_no_socket : 0
-----
Forwarder Event Channel
vmci_conn        : up
app_client_conn  : up
vmci_rx          : 1143
vmci_tx          : 13924
vmci_rx_err      : 0
vmci_tx_err      : 0
vmci_closed_by_peer: 0
vmci_tx_no_socket : 0
app_rx           : 13924
app_tx           : 1143
app_rx_err       : 0
app_tx_err       : 0
app_conn_req     : 1
app_closed_by_peer : 0
app_tx_no_socket : 0
-----
cli_rx           : 1
cli_tx           : 1
cli_tx_err       : 0
counters_reset   : 0

```

在该示例中，输出 `vmci_closed_by_peer: 8` 表示主机代理关闭连接的次数。如果该数字不断增加并且 `vmci conn` 为 `down`，则主机代理无法连接到 `RMQ` 代理。在 `show log follow` 中，在 `Edge` 日志中查找重复的错误：`VmciProxy: [daemon.debug] VMCI Socket is closed by peer`

要在 `ESXi` 主机上诊断该问题，请执行以下操作：

- 要检查 `ESXi` 主机是否连接到 `RMQ` 代理，请运行以下命令：

```

esxcli network ip connection list | grep 5671

tcp    0    0  10.32.43.4:43329  10.32.43.230:5671  ESTABLISHED  35854  newreno
vsfwd
tcp    0    0  10.32.43.4:52667  10.32.43.230:5671  ESTABLISHED  35854  newreno
vsfwd
tcp    0    0  10.32.43.4:20808  10.32.43.230:5671  ESTABLISHED  35847  newreno
vsfwd
tcp    0    0  10.32.43.4:12486  10.32.43.230:5671  ESTABLISHED  35847  newreno vsfwd

```

Edge 诊断和恢复

Edge 诊断

- 使用以下命令检查 `vmtoolsd` 是否正在运行。

```
nsxedge> show process list
Perimeter-Gateway-01-0> show process list
%CPU %MEM    VSZ   RSZ STAT  STARTED    TIME COMMAND
  0.0  0.1   4244   720 Ss      May 16 00:00:15 init [3]
...
  0.0  0.1   4240   640 S       May 16 00:00:00 logger -p daemon debug -t vserrdd
  0.2  0.9  57192  4668 S       May 16 00:23:07 /usr/local/bin/vmtoolsd --plugin-pa
  0.0  0.4   4304  2260 SLs    May 16 00:01:54 /usr/sbin/watchdog
...
```

- 运行以下命令以检查 Edge 是否处于正常状态:

```
nsxedge> show eventmgr
-----
messagebus      : enabled
debug           : 0
profiling       : 0
cfg_rx          : 1
cfg_rx_msgbus   : 0
...
```

可以使用 `show eventmgr` 命令验证是否收到并处理查询命令。

```
nsxedge> show eventmgr
-----
messagebus      : enabled
debug           : 0
profiling       : 0
cfg_rx          : 1
cfg_rx_msgbus   : 0
cfg_rx_err      : 0
cfg_exec_err    : 0
cfg_resp        : 0
cfg_resp_err    : 0
cfg_resp_ln_err : 0
fastquery_rx    : 0 fastquery_err : 0
clearcmd_rx     : 0
clearcmd_err    : 0
ha_rx           : 0
ha_rx_err       : 0
ha_exec_err     : 0
status_rx       : 16
status_rx_err   : 0
status_svr      : 10
status_evt      : 0
```

```

status_evt_push: 0
status_ha      : 0
status_ver     : 1
status_sys     : 5
status_cmd     : 0
status_svr_err : 0
status_evt_err : 0
status_sys_err : 0
status_ha_err  : 0
status_ver_err : 0
status_cmd_err : 0
evt_report     : 1
evt_report_err : 0
hc_report      : 10962
hc_report_err  : 0
cli_rx        : 2
cli_resp      : 1
cli_resp_err  : 0
counter_reset  : 0
----- Health Status -----
system status : good
ha state      : active
cfg version   : 7
generation    : 0
server status : 1
syslog-ng     : 1
haproxy       : 0
ipsec         : 0
sslvpn        : 0
l2vpn         : 0
dns           : 0
dhcp          : 0
heartbeat     : 0
monitor       : 0
gslb          : 0
----- System Events -----

```

Edge 恢复

如果 `vmtoolsd` 未运行或 NSX Edge 处于错误状态，请重新引导 Edge。

要从崩溃中恢复，只需重新引导即可。应当不需要重新部署。

注 执行重新部署时，请记下旧 Edge 中的所有日志记录信息。

要调试内核崩溃，您需要获取以下信息：

- 仍处于崩溃状态的 Edge 虚拟机的 `vmss`（虚拟机挂起）或 `vmxn`（虚拟机快照）文件。如果存在 `vmem` 文件，则还需要此文件。可以使用该文件提取 VMware 支持部门可分析的内核核心转储文件。
- 重新引导（而非重新部署）崩溃的 Edge 后立即生成的 Edge 支持日志。您还可以检查 Edge 日志。请参见 <https://kb.vmware.com/kb/2079380>。
- Edge 控制台的屏幕截图也是非常有用的，但它通常不包含完整的崩溃报告。

防火墙故障排除

5

本节提供了有关解决防火墙问题的信息。

本章讨论了以下主题：

- [关于分布式防火墙](#)
- [身份防火墙](#)

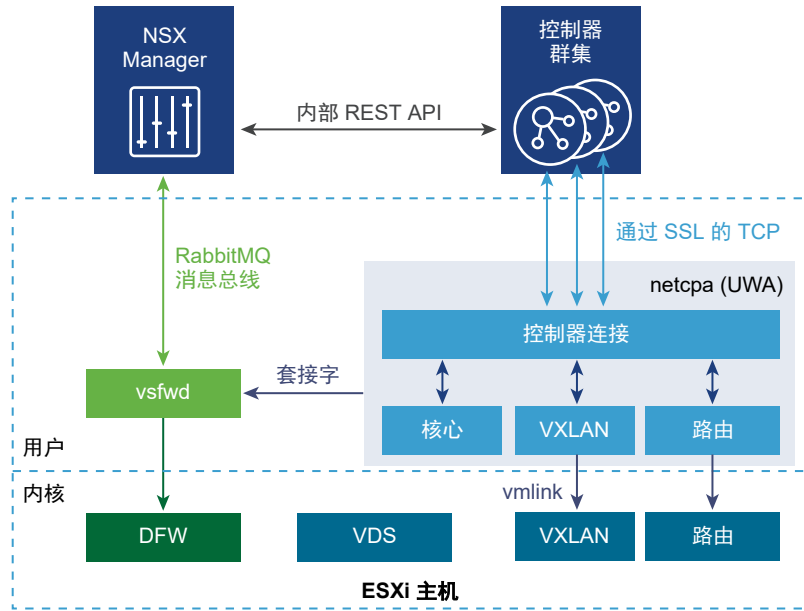
关于分布式防火墙

可以使用 RabbitMQ 消息总线在 vsfwd（RMQ 客户端）和 NSX Manager 上托管的 RMQ 服务器进程之间通信。NSX Manager 使用该消息总线将各种信息发送到 ESXi 主机，包括需要在内核中的分布式防火墙上写入的策略规则。

NSX 分布式防火墙是嵌入在管理程序内核中的防火墙，可以查看和控制虚拟化的工作负载和网络。您可以根据 VMware vCenter 对象（如数据中心和群集）、虚拟机名称和标记、网络结构（如 IP/VLAN/VXLAN 地址）以及 Active Directory 中的用户组标识创建访问控制策略。现在，在通过 vMotion 在不同物理主机之间移动虚拟机时，将会强制实施一致的访问控制策略，而无需重写防火墙规则。由于分布式防火墙嵌入在管理程序中，它可以提供接近线路速率的吞吐量，从而在物理服务器上支持更高的工作负载整合。防火墙的分布式特性使架构具有向外扩展性，可在向数据中心添加更多主机时自动扩展防火墙功能。

ESXi 主机上的 NSX Manager Web 应用程序和 NSX 组件通过 RabbitMQ 代理进程相互通信，该进程在与 NSX Manager Web 应用程序相同的虚拟机上运行。使用的通信协议为 AMQP（高级消息队列协议），并使用 SSL 保护通道安全。在 ESXi 主机上，VSFWD（vShield 防火墙守护程序）进程建立并维护到代理的 SSL 连接，并代表其他组件发送和接收消息，这些组件通过 IPC 与其进行通信。

图 5-1. ESXi 主机用户和内核空间图



适用于 DFW 的 CLI 命令

您可以在 NSX Manager 中央 CLI 上获取有关分布式防火墙的大多数信息。

使用 Show dfw 中央 CLI 命令

获取所需信息的途径如下所示：

- 1 使用 **admin** 凭据登录到 NSX Manager 中央 CLI。
- 2 运行以下命令：
 - a 运行 **show cluster all** 命令以显示所有群集。

```
nsxmgr>show cluster all
```

No.	Cluster Name	Cluster Id	Datacenter Name	Firewall Status
1	Compute Cluster A	domain-c33	Datacenter Site A	Enabled
2	Management & Edge Cluster	domain-c41	Datacenter Site A	Enabled

- b 运行 **show cluster <clusterID>** 命令以显示特定群集中的主机。

```
nsxmgr> show cluster domain-c33
```

Datacenter: Datacenter Site A

Cluster: Compute Cluster A

No.	Host Name	Host Id	Installation Status
1	esx-02a.corp.local	host-32	Enabled
2	esx-01a.corp.local	host-28	Enabled

- c 运行 `show host <hostID>` 以显示主机上的所有虚拟机。

```
nsxmgr> show host host-28
Datacenter: Datacenter Site A
Cluster: Compute Cluster A
Host: esx-01a.corp.local
No.  VM Name    VM Id    Power Status
1    web-02a     vm-219   on
2    web-01a     vm-216   on
3    win8-01a    vm-206   off
4    app-02a     vm-264   on
```

- d 运行 `show vm <vmID>` 命令以显示虚拟机的信息，包括筛选器名称和 vNIC ID：

```
nsxmgr> show vm vm-264
Datacenter: Datacenter Site A
Cluster: Compute Cluster A
Host: esx-01a.corp.local
Host-ID: host-28
VM: app-02a
Virtual Nics List:
1.
Vnic Name      app-02a - Network adapter 1
Vnic Id        502ef2fa-62cf-d178-cb1b-c825fb300c84.000
Filters        nic-79396-eth0-vmware-sfw.2
```

- e 记下 vNIC ID 并运行进一步的命令，例如，`show dfw vnic <vnicID>` 和 `show dfw host <hostID> filter <filter ID> rules`：

```
nsxmgr> show dfw vnic 502ef2fa-62cf-d178-cb1b-c825fb300c84.000
Vnic Name      app-02a - Network adapter 1
Vnic Id        502ef2fa-62cf-d178-cb1b-c825fb300c84.000
Mac Address    00:50:56:ae:6c:6b
Port Group Id  dvportgroup-385
Filters        nic-79396-eth0-vmware-sfw.2

nsxmgr> show dfw host host-28 filter nic-79396-eth0-vmware-sfw.2 rules
ruleset domain-c33 {
  # Filter rules
  rule 1012 at 1 inout protocol any from addrset ip-securitygroup-10 to addrset ip-securitygroup-10 drop with log;
  rule 1013 at 2 inout protocol any from addrset src1013 to addrset src1013 drop;
  rule 1011 at 3 inout protocol tcp from any to addrset dst1011 port 443 accept;
  rule 1011 at 4 inout protocol icmp icmptype 8 from any to addrset dst1011 accept;
  rule 1010 at 5 inout protocol tcp from addrset ip-securitygroup-10 to addrset ip-securitygroup-11 port 8443 accept;
  rule 1010 at 6 inout protocol icmp icmptype 8 from addrset ip-securitygroup-10 to addrset ip-securitygroup-11 accept;
```

```

rule 1009 at 7 inout protocol tcp from addrset ip-securitygroup-11 to addrset ip-
securitygroup-12 port 3306 accept;
rule 1009 at 8 inout protocol icmp icmptype 8 from addrset ip-securitygroup-11 to addrset
ip-securitygroup-12 accept;
rule 1003 at 9 inout protocol ipv6-icmp icmptype 136 from any to any accept;
rule 1003 at 10 inout protocol ipv6-icmp icmptype 135 from any to any accept;
rule 1002 at 11 inout protocol udp from any to any port 67 accept;
rule 1002 at 12 inout protocol udp from any to any port 68 accept;
rule 1001 at 13 inout protocol any from any to any accept;
}

ruleset domain-c33_L2 {
# Filter rules
rule 1004 at 1 inout ethertype any from any to any accept;
}

```

使用 export host-tech-support 中央 CLI 命令

可以使用 `export host-tech-support` 命令将 ESXi 主机支持包导出到指定的服务器中。此外，该命令还在指定的主机上收集与 NSX 相关的输出和文件（不限于以下内容），例如：

- VMKernel 和 vsfwd 日志文件
- 筛选器列表
- DFW 规则列表
- 容器列表
- SpoofGuard 详细信息
- 与主机相关的信息
- 与 IP 发现相关的信息
- RMQ 命令输出
- 安全组、服务配置文件和实例详细信息
- 与 ESX CLI 相关的输出

该命令还会移除 NSX Manager 上的任何临时文件。

要收集与 NSX 相关的输出和文件，请执行以下操作：

- 1 使用 `admin` 凭据登录到 NSX Manager 中央 CLI。
- 2 运行以下命令：
 - a `show cluster all` - 查找所需的主机 ID。
 - b `export host-tech-support host-id scp uid@ip:/path` - 生成 NSX 技术支持包并将其复制到指定的服务器中。

有关详细信息，请参阅：

- [NSX 命令行快速参考](#)。

- NSX 命令行界面参考。

对分布式防火墙进行故障排除

本主题提供了有助于您了解 VMware NSX 6.x 分布式防火墙 (DFW) 和进行故障排除的信息。

问题

- 发布分布式防火墙规则失败。
- 更新分布式防火墙规则失败。

原因

验证下面的每个故障排除步骤是否适用于您的环境。每个步骤提供了相应说明或指向文档的链接，以消除可能的根源并在必要时采取纠正措施。这些步骤按最适当的顺序进行排列，以查找问题并确定相应的解决方案。在执行每个步骤后，再次尝试更新/发布分布式防火墙规则。

解决方案

- 1 验证是否在群集中的每个 ESXi 主机上成功安装了 NSX VIB。为此，在群集中的每个 ESXi 主机上运行以下命令。

```
# esxcli software vib list | grep vsip
esx-vsip                6.0.0-0.0.4744062  VMware  VMwareCertified  2017-01-04

# esxcli software vib list | grep vxlan
esx-vxlan                6.0.0-0.0.4744062  VMware  VMwareCertified  2017-01-04
```

NSX 6.2 之前的 NSX 版本具有一个额外的 VIB：

```
# esxcli software vib list | grep dvfilter
esx-dvfilter-switch-security  5.5.0-0.0.2318233  VMware  VMwareCertified  2015-01-24
```

从 ESXi 6.0 或更高版本上的 NSX 6.3.3 开始，esx-vxlan 和 esx-vsip VIB 将替换为 esx nsxv。

```
# esxcli software vib list | grep nsxv
esx-nsxv                6.0.0-0.0.6216823  VMware  VMwareCertified  2017-08-10
```

- 2 在 ESXi 主机上，验证 vShield-Stateful-Firewall 服务是否处于运行状态。

例如：

```
# /etc/init.d/vShield-Stateful-Firewall status

vShield-Stateful-Firewall is running
```

- 3 验证消息总线是否正确与 NSX Manager 通信。

该进程是由监视程序脚本自动启动的；如果由于未知原因终止，将重新启动该进程。在群集中的每个 ESXi 主机上运行以下命令。

例如：

```
# ps | grep vsfwd

107557 107557 vsfwd /usr/lib/vmware/vsfw/vsfwd
```

应在命令输出中包含运行的至少 12 个或更多的 vsfwd 进程。如果运行的进程较少（很可能只有 2 个），则 vsfwd 未正常运行。

4 验证是否在防火墙配置中打开端口 5671 以进行通信。

以下命令显示到 RabbitMQ 代理的 VSFWD 连接。请在 ESXi 主机上运行以下命令，以查看从 ESXi 主机上的 vsfwd 进程到 NSX Manager 的连接列表。确保在环境中的任何外部防火墙上打开端口 5671 以进行通信。此外，在端口 5671 上应具有至少两个连接。可能在端口 5671 上具有更多连接，因为在 ESXi 主机上部署的 NSX Edge 虚拟机也会建立到 RMQ 代理的连接。

例如：

```
# esxcli network ip connection list |grep 5671

tcp          0      0 192.168.110.51:30133      192.168.110.15:5671  ESTABLISHED
10949155 newreno vsfwd
tcp          0      0 192.168.110.51:39156      192.168.110.15:5671  ESTABLISHED
10949155 newreno vsfwd
```

5 验证是否配置了 VSFWD。

以下命令应显示 NSX Manager IP 地址。

```
# esxcfg-advcfg -g /UserVars/RmqIpAddress
```

6 如果在该 ESXi 主机中使用主机配置文件，请确认未在主机配置文件中设置 RabbitMQ 配置。

请参见：

- <https://kb.vmware.com/kb/2092871>
- <https://kb.vmware.com/kb/2125901>

7 验证 ESXi 主机的 RabbitMQ 凭据是否与 NSX Manager 不同步。请下载 NSX Manager 技术支持日志。在收集所有 NSX Manager 技术支持日志后，在所有日志中搜索类似下面的条目：

将 host-420 替换为可疑主机的 mo-id。

```
PLAIN login refused: user 'uw-host-420' - invalid credentials.
```

- 8 如果在可疑 ESXi 主机的日志中找到这些条目，请重新同步消息总线。

要重新同步消息总线，请使用 REST API。为了更好地了解该问题，请在重新同步消息总线后立即收集日志。

```
HTTP Method : POST
Headers ,
Authorization : base64encoded value of username password
Accept : application/xml
Content-Type : application/xml
Request:

POST https://NSX_Manager_IP/api/2.0/nwfabric/configure?action=synchronize

Request Body:

<nwFabricFeatureConfig>
<featureId>com.vmware.vshield.vsm.messagingInfra</featureId>
<resourceConfig>
<resourceId>{HOST/CLUSTER MOID}</resourceId>
</resourceConfig>
</nwFabricFeatureConfig>
```

- 9 使用 `export host-tech-support <host-id> scp <uid@ip:/path>` 命令收集主机特定的防火墙日志。

例如：

```
nsxmgr# export host-tech-support host-28 scp Administrator@192.168.110.10
Generating logs for Host: host-28...
```

- 10 使用 `show dfw host host-id summarize-dvfilter` 命令验证是否在主机上部署了防火墙规则并将其应用于虚拟机。

在输出中，`module: vsip` 显示已加载 DFW 模块并正在运行。`name` 显示在每个 vNic 上运行的防火墙。

您可以运行 `show dfw cluster all` 命令以获取群集域 ID，然后运行 `show dfw cluster domain-id` 以获取主机 ID。

例如：

```
# show dfw host host-28 summarize-dvfilter

Fastpaths:
agent: dvfilter-faulter, refCount: 1, rev: 0x1010000, apiRev: 0x1010000, module: dvfilter
agent: ESXi-Firewall, refCount: 5, rev: 0x1010000, apiRev: 0x1010000, module: esxfw
agent: dvfilter-generic-vmware, refCount: 1, rev: 0x1010000, apiRev: 0x1010000, module: dvfilter-generic-fastpath
agent: dvfilter-generic-vmware-swsec, refCount: 4, rev: 0x1010000, apiRev: 0x1010000, module: dvfilter-switch-security
```

```

agent: bridgelearningfilter, refCount: 1, rev: 0x1010000, apiRev: 0x1010000, module: vdrb
agent: dvfg-igmp, refCount: 1, rev: 0x1010000, apiRev: 0x1010000, module: dvfg-igmp
agent: vmware-sfw, refCount: 4, rev: 0x1010000, apiRev: 0x1010000, module: vsip

Slowpaths:

Filters:
world 342296 vmm0:2-vm-RHEL63_srv_64-shared-846-3f435476-8f54-4e5a-8d01-59654a4e9979 vcUuid:'3f
43 54 76 8f 54 4e 5a-8d 01 59 65 4a 4e 99 79'
port 50331660 2-vm-RHEL63_srv_64-shared-846-3f435476-8f54-4e5a-8d01-59654a4e9979.eth1
  vNic slot 2
    name: nic-342296-eth1-vmware-sfw.2
    agentName: vmware-sfw
    state: IOChain Attached
    vmState: Detached
    failurePolicy: failClosed
    slowPathID: none
    filter source: Dynamic Filter Creation
  vNic slot 1
    name: nic-342296-eth1-dvfilter-generic-vmware-swsec.1
    agentName: dvfilter-generic-vmware-swsec
    state: IOChain Attached
    vmState: Detached
    failurePolicy: failClosed
    slowPathID: none
    filter source: Alternate Opaque Channel
port 50331661 (disconnected)
  vNic slot 2
    name: nic-342296-eth2-vmware-sfw.2 <===== DFW filter
    agentName: vmware-sfw
    state: IOChain Detached
    vmState: Detached
    failurePolicy: failClosed
    slowPathID: none
    filter source: Dynamic Filter Creation
port 33554441 2-vm-RHEL63_srv_64-shared-846-3f435476-8f54-4e5a-8d01-59654a4e9979
  vNic slot 2
    name: nic-342296-eth0-vmware-sfw.2<===== DFW filter
    agentName: vmware-sfw
    state: IOChain Attached
    vmState: Detached
    failurePolicy: failClosed
    slowPathID: none
    filter source: Dynamic Filter Creation

```

11 运行 show dfw host hostID filter filterID rules 命令。

例如:

```

# show dfw host host-28 filter nic-79396-eth0-vmware-sfw.2 rules

ruleset domain-c33 {
  # Filter rules
  rule 1012 at 1 inout protocol any from addrset ip-securitygroup-10 to addrset ip-

```

```

securitygroup-10 drop with log;
  rule 1013 at 2 inout protocol any from addrset src1013 to addrset src1013 drop;
  rule 1011 at 3 inout protocol tcp from any to addrset dst1011 port 443 accept;
  rule 1011 at 4 inout protocol icmp icmp type 8 from any to addrset dst1011 accept;
  rule 1010 at 5 inout protocol tcp from addrset ip-securitygroup-10 to addrset ip-
securitygroup-11 port 8443 accept;
  rule 1010 at 6 inout protocol icmp icmp type 8 from addrset ip-securitygroup-10 to addrset ip-
securitygroup-11 accept;
  rule 1009 at 7 inout protocol tcp from addrset ip-securitygroup-11 to addrset ip-
securitygroup-12 port 3306 accept;
  rule 1009 at 8 inout protocol icmp icmp type 8 from addrset ip-securitygroup-11 to addrset ip-
securitygroup-12 accept;
  rule 1003 at 9 inout protocol ipv6-icmp icmp type 136 from any to any accept;
  rule 1003 at 10 inout protocol ipv6-icmp icmp type 135 from any to any accept;
  rule 1002 at 11 inout protocol udp from any to any port 67 accept;
  rule 1002 at 12 inout protocol udp from any to any port 68 accept;
  rule 1001 at 13 inout protocol any from any to any accept;
}

ruleset domain-c33_L2 {
  # Filter rules
  rule 1004 at 1 inout ethertype any from any to any accept;
}

```

12 运行 `show dfw host hostID filter filterID addrsets` 命令。

例如：

```

# show dfw host host-28 filter nic-342296-eth2-vmware-sfw.2 addrsets

addrset dst1011 {
ip 172.16.10.10,
ip 172.16.10.11,
ip 172.16.10.12,
ip fe80::250:56ff:feae:3e3d,
ip fe80::250:56ff:feae:f86b,
}
addrset ip-securitygroup-10 {
ip 172.16.10.11,
ip 172.16.10.12,
ip fe80::250:56ff:feae:3e3d,
ip fe80::250:56ff:feae:f86b,
}
addrset ip-securitygroup-11 {
ip 172.16.20.11,
ip fe80::250:56ff:feae:23b9,
}
addrset ip-securitygroup-12 {
ip 172.16.30.11,
ip fe80::250:56ff:feae:d42b,
}
addrset src1013 {
ip 172.16.10.12,
}

```



```
ip 172.17.10.11,
ip fe80::250:56ff:feae:cf88,
ip fe80::250:56ff:feae:f86b,
}
```

- 13** 如果已验证上面的每个故障排除步骤，并且无法将防火墙规则发布到主机虚拟机中，请通过 **NSX Manager UI** 或以下 **REST API** 调用执行主机级别强制同步。

```
URL : [https:]https://<nsx-mgr-ip>/api/4.0/firewall/forceSync/<host-id>
HTTP Method : POST
Headers ,
Authorization : base64encoded value of username password
Accept : application/xml
Content-Type : application/xml
```

解决方案

备注：

- 如果防火墙规则不使用 IP 地址，请确保 **VMware Tools** 正在虚拟机上运行。有关详细信息，请参见 <https://kb.vmware.com/kb/2084048>。

VMware NSX 6.2.0 引入了一个使用 **DHCP** 侦听或 **ARP** 侦听获悉虚拟机 IP 地址的选项。通过使用这些新的发现机制，**NSX** 可以在未安装 **VMware Tools** 的虚拟机上强制实施基于 IP 地址的安全规则。有关详细信息，请参见 **NSX 6.2.0** 发行说明。

在完成主机准备过程后，将会立即激活 **DFW**。如果虚拟机根本不需要使用 **DFW** 服务，可以将其添加到排除列表功能中（默认情况下，自动从 **DFW** 功能中排除 **NSX Manager**、**NSX Controller** 和 **Edge 服务网关**）。在 **DFW** 中创建“全部拒绝”规则后，可能会阻止 **vCenter Server** 访问。有关详细信息，请参见 <https://kb.vmware.com/kb/2079620>。

- 在请求 **VMware** 技术支持部门排除 **VMware NSX 6.x** 分布式防火墙 (**DFW**) 故障时，需要提供以下信息：
 - 群集中的每个 **ESXi** 主机上的 `show dfw host hostID summarize-dvfilter` 命令输出。
 - 从 **网络和安全 > 防火墙 > 常规 (Networking and Security > Firewall > General)** 选项卡中选择分布式防火墙配置，然后单击 **导出配置 (Export Configuration)**。这会将分布式防火墙配置导出为 XML 格式。
 - **NSX Manager** 日志。有关详细信息，请参见 <https://kb.vmware.com/kb/2074678>。
 - **vCenter Server** 日志。有关详细信息，请参见 <https://kb.vmware.com/kb/1011641>。

身份防火墙

问题

发布或更新身份防火墙规则失败。

原因

身份防火墙 (IDFW) 允许使用基于用户的分布式防火墙规则 (DFW)。

基于用户的分布式防火墙规则是由 Active Directory (AD) 组成员的成员资格确定的。IDFW 监控 Active Directory 用户登录到的位置，并将登录名映射到 IP 地址，DFW 使用该地址以应用防火墙规则。IDFW 需要使用 Guest Introspection 框架和/或 Active Directory 事件日志提取。

解决方案

- 1 确保 Active Directory 服务器完全/增量同步可在 NSX Manager 上正常执行。
 - a 在 vSphere Web Client 中，登录到与 NSX Manager 链接的 vCenter。
 - b 导航到主页 > 网络和安全 > NSX Manager (Home > Networking & Security> NSX Managers)，然后从列表中选择您的 NSX Manager。
 - c 依次选择**管理 (Manage)**和**域 (Domains)**选项卡。从列表中选择您的域。确认**上次同步状态 (Last Synchronization Status)**列显示“成功”，且**上次同步时间 (Last Synchronization Time)**为当前时间。
- 2 如果您的防火墙环境使用事件日志提取方法进行登录检测，请按照以下步骤确认已为您的域配置事件日志服务器：
 - a 在 vSphere Web Client 中，登录到与 NSX Manager 链接的 vCenter。
 - b 导航到主页 > 网络和安全 > NSX Manager (Home > Networking & Security> NSX Managers)，然后从列表中选择您的 NSX Manager。
 - c 依次选择**管理 (Manage)**和**域 (Domains)**选项卡。从列表中选择您的域。您可以在此处查看和编辑详细的域配置。
 - d 从域详细信息中选择**事件日志服务器 (Event Log Servers)**，并确认已添加您的事件日志服务器。
 - e 选择您的事件日志服务器，并确认**上次同步状态 (Last Sync Status)**列显示“成功”，且**上次同步时间 (Last Sync Time)**为当前时间。
- 3 如果您的防火墙环境使用 Guest Introspection，则必须将该框架部署到受 IDFW 保护的虚拟机所在的计算群集中。UI 上的服务运行状况应为绿色。有关 Guest Introspection 诊断信息，请查看以下知识库文章：“排除 vShield Endpoint/NSX Guest Introspection 故障” (<https://kb.vmware.com/kb/2094261>) 和“在 VMware NSX for vSphere 6.x Guest Introspection 通用服务虚拟机中收集日志” (<https://kb.vmware.com/kb/2144624>)。

- 4 在确认登录检测方法的配置正确无误后，请确保 NSX Manager 可接收登录事件；
 - a 以 Active Directory 用户身份登录。
 - b 运行以下命令以查询登录事件。确认结果中返回了您的用户。GET `https://<nsxmgr-ip>/1.0/identity/userIpMapping`。

```
Example output:
<UserIpMappings>
  <UserIpMapping>
    <ip>50.1.111.192</ip>
    <userName>user1_group20</userName>
    <displayName>user1_group20</displayName>
    <domainName>cd.ad1.db.com</domainName>
    <startTime class="sql-timestamp">2017-05-11 22:30:51.0</startTime>
    <startType>EVENTLOG</startType>
    <lastSeenTime class="sql-timestamp">2017-05-11 22:30:52.0</lastSeenTime>
    <lastSeenType>EVENTLOG</lastSeenType>
  </UserIpMapping>
</UserIpMappings>
```

- 5 确认您的安全组已在防火墙规则中使用，或分配有安全策略。除非满足这两个条件之一，否则不会在 IDFW 中处理安全组。
- 6 在确认 IDFW 可正确检测登录后，请确认您的桌面虚拟机所在的 ESXi 主机可接收正确的配置。这些步骤将使用 NSX Manager 集中式 CLI。要检查 **ip-securitygroup** 列表中填充的桌面虚拟机 IP 地址，请执行以下操作：
 - a 参见[适用于 DFW 的 CLI 命令](#)以检索对桌面虚拟机应用的筛选器名称。
 - b 运行 `show dfw host hostID filter filterID rules` 命令以查看本地 DFW 规则项目。
 - c 运行 `show dfw host hostID filter filterID addrsets` 命令以查看 **ip-securitygroup** 列表中填充的 IP 地址。确认您的 IP 显示在列表中。

解决方案

注意：在请求 VMware 技术支持人员对身份 IDFW 进行故障排除时，以下数据很有用：

- 如果使用事件日志提取 Active Directory 规模数据：

- 单个 NSX Manager 的域数量
 - 林数量
 - 每个林的用户数量
 - 每个域的用户数量
 - 每个域的 Active Directory 组数量
 - 每个 Active Directory 组的用户数量
 - 每个用户的 Active Directory 数量
 - 域控制器数量

Active Directory 日志服务器的数量

- 用户登录规模数据:
 - 每分钟的平均用户数量
- 使用 IDFW 通过 VDI 进行部署的详细信息:
 - 每个 VC 的 VDI 桌面数量
 - 每个 VC 的主机数量
 - 每个主机的 VDI 桌面数量
- 如果使用 Guest Introspection:
 - VMTools (Guest Introspection 驱动程序) 的版本
 - Windows 客户机操作系统的版本

负载均衡故障排除

6

NSX Edge 负载均衡器使网络流量可以沿着多个路径流向特定目标。它将入站服务请求均匀分布在多个服务器中，从方式上确保负载分配对用户透明。可以在 **NSX** 中配置两种类型的负载均衡服务：单路并联式模式（也称为代理模式）或串联式模式（也称为透明模式）。有关详细信息，请参阅 **NSX** 管理指南。

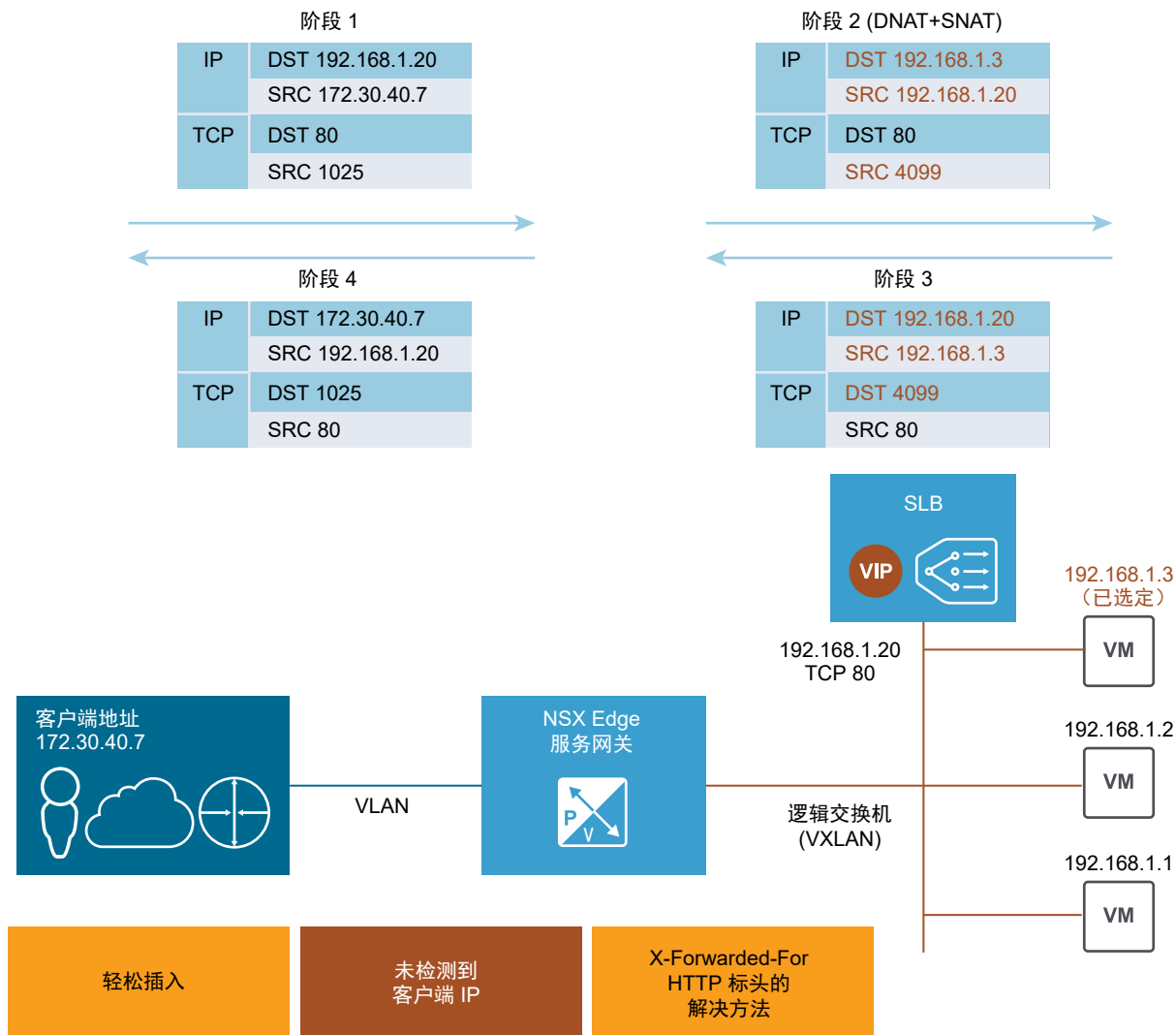
在开始故障排除和配置验证之前，请获取准确的错误描述，创建与客户端、虚拟服务器和后端服务器相关的拓扑图以及了解应用程序要求。例如，客户端无法连接与连接后的随机会话错误不同。在排除负载均衡器故障时，始终从验证连接错误开始。

本章讨论了以下主题：

- [配置单路并联式负载均衡器](#)
- [负载均衡器的故障排除流程图](#)
- [使用 UI 验证负载均衡器配置和排除故障](#)
- [使用 CLI 的负载均衡器故障排除](#)
- [常见的负载均衡器问题](#)

配置单路并联式负载均衡器

可以将 Edge 服务网关 (ESG) 视为传入客户端流量的代理。



在代理模式下，负载均衡器将自己的 IP 地址作为源地址，以将请求发送到后端服务器。后端服务器查看从负载均衡器中发送的所有流量，并直接响应负载均衡器。这种模式也称为 **SNAT 模式**或非透明模式。有关详细信息，请参阅 **NSX 管理指南**。

典型的 **NSX** 单路并联式负载均衡器部署在具有其后端服务器的相同子网上，与逻辑路由器分开。**NSX** 负载均衡器虚拟服务器侦听虚拟 IP 以查找来自客户端的传入请求，并将这些请求发送到后端服务器。对于返回流量，需要使用反向 **NAT** 以将源 IP 地址从后端服务器更改为虚拟 IP (**VIP**) 地址，然后将虚拟 IP 地址发送到客户端。如果不执行该操作，到客户端的连接将中断。

在 **ESG** 收到流量后，它执行两个操作：目标网络地址转换 (**DNAT**) 以将 **VIP** 地址更改为某个负载均衡计算机的 IP 地址；以及源网络地址转换 (**SNAT**) 以将客户端 IP 地址与 **ESG** IP 地址调换。

然后，**ESG** 服务器将流量发送到负载均衡服务器，负载均衡服务器将响应发送回 **ESG**，然后发送回客户端。该选项比串联式模式容易配置得多，但具有两个潜在问题。第一个问题是，该模式需要使用专用的 **ESG** 服务器，第二个问题是，负载均衡器服务器不知道原始客户端 IP 地址。**HTTP/HTTPS** 应用程序的一种解决方法是，在 **HTTP** 应用程序配置文件中启用“插入 **X-Forwarded-For**”，以便在发送到后端服务器的请求的 **X-Forwarded-For** **HTTP** 标头中包含客户端 IP 地址。

如果 **HTTP/HTTPS** 以外的应用程序要求在后端服务器上看到客户端 IP 地址，您可以将 IP 池配置为透明的。如果客户端没有位于与后端服务器相同的子网上，建议使用串联式模式。否则，您必须将负载均衡器 IP 地址作为后端服务器的默认网关。

注 通常，可以使用三种方法确保连接完整性：

- 串联式/透明模式
- **SNAT/代理/非透明模式**（如上所述）
- 直接服务器返回 (**Direct Server Return, DSR**) - 目前，不支持这种模式

在 **DSR** 模式下，后端服务器直接响应客户端。目前，**NSX** 负载均衡器不支持 **DSR**。

步骤

- 1 例如，让我们配置一个具有 **SSL** 卸载的单路并联式虚拟服务器。双击 **Edge**，然后选择**管理 > 设置 > 证书 (Manage > Settings > Certificate)**以创建一个证书。

- 2 选择**管理 > 负载均衡器 > 全局配置 > 编辑 (Manage > Load Balancer > Global Configuration > Edit)**以启用负载均衡器服务。

Edit Load balancer global configuration

☒ Enable Load Balancer

☐ Enable Acceleration

☐ Logging

Log Level: **Info** ▼

☐ Enable Service Insertion

Service Definition:

Service Configuration:

Deployment Specification:

- 3 选择**管理 > 负载均衡器 > 应用程序配置文件 (Manage > Load Balancer > Application Profiles)**以创建一个 HTTPS 应用程序配置文件。

New Profile ?

Name:

Type: **HTTPS** ▼

☐ Enable SSL Passthrough

HTTP Redirect URL:

Persistence: **None** ▼

Cookie Name:

Mode:

Expires in (Seconds):

☐ Insert X-Forwarded-For HTTP header

☐ Enable Pool Side SSL

Virtual Server Certifica... Pool Certificates

Service Certificates CA Certificates CRL

☒ Configure Service Certificate

	Common Name	Issuer	Validity
<input checked="" type="radio"/>	VSM_SOLUTION_71	VSM_SOLUTION_71	Tue Sep 8 2015 - Thu
<input type="radio"/>	VSM_SOLUTION_71	VSM_SOLUTION_71	Tue Sep 8 2015 - Thu

注 上面的屏幕截图使用自签名证书以仅用于说明目的。

- 4 (可选) 单击**管理 > 负载均衡器 > 服务监控器 (Manage > Load Balancer > Service Monitoring)**, 然后编辑默认服务监控以将其从基本 HTTP/HTTPS 更改为特定的 URL/URI (如果需要)。

5 选择**管理 > 负载均衡器 > 池 (Manage > Load Balancer > Pools)**以创建服务器池。

要使用 **SNAT** 模式，请在池配置中取消选中**透明 (Transparent)**复选框。

Enabled	Name	IP Address / VC Container	Weight	Monitor Port	Port	Max Connections	Min Connections
✓	web-01a	172.16.10.11	1	443	443	0	0
✓	web-02a	172.16.10.12	1	443	443	0	0

确保已列出并启用虚拟机。

6 （可选）单击**管理 > 负载均衡器 > 池 > 显示池统计信息 (Manage > Load Balancer > Pools > Show Pool Statistics)**以检查状态。

确保成员处于已启动状态。

7 选择**管理 > 负载均衡器 > 虚拟服务器 (Manage > Load Balancer > Virtual Servers)**以创建一个虚拟服务器。

如果要将 **L4** 负载均衡器用于 **UDP** 或更高性能的 **TCP**，请选中**启用加速 (Enable Acceleration)**。如果选中**启用加速 (Enable Acceleration)**，请确保负载均衡器 **NSX Edge** 上的防火墙状态为已启用 (**Enabled**)，因为 **L4 SNAT** 需要使用防火墙。

确保 **IP** 地址绑定到服务器池。

- 8 (可选) 如果使用一个应用程序规则, 请在**管理 > 负载均衡器 > 应用程序规则 (Manage > Load Balancer > Application Rules)**中检查配置。

Add Application Rule

Name: App-Rule-1

Script: # A sample application rule to log the name of the virtual server
capture request header Host len 32

- 9 如果使用一个应用程序规则, 请在**管理 > 负载均衡器 > 虚拟服务器 > 高级 (Manage > Load Balancer > Virtual Servers > Advanced)**中确保该应用程序规则与虚拟服务器相关联。

有关支持的示例, 请参见: <https://communities.vmware.com/docs/DOC-31772>。

Edit Virtual Server

General Advanced

Application Rules:

+ × ↕ ⇅ Q Filter

Rule Id	Name	Script
applicationRule-1	App-rule-1	capture request he...

在非透明模式下, 后端服务器无法看到客户端 IP, 但可以看到负载均衡器内部 IP 地址。作为 HTTP/HTTPS 流量的解决方法, 请选中**插入 X-Forwarded-For HTTP 标头 (Insert X-Forwarded-For HTTP header)**。如果选中该选项, Edge 负载均衡器将添加 “X-Forwarded-For” 标头并且值为客户端源 IP 地址。

Edit Profile

Name: http_application_profile

Type: HTTP

☒ Enable SSL Passthrough

HTTP Redirect URL:

Persistence: None

Cookie Name:

Mode:

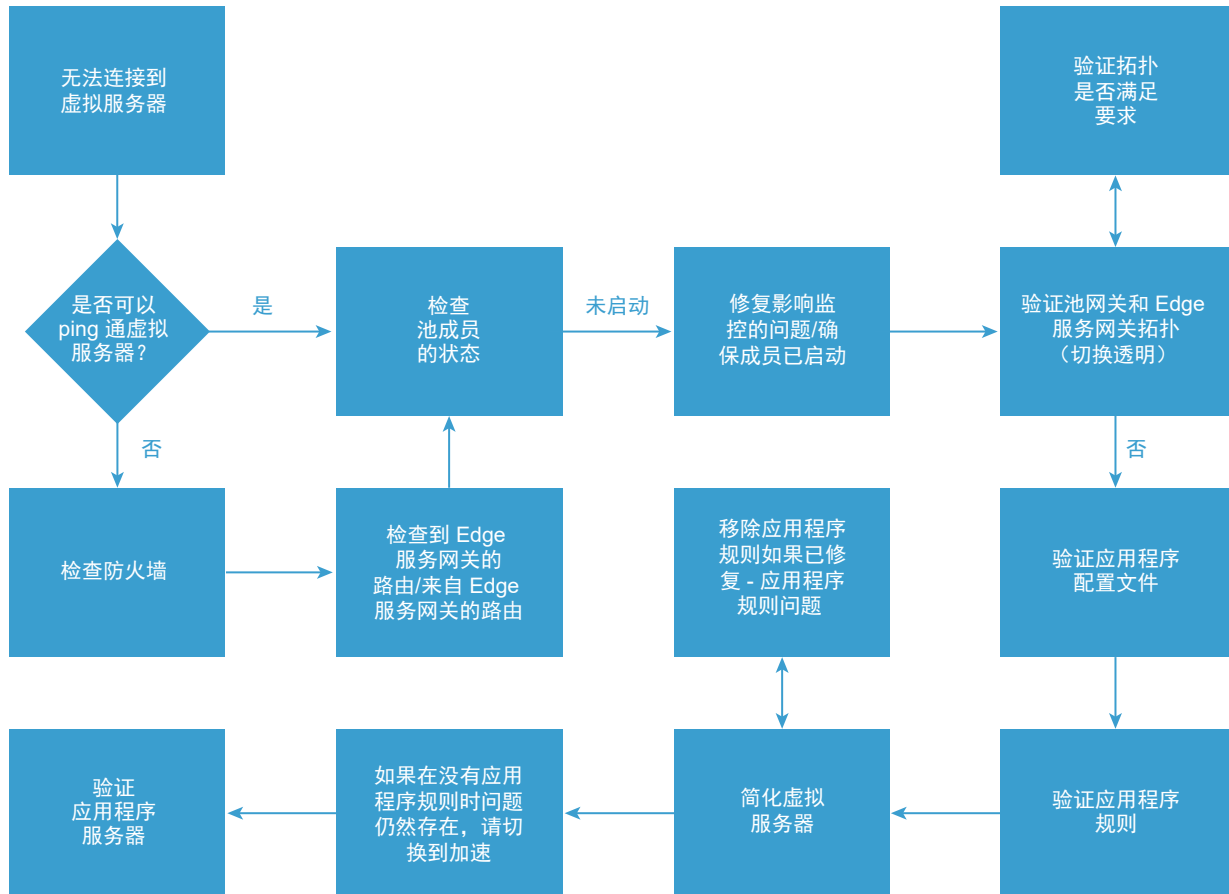
Expires in (Seconds):

☒ Insert X-Forwarded-For HTTP header

☐ Enable Pool Side SSL

负载均衡器的故障排除流程图

以下流程图简要说明了如何解决负载均衡器问题。



使用 UI 验证负载均衡器配置和排除故障

您可以通过 vSphere Web Client 验证负载均衡器配置。您可以使用 UI 执行某些负载均衡器故障排除。

在了解哪些设备应该正常工作并定义问题后，通过 UI 按以下方式验证配置。

前提条件

请记住以下详细信息：

- 虚拟服务器的 IP、协议和端口。
- 后端应用程序服务器的 IP 和端口。
- 预期的拓扑 - 串联式或单路并联式。有关详细信息，请参阅《NSX 管理指南》中的“逻辑负载均衡器”主题。
- 验证跟踪路由，并使用其他网络连接工具查看是否将数据包传输到正确的位置（Edge 服务网关）。
- 验证任何上游防火墙是否正确允许流量通过。

- 定义遇到的问题。例如，虚拟服务器的 **DNS** 记录正确，但未获取任何内容或内容不正确，等等。

问题

负载均衡器未正常工作。

解决方案

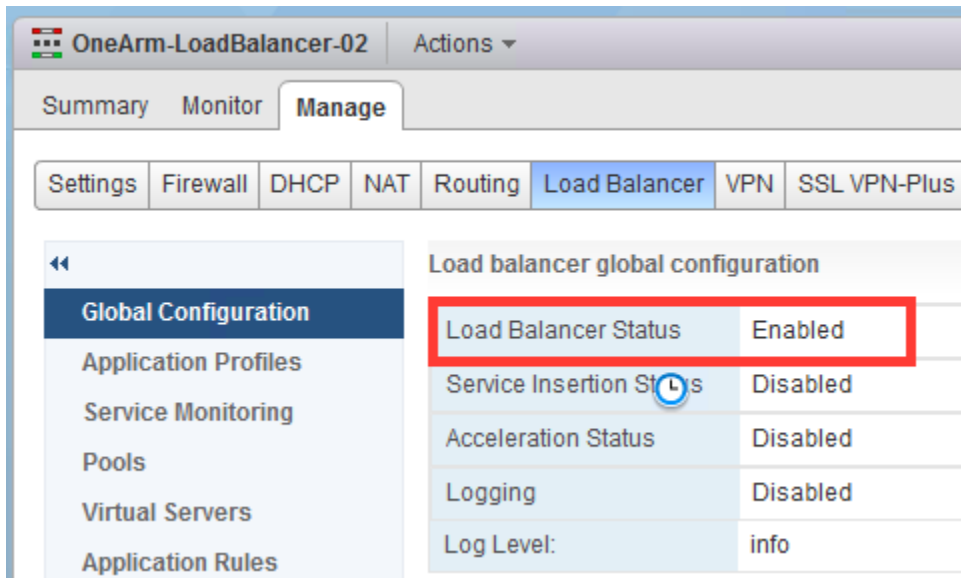
- 1 验证以下应用程序要求 - 需要在负载均衡器上支持的协议（**TCP**、**UDP**、**HTTP**、**HTTPS**）、端口、持久性要求以及池成员。
 - 是否启用了负载均衡器和防火墙，以及 **Edge** 服务网关是否具有正确的路由？
 - 虚拟服务器应该侦听哪些 **IP** 地址、端口和协议？
 - 是否使用 **SSL** 卸载？在与后端服务器通信时是否需要使用 **SSL**？
 - 是否使用应用程序规则？
 - 采用什么拓扑？**NSX** 负载均衡器需要分析来自客户端和服务器的所有流量。
 - **NSX** 负载均衡器是否为串联式负载均衡器，或者是否转换客户端源地址以确保返回流量传回到负载均衡器？

2 导航到 NSX Edge，并按以下方式验证启用负载平衡并允许流量通过所需的配置：

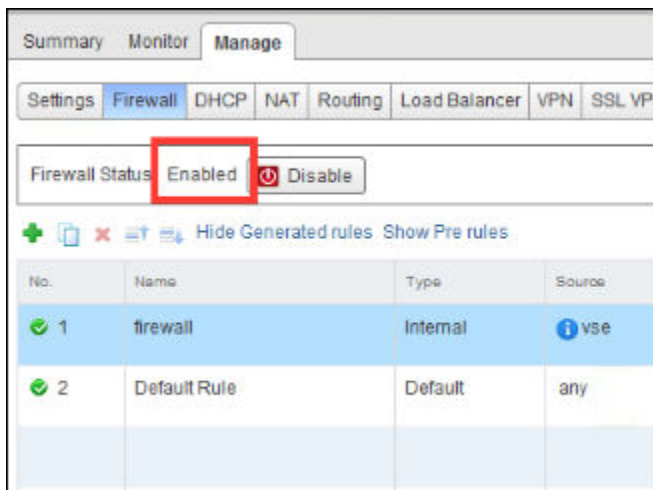
- a 验证负载平衡器是否列出为已启动 (Up)。

The screenshot displays the NSX Edge configuration interface. At the top, there are tabs for 'Summary', 'Monitor', and 'Manage'. Below these, there are sub-tabs for 'Settings', 'Firewall', 'DHCP', and 'NAT'. The 'Settings' tab is selected, and a sidebar on the left shows a menu with 'Configuration' (highlighted), 'Interfaces', and 'Certificates'. The main content area shows a table of configuration items with their status. The 'Load Balancer' row is highlighted in blue, and its status is 'Up', which is also highlighted with a yellow box. The status was last updated on Wednesday, August 17, 2016, at 11:28:58 AM.

Status last updated on: Wednesday, August 17, 2016 11:28:58 AM	
High Availability	Not Configured
Routing	Applied
Syslog	Up
SSL VPN-Plus	Not Configured
DNS	Not Configured
Load Balancer	Up



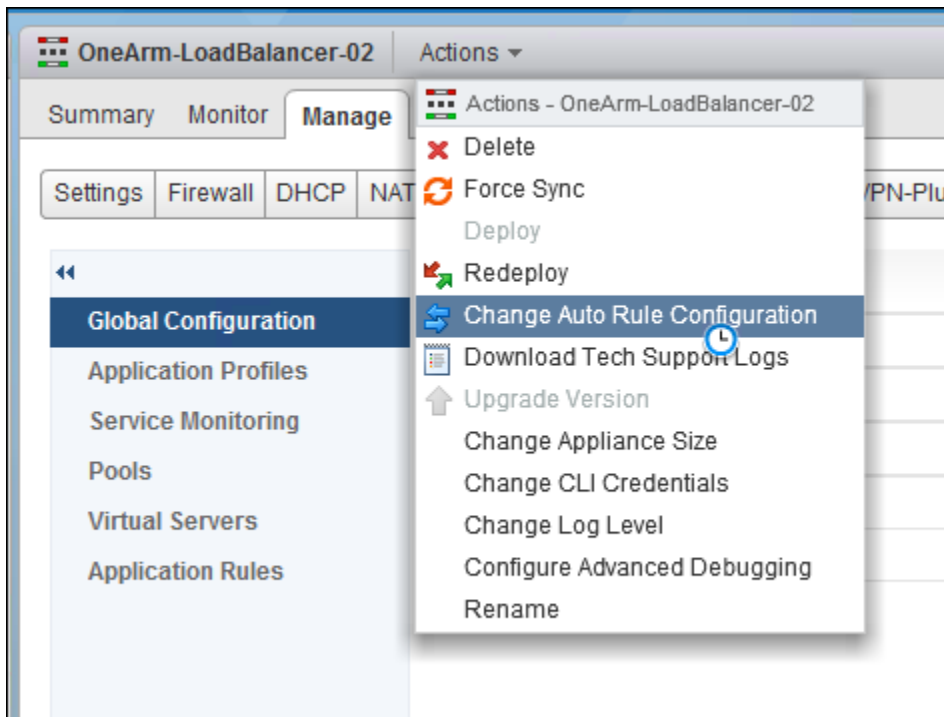
- b 验证防火墙是否为已启用 (**Enabled**)。必须为加速的虚拟服务器启用防火墙。非加速 TCP 和 L7 HTTP/HTTPS VIP 必须具有允许流量通过的策略。请注意，防火墙筛选器不影响加速的虚拟服务器。



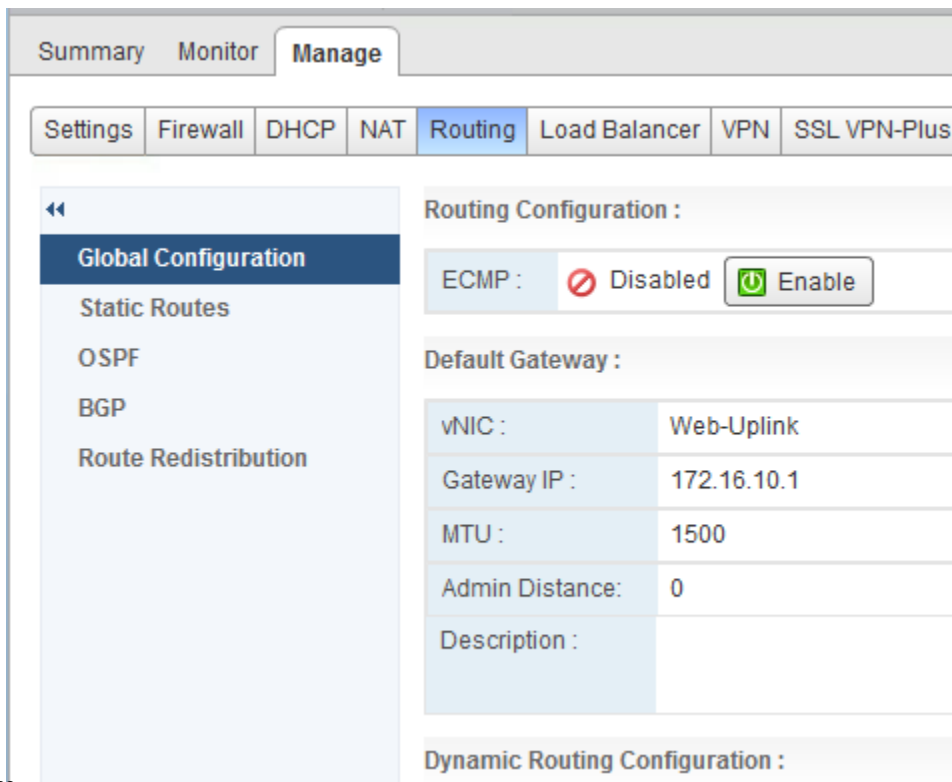
- c 验证是否为虚拟服务器创建了 NAT 规则。在 NAT 选项卡上，单击[隐藏内部规则 \(Hide internal rules\)](#)或[取消隐藏内部规则 \(Unhide internal rules\)](#)链接以进行验证。

注 如果启用了负载均衡并配置了服务，但未配置任何 NAT 规则，则意味着未启用自动规则配置。

- d 您可以更改自动规则配置。有关详细信息，请参阅《NSX 管理指南》中的“更改自动规则配置”主题。在部署 NSX Edge 服务网关时，您可以选择配置自动规则配置。如果在部署 Edge 服务网关时未选择该选项，您必须启用该选项以使负载均衡器正常工作。通过 UI 检查池成员状态。



- e 验证路由，并验证 Edge 服务网关是否具有到客户端系统和后端服务器的默认路由或静态路由。如果没有到服务器的路由，运行状况检查将失败。如果使用动态路由协议，则可能需要使用 CLI。有关详细信息，请参阅 [NSX 路由 CLI](#)。
- a 验证默认路由。



- b 验证连接的路由。在这些路由中，Edge 服务网关在子网中具有一个接口。很多时候，应用程序服

务器连接到这些服务器。

⚙️ 0 Job(s) In Progress
❗ 0 Job(s) Failed

aces of this NSX Edge.

⚙️ Actions

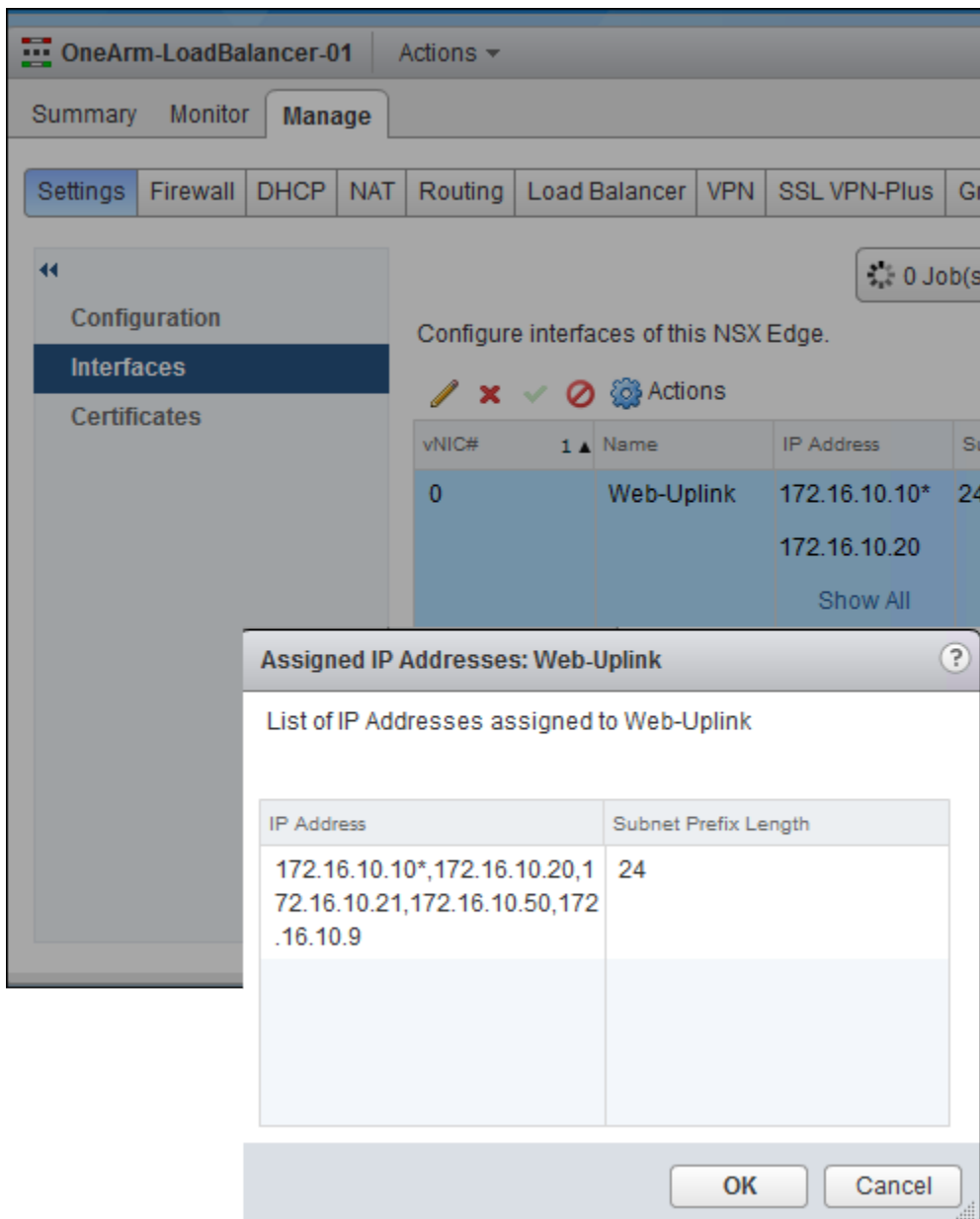
🔍 Filter

Name	IP Address	Subnet Prefix Length	Connected To	Type	Status
Web-Uplink	172.16.10.10*	24	Web-Tier-01	Uplink	✓
	172.16.10.20				
	Show All				
INLINE_SUBNI	172.16.100.1*	24	INLINE_SUBNI	Internal	✓
vnic2				Internal	❌
vnic3				Internal	❌
vnic4				Internal	❌
vnic5				Internal	❌

- c 从路由 (Routing)选项卡 > 静态路由 (Static Routes)中验证静态路由。

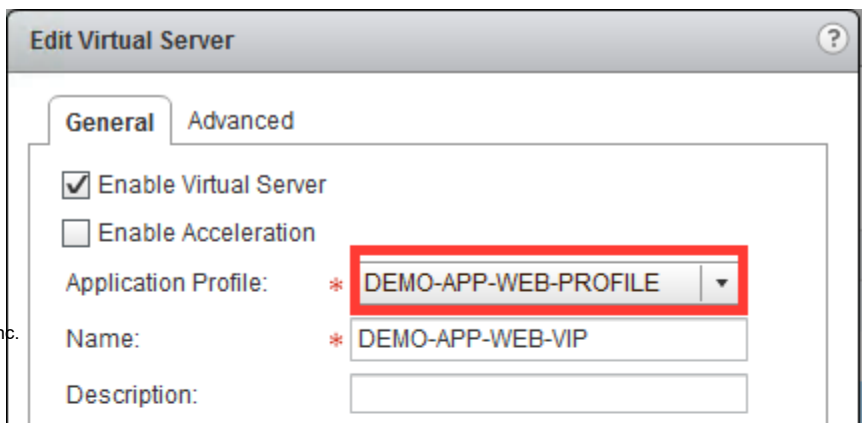
3 验证虚拟服务器的 IP 地址、端口和协议。

- a 双击一个 NSX Edge，然后导航到**管理 (Manage)** > **设置 (Settings)** > **接口 (Interfaces)**。验证是否将虚拟服务器的 IP 地址添加到接口中。



- b 验证虚拟服务器是否配置了正确的 IP 地址、端口和协议以支持应用程序。

- a 验证虚拟服务器使用的应用程序配置文件。



- c 验证应用程序配置文件是否符合支持的持久方法、类型（协议）和 **SSL**（如果需要）。如果使用 SSL，请确保使用具有正确名称和过期日期的证书。

Edit Profile

Name: DEMO-APP-WEB-PROFILE

Type: HTTPS

☐ Enable SSL Passthrough

HTTP Redirect URL:

Persistence: Source IP

Cookie Name:

Mode:

Expires in (Seconds):

☒ Insert X-Forwarded-For HTTP header

☐ Enable Pool Side SSL

Virtual Server Certificates Pool Certificates

Service Certificates CA Certificates CRL

☒ Configure Service Certificate

	Common Name	Issuer	Validity
<input checked="" type="radio"/>	DEMO.WEB.APP.CO	DEMO.WEB.APP.CO	Wed Apr 27 2016 - Sat
<input type="radio"/>	VSM_SOLUTION_71	VSM_SOLUTION_71	Tue Sep 8 2015 - Thu
<input type="radio"/>	VSM_SOLUTION_71	VSM_SOLUTION_71	Tue Sep 8 2015 - Thu
<input type="radio"/>	psc-01a.corp.local	CA	Thu Mar 12 2015 - Thu
<input type="radio"/>	VSM_SOLUTION_49	VSM_SOLUTION_49	Tue Sep 8 2015 - Thu
<input type="radio"/>	VSM_SOLUTION_49	VSM_SOLUTION_49	Tue Sep 8 2015 - Thu

Cipher: Default

Client Authentication: Ignore

OK Cancel

- d 验证客户端是否使用正确的证书进行连接。

Edit Profile

Name: DEMO-APP-WEB-PROFILE

Type: HTTPS

☐ Enable SSL Passthrough

HTTP Redirect URL:

Persistence: Source IP

Cookie Name:

Mode:

Expires in (Seconds):

☒ Insert X-Forwarded-For HTTP header

☒ Enable Pool Side SSL

Virtual Server Certificates **Pool Certificates**

Service Certificates CA Certificates CRL

☒ Configure Service Certificate

	Common Name	Issuer	Validity
<input checked="" type="radio"/>	DEMO.WEB.APP.COF	DEMO.WEB.APP.COF	Wed Apr 27 2016 - Sa
<input type="radio"/>	VSM_SOLUTION_71f	VSM_SOLUTION_71f	Tue Sep 8 2015 - Thu
<input type="radio"/>	VSM_SOLUTION_71f	VSM_SOLUTION_71f	Tue Sep 8 2015 - Thu
<input type="radio"/>	psc-01a.corp.local	CA	Thu Mar 12 2015 - Th
<input type="radio"/>	VSM_SOLUTION_49c	VSM_SOLUTION_49c	Tue Sep 8 2015 - Thu
<input type="radio"/>	VSM_SOLUTION_49c	VSM_SOLUTION_49c	Tue Sep 8 2015 - Thu

Cipher: Default

Client Authentication: Ignore

OK Cancel

- e 验证是否需要使用客户端证书，但未配置客户端。此外，还要验证选择的密码列表是否太窄（例如，使用旧浏览器的客户端）。

Edit Profile

Name: DEMO-APP-WEB-PROFILE

Type: HTTPS

☐ Enable SSL Passthrough

HTTP Redirect URL:

Persistence: Source IP

Cookie Name:

Mode:

Expires in (Seconds):

☒ Insert X-Forwarded-For HTTP header

☐ Enable Pool Side SSL

Virtual Server Certificates | Pool Certificates

Service Certificates | CA Certificates | CRL

☒ Configure Service Certificate

	Common Name	Issuer	Validity
<input checked="" type="radio"/>	DEMO.WEB.APP.CO	DEMO.WEB.APP.CO	Wed Apr 27 2016 - Sat
<input type="radio"/>	VSM_SOLUTION_71	VSM_SOLUTION_71	Tue Sep 8 2015 - Thu
<input type="radio"/>	VSM_SOLUTION_71	VSM_SOLUTION_71	Tue Sep 8 2015 - Thu
<input type="radio"/>	psc-01a.corp.local	CA	Thu Mar 12 2015 - Thu
<input type="radio"/>	VSM_SOLUTION_49	VSM_SOLUTION_49	Tue Sep 8 2015 - Thu
<input type="radio"/>	VSM_SOLUTION_49	VSM_SOLUTION_49	Tue Sep 8 2015 - Thu

Cipher: Default

Client Authentication: Ignore

OK Cancel

- f 验证是否需要通过 SSL 访问后端服务器。

Edit Profile

Name: DEMO-APP-WEB-PROFILE

Type: HTTPS

☐ Enable SSL Passthrough

HTTP Redirect URL:

4 按以下方式检查池状态和配置：

- a 验证池状态，必须启动至少一个成员以处理流量，但一个成员可能不足以处理所有流量。如果启动了零个或有限数量的池成员，请尝试按照后续步骤中所述纠正该问题。

Pool ID

Name

Pool and Member Status

Pool Status and Statistics:

Pool ID	Name	Status
pool-1	TENANT-1-TCP-P...	UP

Member Status and Statistics:

Name	IP Address / VC Container	Status	Member ID
SERVER-1	10.10.10.100	UP	member-1
SERVER-2	10.10.10.101	UP	member-2

- b 验证拓扑是否正确。SNAT 客户端流量是在池配置中控制的。如果托管负载均衡器功能的 Edge 服务网关不是串联式网关以查看所有流量，该网关将失败。要保留客户端源 IP，请选择**透明 (Transparent)**模式。有关信息，请参阅《NSX 管理指南》。

Edit Pool

Name:

* DEMO_APP_WEB_POOL

Description:

Algorithm:

ROUND-ROBIN

Algorithm Parameters:

Monitors:

default_http_monitor

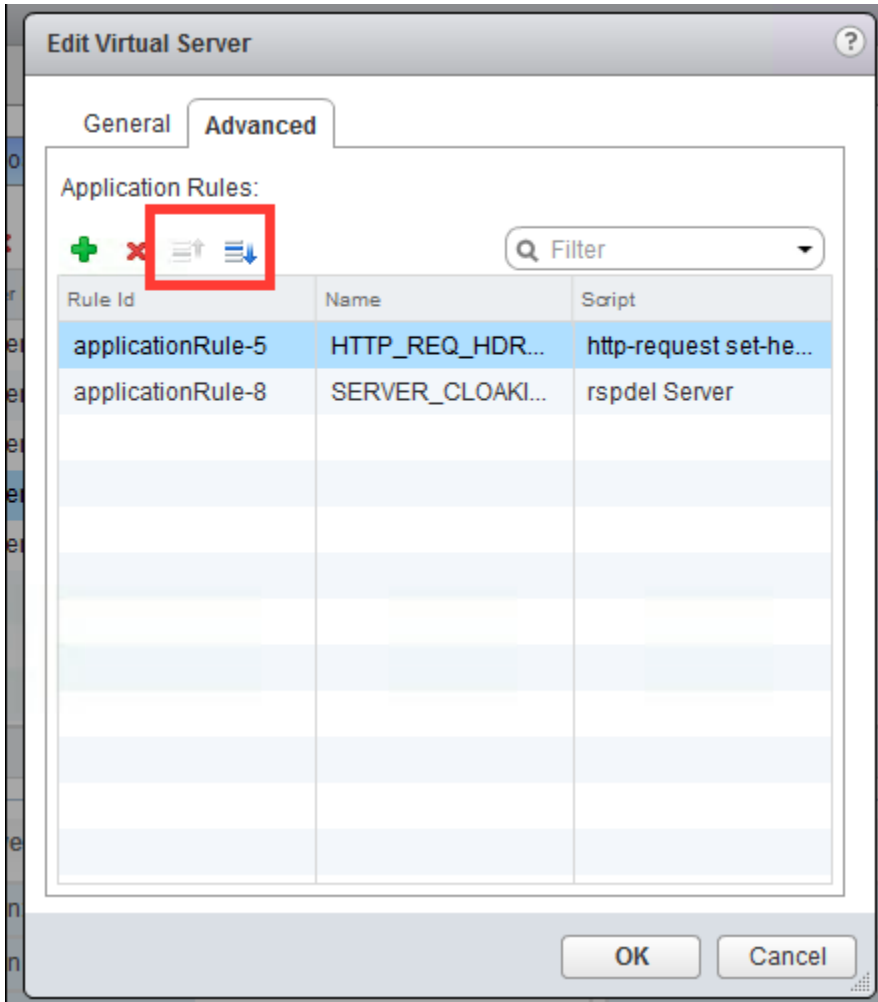
Members:

+

×

Enabled	Name	IP Address / VC Container	Weight	Monitor Port	Port	Max Connections	M C
✓	Server-2	172.16.1...	1	80		0	(
✓	Server-1	172.16.1...	5	80		0	(

- 5 如果使用应用程序规则，请验证这些规则。如果需要，请移除这些规则以查看是否传输流量。
- a 对这些规则进行重新排序，以查看规则顺序是否导致逻辑中断流量传输。有关如何添加应用程序规则和查看应用程序规则示例的信息，请参阅《NSX 管理指南》中的“添加应用程序规则”主题。



后续步骤

如果找不到问题，您可能需要使用 CLI（命令行界面）查明发生的情况。有关详细信息，请参阅[使用 CLI 的负载均衡器故障排除](#)。

使用 CLI 的负载均衡器故障排除

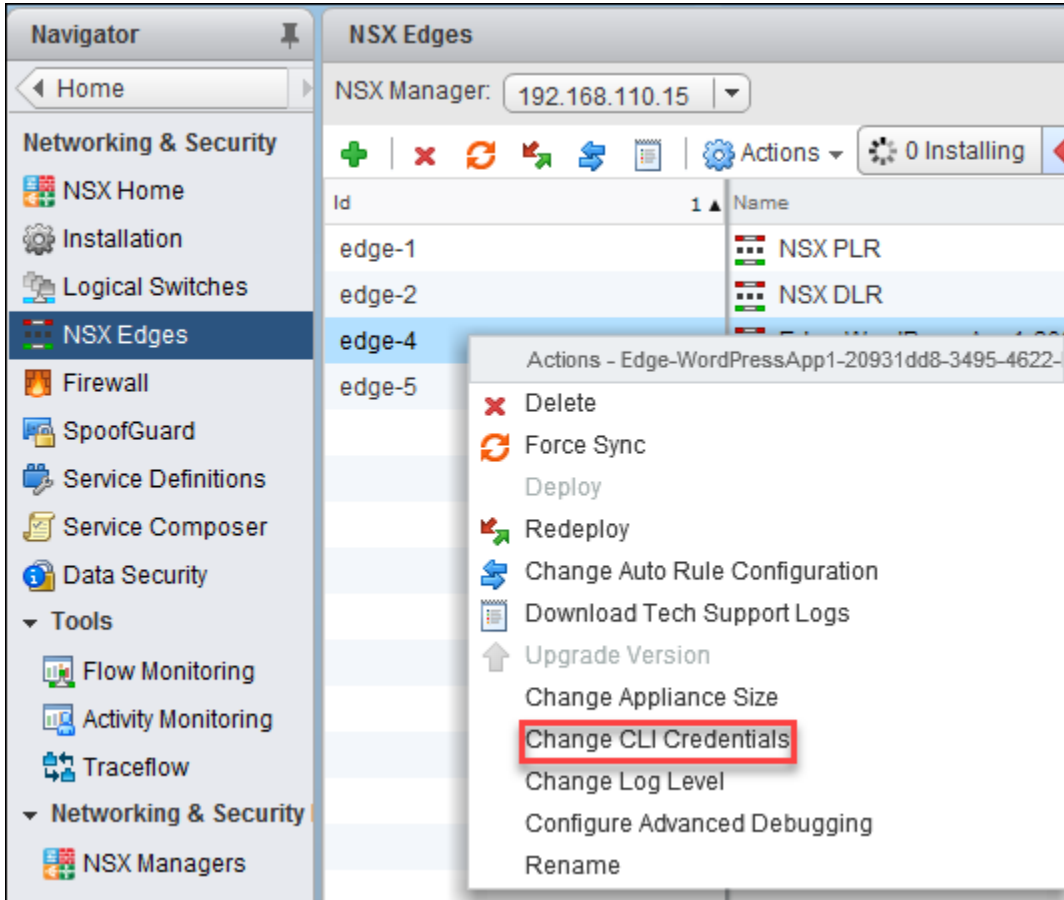
可以使用 NSX CLI 获取详细的跟踪日志，捕获数据包以及查看衡量指标以排除负载均衡器故障。

问题

负载均衡未正常工作。

解决方案

- 1 启用或确认您可以通过 SSH 访问虚拟设备。Edge 服务网关是一个虚拟设备，它可以在部署时选择启用 SSH。如果您需要启用 SSH，请选择所需的设备，然后在**操作 (Actions)**菜单中单击**更改 CLI 凭据 (Change CLI Credentials)**。



- 2 Edge 服务网关具有多个 show 命令以查看运行时状态和配置状态。请使用这些命令显示配置和统计信息。

```
nsxedge> show configuration loadbalancer
nsxedge> show configuration loadbalancer virtual [virtual-server-name]
nsxedge> show configuration loadbalancer pool [pool-name]
nsxedge> show configuration loadbalancer monitor [monitor-name]
nsxedge> show configuration loadbalancer profile [profile-name]
nsxedge> show configuration loadbalancer rule [rule-name]
```

- 3 要使负载平衡和 NAT 正常工作，应启用防火墙。请使用 `#show firewall` 命令。如果使用该命令没有看到任何有意义的输出，请参阅[使用 UI 验证负载平衡器配置和排除故障](#)部分。

```

ISX-edge-8-0> show firewall
Chain PREROUTING (policy ACCEPT 21947 packets, 7809K bytes)
:cid  pkts bytes target    prot opt in     out     source      destination

Chain INPUT (policy ACCEPT 0 packets, 0 bytes)
:cid  pkts bytes target    prot opt in     out     source      destination
)      348 67915 ACCEPT    all  --  lo     *       0.0.0.0/0    0.0.0.0/0
)      134 5360 DROP      all  --  *     *       0.0.0.0/0    0.0.0.0/0    state INVALID
)     21482 7736K block_in all  --  *     *       0.0.0.0/0    0.0.0.0/0
)     20545 7671K ACCEPT    all  --  *     *       0.0.0.0/0    0.0.0.0/0    state RELATED
)       937 65139 usr_rules all  --  *     *       0.0.0.0/0    0.0.0.0/0
)         0 0 DROP      all  --  *     *       0.0.0.0/0    0.0.0.0/0

Chain FORWARD (policy ACCEPT 0 packets, 0 bytes)
:cid  pkts bytes target    prot opt in     out     source      destination

Chain OUTPUT (policy ACCEPT 20673 packets, 1248K bytes)
:cid  pkts bytes target    prot opt in     out     source      destination

Chain POSTROUTING (policy ACCEPT 0 packets, 0 bytes)
:cid  pkts bytes target    prot opt in     out     source      destination
)      348 67915 ACCEPT    all  --  *     lo     0.0.0.0/0    0.0.0.0/0
)       34 1360 DROP      all  --  *     *     0.0.0.0/0    0.0.0.0/0    state INVALID
)     20295 1179K block_out all  --  *     *     0.0.0.0/0    0.0.0.0/0
)         0 0 ACCEPT    all  --  *     *     0.0.0.0/0    0.0.0.0/0    PHYSDEV match
)         0 0 ACCEPT    all  --  *     *     0.0.0.0/0    0.0.0.0/0    PHYSDEV match
)         0 0 ACCEPT    all  --  *     *     0.0.0.0/0    0.0.0.0/0    PHYSDEV match
)         0 0 ACCEPT    all  --  *     *     0.0.0.0/0    0.0.0.0/0    PHYSDEV match
)     14599 802K ACCEPT    all  --  *     *     0.0.0.0/0    0.0.0.0/0    state RELATED
)     5696 377K usr_rules all  --  *     *     0.0.0.0/0    0.0.0.0/0
)         0 0 DROP      all  --  *     *     0.0.0.0/0    0.0.0.0/0

Chain block_in (1 references)
:cid  pkts bytes target    prot opt in     out     source      destination

Chain block_out (1 references)
:cid  pkts bytes target    prot opt in     out     source      destination

Chain usr_rules (2 references)
:cid  pkts bytes target    prot opt in     out     source      destination
133137 4861 333K ACCEPT    all  --  *     *     0.0.0.0/0    0.0.0.0/0    match-set 0_
133138 0 0 ACCEPT    all  --  *     *     0.0.0.0/0    0.0.0.0/0    match-set 1_
133139 936 65099 ACCEPT    all  --  *     *     0.0.0.0/0    0.0.0.0/0    match-set 2_
133141 835 43459 ACCEPT    all  --  *     *     0.0.0.0/0    0.0.0.0/0    match-set 3_
133131 1 40 LOG      all  --  *     *     0.0.0.0/0    0.0.0.0/0    LOG flags 0
133131 1 40 ACCEPT    all  --  *     *     0.0.0.0/0    0.0.0.0/0

```


- 4 负载均衡器需要使用 NAT 才能正常工作。请使用 `show nat` 命令。如果使用该命令没有看到任何有意义的输出，请参阅[使用 UI 验证负载均衡器配置和排除故障](#)部分。

```

NSX-edge-8-0> show nat
Chain PREROUTING (policy ACCEPT 568 packets, 40044 bytes)
rid  pkts bytes target    prot opt in     out     source        destination
0      568 40044 int_dnat  all  --  *      *        0.0.0.0/0      0.0.0.0/0
0      568 40044 usr_dnat  all  --  *      *        0.0.0.0/0      0.0.0.0/0

Chain INPUT (policy ACCEPT 568 packets, 40044 bytes)
rid  pkts bytes target    prot opt in     out     source        destination

Chain OUTPUT (policy ACCEPT 896 packets, 46706 bytes)
rid  pkts bytes target    prot opt in     out     source        destination
0      896 46706 int_dnat  all  --  *      *        0.0.0.0/0      0.0.0.0/0
0      896 46706 usr_dnat  all  --  *      *        0.0.0.0/0      0.0.0.0/0

Chain POSTROUTING (policy ACCEPT 896 packets, 46706 bytes)
rid  pkts bytes target    prot opt in     out     source        destination
0      896 46706 int_snat  all  --  *      *        0.0.0.0/0      0.0.0.0/0
0      896 46706 usr_snat  all  --  *      *        0.0.0.0/0      0.0.0.0/0

Chain int_dnat (2 references)
rid  pkts bytes target    prot opt in     out     source        destination

Chain int_snat (1 references)
rid  pkts bytes target    prot opt in     out     source        destination
0      0      0 ACCEPT    all  --  *      *        0.0.0.0/0      0.0.0.0/0

Chain usr_dnat (2 references)
rid  pkts bytes target    prot opt in     out     source        destination
0      0      0 DNAT      tcp  --  vNic_2  *        0.0.0.0/0      192.168.8.20
0      0      0 LOG       all  --  vNic_2  *        0.0.0.0/0      192.168.8.11
0      0      0 DNAT      all  --  vNic_2  *        0.0.0.0/0      192.168.8.11

Chain usr_snat (1 references)
rid  pkts bytes target    prot opt in     out     source        destination
0      0      0 LOG       all  --  *      vNic_2  10.10.10.101   0.0.0.0/0
0      0      0 SNAT      all  --  *      vNic_2  10.10.10.101   0.0.0.0/0
0      0      0 LOG       all  --  *      vNic_2  10.10.10.0/24  0.0.0.0/0
0      0      0 SNAT      all  --  *      vNic_2  10.10.10.0/24  0.0.0.0/0
NSX-edge-8-0>

```

- 5 除了启用的防火墙以及具有 NAT 规则的负载均衡器以外，您还应该确保启用负载均衡进程。请使用 `show service loadbalancer` 命令检查负载均衡器引擎状态 (L4/L7)。

```

nsxedge> show service loadbalancer
haIndex:          0
-----
Loadbalancer Services Status:

L7 Loadbalancer   : running
-----
L7 Loadbalancer Statistics:
STATUS    PID      MAX_MEM_MB  MAX SOCK  MAX_CONN  MAX_PIPE  CUR_CONN  CONN_RATE
CONN_RATE_LIMIT MAX_CONN_RATE
running   1580      0          2081      1024      0         0         0
0         0
-----

```

```

L4 Loadbalancer Statistics:
MAX_CONN  ACT_CONN  INACT_CONN  TOTAL_CONN
0          0          0            0

Prot LocalAddress:Port Scheduler Flags
-> RemoteAddress:Port      Forward Weight ActiveConn InActConn

```

- a 使用 `show service loadbalancer session` 命令查看负载平衡器会话表。如果在系统上具有流量，则会看到会话。

```

nsxedge> show service loadbalancer session

-----
L7 Loadbalancer Statistics:
STATUS      PID      MAX_MEM_MB  MAX SOCK    MAX_CONN    MAX_PIPE    CUR_CONN    CONN_RATE
CONN_RATE_LIMIT MAX_CONN_RATE
running     1580      0           2081        1024        0           0           0
0           0

-----L7 Loadbalancer Current Sessions:

0x2192df1f300: proto=unix_stream src=unix:1 fe=GLOBAL be=<NONE> srv=<none> ts=09 age=0s
calls=2 rq[f=c08200h,
i=0,an=00h,rx=20s,wx=,ax=] rp[f=008000h,i=0,an=00h,rx=wx,ax=] s0=[7,8h,fd=1,ex=]
s1=[7,0h,fd=-1,ex=] exp=19s

-----
L4 Loadbalancer Statistics:
MAX_CONN  ACT_CONN  INACT_CONN  TOTAL_CONN
0          0          0            0

L4 Loadbalancer Current Sessions:

pro expire state      source      virtual      destination

```

- b 检查 `show service loadbalancer` 命令以查看负载平衡器第 7 层粘性表状态。请注意，该表不显示有关加速的虚拟服务器的信息。

```

nsxedge> show service loadbalancer table

-----
L7 Loadbalancer Sticky Table Status:

TABLE      TYPE      SIZE(BYTE)  USED(BYTE)

```

- 6 如果所需的所有服务正常运行，请查看路由表并需要具有到客户端和服务器的路由。请使用 `show ip route` 和 `show ip forwarding` 命令以将路由映射到接口。

```

NSX-edge-8-0> sh ip route

Codes: O - OSPF derived, i - IS-IS derived, B - BGP derived,
C - connected, S - static, L1 - IS-IS level-1, L2 - IS-IS level-2,
IA - OSPF inter area, E1 - OSPF external type 1, E2 - OSPF external type 2,
N1 - OSPF NSSA external type 1, N2 - OSPF NSSA external type 2

Total number of routes: 4

S      0.0.0.0/0          [1/1]          via 192.168.8.2
C      10.10.10.0/24      [0/0]          via 10.10.10.1
C      169.254.1.4/30     [0/0]          via 169.254.1.5
C      192.168.8.0/24     [0/0]          via 192.168.8.3
NSX-edge-8-0> sh ip forwarding
Codes: C - connected, R - remote,
> - selected route, * - FIB route

R>* 0.0.0.0/0 via 192.168.8.2, vNic_2
C>* 10.10.10.0/24 is directly connected, vNic_0
C>* 169.254.1.4/30 is directly connected, vNic_0
C>* 192.168.8.0/24 is directly connected, vNic_2
NSX-edge-8-0>

```

- 7 确保使用 `show arp` 命令获取系统（例如，网关或下一跃点）和后端服务器的 ARP 条目。

```

OneArm-LoadBalancer-01-0> show arp
-----
vShield Edge ARP Cache:
IP Address                Interface  MAC Address      State
fe80::250:56ff:feae:f86b  vNic_0    00:50:56:ae:f8:6b STALE
fe80::250:56ff:feae:5066  vNic_1    00:50:56:ae:50:66 STALE
fe80::250:56ff:feae:3e3d  vNic_0    00:50:56:ae:3e:3d STALE
172.16.100.11             vNic_1    00:50:56:ae:50:66 REACHABLE
172.16.10.1               vNic_0    02:50:56:56:44:52 REACHABLE
172.16.10.11              vNic_0    00:50:56:ae:3e:3d REACHABLE
OneArm-LoadBalancer-01-0>

```

- 8 这些日志提供了相应的信息以帮助查找流量，这可能有助于诊断问题。请使用 `show log` 或 `show log follow` 命令跟踪日志以帮助查找流量。请注意，您必须运行负载均衡器，启用日志记录 (Logging) 并将其设置为信息 (Info) 或调试 (Debug)。

```

nsxedge> show log
2016-04-20T20:15:36+00:00 vShieldEdge kernel: Initializing cgroup subsys cpuset
2016-04-20T20:15:36+00:00 vShieldEdge kernel: Initializing cgroup subsys cpu
2016-04-20T20:15:36+00:00 vShieldEdge kernel: Initializing cgroup subsys cpuacct
...

```

- 9 在确认基本服务使用正确的客户端路径运行后，让我们查看在应用程序层中发生的情况。请使用 `show service loadbalancer pool` 命令查看负载均衡器池状态 (L4/L7)。必须启动一个池成员以处理内容，通常需要使用多个池成员，因为请求数量超过单个工作负载的容量。如果运行状况监控是由内置运行状况检查提供的，在运行状况检查失败时，输出将显示 `last state change time` 和 `failure reason`。如果监控服务提供运行状况监控，除了上述两个输出以外，还会显示 `last check time`。

```
nsxedge> show service loadbalancer pool
-----
Loadbalancer Pool Statistics:

POOL Web-Tier-Pool-01
| LB METHOD round-robin
| LB PROTOCOL L7
| Transparent disabled
| SESSION (cur, max, total) = (0, 0, 0)
| BYTES in = (0), out = (0)
+-->POOL MEMBER: Web-Tier-Pool-01/web-01a, STATUS: UP
| | HEALTH MONITOR = BUILT-IN, default_https_monitor:L7OK
| | | LAST STATE CHANGE: 2016-05-16 07:02:00
| | SESSION (cur, max, total) = (0, 0, 0)
| | BYTES in = (0), out = (0)
+-->POOL MEMBER: Web-Tier-Pool-01/web-02a, STATUS: UP
| | HEALTH MONITOR = BUILT-IN, default_https_monitor:L7OK
| | | LAST STATE CHANGE: 2016-05-16 07:02:01
| | SESSION (cur, max, total) = (0, 0, 0)
| | BYTES in = (0), out = (0)
```

- 10 检查服务监控器状态（“正常”、“警告”、“严重”）以查看所有配置的后端服务器的运行状况。

```
nsxedge> show service loadbalancer monitor
-----
Loadbalancer Health Check Statistics:

MONITOR PROVIDER    POOL          MEMBER        HEALTH STATUS
built-in            Web-Tier-Pool-01  web-01a      default_https_monitor:L7OK
built-in            Web-Tier-Pool-01  web-02a      default_https_monitor:L7OK
```

对于 `show service load balancer monitor` 命令，将在 CLI 输出中显示三种类型的运行状况监控值：

- 内置：已启用运行状况检查并且由 L7 引擎（HA 代理）执行。
- 监控服务：已启用运行状况检查并由监控服务引擎 (NAGIOS) 执行。可以使用 `show service monitor` 和 `show service monitor service` CLI 命令检查监控服务运行状态。状态 (Status) 字段应该为正常、警告或严重。
- 未定义：已禁用运行状况检查。

输出的最后一列是池成员的运行状况。将显示以下状态：

表 6-1. 运行状况和说明

运行状况	说明
内置	<ul style="list-style-type: none"> ■ UNK: 未知 ■ INI: 正在初始化 ■ SOCKERR: 套接字错误 ■ L4OK: 已在第 4 层上通过检查, 未启用上面的层测试 ■ L4TOUT: 第 1-4 层超时 ■ L4CON: 第 1-4 层连接问题。例如, “Connection refused” (tcp rst) 或 “No route to host” (icmp) ■ L6OK: 已在第 6 层上通过检查 ■ L6TOUT: 第 6 层 (SSL) 超时 ■ L6RSP: 第 6 层响应无效 - 协议错误。原因可能是: <ul style="list-style-type: none"> ■ 后端服务器仅支持 “SSLv3” 或 “TLSv1.0”, ■ 后端服务器的证书无效, 或者 ■ 密码协商失败, 等等 ■ L7OK: 已在第 7 层上通过检查 ■ L7OKC: 已在第 7 层上有条件地通过检查。例如, 404 with disable-on-404 ■ L7TOUT: 第 7 层 (HTTP/SMTP) 超时 ■ L7RSP: 第 7 层响应无效 - 协议错误。 ■ L7STS: 第 7 层响应错误。例如, HTTP 5xx
严重	<ul style="list-style-type: none"> ■ 您的 SSL 库不支持 SSL 协议版本 2 ■ 不支持的 SSL 协议版本 ■ 无法创建 SSL 上下文 ■ 无法建立 SSL 连接 ■ 无法启动 SSL 握手 ■ 无法检索服务器证书 ■ 无法检索证书使用者 ■ 证书中的时间格式不正确 ■ 证书 “<cn>” 在 <expire time of certificate> 时过期 ■ 证书 “<cn>” 在今天的 <expire time of certificate> 时过期
警告/严重	证书 “<cn>” 在 <days_left/expire time of certificate> 天后过期

表 6-1. 运行状况和说明（续）

运行状况	说明
ICMP	<ul style="list-style-type: none"> ■ 无法访问网络 ■ 无法访问主机 ■ 无法访问协议 ■ 无法访问端口 ■ 源路由失败 ■ 已隔离源主机 ■ 未知的网络 ■ 未知的主机 ■ 网络被拒绝 ■ 主机被拒绝 ■ 网络的服务类型 (ToS) 不正确 ■ 主机的服务类型 (ToS) 不正确 ■ 已按筛选器禁止 ■ 主机优先级冲突 ■ 优先级临界值。该操作所需的最低优先级 ■ 无效的代码
UDP/TCP	<ul style="list-style-type: none"> ■ 套接字创建失败 ■ 连接到地址 <code>xxxx</code> 和端口 <code>xxx</code>: [请参阅 Linux 错误代码] ■ 未从主机收到数据 ■ 从主机/套接字收到意外响应
HTTP/HTTPS	<ul style="list-style-type: none"> ■ HTTP UNKNOWN: 内存分配错误 ■ HTTP CRITICAL: 无法打开 TCP 套接字（创建套接字或连接到服务器失败） ■ HTTP CRITICAL: 在接收数据时出错 ■ HTTP CRITICAL: 未从主机收到数据 ■ HTTP CRITICAL: 从主机收到的 HTTP 响应无效: <code><status line></code>（不正确的预期状态行格式） ■ HTTP CRITICAL: 无效的状态行 <code><status line></code>（状态代码不是 3 个数字: XXX） ■ HTTP CRITICAL: 无效的状态 <code><status line></code>（状态代码 ≥ 600 或 < 100） ■ HTTP CRITICAL: 找不到字符串 ■ HTTP CRITICAL: 找不到模式 ■ HTTP WARNING: 页面大小 <code><page_length></code> 太大 ■ HTTP WARNING: 页面大小 <code><page_length></code> 太小

11 如果错误代码为 L4TOUT/L4CON，这通常是底层网络上的连接问题。Duplicate IP 通常是此类问题的根本原因。在出现该错误时，请按以下方式进行故障排除：

- a 在两个 Edge 上使用 `show service highavailability` 命令启用高可用性 (High Availability, HA) 时，检查这些 Edge 的 HA 状态。检查 HA 链路是否为 DOWN 并且所有 Edge 为 Active，因此，在网络上没有重复的 Edge IP。
- b 通过 `show arp` 命令检查 Edge ARP 表，并验证后端服务器的 ARP 条目是否在两个 MAC 地址之间变化。
- c 检查后端服务器 ARP 表，或者使用 `arp-ping` 命令并检查任何其他计算机是否具有与 Edge IP 类似的相同 IP。

- 12** 检查负载均衡器对象统计信息（VIP、池、成员）。查看特定的池并确认成员已启动并正在运行。检查是否启用了透明模式。如果启用，Edge 服务网关应该是客户端和服务端之间的串联式网关。验证服务器是否显示会话计数器增加值。

```
nsxedge> show service loadbalancer pool Web-Tier-VIP-01
```

TIMESTAMP	SESSIONS	BYTESIN	BYTESOUT	SESSIONRATE	HTTPREQS
2016-04-27 19:56:40	00	00	00	00	00
2016-04-27 19:55:00	00	32	100	00	00

```
nsxedge> show service loadbalancer pool Web-Tier-VIP-01 | MEMBER
+--> POOL MEMBER: TENANT-1-TCP-POOL-80/SERVER-1, STATUS:  UP
+--> POOL MEMBER: TENANT-1-TCP-POOL-80/SERVER-2, STATUS:  UP
```

- 13** 现在，查看虚拟服务器，验证是否具有默认池并查看该池是否还绑定到该服务器。如果通过应用程序规则使用池，您需要查看 `#show service loadbalancer pool` 命令中显示的特定池。指定虚拟服务器的名称。

```
nsxedge> show service loadbalancer virtual Web-Tier-VIP-01
```

Loadbalancer VirtualServer Statistics:

```

VIRTUAL Web-Tier-VIP-01
| ADDRESS [172.16.10.10]:443
| SESSION (cur, max, total) = (0, 0, 0)
| RATE (cur, max, limit) = (0, 0, 0)
| BYTES in = (0), out = (0)
+-->POOL Web-Tier-Pool-01
| LB METHOD round-robin
| LB PROTOCOL L7
| Transparent disabled
| SESSION (cur, max, total) = (0, 0, 0)
| BYTES in = (0), out = (0)
+-->POOL MEMBER: Web-Tier-Pool-01/web-01a, STATUS: UP
| | HEALTH MONITOR = BUILT-IN, default_https_monitor:L7OK
| | | LAST STATE CHANGE: 2016-05-16 07:02:00
| | SESSION (cur, max, total) = (0, 0, 0)
| | BYTES in = (0), out = (0)
+-->POOL MEMBER: Web-Tier-Pool-01/web-02a, STATUS: UP
| | HEALTH MONITOR = BUILT-IN, default_https_monitor:L7OK
| | | LAST STATE CHANGE: 2016-05-16 07:02:01
| | SESSION (cur, max, total) = (0, 0, 0)
| | BYTES in = (0), out = (0)

```

- 14** 如果正确配置了所有内容，但仍出现错误，您应该捕获流量以了解发生的情况。共有两个连接：客户端到虚拟服务器以及 **Edge** 服务网关到后端池（在池级别具有或没有透明配置）。**#show ip forwarding** 命令列出了 vNic 接口，并且您可以使用该数据。

例如，假设客户端计算机位于 **vNic_0** 上，而服务器位于 **vNic_1** 上。您使用在端口 **80** 上运行的客户端 IP 地址 **192.168.1.2** 和 VIP IP **192.168.2.2**。负载均衡器接口 IP 为 **192.168.3.1**，后端服务器 IP 为 **192.168.3.3**。共有两个不同的数据包捕获命令：一个命令显示数据包，另一个命令将数据包捕获到一个可下载的文件中。请捕获数据包以检测负载均衡器异常故障。您可以捕获来自两个方向的数据包：

- 捕获来自客户端的数据包。
- 捕获发送到后端服务器的数据包。

```
#debug packet capture interface interface-name [filter using _ for space]- creates a packet
capture file that you can download
#debug packet display interface interface-name [filter using _ for space]- outputs packet data to
the console
#debug show files - to see a list of packet capture
#debug copy scp user@url:path file-name/all - to download the packet capture
```

例如：

- 在 vNIC_0 上捕获：debug packet display interface vNic_0
- 在所有接口上捕获：debug packet display interface any
- 使用筛选器在 vNIC_0 上捕获：debug packet display interface vNic_0
host_192.168.11.3_and_host_192.168.11.41
- 捕获客户端到虚拟服务器流量的数据包：#debug packet display|capture interface vNic_0
host_192.168.1.2_and_host_192.168.2.2_and_port_80
- 在 Edge 服务网关和池处于透明模式的服务器之间捕获数据包：#debug packet display|capture
interface vNic_1 host 192.168.1.2_and_host_192.168.3.3_and_port_80
- 在 Edge 服务网关和池未处于透明模式的服务器之间捕获数据包：#debug packet display|
capture interface vNic_1 host 192.168.3.1_and_host_192.168.3.3_and_port_80

常见的负载均衡器问题

本主题讨论了一些问题以及如何解决这些问题。

在使用 NSX 负载均衡时，通常会出现以下问题：

- TCP 端口（例如，端口 **443**）上的负载均衡无法正常工作。
 - 验证拓扑。有关详细信息，请参见《NSX 管理指南》。
 - 验证是否可以执行 ping 操作以访问虚拟服务器 IP 地址，或者查看上游路由器以确保填充了 ARP 表。
 - [使用 UI 验证负载均衡器配置和排除故障。](#)
 - [使用 CLI 的负载均衡器故障排除。](#)

- 捕获数据包。
- 未利用负载平衡池的某个成员。
 - 验证服务器是否位于池中并启用以及监控运行状况。
- Edge 流量未实现负载平衡。
 - 验证池和持久性配置。如果配置了持久性并使用少量客户端，您可能会发现未将连接平均分配到后端池成员。
- 第 7 层负载平衡引擎停止。
- 运行状况监控引擎停止。
 - 启用负载平衡器服务。请参阅《NSX 管理指南》。
- 池成员监控状态为“警告/严重”。
 - 验证是否可以从负载平衡器中访问应用程序服务器。
 - 验证应用程序服务器防火墙或 DFW 是否允许流量通过。
 - 确保应用程序服务器能够响应指定的运行状况探查。
- 池成员处于“非活动”状态。
 - 验证是否在池配置中启用了池成员。
- 第 7 层粘性表与备用 Edge 不同步。
 - 确保配置了 HA。
- 具有客户端连接，但无法完成应用程序事务。
 - 验证是否在应用程序配置文件中配置了正确的持久性。
 - 如果应用程序仅使用池中的一个服务器（而不是两个），则很可能会出现持久性问题。

基本故障排除

- 1 在 vSphere Web Client 中检查负载平衡器配置状态：
 - a 单击**网络和安全 > NSX Edge (Networking & Security > NSX Edges)**。
 - b 双击一个 NSX Edge。
 - c 单击**管理 (Manage)**，然后单击**负载平衡器 (Load Balancer)**选项卡。
 - d 检查负载平衡器状态和配置的日志记录级别。
- 2 在排除负载平衡器服务故障之前，请在 NSX Manager 上运行以下命令以确保该服务已启动并正在运行：

```
nsxmgr> show edge edge-4 service loadbalancer
haIndex:          0
-----
Loadbalancer Services Status:
```

```

L7 Loadbalancer      : running
-----
L7 Loadbalancer Statistics:
STATUS      PID      MAX_MEM_MB  MAX SOCK   MAX_CONN  MAX_PIPE  CUR_CONN  CONN_RATE
CONN_RATE_LIMIT MAX_CONN_RATE
running     1580      0          2081      1024      0         0         0
0           0
-----
L4 Loadbalancer Statistics:
MAX_CONN  ACT_CONN  INACT_CONN  TOTAL_CONN
0         0         0           0

Prot LocalAddress:Port Scheduler Flags
-> RemoteAddress:Port      Forward Weight ActiveConn InActConn

```

注 您可以运行 `show edge all` 以查找 NSX Edge 名称。

解决配置问题

如果 NSX 用户界面或 REST API 调用拒绝负载平衡器配置操作，则会将其归类为配置问题。

解决数据层面问题

NSX Manager 接受负载平衡器配置，但在客户端和 Edge 负载平衡服务器之间出现连接或性能问题。数据层面问题还包括负载平衡器运行时 CLI 问题以及负载平衡器系统事件问题。

- 1 使用以下 REST API 调用将 NSX Manager 中的 Edge 日志记录级别从 INFO 更改为 TRACE 或 DEBUG。

```

URL: https://NSX_Manager_IP/api/1.0/services/debug/loglevel/com.vmware.vshield.edge?level=TRACE
Method: POST

```

- 2 在 vSphere Web Client 中检查池成员状态。
 - a 单击**网络和安全 > NSX Edge (Networking & Security > NSX Edges)**。
 - b 双击一个 NSX Edge。
 - c 单击**管理 (Manage)**，然后单击**负载平衡器 (Load Balancer)**选项卡。
 - d 单击**池 (Pools)**以查看配置的负载平衡器池摘要。
 - e 选择负载平衡器池。单击**显示池统计信息 (Show Pool Statistics)**，然后验证池状态是否为“已启动”。
- 3 通过使用以下 REST API 调用，您可以从 NSX Manager 中获取更详细的负载平衡器池配置统计信息：

```

URL: https://NSX_Manager_IP/api/4.0/edges/{edgeId}/loadbalancer/statistics
Method: GET

<?xml version="1.0" encoding="UTF-8"?>

```

```

<loadBalancerStatusAndStats>
  <timeStamp>1463507779</timeStamp>
  <pool>
    <poolId>pool-1</poolId>
    <name>Web-Tier-Pool-01</name>
    <member>
      <memberId>member-1</memberId>
      <name>web-01a</name>
      <ipAddress>172.16.10.11</ipAddress>
      <status>UP</status>
      <lastStateChangeTime>2016-05-16 07:02:00</lastStateChangeTime>
      <bytesIn>0</bytesIn>
      <bytesOut>0</bytesOut>
      <curSessions>0</curSessions>
      <httpReqTotal>0</httpReqTotal>
      <httpReqRate>0</httpReqRate>
      <httpReqRateMax>0</httpReqRateMax>
      <maxSessions>0</maxSessions>
      <rate>0</rate>
      <rateLimit>0</rateLimit>
      <rateMax>0</rateMax>
      <totalSessions>0</totalSessions>
    </member>
    <member>
      <memberId>member-2</memberId>
      <name>web-02a</name>
      <ipAddress>172.16.10.12</ipAddress>
      <status>UP</status>
      <lastStateChangeTime>2016-05-16 07:02:01</lastStateChangeTime>
      <bytesIn>0</bytesIn>
      <bytesOut>0</bytesOut>
      <curSessions>0</curSessions>
      <httpReqTotal>0</httpReqTotal>
      <httpReqRate>0</httpReqRate>
      <httpReqRateMax>0</httpReqRateMax>
      <maxSessions>0</maxSessions>
      <rate>0</rate>
      <rateLimit>0</rateLimit>
      <rateMax>0</rateMax>
      <totalSessions>0</totalSessions>
    </member>
    <status>UP</status>
    <bytesIn>0</bytesIn>
    <bytesOut>0</bytesOut>
    <curSessions>0</curSessions>
    <httpReqTotal>0</httpReqTotal>
    <httpReqRate>0</httpReqRate>
    <httpReqRateMax>0</httpReqRateMax>
    <maxSessions>0</maxSessions>
    <rate>0</rate>
    <rateLimit>0</rateLimit>
    <rateMax>0</rateMax>
    <totalSessions>0</totalSessions>
  </pool>
</virtualServer>

```

```

<virtualServerId>virtualServer-1</virtualServerId>
<name>Web-Tier-VIP-01</name>
<ipAddress>172.16.10.10</ipAddress>
<status>OPEN</status>
<bytesIn>0</bytesIn>
<bytesOut>0</bytesOut>
<curSessions>0</curSessions>
<httpReqTotal>0</httpReqTotal>
<httpReqRate>0</httpReqRate>
<httpReqRateMax>0</httpReqRateMax>
<maxSessions>0</maxSessions>
<rate>0</rate>
<rateLimit>0</rateLimit>
<rateMax>0</rateMax>
<totalSessions>0</totalSessions>
</virtualServer>
</loadBalancerStatusAndStats>

```

- 4 要从命令行中检查负载均衡器统计信息，请在 NSX Edge 上运行以下命令。

对于特定的虚拟服务器：先运行 `show service loadbalancer virtual` 以获取虚拟服务器名称，然后运行 `show statistics loadbalancer virtual <virtual-server-name>`。

对于特定的 TCP 池：先运行 `show service loadbalancer pool` 以获取池名称，然后运行 `show statistics loadbalancer pool <pool-name>`。

- 5 查看负载均衡器统计信息以查找故障迹象。

虚拟专用网络 (VPN) 故障排除

7

NSX Edge 支持多种类型的 VPN。本故障排除部分介绍了如何对 L2 VPN 和 SSL VPN 问题进行故障排除。

本章讨论了以下主题：

- L2 VPN
- SSL VPN
- IPSEC VPN

L2 VPN

通过使用 L2 VPN，您可以将多个逻辑 L2 网络（VLAN 和 VXLAN）延伸到 L3 边界以外，这些网络是在 SSL VPN 中建立隧道的。此外，您还可以在 L2 VPN 服务器上配置多个站点。当虚拟机在站点之间移动时，它们仍位于同一子网上，且其 IP 地址保持不变。您还可以选择在远程站点上部署单独的 Edge，而不在该站点上启用 NSX。输出优化允许 Edge 将发送的任何数据包路由到本地输出优化 IP 地址，并桥接任何其他内容。

因此，L2 VPN 允许企业在分隔的不同地理位置之间无缝迁移受 VXLAN 或 VLAN 支持的工作负载。对于云提供程序，L2 VPN 向加入租户提供一种机制，这种机制无需修改工作负载和应用程序的现有 IP 地址。

L2 VPN 常见配置问题

本主题讨论了与 L2 VPN 相关的常见配置问题。

问题

下面是常见的配置问题：

- 配置了 L2 VPN 客户端，但面向 Internet 的防火墙不允许流量通过目标端口 443 在隧道中流动。
- 将 L2 VPN 客户端配置为验证服务器证书，但没有为该客户端配置正确的 CA 证书或 FQDN。
- 配置了 L2 VPN 服务器，但未在面向 Internet 的防火墙上创建 NAT/防火墙规则。

- 分布式端口组或标准端口组不支持中继接口。

注 默认情况下，L2 VPN 服务器侦听端口 443。可以从 L2 VPN 服务器设置中配置该端口。

默认情况下，L2 VPN 客户端建立到端口 443 的出站连接。可以从 L2 VPN 客户端设置中配置该端口。

解决方案

- 1 检查 L2 VPN 服务器进程是否正在运行。
 - a 登录到 NSX Edge 虚拟机。
 - b 运行 `show process monitor` 命令，并验证您是否可以找到名为 `l2vpn` 的进程。
 - c 运行 `show service network-connections` 命令，并验证 `l2vpn` 进程是否侦听端口 443。
- 2 检查 L2 VPN 客户端进程是否正在运行。
 - a 登录到 NSX Edge 虚拟机。
 - b 运行 `show process monitor` 命令，并验证您是否可以找到名为 `naclientd` 的进程。
 - c 运行 `show service network-connections` 命令，并验证 `naclientd` 进程是否侦听端口 443。
- 3 检查是否可以从 Internet 中访问 L2 VPN 服务器。
 - a 打开浏览器，然后访问 `https://<l2vpn-public-ip>`。
 - b 将显示一个门户登录页面。如果显示门户页面，则表示可以通过 Internet 访问 L2 VPN 服务器。
- 4 检查分布式端口组或标准端口组是否支持中继接口。
 - a 如果分布式端口组支持中继接口，则会自动设置一个池端口。
 - b 如果标准端口组支持中继接口，您应该按以下方式手动配置 vSphere Distributed Switch:
 - 将端口设置为**混杂 (promiscuous)**模式。
 - 将**伪传输 (Forged Transmits)**设置为**接受 (Accept)**。
- 5 缓解 L2 VPN 循环问题。
 - a 如果未正确配置网卡绑定，则会发现两个严重问题 - MAC 地址漂移和重复的数据包。请按照[用于缓解循环的 L2VPN 选项](#)中所述验证配置。
- 6 检查跨 L2 VPN 的虚拟机是否可以相互通信。
 - a 登录到 L2 VPN 服务器 CLI，并在相应的 tap 接口上捕获数据包：`debug packet capture interface name`。
 - b 登录到 L2 VPN 客户端，并在相应的 tap 接口上捕获数据包：`debug packet capture interface name`。
 - c 分析这些捕获结果以检查是否解析了 ARP 并传输数据流量。

- d 检查 Allow Forged Transmits: dvSwitch 属性是否设置为 *L2 VPN 中继端口*。
- e 检查池端口是否设置为 *L2 VPN 中继端口*。为此，请登录到主机并发出 `net-dvs -l` 命令。检查为 L2 VPN Edge 内部端口设置的 sink 属性 (`com.vmware.etherswitch.port.extraEthFRP = SINK`)。内部端口是指 NSX Edge 中继连接到的 *dvPort*。

net-dvs -l
ESXi

```

port 939:
  com.vmware.common.port.alias = , propType = CONFIG
  com.vmware.common.port.connectid = 323234212 , propType = CONFIG
  com.vmware.common.port.portgroupid = dvportgroup-181 , propType = CONFIG
  com.vmware.common.port.block = false , propType = CONFIG
  com.vmware.common.port.dvfilter = filters (num = 0):
    propType = CONFIG
  com.vmware.common.port.ptAllowed = 0x 0. 0. 0. 0
    propType = CONFIG
  com.vmware.etherswitch.port.txUplink = normal , propType = CONFIG
  com.vmware.common.port.volatile.persist = /vmfs/volumes/9ec6ae8b-38b8e621/.dvsData/1c ec 0e 50 02 9c a9 21-b6 d5
fc 73 e5 79 69/939 , propType = CONFIG
  com.vmware.common.port.ptAllowedRT = 0x 0. 0. 0. 0
    propType = RUNTIME
  com.vmware.net.vxlan.trunkcfg = 0x63.6f.6e.66.69.67.56.65.72.73.69.6f.6e.3d.30.2e.31.3b.61.6c.6c.6f.77.47.75.65.7
74.56.6c.61.6e.3d.30.3b.6e.75.6d.54.72.75.6e.6b.4d.65.6d.62.65.72.73.3d.31.3b.74.72.75.6e.6b.4d.65.6d.5f.30.5f.43.70.45.6e.61.62.
.65.64.3d.31.3b.74.72.75.6e.6b.4d.65.6d.5f.30.5f.56.6e.69.3d.35.30.30.31.3b.74.72.75.6e.6b.4d.65.6d.5f.30.5f.4d.63.61.73.74.49.70
d.30.2e.30.2e.30.2e.31.3b
    propType = CONFIG POLICY
  com.vmware.etherswitch.port.extraEthFRP = SINK
    propType = CONFIG POLICY
  com.vmware.etherswitch.port.teaming:
    load balancing = first uplink (i.e. explicit)
    link selection = link state up;
    link behavior = notify switch; best effort on failure; shotgun on failure;
    active = dvUplink1;
    standby =
    propType = CONFIG
  com.vmware.etherswitch.port.security = deny promiscuous; deny mac change; allow forged frames
    propType = CONFIG
  com.vmware.etherswitch.port.vlan = Guest VLAN tagging
    ranges = 0
    propType = CONFIG
  com.vmware.common.port.statistics:
    pktsInUnicast = 0
    bytesInUnicast = 0
    pktsInMulticast = 6
    bytesInMulticast = 620
  
```

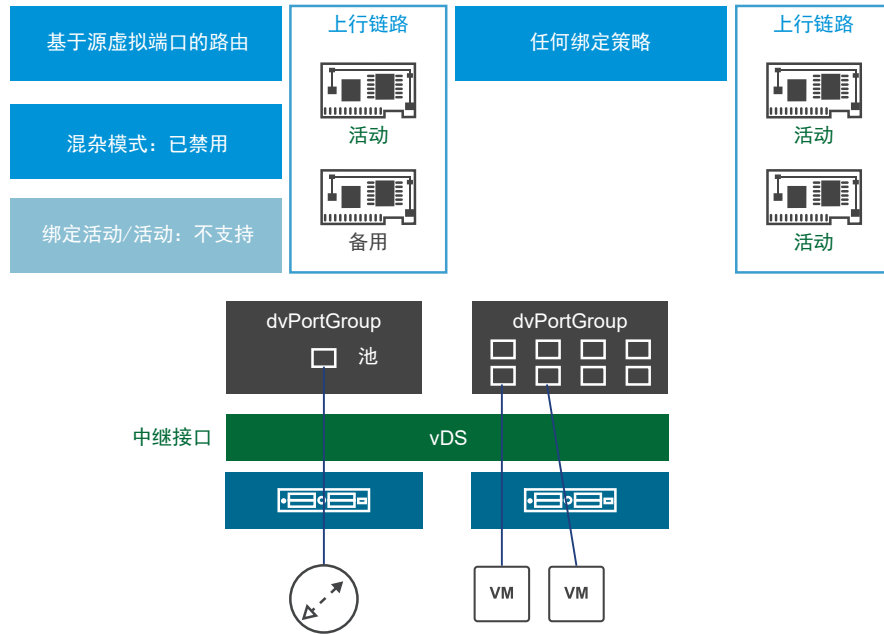
Sink port should be enabled for
the dvPort where the Edge trunk
is connected to

用于缓解循环的 L2VPN 选项

以下两个选项可用于缓解循环。NSX Edge 和虚拟机可以位于不同的 ESXi 主机上，也可以位于同一 ESXi 主机上。

选项 1: 将 L2VPN Edge 和虚拟机的各自 ESXi 主机分开

1. 在单独的 ESXi 主机上部署 L2VPN Edge 和虚拟机



- 1 将 Edge 和虚拟机部署在单独的 ESXi 主机上。
- 2 为与 Edge 的中继虚拟网卡关联的分布式端口组配置绑定策略和故障切换策略，如下所示：
 - a 以“基于源虚拟端口的路由”方式开展负载均衡。
 - b 仅将一个上行链路配置为“活动”，而将另一个上行链路配置为“备用”。
- 3 为与虚拟机关联的分布式端口组配置绑定策略和故障切换策略，如下所示：
 - a 任何绑定策略均可。
 - b 可以配置多个活动上行链路。

4 将 Edge 配置为使用池端口模式，并在中继虚拟网卡上禁用混杂模式。

注

- 禁用混杂模式：如果使用 vSphere Distributed Switch。
- 启用混杂模式：如果使用虚拟交换机配置中继接口。

如果虚拟交换机已启用混杂模式，则不会丢弃来自混杂端口当前未使用的上行链路的某些数据包。您应该为混杂端口启用 `ReversePathFwdCheckPromisc`，然后将其禁用，这会明确丢弃来自当前未使用的上行链路的所有数据包。

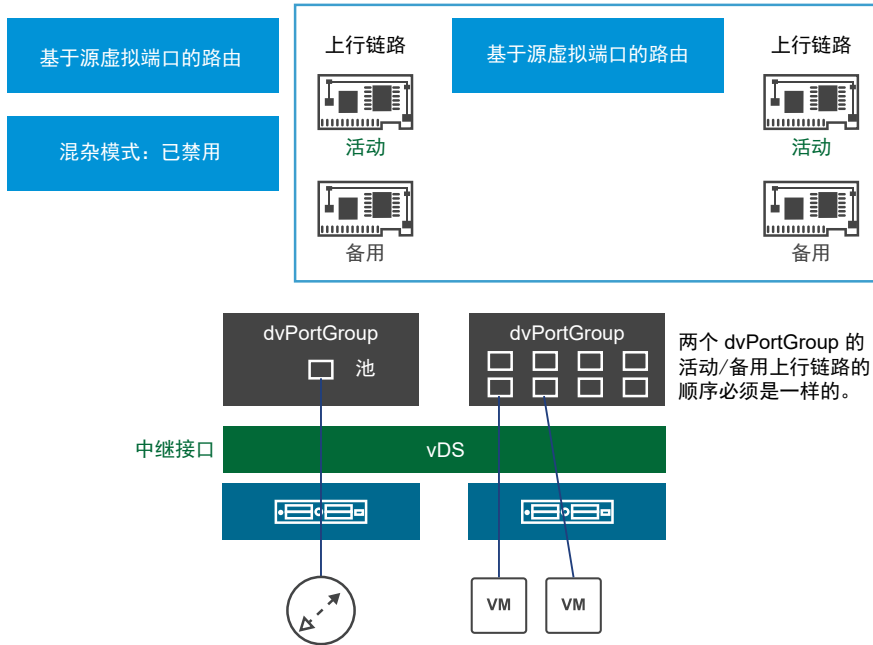
要阻止重复的数据包，请从 NSX Edge 所在的 ESXi CLI 中为混杂模式激活 RPF 检查：

```
esxcli system settings advanced set -o /Net/ReversePathFwdCheckPromisc -i 1
esxcli system settings advanced list -o /Net/ReversePathFwdCheckPromisc
Path: /Net/ReversePathFwdCheckPromisc
Type: integer
Int Value: 1
Default Int Value: 0
Max Value: 1
Min Value: 0
String Value:
Default String Value:
Valid Characters:
Description: Block duplicate packet in a teamed environment when the virtual switch is set to
Promiscuous mode.
```

在 **PortGroup** 安全策略中，将 **PromiscuousMode** 从 **Accept** 更改为 **Reject**，然后改回到 **Accept** 以激活配置的更改。

- 选项 2: Edge 和虚拟机位于同一 ESXi 主机上

2. 在同一主机上部署 L2VPN Edge 和虚拟机



- a 为与 Edge 的中继虚拟网卡关联的分布式端口组配置绑定策略和故障切换策略，如下所示：
 - 1 以“基于源虚拟端口的路由”方式开展负载平衡。
 - 2 将一个上行链路配置为“活动”，而将另一个上行链路配置为“备用”。
- b 为与虚拟机关联的分布式端口组配置绑定策略和故障切换策略，如下所示：
 - 1 任何绑定策略均可。
 - 2 只能有一个上行链路处于活动状态。
 - 3 虚拟机分布式端口组和 Edge 中继虚拟网卡分布式端口组的活动/备用上行链路的顺序必须相同。
- c 将客户端侧的独立 Edge 配置为使用池端口模式，并在中继虚拟网卡上禁用混杂模式。

使用 CLI 进行故障排除

您可以使用 NSX 命令行界面 (Command Line Interface, CLI) 执行一些 L2 VPN 故障排除。

问题

L2 VPN 无法正常工作。

解决方案

- 1 使用以下中央 CLI 命令查看配置问题：


```
show edge <edgeID> configuration l2vpn.
```

 例如，`show edge edge-1 configuration l2vpn.`

2 在客户端和服务端 Edge 上使用以下命令：

- show configuration l2vpn - 检查以下四个键值以验证服务器。

show configuration l2vpn

```
vShield Edge L2 VPN Config:
{
  "l2vpn" : {
    "cipher" : {
      "RC4-MD5"
    },
    "listenerPort" : 443,
    "clientVnicIndex" : null,
    "filters" : [],
    "serverPort" : null,
    "caCertificate" : null,
    "peerSiteAlgorithm" : null,
    "listenerIp" : "192.168.100.3",
    "peerSites" : [
      {
        "vseVnicNames" : [
          "vNic_10"
        ],
        "name" : "L2VPN-Site1",
        "filters" : [],
        "l2vpnUser" : {
          "password" : "*****",
          "userId" : "vpnuser1"
        }
      }
    ],
    "clientProxySetting" : null,
    "enable" : true,
    "trunkedVnicIndexes" : [
      2
    ],
    "serverVnicIndex" : null,
    "l2vpnUsers" : [],
    "serverAddress" : null,
    "logging" : {
      "enable" : false,
      "logLevel" : "info"
    },
    "vseVnicNames" : null,
    "serverCertificate" : null
  }
}
```

Cipher

Port

Server IP

Peer Site Configuration

- show service l2vpn bridge - 接口数取决于 L2 VPN 客户端数。在下面的输出中，配置了单个 L2 VPN 客户端 (na1)。Port1 是指 vNic_2。在 vNic_2 接口上发现了 MAC 地址 02:50:56:56:44:52，并且该地址不是 Edge (L2 VPN 服务器) 的本地地址。以下示例中的第 3 行是指 na1 接口。

```
plr01-0> show service l2vpn bridge
```

bridge name	bridge id	STP enabled	interfaces
br-sub	8000.0050568e19fb	no	vNic_2 na1

List of learned MAC addresses for L2 VPN bridge br-sub

port no	mac addr	is local?	vlanid	ageing timer
1	00:50:56:8e:19:fb	yes	0	0.00
1	02:50:56:56:44:52	no	1	0.87
2	2a:56:30:31:7e:3b	yes	0	0.00

- show service l2vpn trunk table

- `show service l2vpn conversion table` - 在以下示例中，到达隧道 1 的以太网帧将其 VLAN ID 1 转换为具有 VLAN 号 5001 的 VXLAN，然后再将数据包传送到 VDS。



- `show process monitor` - 确定 l2vpn（服务器）和 naclentd（客户端）进程是否正在运行。
- `show service network-connections` - 确定 l2vpn（服务器）和 naclentd（客户端）进程是否正在侦听端口 443。

SSL VPN

您可以使用此信息对设置中出现的问题进行故障排除。

无法打开 SSL VPN Web 门户

SSL VPN 用户无法打开 SSL VPN Web 门户登录页面以下载并安装 SSL VPN-Plus 客户端安装软件包。

问题

无法打开 SSL VPN Web 门户登录页面，或者在系统浏览器中未正确呈现该页面。

原因

以下原因之一可能会导致该问题：

- 系统使用不支持的浏览器版本。
- 在浏览器中未启用 Cookie 和 JavaScript。

解决方案

- 1 确保在任何以下支持的浏览器中打开 SSL VPN Web 门户登录页面。

浏览器	支持的最低版本
Internet Explorer	9.0.8112.16421
Chrome	67.03396
Safari	10.x

- 2 打开浏览器设置，并确保启用了 Cookie 和 JavaScript。

- 3 如果浏览器语言未设置为英语，请将语言设置为英语，并查看问题是否仍然存在。
- 4 检查是否在 SSL VPN 服务器上选择了 AES 密码。某些浏览器不支持 AES 加密。

SSL VPN-Plus: 安装故障

可以使用本主题了解可能的 SSL VPN-Plus 客户端特定安装问题以及如何解决这些问题。

问题

与 SSL VPN-Plus 客户端安装相关的常见问题如下所示：

- 已成功安装 SSL VPN-Plus 客户端，但客户端无法正常工作。
- 在 Mac 计算机上，显示内核扩展警告消息。
- 在 Mac OS High Sierra 上，显示以下安装错误消息：

```
/opt/sslvpn-plus/naclient/signed_kext/tap.kext failed to load - (libkern/kext)system policy prevents loading; check the system/kernel logs for errors or try kextutil(8).
Error: Could not load /opt/sslvpn-plus/naclient/signed_kext/tap.kext
```

```
installer[4571] <Debug>: install:didFailWithError:Error Domain=
PKInstallErrorDomain Code=112 "An error occurred while running scripts from the package
"naclient.pkg" .
" UserInfo={NSFilePath=./postinstall,NSURL=file:///<pathtofile>/
naclient.pkg,PKInstallPackageIdentifier=
com.vmware.sslvpn,NSLocalizedString=An error occurred while running scripts from the
package "naclient.pkg" .}

installer[4571] <Error>: Install failed: The Installer encountered an error that caused the
installation to fail. Contact the software manufacturer for assistance.
installer: The install failed (The Installer encountered an error that caused the installation to
fail.
Contact the software manufacturer for assistance.)
```

- 在 Windows 计算机上，显示以下错误消息：由于 E000024B，驱动程序安装失败。请尝试重新引导计算机。(Driver installation failed for reason E000024B: please try rebooting the machine.)

原因

以下原因之一可能会导致 SSL VPN-Plus 客户端失败，甚至在计算机上成功安装该客户端后：

- 缺少配置文件 (naclient.cfg) 或该配置文件无效。
- 目录权限或用户权限不正确。
- 无法访问 SSL VPN 服务器。
- 在 Mac 和 Linux 计算机上，未加载 TAP 驱动程序。

在 Mac 计算机上，由于系统阻止加载内核扩展，显示内核扩展警告消息。

在 Mac OS High Sierra 上，在 Mac 计算机不允许使用 `kext` 并且也未提示您加载 `kext` 时，显示安装错误。

在 Windows 计算机上，由于在 Edge SSL VPN-Plus 客户端安装程序中启用了**隐藏 SSL 客户端网络适配器 (Hide SSL client network adapter)**选项，显示驱动程序安装失败 (E000024B)。

解决方案

- 1 确保在支持的操作系统上安装 SSL VPN-Plus 客户端。有关支持的操作系统的信息，请参见 NSX 管理指南中的“SSL VPN-Plus 概览”主题。
- 2 在 Windows 计算机上，确保安装 SSL VPN-Plus 客户端的用户具有**管理员**特权。在 Mac 和 Linux 计算机上，用户必须具有 **root** 特权才能安装 SSL VPN-Plus 客户端。此外，要在 Mac 计算机上成功启动并运行 SSL VPN-Plus 客户端，用户必须在 `usr/local/lib` 目录中具有**执行**权限。
- 3 在 Linux 计算机上，确保安装了以下库。需要使用这些库，UI 才能正常工作。

- TCL
- TK
- NSS

- 4 如果在 Mac 和 Linux 计算机上未加载 TAP 驱动程序，请运行 Shell 脚本以加载该驱动程序。

操作系统	说明
Mac	使用 sudo 特权从 <code>/opt/sslvpn-plus/naclient/</code> 目录中运行 <code>Naclient.sh</code> Shell 脚本。
Linux	使用 sudo 特权运行 <code>naclient.sh</code> Shell 脚本。您可以在 <code>linux_phat_client/</code> <code>linux_phat_client</code> 目录中找到该脚本。

- 5 要解决在具有 macOS High Sierra 或更高版本的计算机上显示的内核扩展警告消息，您必须明确批准用户加载内核扩展 (`kext`)。执行以下步骤：
 - a 在 Mac 计算机上，打开**系统偏好设置 (System Preferences) > 安全性与隐私 (Security & Privacy)**窗口。
 - b 在窗口底部，您可能会看到类似于“已阻止加载某些系统软件，请单击“允许”按钮。”(Some system software was blocked from loading, Click the "Allow" button.) 的消息。
 - c 要继续进行安装，请单击**允许 (Allow)**。

有关批准用户加载内核扩展的详细信息，请参见 https://developer.apple.com/library/content/technotes/tn2459/_index.html。
 - d 在加载内核扩展时，SSL VPN-Plus 客户端安装进程继续在后台运行。安装了 SSL VPN-Plus 客户端，但显示以下错误消息：安装失败。安装程序遇到错误，从而导致安装失败。请联系软件制造商以获取帮助 (The installation failed. The installer encountered an error that cause the installation to fail. Contact the software manufacturer for assistance)。
 - e 要解决该错误，请卸载 SSL VPN-Plus 客户端并重新安装。

6 要解决在 Mac OS High Sierra 上显示的安装错误消息，请执行以下步骤。

- a 确保启用了通知。转到**系统偏好设置 (System Preferences) > 安全性与隐私 (Security & Privacy) > 允许通知 (Allow Notifications)**。

注 首次在 Mac OS High Sierra 上安装 SSL VPN-Plus 客户端时，通知窗口将提示您允许安装。该通知通常持续 30 分钟。如果在单击**允许 (Allow)**之前通知消失，请重新启动计算机并重新安装 SSL VPN-Plus 客户端。

如果安装仍失败，这意味着您的系统不允许使用内核扩展 (**kext**)，也不会提示您加载 **kext**。完成其余子步骤以将 **tuntap kext team id** 添加到预先批准的 **kext** 列表中。

- b 在恢复模式下重新启动 Mac 计算机。
 - 1 单击屏幕左上角的 Apple 徽标。
 - 2 单击**重新启动 (Restart)**。
 - 3 立即按 **Command** 和 **R** 键，直至您看到 Apple 徽标或旋转地球仪。由于无法通过内置恢复系统启动，在 Mac 计算机尝试连接到 Internet 以启动 macOS 恢复时，将显示旋转地球仪。现在，将在恢复模式下启动 Mac。
 - c 在顶部栏上，单击**实用工具 (Utilities) > 终端 (Terminal)**。
 - d 要将 **tuntap kext team id** 添加到预先批准的 **kext** 列表中，请运行 **- spctl kext-consent add KS8XL6T9FZ** 命令。
 - e 在正常模式下重新启动 Mac 计算机。
 - f 要验证是否在预先批准的 **kext** 列表中显示该 **team id**，请运行 **- spctl kext-consent list** 命令。
 - g 安装 SSL VPN-Plus 客户端软件包。
- 7 在 Windows 计算机上，如果您看到驱动程序安装失败错误 (E00024B)，请在 Edge SSL VPN-Plus 客户端安装程序中禁用**隐藏 SSL 客户端网络适配器 (Hide SSL client network adapter)**选项。有关禁用该选项的说明，请参见 VMware 知识库文章：<https://kb.vmware.com/s/article/2108766>。

SSL VPN-Plus: 通信问题

可以使用本主题了解可能的 SSL VPN 连接和数据路径问题以及如何解决这些问题。

问题

与 SSL VPN 连接和数据路径相关的常见问题如下所示：

- SSL VPN-Plus 客户端无法连接到 SSL VPN 服务器。
- 安装了 SSL VPN-Plus 客户端，但 SSL VPN-Plus 服务未运行。
- 已达到最大登录用户数。SSL VPN Web 门户或 SSL VPN-Plus 客户端显示以下消息：

根据 SSL VPN 许可证，已达到最大用户数/已达到最大登录用户数。请过一段时间之后重试 (Maximum users reached/Maximum count of logged in user reached as per SSL VPN license. Please try after some time) 或 SSL 读取失败 (SSL read has failed)。

- SSL VPN 服务正在运行，但数据路径无效。
- 已建立 SSL VPN 连接，但无法访问专用网络中的应用程序。

解决方案

- 1 如果 SSL VPN-Plus 客户端无法连接到 SSL VPN 服务器，请执行以下操作：
 - 确保 SSL VPN 用户使用正确的用户名和密码登录。
 - 查 SSL VPN 用户是否有效。
 - 验证 SSL VPN 用户是否可以使用 Web 门户访问 SSL VPN 服务器。
- 2 在 NSX Edge 上，执行以下步骤以验证 SSL VPN 进程是否正在运行。
 - a 从 CLI 中登录到 NSX Edge。有关登录到 Edge CLI 的详细信息，请参见 NSX 命令行界面参考。
 - b 运行 `show process monitor` 命令，并找到 `sslvpn` 进程。
 - c 运行 `show service network-connections` 命令，并检查是否在端口 443 上列出了 `sslvpn` 进程。

注 默认情况下，系统使用端口 443 传输 SSL 流量。不过，如果为 SSL 流量配置了不同的 TCP 端口，请确保在该 TCP 端口号上列出了 `sslvpn` 进程。

- 3 在 SSL VPN-Plus 客户端上，验证 SSL VPN-Plus 服务是否正在运行。

操作系统	说明
Windows	打开 任务管理器 ，并检查是否启动了 SSL VPN-Plus 客户端服务。
Mac	<ul style="list-style-type: none"> ■ 确保为守护进程启动了 <code>naclntd</code> 进程。 ■ 确保为 GUI 启动了 <code>naclnt</code> 进程。 要检查该进程是否正在运行，请运行 <code>ps -ef grep "naclnt"</code> 命令。
Linux	<ul style="list-style-type: none"> ■ 确保启动了 <code>naclntd</code> 和 <code>naclnt_poll</code> 进程。 ■ 要检查该进程是否正在运行，请运行 <code>ps -ef grep "naclnt"</code> 命令。

如果这些服务没有运行，请运行以下命令以启动这些服务。

操作系统	命令
Mac	运行 <code>sudo launchctl load -w /Library/LaunchDaemons/com.vmware.naclntd.plist</code> 命令。
Linux	运行 <code>sudo service naclnt start</code> 命令。

- 4 如果达到了最大登录 SSL VPN 用户数，请增加 NSX Edge 规格以增加并发用户 (CCU) 数。
有关详细信息，请参见 NSX 管理指南。请注意，执行此操作时已连接的用户会断开与 VPN 的连接。

- 5 如果 SSL VPN 服务正在运行，但数据路径无效，请执行以下步骤：
 - a 检查在成功连接后是否分配了虚拟 IP。
 - b 验证是否添加了路由。
- 6 如果无法访问专用（后端）网络中的应用程序，请执行以下步骤以解决该问题：
 - a 确保专用网络和 IP 池没有位于同一子网中。
 - b 如果管理员未定义 IP 池，或者 IP 池用尽，请执行以下步骤：
 - 1 登录到 vSphere Web Client。
 - 2 单击**网络和安全 (Networking & Security)**，然后单击 **NSX Edge (NSX Edges)**。
 - 3 双击一个 NSX Edge，然后单击 **SSL VPN-Plus** 选项卡。
 - 4 按照 NSX 管理指南中的“添加 IP 池”主题所述添加一个静态 IP 池。请确保在**网关 (Gateway)**文本框中添加 IP 地址。该网关 IP 地址将分配给 *na0* 接口。所有非 TCP 流量都将流过名为 *na0* 接口的虚拟适配器。您可以创建多个具有不同网关 IP 地址的 IP 池，但将其分配给同一个 *na0* 接口。
 - 5 使用 `show interface na0` 命令验证提供的 IP 地址，并检查是否所有 IP 池都分配给同一个 *na0* 接口。
 - 6 登录到客户端计算机，转到 **SSL VPN-Plus 客户端 - 统计信息 (SSL VPN-Plus Client - Statistics)**屏幕并验证分配的虚拟 IP 地址。
 - c 登录到 NSX Edge 命令行界面 (CLI)，然后运行 `debug packet capture interface na0` 命令，以在 *na0* 接口上捕获数据包。您也可以使用**数据包捕获 (Packet Capture)**工具捕获数据包。有关详细信息，请参见 NSX 管理指南。

注 数据包捕获将继续在后台运行，直到您通过运行 `no debug packet capture interface na0` 命令停止捕获。

 - d 如果未启用 TCP 优化，请验证防火墙规则。
 - e 对于非 TCP 流量，请确保后端网络的默认网关设置为 Edge 的内部接口。
 - f 对于 Mac 和 Linux 客户端，请登录到安装了 SSL VPN 客户端的系统，然后运行 `tcpdump -i tap0 -s 1500 -w filepath` 命令，以在 *tap0* 接口或虚拟适配器上捕获数据包。在 Windows 客户端上，使用数据包分析器工具（如 Wireshark），并在 SSL VPN-Plus 客户端适配器上捕获数据包。
- 7 如果上述步骤没有解决该问题，请使用以下 NSX Edge CLI 命令进一步进行故障排除。

用途	命令
检查 SSL VPN 状态。	<code>show service sslvpn-plus</code>
检查 SSL VPN 统计信息。	<code>show service sslvpn-plus stats</code>
检查连接的 VPN 客户端。	<code>show service sslvpn-plus tunnels</code>
检查 SSL VPN-Plus 会话。	<code>show service sslvpn-plus sessions</code>

SSL VPN-Plus: 身份验证问题

您遇到了 SSL VPN-Plus 身份验证问题。

问题

SSL VPN-Plus 身份验证失败。

解决方案

◆ 对于身份验证问题，请验证以下设置：

- a 确保可从 NSX Edge 中访问外部身份验证服务器。从 NSX Edge 中，对身份验证服务器执行 ping 操作，验证该服务器是否可访问。
- b 使用 LDAP 浏览器等工具检查外部身份验证服务器配置，查看配置是否可正常工作。使用 LDAP 浏览器只能检查 LDAP 和 AD 身份验证服务器。
- c 如果在身份验证过程中配置了本地身份验证服务器，请确保将其设置为最低优先级。
- d 如果使用 Active Directory (AD)，请将其设置为 no-ssl 模式，并在可从中访问 AD 服务器的接口上捕获数据包。
- e 如果在 syslog 服务器中成功通过身份验证，您会看到类似以下内容的消息：Log Output - SVP_LOG_NOTICE,
10-28-2013,09:28:39,Authentication,a,-,-,10.112.243.61,-,PHAT,,SUCCESS,,10-28-2013,09:28:39,-,-,,,,,,,,,-,-,
- f 如果在 syslog 服务器中未能通过身份验证，您会看到类似以下内容的消息：Log Output - SVP_LOG_NOTICE,
10-28-2013,09:28:39,Authentication,a,-,-,10.112.243.61,-,PHAT,,FAILURE,,10-28-2013,09:28:39,-,-,,,,,,,,,-,-,

SSL VPN-Plus 客户端停止响应

在启用 TCP 优化后，SSL VPN-Plus 客户端停止响应。

问题

您将 SSL VPN-Plus 服务配置为在 NSX Edge 上运行，并启用 TCP 优化以通过隧道发送流量。在 SSL VPN-Plus 客户端上运行任何网络性能测量和优化工具（如 iperf3）时，SSL VPN-Plus 客户端停止响应。

原因

从 SSL VPN-Plus 客户端中发送数据时，以下两种情况之一可能会导致隧道读取错误：

- 通过发送 TCP FIN 序列，后端服务器关闭与 SSL VPN 服务器之间的 TCP 连接。
- 在将数据转发到后端服务器时，隧道写入操作失败。

隧道读取错误为未知协议 ID (unknown protocol ID)。该错误清除 SSL VPN 服务器和 SSL VPN-Plus 客户端之间的隧道，这会导致 SSL 读取/写入操作在客户端上失败，并且 SSL VPN-Plus 客户端停止响应。

解决方案

- ◆ 要解决该问题，请在 vSphere Web Client 中执行以下步骤，以便为通过 SSL VPN 通道的专用网络流量禁用 TCP 优化。
 - a 双击在其中配置了 SSL VPN-Plus 服务的 NSX Edge 虚拟机。
 - b 单击 **SSL VPN-Plus** 选项卡，然后选择专用网络。
 - c 清除启用 TCP 优化 (Enable TCP Optimization) 复选框。

基本日志分析

SSL VPN-Plus 网关日志将发送到在 NSX Edge 设备上配置的 syslog 服务器。SSL VPN-Plus 客户端日志存储在远程用户计算机上的以下目录中：C:\Users\username\AppData\Local\VMware\vpn\svpn_client.log。

基本日志分析 - 身份验证

身份验证成功

- 以下日志输出显示用户 a 在 2016 年 10 月 28 日 9 时 28 分成功通过网络访问客户端进行身份验证。

```
SVP_LOG_NOTICE,10-28-2016,09:28:39,Authentication,a,-,-,10.112.243.61,-,PHAT,,SUCCESS,,10-28-2016,09:28:39,-,-,,,,,,,,,,,,,-,-,-
```

身份验证失败

- 以下日志输出显示用户 a 在 2016 年 10 月 28 日 9 时 28 分无法通过网络访问客户端进行身份验证。

```
SVP_LOG_NOTICE,10-28-2016,09:28:39,Authentication,a,-,-,10.112.243.61,-,PHAT,,FAILURE,,10-28-2016,09:28:39,-,-,,,,,,,,,,,,,-,-,-
```

要解决身份验证问题，请参阅 [SSL VPN-Plus: 安装故障](#)。

基本日志分析 - 数据路径

数据路径成功

- 以下日志输出显示用户 a 在 2016 年 10 月 28 日 9 时 41 分成功使用网络访问客户端通过 TCP 连接到后端 Web 服务器 192.168.10.8。

```
SVP_LOG_INFO,10-28-2016,09:41:03,TCP
Connect,a,-,-,10.112.243.61,-,PHAT,,SUCCESS,,10-28-2013,09:41:03,-,-,192.168.10.8,80,,,,,,,,,-,-,-
```

数据路径失败

- 以下日志输出显示用户 a 在 2016 年 10 月 28 日 9 时 41 分无法使用网络访问客户端通过 TCP 连接到后端 Web 服务器 192.168.10.8。

```
SVP_LOG_INFO,10-28-2016,09:41:03,TCP
Connect,a,-,-,10.112.243.61,-,PHAT,,FAILURE,,10-28-2013,09:41:03,-,-,192.168.10.8,80,,,,,,,,,-,-,-
```

IPSEC VPN

使用此处的信息可帮助您解决设置中的协商问题。

成功协商（阶段 1 和阶段 2）

下例显示了 NSX Edge 和 Cisco 设备之间一次成功协商的结果。

NSX Edge

在 NSX Edge 命令行界面（`ipsec auto -status`，即 `show service ipsec` 命令的一部分）中：

```
000 #2: "s1-c1":500 STATE_QUICK_I2 (sent QI2, IPsec SA established);
      EVENT_SA_REPLACE in 2430s; newest IPSEC; eroute owner; isakmp#1; idle;
      import:admin initiate
000 #2: "s1-c1" esp.f5f6877d@10.20.131.62 esp.7aaf335f@10.20.129.80
      tun.0@10.20.131.62 tun.0@10.20.129.80 ref=0 refhim=4294901761
000 #1: "s1-c1":500 STATE_MAIN_I4 (ISAKMP SA established); EVENT_SA_REPLACE in
      27623s; newest ISAKMP; lastdpd=0s(seq in:0 out:0); idle;
      import:admin initiate
```

Cisco

```
ciscoasa# show crypto isakmp sa detail

Active SA: 1
Rekey SA: 0 (A tunnel will report 1 Active and 1 Rekey SA during rekey)
Total IKE SA: 1

IKE Peer: 10.20.129.80
Type : L2L      Role   : responder
Rekey : no      State  : MM_ACTIVE
Encrypt : 3des  Hash   : SHA
Auth : preshared Lifetime: 28800
Lifetime Remaining: 28379
```

阶段 1 策略不匹配

下面列出了阶段 1 策略不匹配错误日志。

NSX Edge

NSX Edge 在 `STATE_MAIN_I1` 状态下挂起。在 `/var/log/messages` 中查找显示对等站点返回包含“`NO_PROPOSAL_CHOSEN`”集的 IKE 消息的信息。

```
000 #1: "s1-c1":500 STATE_MAIN_I1 (sent MI1,
      expecting MR1); EVENT_RETRANSMIT in 7s; nodpd; idle;
      import:admin initiate
000 #1: pending Phase 2 for "s1-c1" replacing #0
Aug 26 12:31:25 weiqing-desktop ipsec[6569]:
      | got payload 0x800(ISAKMP_NEXT_N) needed: 0x0 opt: 0x0
Aug 26 12:31:25 weiqing-desktop ipsec[6569]:
```

```

    | ***parse ISAKMP Notification Payload:
Aug 26 12:31:25 weiqing-desktop ipsec[6569]:
    |   next payload type: ISAKMP_NEXT_NONE
Aug 26 12:31:25 weiqing-desktop ipsec[6569]: |   length: 96
Aug 26 12:31:25 weiqing-desktop ipsec[6569]:
    |   DOI: ISAKMP_DOI_IPSEC
Aug 26 12:31:25 weiqing-desktop ipsec[6569]: |   protocol ID: 0
Aug 26 12:31:25 weiqing-desktop ipsec[6569]: |   SPI size: 0
Aug 26 12:31:25 weiqing-desktop ipsec[6569]:
    |   Notify Message Type: NO_PROPOSAL_CHOSEN
Aug 26 12:31:25 weiqing-desktop ipsec[6569]:
    "s1-c1" #1: ignoring informational payload,
    type NO_PROPOSAL_CHOSEN msgid=00000000

```

Cisco

如果已启用调试加密，则会打印错误消息以显示未接受任何建议。

```

ciscoasa# Aug 26 18:17:27 [IKEv1]:
    IP = 10.20.129.80, IKE_DECODE RECEIVED
    Message (msgid=0) with payloads : HDR + SA (1)
    + VENDOR (13) + VENDOR (13) + NONE (0) total length : 148
Aug 26 18:17:27 [IKEv1 DEBUG]: IP = 10.20.129.80,
    processing SA payload
Aug 26 18:17:27 [IKEv1]: Phase 1 failure: Mismatched attribute
    types for class Group Description: Rcv'd: Group 5
    Cfg'd: Group 2
Aug 26 18:17:27 [IKEv1]: Phase 1 failure: Mismatched attribute
    types for class Group Description: Rcv'd: Group 5
    Cfg'd: Group 2
Aug 26 18:17:27 [IKEv1]: IP = 10.20.129.80, IKE_DECODE SENDING
    Message (msgid=0) with payloads : HDR + NOTIFY (11)
    + NONE (0) total length : 124
Aug 26 18:17:27 [IKEv1 DEBUG]: IP = 10.20.129.80,
    All SA proposals found unacceptable
Aug 26 18:17:27 [IKEv1]: IP = 10.20.129.80, Error processing
    payload: Payload ID: 1
Aug 26 18:17:27 [IKEv1 DEBUG]: IP = 10.20.129.80, IKE MM Responder
    FSM error history (struct &0xd8355a60) <state>, <event>:
    MM_DONE, EV_ERROR-->MM_START, EV_RCV_MSG-->MM_START,
    EV_START_MM-->MM_START, EV_START_MM-->MM_START,
    EV_START_MM-->MM_START, EV_START_MM-->MM_START,
    EV_START_MM-->MM_START, EV_START_MM
Aug 26 18:17:27 [IKEv1 DEBUG]: IP = 10.20.129.80, IKE SA
    MM:9e0e4511 terminating: flags 0x01000002, refcnt 0,
    tuncnt 0
Aug 26 18:17:27 [IKEv1 DEBUG]: IP = 10.20.129.80, sending
    delete/delete with reason message

```

阶段 2 不匹配

下面列出了阶段 2 策略不匹配错误日志。

NSX Edge

NSX Edge 在 STATE_QUICK_I1 状态下挂起。日志消息显示对等站点发送了一条 NO_PROPOSAL_CHOSEN 消息。

```
000 #2: "s1-c1":500 STATE_QUICK_I1 (sent QI1, expecting
      QR1); EVENT_RETRANSMIT in 11s; lastdpd=-1s(seq in:0 out:0);
      idle; import:admin initiate
Aug 26 12:33:54 weiqing-desktop ipsec[6933]: | got payload
      0x800(ISAKMP_NEXT_N) needed: 0x0 opt: 0x0
Aug 26 12:33:54 weiqing-desktop ipsec[6933]: | ***parse
      ISAKMP Notification Payload:
Aug 26 12:33:54 weiqing-desktop ipsec[6933]: |      next payload
      type: ISAKMP_NEXT_NONE
Aug 26 12:33:54 weiqing-desktop ipsec[6933]: |      length: 32
Aug 26 12:33:54 weiqing-desktop ipsec[6933]: |
      |      DOI: ISAKMP_DOI_IPSEC
Aug 26 12:33:54 weiqing-desktop ipsec[6933]: |      protocol ID: 3
Aug 26 12:33:54 weiqing-desktop ipsec[6933]: |      SPI size: 16
Aug 26 12:33:54 weiqing-desktop ipsec[6933]: |      Notify Message
      Type: NO_PROPOSAL_CHOSEN
Aug 26 12:33:54 weiqing-desktop ipsec[6933]: "s1-c1" #3:
      ignoring informational payload, type NO_PROPOSAL_CHOSEN
      msgid=00000000
```

Cisco

调试消息显示阶段 1 已完成，但阶段 2 因为策略协商失败而失败。

```
Aug 26 16:03:49 [IKEv1]: Group = 10.20.129.80,
      IP = 10.20.129.80, PHASE 1 COMPLETED
Aug 26 16:03:49 [IKEv1]: IP = 10.20.129.80, Keep-alive type
      for this connection: DPD
Aug 26 16:03:49 [IKEv1 DEBUG]: Group = 10.20.129.80,
      IP = 10.20.129.80, Starting P1 rekey timer: 21600 seconds
Aug 26 16:03:49 [IKEv1]: IP = 10.20.129.80, IKE_DECODE RECEIVED
      Message (msgid=b2cdcb13) with payloads : HDR + HASH (8)
      + SA (1) + NONCE (10) + KE (4) + ID (5) + ID (5) + NONE (0)
      total length : 288
.
.
.
Aug 26 16:03:49 [IKEv1]: Group = 10.20.129.80, IP = 10.20.129.80,
      Session is being torn down. Reason: Phase 2 Mismatch
```

PFS 不匹配

下面列出了 PFS 不匹配错误日志。

NSX Edge

PFS 协商为阶段 2 的一部分。如果 PFS 不匹配，则该行为类似于[阶段 2 不匹配](#)中所述的失败案例。

```
000 #4: "s1-c1":500 STATE_QUICK_I1 (sent QI1, expecting
      QR1); EVENT_RETRANSMIT in 8s; lastdpd=-1s(seq in:0 out:0);
      idle; import:admin initiate
Aug 26 12:35:52 weiqing-desktop ipsec[7312]: | got payload 0x800
      (ISAKMP_NEXT_N) needed: 0x0 opt: 0x0
Aug 26 12:35:52 weiqing-desktop ipsec[7312]:
      | ***parse ISAKMP Notification Payload:
Aug 26 12:35:52 weiqing-desktop ipsec[7312]: |      next payload
      type: ISAKMP_NEXT_NONE
Aug 26 12:35:52 weiqing-desktop ipsec[7312]: |      length: 32
Aug 26 12:35:52 weiqing-desktop ipsec[7312]: |
      |      DOI: ISAKMP_DOI_IPSEC
Aug 26 12:35:52 weiqing-desktop ipsec[7312]: |      protocol ID: 3
Aug 26 12:35:52 weiqing-desktop ipsec[7312]: |      SPI size: 16
Aug 26 12:35:52 weiqing-desktop ipsec[7312]: |      Notify Message
      Type: NO_PROPOSAL_CHOSEN
Aug 26 12:35:52 weiqing-desktop ipsec[7312]: "s1-c1" #1: ignoring
      informational payload, type NO_PROPOSAL_CHOSEN
      msgid=00000000
Aug 26 12:35:52 weiqing-desktop ipsec[7312]: | info:  fa 16 b3 e5
      91 a9 b0 02  a3 30 e1 d9  6e 5a 13 d4
Aug 26 12:35:52 weiqing-desktop ipsec[7312]: | info:  93 e5 e4 d7
Aug 26 12:35:52 weiqing-desktop ipsec[7312]:
      | processing informational NO_PROPOSAL_CHOSEN (14)
```

Cisco

```
<BS>Aug 26 19:00:26 [IKEv1 DEBUG]: Group = 10.20.129.80,
      IP = 10.20.129.80, sending delete/delete with
      reason message
Aug 26 19:00:26 [IKEv1 DEBUG]: Group = 10.20.129.80,
      IP = 10.20.129.80, constructing blank hash payload
Aug 26 19:00:26 [IKEv1 DEBUG]: Group = 10.20.129.80,
      IP = 10.20.129.80, constructing blank hash payload
Aug 26 19:00:26 [IKEv1 DEBUG]: Group = 10.20.129.80,
      IP = 10.20.129.80, constructing IKE delete payload
Aug 26 19:00:26 [IKEv1 DEBUG]: Group = 10.20.129.80,
      IP = 10.20.129.80, constructing qm hash payload
Aug 26 19:00:26 [IKEv1]: IP = 10.20.129.80, IKE_DECODE SENDING
      Message (msgid=19eb1e59) with payloads : HDR + HASH (8)
      + DELETE (12) + NONE (0) total length : 80
Aug 26 19:00:26 [IKEv1]: Group = 10.20.129.80, IP = 10.20.129.80,
      Session is being torn down. Reason: Phase 2 Mismatch
```

PSK 不匹配

下面列出了 PSK 不匹配错误日志

NSX Edge

在阶段 1 的最后一轮中协商 PSK。如果 PSK 协商失败，则 NSX Edge 状态为 STATE_MAIN_I4。对等站点将发送包含 INVALID_ID_INFORMATION 的消息。

```
Aug 26 11:55:55 weiqing-desktop ipsec[3855]:
    "s1-cl" #1: transition from state STATE_MAIN_I3 to
    state STATE_MAIN_I4
Aug 26 11:55:55 weiqing-desktop ipsec[3855]: "s1-cl" #1:
    STATE_MAIN_I4: ISAKMP SA established
    {auth=OAKLEY_PRESHARED_KEY
    cipher=oakley_3des_cbc_192 prf=oakley_sha group=modp1024}
Aug 26 11:55:55 weiqing-desktop ipsec[3855]: "s1-cl" #1: Dead Peer
    Detection (RFC 3706): enabled
Aug 26 11:55:55 weiqing-desktop ipsec[3855]: "s1-cl" #2:
    initiating Quick Mode PSK+ENCRYPT+TUNNEL+PFS+UP+SAREFTRACK
    {using isakmp#1 msgid:e8add10e proposal=3DES(3)_192-SHA1(2)_160
    pfsgroup=OAKLEY_GROUP_MODP1024}
Aug 26 11:55:55 weiqing-desktop ipsec[3855]: "s1-cl" #1:
    ignoring informational payload, type INVALID_ID_INFORMATION
    msgid=00000000
```

Cisco

```
Aug 26 15:27:07 [IKEv1]: IP = 10.115.199.191,
    IKE_DECODE SENDING Message (msgid=0) with payloads : HDR
    + KE (4) + NONCE (10) + VENDOR (13) + VENDOR (13)
    + VENDOR (13) + VENDOR (13) + NAT-D (130) + NAT-D (130)
    + NONE (0) total length : 304
Aug 26 15:27:07 [IKEv1]: Group = 10.115.199.191,
    IP = 10.115.199.191, Received encrypted Oakley Main Mode
    packet with invalid payloads, MessID = 0
Aug 26 15:27:07 [IKEv1]: IP = 10.115.199.191, IKE_DECODE SENDING
    Message (msgid=0) with payloads : HDR + NOTIFY (11)
    + NONE (0) total length : 80
Aug 26 15:27:07 [IKEv1]: Group = 10.115.199.191,
    IP = 10.115.199.191, ERROR, had problems decrypting
    packet, probably due to mismatched pre-shared key.
    Aborting
```

成功协商的数据包捕获

下面列出了 NSX Edge 和 Cisco 设备之间成功协商的数据包捕获会话。

No.	Time	Source	Destination	Protocol	Info
9203	768.394800	10.20.129.80	10.20.131.62	ISAKMP	Identity Protection (Main Mode)
Frame 9203 (190 bytes on wire, 190 bytes captured)					
Ethernet II, Src: Vmware_9d:2c:dd (00:50:56:9d:2c:dd),					


```

    Dst: Cisco_80:70:f5 (00:13:c4:80:70:f5)
Internet Protocol, Src: 10.20.129.80 (10.20.129.80),
    Dst: 10.20.131.62 (10.20.131.62)
User Datagram Protocol, Src Port: isakmp (500), Dst Port: isakmp (500)
Internet Security Association and Key Management Protocol
Initiator cookie: 92585D2D797E9C52
Responder cookie: 0000000000000000
Next payload: Security Association (1)
Version: 1.0
Exchange type: Identity Protection (Main Mode) (2)
Flags: 0x00
Message ID: 0x00000000
Length: 148
Security Association payload
  Next payload: Vendor ID (13)
  Payload length: 84
  Domain of interpretation: IPSEC (1)
  Situation: IDENTITY (1)
  Proposal payload # 0
    Next payload: NONE (0)
    Payload length: 72
    Proposal number: 0
    Protocol ID: ISAKMP (1)
    SPI Size: 0
    Proposal transforms: 2
    Transform payload # 0
      Next payload: Transform (3)
      Payload length: 32
      Transform number: 0
      Transform ID: KEY_IKE (1)
      Life-Type (11): Seconds (1)
      Life-Duration (12): Duration-Value (28800)
      Encryption-Algorithm (1): 3DES-CBC (5)
      Hash-Algorithm (2): SHA (2)
      Authentication-Method (3): PSK (1)
      Group-Description (4): 1536 bit MODP group (5)
    Transform payload # 1
      Next payload: NONE (0)
      Payload length: 32
      Transform number: 1
      Transform ID: KEY_IKE (1)
      Life-Type (11): Seconds (1)
      Life-Duration (12): Duration-Value (28800)
      Encryption-Algorithm (1): 3DES-CBC (5)
      Hash-Algorithm (2): SHA (2)
      Authentication-Method (3): PSK (1)
      Group-Description (4): Alternate 1024-bit MODP group (2)
  Vendor ID: 4F456C6A405D72544D42754D
  Next payload: Vendor ID (13)
  Payload length: 16
  Vendor ID: 4F456C6A405D72544D42754D
Vendor ID: RFC 3706 Detecting Dead IKE Peers (DPD)
  Next payload: NONE (0)
  Payload length: 20
  Vendor ID: RFC 3706 Detecting Dead IKE Peers (DPD)

```

No.	Time	Source	Destination	Protocol Info
9204	768.395550	10.20.131.62	10.20.129.80	ISAKMP Identity Protection (Main Mode)

Frame 9204 (146 bytes on wire, 146 bytes captured)
 Ethernet II, Src: Cisco_80:70:f5 (00:13:c4:80:70:f5),
 Dst: Vmware_9d:2c:dd (00:50:56:9d:2c:dd)
 Internet Protocol, Src: 10.20.131.62 (10.20.131.62),
 Dst: 10.20.129.80 (10.20.129.80)
 User Datagram Protocol, Src Port: isakmp (500), Dst Port: isakmp (500)
 Internet Security Association and Key Management Protocol
 Initiator cookie: 92585D2D797E9C52
 Responder cookie: 34704CFC8C8DBD09
 Next payload: Security Association (1)
 Version: 1.0
 Exchange type: Identity Protection (Main Mode) (2)
 Flags: 0x00
 Message ID: 0x00000000
 Length: 104
 Security Association payload
 Next payload: Vendor ID (13)
 Payload length: 52
 Domain of interpretation: IPSEC (1)
 Situation: IDENTITY (1)
 Proposal payload # 1
 Next payload: NONE (0)
 Payload length: 40
 Proposal number: 1
 Protocol ID: ISAKMP (1)
 SPI Size: 0
 Proposal transforms: 1
 Transform payload # 1
 Next payload: NONE (0)
 Payload length: 32
 Transform number: 1
 Transform ID: KEY_IKE (1)
 Encryption-Algorithm (1): 3DES-CBC (5)
 Hash-Algorithm (2): SHA (2)
 Group-Description (4): Alternate 1024-bit MODP group (2)
 Authentication-Method (3): PSK (1)
 Life-Type (11): Seconds (1)
 Life-Duration (12): Duration-Value (28800)
 Vendor ID: Microsoft L2TP/IPSec VPN Client
 Next payload: NONE (0)
 Payload length: 24
 Vendor ID: Microsoft L2TP/IPSec VPN Client

No.	Time	Source	Destination	Protocol Info
9205	768.399599	10.20.129.80	10.20.131.62	ISAKMP Identity Protection (Main Mode)

Frame 9205 (222 bytes on wire, 222 bytes captured)
 Ethernet II, Src: Vmware_9d:2c:dd (00:50:56:9d:2c:dd),
 Dst: Cisco_80:70:f5 (00:13:c4:80:70:f5)

```

Internet Protocol, Src: 10.20.129.80 (10.20.129.80),
    Dst: 10.20.131.62 (10.20.131.62)
User Datagram Protocol, Src Port: isakmp (500), Dst Port: isakmp (500)
Internet Security Association and Key Management Protocol
    Initiator cookie: 92585D2D797E9C52
    Responder cookie: 34704CFC8C8DBD09
    Next payload: Key Exchange (4)
    Version: 1.0
    Exchange type: Identity Protection (Main Mode) (2)
    Flags: 0x00
    Message ID: 0x00000000
    Length: 180
    Key Exchange payload
        Next payload: Nonce (10)
        Payload length: 132
        Key Exchange Data (128 bytes / 1024 bits)
    Nonce payload
        Next payload: NONE (0)
        Payload length: 20
        Nonce Data

```

No.	Time	Source	Destination	Protocol Info
9206	768.401192	10.20.131.62	10.20.129.80	ISAKMP Identity Protection (Main Mode)

```

Frame 9206 (298 bytes on wire, 298 bytes captured)
Ethernet II, Src: Cisco_80:70:f5 (00:13:c4:80:70:f5),
    Dst: Vmware_9d:2c:dd (00:50:56:9d:2c:dd)
Internet Protocol, Src: 10.20.131.62 (10.20.131.62),
    Dst: 10.20.129.80 (10.20.129.80)
User Datagram Protocol, Src Port: isakmp (500), Dst Port: isakmp (500)
Internet Security Association and Key Management Protocol
    Initiator cookie: 92585D2D797E9C52
    Responder cookie: 34704CFC8C8DBD09
    Next payload: Key Exchange (4)
    Version: 1.0
    Exchange type: Identity Protection (Main Mode) (2)
    Flags: 0x00
    Message ID: 0x00000000
    Length: 256
    Key Exchange payload
        Next payload: Nonce (10)
        Payload length: 132
        Key Exchange Data (128 bytes / 1024 bits)
    Nonce payload
        Next payload: Vendor ID (13)
        Payload length: 24
        Nonce Data
    Vendor ID: CISCO-UNITY-1.0
        Next payload: Vendor ID (13)
        Payload length: 20
        Vendor ID: CISCO-UNITY-1.0
    Vendor ID: draft-beaulieu-ike-xauth-02.txt
        Next payload: Vendor ID (13)
        Payload length: 12

```

```

Vendor ID: draft-beaulieu-ike-xauth-02.txt
Vendor ID: C1B7EBE18C8CBD099E89695E2CB16A4A
Next payload: Vendor ID (13)
Payload length: 20
Vendor ID: C1B7EBE18C8CBD099E89695E2CB16A4A
Vendor ID: CISCO-CONCENTRATOR
Next payload: NONE (0)
Payload length: 20
Vendor ID: CISCO-CONCENTRATOR

```

No.	Time	Source	Destination	Protocol	Info
9207	768.404990	10.20.129.80	10.20.131.62	ISAKMP	Identity Protection (Main Mode)

```

Frame 9207 (110 bytes on wire, 110 bytes captured)
Ethernet II, Src: Vmware_9d:2c:dd (00:50:56:9d:2c:dd),
      Dst: Cisco_80:70:f5 (00:13:c4:80:70:f5)
Internet Protocol, Src: 10.20.129.80 (10.20.129.80),
      Dst: 10.20.131.62 (10.20.131.62)
User Datagram Protocol, Src Port: isakmp (500), Dst Port: isakmp (500)
Internet Security Association and Key Management Protocol
  Initiator cookie: 92585D2D797E9C52
  Responder cookie: 34704CFC8C8DBD09
  Next payload: Identification (5)
  Version: 1.0
  Exchange type: Identity Protection (Main Mode) (2)
  Flags: 0x01
  Message ID: 0x00000000
  Length: 68
  Encrypted payload (40 bytes)

```

No.	Time	Source	Destination	Protocol	Info
9208	768.405921	10.20.131.62	10.20.129.80	ISAKMP	Identity Protection (Main Mode)

```

Frame 9208 (126 bytes on wire, 126 bytes captured)
Ethernet II, Src: Cisco_80:70:f5 (00:13:c4:80:70:f5),
      Dst: Vmware_9d:2c:dd (00:50:56:9d:2c:dd)
Internet Protocol, Src: 10.20.131.62 (10.20.131.62),
      Dst: 10.20.129.80 (10.20.129.80)
User Datagram Protocol, Src Port: isakmp (500), Dst Port: isakmp (500)
Internet Security Association and Key Management Protocol
  Initiator cookie: 92585D2D797E9C52
  Responder cookie: 34704CFC8C8DBD09
  Next payload: Identification (5)
  Version: 1.0
  Exchange type: Identity Protection (Main Mode) (2)
  Flags: 0x01
  Message ID: 0x00000000
  Length: 84
  Encrypted payload (56 bytes)

```

No.	Time	Source	Destination	Protocol	Info
9209	768.409799	10.20.129.80	10.20.131.62	ISAKMP	Quick Mode

```

Frame 9209 (334 bytes on wire, 334 bytes captured)

```

```

Ethernet II, Src: Vmware_9d:2c:dd (00:50:56:9d:2c:dd),
      Dst: Cisco_80:70:f5 (00:13:c4:80:70:f5)
Internet Protocol, Src: 10.20.129.80 (10.20.129.80),
      Dst: 10.20.131.62 (10.20.131.62)
User Datagram Protocol, Src Port: isakmp (500), Dst Port: isakmp (500)
Internet Security Association and Key Management Protocol
  Initiator cookie: 92585D2D797E9C52
  Responder cookie: 34704CFC8C8DBD09
  Next payload: Hash (8)
  Version: 1.0
  Exchange type: Quick Mode (32)
  Flags: 0x01
  Message ID: 0x79a63fb1
  Length: 292
  Encrypted payload (264 bytes)

```

No.	Time	Source	Destination	Protocol	Info
9210	768.411797	10.20.131.62	10.20.129.80	ISAKMP	Quick Mode

```

Frame 9210 (334 bytes on wire, 334 bytes captured)
Ethernet II, Src: Cisco_80:70:f5 (00:13:c4:80:70:f5),
      Dst: Vmware_9d:2c:dd (00:50:56:9d:2c:dd)
Internet Protocol, Src: 10.20.131.62 (10.20.131.62),
      Dst: 10.20.129.80 (10.20.129.80)
User Datagram Protocol, Src Port: isakmp (500), Dst Port: isakmp (500)
Internet Security Association and Key Management Protocol
  Initiator cookie: 92585D2D797E9C52
  Responder cookie: 34704CFC8C8DBD09
  Next payload: Hash (8)
  Version: 1.0
  Exchange type: Quick Mode (32)
  Flags: 0x01
  Message ID: 0x79a63fb1
  Length: 292
  Encrypted payload (264 bytes)

```

No.	Time	Source	Destination	Protocol	Info
9211	768.437057	10.20.129.80	10.20.131.62	ISAKMP	Quick Mode

```

Frame 9211 (94 bytes on wire, 94 bytes captured)
Ethernet II, Src: Vmware_9d:2c:dd (00:50:56:9d:2c:dd),
      Dst: Cisco_80:70:f5 (00:13:c4:80:70:f5)
Internet Protocol, Src: 10.20.129.80 (10.20.129.80),
      Dst: 10.20.131.62 (10.20.131.62)
User Datagram Protocol, Src Port: isakmp (500), Dst Port: isakmp (500)
Internet Security Association and Key Management Protocol
  Initiator cookie: 92585D2D797E9C52
  Responder cookie: 34704CFC8C8DBD09
  Next payload: Hash (8)
  Version: 1.0
  Exchange type: Quick Mode (32)
  Flags: 0x01
  Message ID: 0x79a63fb1
  Length: 52
  Encrypted payload (24 bytes)

```

NSX Controller 故障排除

8

本节提供了有关确定 NSX Controller 故障原因和排除控制器故障的信息。

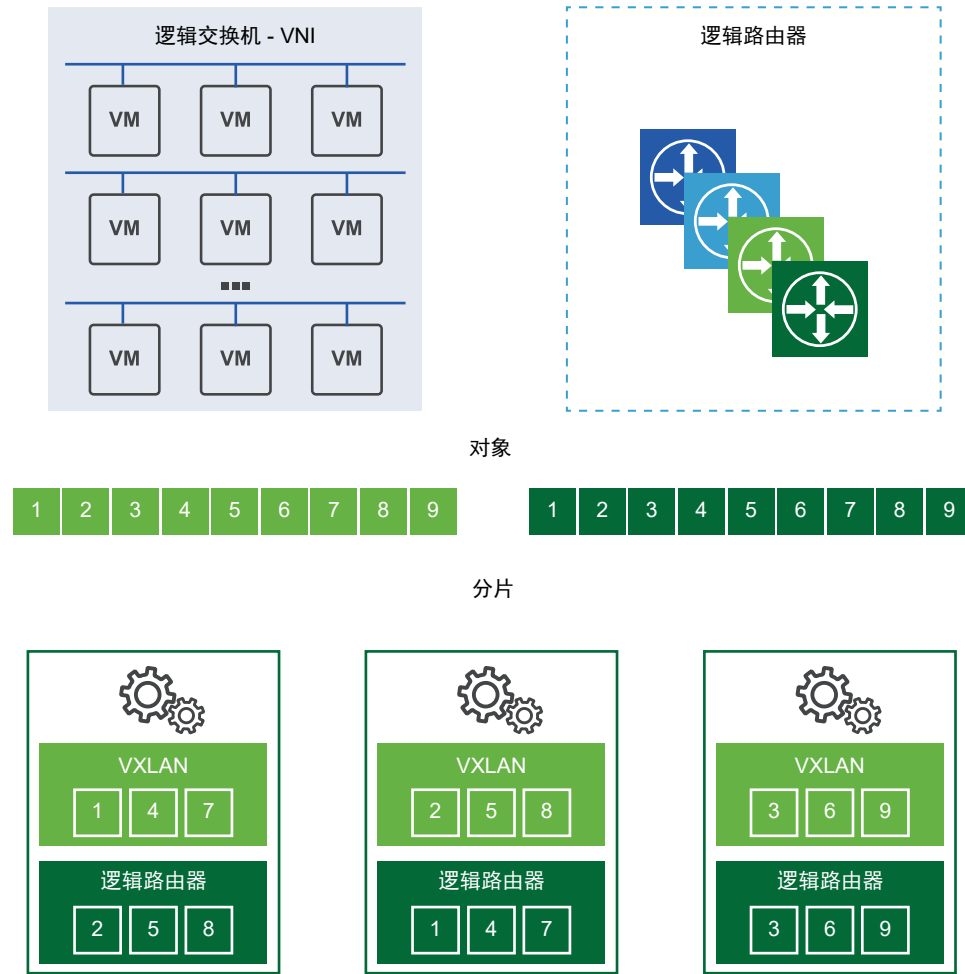
本章讨论了以下主题：

- [了解控制器群集架构](#)
- [NSX Controller 部署问题](#)
- [磁盘延迟故障排除](#)
- [NSX Controller 群集故障](#)
- [NSX Controller 已断开连接](#)
- [控制层面代理 \(netcpa\) 问题](#)

了解控制器群集架构

NSX Controller 群集表示一个横向扩展分布式系统，将为其中的每个控制器节点分配一组角色，这些角色定义了该节点可以执行的任务类型。为了提高弹性和性能，应该在三个不同的主机中部署控制器虚拟机。

分片用于在 NSX Controller 群集节点之间分配工作负载。分片是指将 NSX Controller 工作负载拆分为不同分片的操作，以便每个 NSX Controller 实例具有相等的工作量。



这说明了不同的控制器节点如何作为给定实体的主控制器节点，例如，逻辑交换、逻辑路由和其他服务。在为某个角色选择主 **NSX Controller** 实例后，该 **NSX Controller** 在群集中的所有可用 **NSX Controller** 实例之间拆分不同的逻辑交换机和路由器。

分片上的每个编号框表示主控制器用于拆分工作负载的分片。逻辑交换机主控制器将逻辑交换机拆分为分片，并将这些分片分配给不同的 **NSX Controller** 实例。逻辑路由器的主控制器还会将逻辑路由器拆分为分片，并将这些分片分配给不同的 **NSX Controller** 实例。

这些分片将分配给该群集中的不同 **NSX Controller** 实例。角色的主控制器决定了将哪些 **NSX Controller** 实例分配给哪个分片。如果将请求传送到路由器分片 3，则指示该分片连接到第三个 **NSX Controller** 实例。如果将请求传送到逻辑交换机分片 2，则第二个 **NSX Controller** 实例处理该请求。

在群集中的某个 **NSX Controller** 实例发生故障时，角色的主控制器将分片重新分配给其余的可用实例。将选举某个控制器节点以作为每个角色的主控制器。主控制器负责将分片分配给各个控制器节点，确定节点何时发生故障以及将这些分片重新分配给其他节点。主控制器还会向 **ESXi** 主机通知群集节点故障情况。

要为每个角色选举主控制器，需要获得群集中的所有活动和非活动节点的多数选票。这就是必须始终为控制器群集部署奇数节点的主要原因。

ZooKeeper

ZooKeeper 是负责 **NSX Controller** 群集机制的客户端服务器架构。控制器群集是使用 **ZooKeeper** 发现和创建的。在群集启动时，实际表示在所有节点之间启动 **ZooKeeper**。**ZooKeeper** 节点执行选举过程以形成控制群集。在群集中必须具有一个 **ZooKeeper** 主节点。这是通过节点间的选举完成的。

在创建新的控制器节点时，**NSX Manager** 将节点信息以及节点 IP 和 ID 传播到当前群集。因此，每个节点了解可用于创建群集的总节点数。在 **ZooKeeper** 主节点选举期间，每个节点投一票以选举主节点。将再次触发选举，直到某个节点获得多数选票。例如，在三节点群集中，主节点必须至少获得两票。

注 为了防止出现无法选举 **ZooKeeper** 主节点的情况，群集中的节点数必须为三个。

- 在部署第一个控制器时，这是一种特殊情况，并且第一个控制器变为主控制器。因此，在部署控制器时，第一个节点必须完成部署，然后再添加任何其他节点。
- 在添加第二个控制器时，这也是一种特殊情况，因为此时的节点数为偶数。
- 在添加第三个节点时，群集达到支持的稳定状态。

ZooKeeper 每次只能承受一个故障。这意味着，如果一个控制器节点发生故障，必须在发生任何其他故障之前恢复该节点。否则，可能会出现群集损坏问题。

中央控制层面 (Central Control Plane, CCP) 域管理器

这是 **ZooKeeper** 上面的层，它提供配置以启动所有节点上的 **ZooKeeper**。域管理器在群集中的所有节点之间更新配置，然后进行远程过程调用以启动 **ZooKeeper** 进程。

域管理器负责启动所有域。要加入群集，**CCP** 域与其他计算机上的 **CCP** 域进行通信。帮助进行群集初始化的 **CCP** 域组件为 *zk-cluster-bootstrap*。

控制器与其他组件的关系

控制器群集负责维护有关逻辑交换机、逻辑路由器和 **VTEP** 的信息，并向 **ESXi** 主机提供这些信息。

在创建逻辑交换机时，群集中的控制器节点确定哪个节点是该逻辑交换机的主节点或所有者。在添加逻辑路由器时，这同样适用。

在为逻辑交换机或逻辑路由器确定所有权后，该节点将该所有权信息发送到属于该交换机或路由器的传输区域的 **ESXi** 主机。选举所有权以及将所有权信息传播到主机的整个过程称为“分片”。请注意，所有权表示节点负责该逻辑交换机或逻辑路由器的所有 **NSX** 相关操作。其他节点不会为该逻辑交换机执行任何操作。

由于只有一个所有者必须是逻辑交换机和逻辑路由器的真实数据源，因此，只要控制器群集以某种方式损坏以使两个或更多个节点选举为逻辑交换机或逻辑路由器的所有者，网络中的每个主机可能具有有关该逻辑交换机或逻辑路由器的真实数据源的不同信息。如果发生这种情况，网络将会发生故障，因为网络控制 and 数据层面操作只能具有一个真实数据源。

如果控制器节点发生故障，群集中的其余节点将重新运行分片以确定逻辑交换机和逻辑路由的所有权。

NSX Controller 部署问题

NSX Controller 是 NSX Manager 使用 OVA 格式部署的。具有控制器群集可以提供高可用性。部署控制器要求 NSX Manager、vCenter Server 和 ESXi 主机配置了 DNS 和 NTP。必须使用静态 IP 池为每个控制器分配 IP 地址。

建议您实施 DRS 反关联性规则以使 NSX Controller 位于单独的主机上。您必须部署三个 NSX Controller。

常见的控制器问题

在部署 NSX Controller 期间，可能遇到的典型问题如下所示：

- NSX Controller 部署失败。
- NSX Controller 无法加入群集。
- 运行 `show control-cluster status` 命令将显示 Majority status 在 Connected to cluster majority 和 Interrupted connection to cluster majority 之间变化。
- NSX 仪表板显示连接状态问题。
- `show control-cluster status` 命令是建议用于查看控制器是否加入控制群集的命令。您需要在每个控制器上运行该命令以确定总体群集状态。

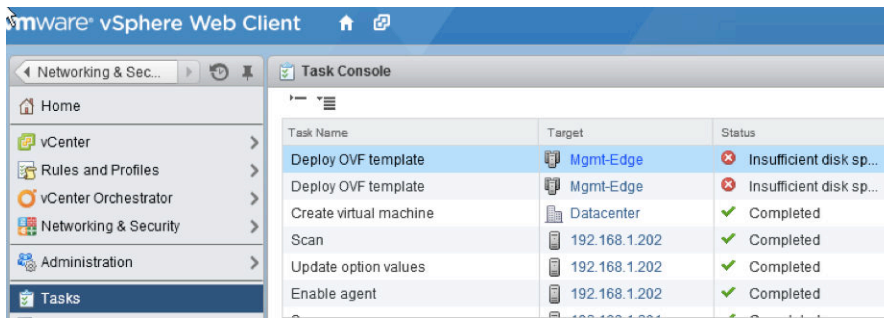
```
controller # show control-cluster status
Type                Status                                     Since
-----
Join status:        Join complete                               10/17 18:16:58
Majority status:    Connected to cluster majority                       10/17 18:16:46
Restart status:     This controller can be safely restarted                 10/17 18:16:51
Cluster ID:         af2e9dec-19b9-4530-8e68-944188584268
Node UUID:          af2e9dec-19b9-4530-8e68-944188584268
Role                Configured status  Active status
-----
api_provider        enabled           activated
persistence_server  enabled           activated
switch_manager      enabled           activated
logical_manager     enabled           activated
dht_node            enabled           activated
```

注 如果发现控制器节点断开连接，请不要使用 `join cluster` 或 `force join` 命令。该命令不用于将节点加入群集。如果这样做，群集可能会进入完全不确定的状态。

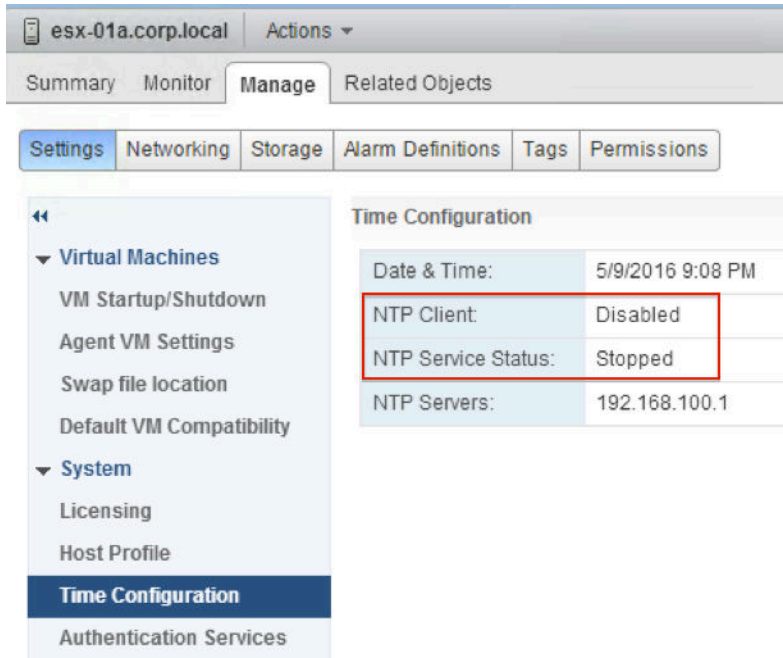
群集启动节点仅仅指的是，在其中查看成员何时启动的群集成员。如果该列表包含不再使用的群集成员，请不要担心。这不会影响群集功能。

所有群集成员应具有相同的群集 ID。否则，群集将处于损坏状态，您应该与 VMware 技术支持人员一起修复该问题。

- `show control-cluster startup-nodes` 命令不用于显示群集中的所有当前节点，而是显示在控制器进程重新启动时该节点使用哪些其他控制器节点以在群集中添加成员。因此，命令输出可能显示已关闭或以其他方式从群集中删除的某些节点。
- 此外，`show control-cluster network ipsec status` 命令还允许检查 Internet 协议安全 (Internet Protocol Security, IPsec) 状态。如果您发现控制器在几分钟到几小时内无法相互通信，请运行 `cat /var/log/syslog | egrep "sending DPD request|IKE_SA"` 命令并查看日志消息是否指示没有流量。也可以运行 `ipsec statusall | egrep "bytes_i|bytes_o"` 命令并确认没有建立两个 IPsec 隧道。在向 VMware 技术支持代表报告可疑的控制群集问题时，请提供这些命令的输出和控制器日志。
- NSX Manager 和 NSX Controller 之间出现 IP 连接问题。这通常是物理网络连接问题或防火墙阻止通信造成的。
- 在 vSphere 上没有足够的资源（如存储）以托管控制器。在控制器部署期间，可以查看 vCenter 事件和任务日志以发现这些问题。



- “恶意”控制器出现异常或升级的控制器处于**已断开连接 (Disconnected)**状态。
- 未正确配置 ESXi 主机和 NSX Manager 上的 DNS。
- ESXi 主机和 NSX Manager 上的 NTP 不同步。



- 在新连接的虚拟机无法访问网络时，这可能是控制层面问题造成的。检查控制器状态。

还要尝试在 ESXi 主机上运行 `esxcli network vswitch dvs vmware vxlan network list --vds-name <name>` 命令以检查控制层面状态。请注意，控制器连接中断。

```
/etc/vmware/netcpa # esxcli network vswitch dvs vmware vxlan network list --vds-name Compute_VDS
VXLAN ID Multicast IP Control Plane Controller Connection
ARP Entry Count MTEP Count
-----
5000 N/A (headend replication) Enabled (multicast proxy, ARP proxy) 192.168.110.203 (down)
```

- 运行 `show log manager followNSX Manager CLI` 命令可以找出无法部署控制器的任何其他原因。

```
2014-02-26 10:09:44.931 GMT INFO taskScheduler-25 VcConnection$VimClient:1219 - Create stub for com.vmware.vim.binding
28c5157-abf3-718e-88c5-42209f389211
2014-02-26 10:09:44.932 GMT DEBUG VcEventsReaderThread VcEventsReader$VcEventsReaderThread:301 - got prop collector up
ctReference: type = PropertyFilter, value = session[d46b86a2-7a10-c17e-6ebe-8ab252ee4efd]527420f2-bdd7-529b-8ab6-17d16
6E3-4A64-96D7-5833C287588F
2014-02-26 10:09:44.937 GMT ERROR taskScheduler-25 VCUtils:184 - Error while waiting for property collector updates.
com.vmware.vim.binding.vim.fault.NoDiskSpace:
datastore = datastore1 (1)
inherited from com.vmware.vim.binding.vim.fault.FileFault:
file = [datastore1 (1)] NSX_Controller_1c3dd18d-0cd3-4d7d-896b-51247176ae77/NSX_Controller_1c3dd18d-0cd3-4d7d-896b-512
inherited from com.vmware.vim.binding.vim.fault.VimFault:
inherited from com.vmware.vim.binding.vim.fault.NoDiskSpace: Insufficient disk space on datastore 'datastore1 (1)'.
```

主机连接问题

可以使用以下命令检查主机连接错误。请在每个控制器节点上运行这些命令。

- 使用 `show log cloudnet/cloudnet_java-vnet-controller*.log filtered-by host_IP` 命令检查任何异常错误统计信息。
- 使用以下命令验证逻辑交换机/路由器消息统计信息或高消息速率：
 - `show control-cluster core stats:overall stats`
 - `show control-cluster core stats-sample:latest stats samples`

- `show control-cluster core connection-stats ip: per connection stats`
- `show control-cluster logical-switches stats`
- `show control-cluster logical-routers stats`
- `show control-cluster logical-switches stats-sample`
- `show control-cluster logical-routers stats-sample`
- `show control-cluster logical-switches vni-stats vni`
- `show control-cluster logical-switches vni-stats-sample vni`
- `show control-cluster logical-switches connection-stats ip`
- `show control-cluster logical-routers connection-stats ip`
- 您可以使用 `show host hostID health-status` 命令检查准备的群集中的主机的运行状况。对于控制器故障排除，支持以下运行状况检查：
 - 检查 `net-config-by-vsm.xml` 是否同步到控制器列表。
 - 检查是否具有到控制器的套接字连接。
 - 检查是否创建了 VXLAN 网络标识符 (VXLAN Network Identifier, VNI) 以及配置是否正确。
 - 检查 VNI 是否连接到主控制器（如果启用了控制层面）。

安装和部署问题

- 验证是否在群集中部署了至少三个控制器节点。VMware 建议使用本机 vSphere 反关联性规则，以避免在同一 ESXi 主机上部署多个控制器节点。
- 验证所有 NSX Controller 是否显示 **Connected** 状态。如果任何控制器节点显示 **Disconnected** 状态，请在所有控制器节点上运行 `show control-cluster status` 命令以确保以下信息是一致的：

类型	状态
Join status	Join complete
Majority status	Connected to cluster majority
Cluster ID	在所有控制器节点上具有相同的信息

- 确保所有角色在所有控制器节点上是一致的：

角色	配置状态	活动状态
api_provider	enabled	activated
persistence_server	enabled	activated
switch_manager	enabled	activated
logical_manager	enabled	activated
directory_server	enabled	activated

- 验证 `vnet-controller` 进程是否正在运行。在所有控制器节点上运行 `show process` 命令并确保 `java-dir-server` 服务正在运行。
- 验证群集历史记录，并确保没有迹象表明主机连接发生变化，VNI 加入失败或群集成员资格发生异常变化。要验证这种情况，请运行 `show control-cluster history` 命令。这些命令还会显示节点是否频繁重新启动。确认没有很多具有零 (0) 大小和不同进程 ID 的日志文件。
- 验证是否配置了 VXLAN 网络标识符 (VXLAN Network Identifier, VNI)。有关详细信息，请参阅《VMware VXLAN Deployment Guide》的 VXLAN 准备步骤部分。
- 验证是否在控制器群集上启用了 SSL。请在每个控制器节点上运行 `show log cloudnet/cloudnet_java-vnet-controller*.log filtered-by sslEnabled` 命令。

磁盘延迟故障排除

您可以从**管理 (Management)**选项卡中查看磁盘延迟警示。NSX Controller 必须在具有较低延迟的磁盘上运行。

查看磁盘延迟警报

磁盘延迟警报监控和报告磁盘可用性或延迟问题。您可以查看每个 NSX Controller 的磁盘延迟详细信息。读取延迟和写入延迟计算结果计入 5 秒（默认）移动平均值，而该结果用于在违反延迟限制时触发警报。在平均值降低到低水位线后，将关闭警报。默认情况下，高水位线设置为 200 毫秒，低水位线设置为 100 毫秒。较高的延迟影响每个控制器节点上的分布式群集应用程序运行。

要查看 NSX Controller 的磁盘延迟警报，请执行以下过程：







前提条件



已达到延迟限制。

步骤

- 1 登录到 vSphere Web Client。
- 2 单击**网络和安全 (Networking & Security)**，然后单击**安装 (Installation)**。

- 3 在**管理 (Management)**下面，转到所需的控制器，然后单击**磁盘警报 (Disk Alert)**链接。
将显示“磁盘延迟警报”窗口。

192.168.110.33 - Disk Latency Alerts				
Device Name	Latency Type	Refresh Time	Last Latency (ms)	Average Latency (ms)
 sda	Write	9/26/2016 2:15 PM	3.2	7.906
 sda	Read	9/26/2016 1:08 PM	0.0	0.0
 dm-1	Write	9/16/2016 5:11 PM	0.0	0.0
 dm-1	Read	9/22/2016 4:31 PM	0.0	0.0
 dm-0	Write	9/26/2016 2:15 PM	3.64	9.822
 dm-0	Read	9/26/2016 10:05 AM	0.0	33.334
6 items				

5	 Disk Alert	 Disk Alert
---	---	--

结果

您可以查看选定控制器的延迟详细信息。警报日志在 `cloudnet/run/iostat/iostat_alert.log` 文件中存储 7 天。您可以使用 `show log cloudnet/run/iostat/iostat_alert.log` 命令显示日志文件。

后续步骤

有关磁盘延迟的详细故障排除信息，请参阅[磁盘延迟问题](#)。

有关日志消息的详细信息，请参阅 **NSX** 日志记录和系统事件。

磁盘延迟问题

控制器必须在具有较低延迟的磁盘上运行。群集要求每个节点的磁盘存储系统具有小于 300 毫秒的峰值写入延迟以及小于 100 毫秒的平均写入延迟。

问题

- 部署的 NSX Controller 与控制器群集断开连接。
- 由于磁盘分区已满，无法收集任何控制器日志。
- 如果存储系统不满足这些要求，则群集可能变得不稳定，并且导致系统停机时间。

- 在 `show network connections of-type tcp` 命令输出中不再显示适用于正常工作的 NSX Controller 的 TCP 侦听器。
- 断开连接的控制器尝试使用无效的全零 UUID 加入群集。
- `show control-cluster history` 命令显示类似下面的消息：

```
INFO.20150530-000550.1774:D0530 13:25:29.452639 1983 zookeeper_client.cc:774] Zookeeper client disconnected!
```
- 在 NSX Controller 控制台中运行 `show log cloudnet/cloudnet_java-zookeeper*.log` 命令将包含类似下面的条目：

```
cloudnet_java-zookeeper.20150530-000550.1806.log-2015-05-30
13:25:07,382 47956539 [SyncThread:1] WARN
org.apache.zookeeper.server.persistence.FileTxnLog - fsync-ing the write ahead
log in SyncThread:1 took 3219ms which will adversely effect operation latency.
See the ZooKeeper troubleshooting guide
```

- NSX Controller 日志包含类似下面的条目：

```
D0525 13:46:07.185200 31975
rpc-broker.cc:369] Registering address resolution for: 20.5.1.11:7777
D0525 13:46:07.185246 31975
rpc-tcp.cc:548] Handshake complete, both peers support the same
protocol
D0525 13:46:07.197654 31975
rpc-tcp.cc:1048] Rejecting a connection from peer
20.5.1.11:42195/ef447643-f05d-4862-be2c-35630df39060, cluster
9f7ea8ff-ab80-4c0c-825e-628e834aa8a5, which doesn't match our cluster
(00000000-0000-0000-0000-000000000000)
D0525 13:46:07.222869 31975
rpc-tcp.cc:1048] Rejecting a connection from peer
20.5.1.11:42195/ef447643-f05d-4862-be2c-35630df39060, cluster
9f7ea8ff-ab80-4c0c-825e-628e834aa8a5, which doesn't match our cluster
(00000000-0000-0000-0000-000000000000)
```

原因

出现该问题是因为磁盘速度较慢，这会对 NSX Controller 群集造成不利的影响。

- 在 `/var/log/cloudnet/cloudnet_java-zookeeper log` 文件中查找 *fsync* 消息以检查较慢的磁盘。如果 *fsync* 所需的时间超过 1 秒，Zookeeper 将显示 *fsync* 警告消息，这清楚地指明了磁盘太慢。VMware 建议将一个逻辑单元号 (Logical Unit Number, LUN) 明确专用于控制群集以及/或者将存储阵列移到离控制群集更近的位置以降低延迟。

- 您可以查看读取延迟和写入延迟计算结果（默认情况下，计入 5 秒移动平均值），而该结果用于在违反延迟限制时触发警报。在平均值降低到低水位线后，将关闭警报。默认情况下，高水位线设置为 200 毫秒，低水位线设置为 100 毫秒。您可以使用 `show disk-latency-alert config` 命令。输出如下所示：

```
enabled=True    low-wm=51      high-wm=150
nsx-controller # set disk-latency-alert enabled yes
nsx-controller # set disk-latency-alert low-wm 100
nsx-controller # set disk-latency-alert high-wm 200
```

- 可以使用 `GET /api/2.0/vdn/controller/<controller-id>/systemStats` REST API 获取控制器节点的延迟警报状态。
- 可以使用 `GET /api/2.0/vdn/controller` REST API 指示是否在控制器节点上检测到磁盘延迟警报。

解决方案

- 1 在低延迟磁盘上部署 NSX Controller。
- 2 每个控制器应使用自己的磁盘存储服务器。不要在两个控制器之间共享同一磁盘存储服务器。

后续步骤

有关如何查看警报的详细信息，请参阅[查看磁盘延迟警报](#)。

NSX Controller 群集故障

当群集中的一个 NSX Controller 节点出现故障时，仍会有两个控制器在工作。此时群集的大多数节点保持工作，并且控制层面仍继续正常工作。

问题

NSX Controller 群集出现故障。

解决方案

- 1 登录到 vSphere Web Client。
- 2 从网络和安全性 (Networking & Security) 中，单击安装 > 管理 (Installation > Management)。
- 3 在“NSX Controller 节点”部分中，查看“对等”列。如果“对等”列显示绿色框，则表示群集中的对等控制器连接没有出现错误。红色框表示对等控制器连接出现错误。请单击该框以查看详细信息。
- 4 如果“对等”列显示控制器群集出现问题，请登录到每个 NSX Controller CLI 以执行详细诊断。请运行 `show control-cluster status` 命令以诊断每个控制器的状态。群集中的所有控制器必须具有相同的群集 UUID，但群集 UUID 可能与主控制器的 UUID 不同。您可以按照 [NSX Controller 部署问题](#) 中所述查找有关部署问题的信息。

5 在重新部署控制器节点或控制器群集之前，您可以尝试使用以下步骤来解决问题：

- a 检查是否打开了控制器电源。
- b 尝试在受影响的控制器与其他节点和管理器之间执行 ping 操作以检查网络路径。如果发现任何网络问题，请按照 [NSX Controller 部署问题](#) 中所述解决这些问题。
- c 使用以下 CLI 命令检查 Internet 协议安全 (IPSec) 状态。
 - 使用 `show control-cluster network ipsec status` 命令验证是否启用了 IPSec。
 - 使用 `show control-cluster network ipsec tunnels` 命令验证 IPSec 隧道的状态。

也可以使用 IPSec 状态信息向 VMware 技术支持人员提交凭单。

- d 如果该问题不是网络问题，您可以选择是重新引导还是重新部署。

如果要重新引导节点，请确保每次只重新引导一个控制器。不过，如果控制器群集处于某种状态，其中的多个控制器节点发生故障，请同时重新引导所有这些节点。在重新引导正常运行的群集中的节点时，请始终确认随后正确重建了该群集，然后确认已正确完成群集重新分片。

6 如果决定重新部署控制器，请使用以下两种方法之一：

- 方法 1：删除已损坏的控制器节点并重新部署新的控制器节点。
- 方法 2：删除控制器群集并重新部署新的控制器群集。

VMware 建议采用第二种方法。

后续步骤

选择任意一种方法：

- [方法 1：删除已损坏的控制器并重新部署新的控制器](#)
- [方法 2：重新部署 NSX Controller 群集](#)

方法 1：删除已损坏的控制器并重新部署新的控制器

您可以先尝试在不重新部署新 NSX Controller 群集的情况下解决该问题。在此方法中，先删除已损坏的 NSX Controller 节点，然后再部署新的 NSX Controller 节点。

步骤

1 删除 NSX Controller

您可以强制或正常删除 NSX Controller。正常移除过程在移除节点之前检查以下情况：

2 重新部署 NSX Controller

删除已损坏的控制器节点后，部署新的控制器节点。

删除 NSX Controller

您可以强制或正常删除 NSX Controller。正常移除过程在移除节点之前检查以下情况：

- 当前未执行 NSX Controller 节点升级操作。

- 控制器群集正常运行，并且可以处理控制器群集 API 请求。
- 主机状态（从 vCenter Server 清单中获取）显示已连接并打开电源。
- 这不是最后一个控制器节点。

强制移除过程不会在移除控制器节点之前检查上述情况。

- 在删除控制器时注意的事项：
 - 在通过 vSphere Web Client UI 或 API 删除之前，不要尝试删除控制器虚拟机。当 UI 不可使用时，请使用 `DELETE /2.0/vdn/controller/{controllerId}` API 删除控制器。
 - 在删除节点后，请确保现有的群集保持稳定。
 - 在删除群集中的所有节点时，必须使用**强制移除控制器 (Forcefully remove the controller)**选项删除剩下的最后一个节点。始终验证是否成功删除了控制器虚拟机。如果不成功，请手动关闭虚拟机电源并使用此 UI 删除控制器虚拟机。
 - 如果删除操作失败，则意味着无法删除虚拟机。在这种情况下，请通过 UI 使用**强制移除控制器 (Forcefully remove the controller)**选项调用控制器删除。对于 API，将 `forceRemoval` 参数设置为 `true`。在强制移除之后，请手动关闭虚拟机电源并使用此 UI 删除控制器虚拟机。
 - 由于多节点群集只能承受一个故障，并且删除操作计为一个故障，因此，必须在发生另一个故障之前重新部署删除的节点。
- 对于跨 vCenter NSX 环境：
 - 直接在 vCenter Server 中删除控制器虚拟机或关闭虚拟机电源是不支持的操作。**状态 (Status)**列将显示**不同步 (Out of sync)**状态。
 - 如果控制器删除操作仅部分成功，并在跨 vCenter NSX 环境的 NSX Manager 数据库中遗留一个条目，请使用 `DELETE api/2.0/vdn/controller/external` API。
 - 如果控制器是通过 NSX Manager API 导入的，请将 `removeExternalControllerReference` API 与 `forceRemoval` 选项一起使用。
 - 在删除控制器时，NSX 请求通过 vCenter Server 使用控制器虚拟机的受管对象 ID (Managed Object ID, MOID) 删除该虚拟机。如果 vCenter Server 按 MOID 找不到虚拟机，NSX 将报告控制器删除请求失败并中止该操作。

如果选择**强制删除 (Forcefully Delete)**选项，NSX 不会中止控制器删除操作并清除控制器的信息。NSX 还会更新所有主机以不再信任已删除的控制器。不过，如果控制器虚拟机仍处于活动状态并使用不同的 MOID 运行，它仍具有以控制器群集成员身份加入的凭据。在这种场景下，分配给此控制器节点的任何逻辑交换机或路由器都将无法正常运行，因为 ESXi 主机不再信任已删除的控制器。

要删除 NSX Controller，请执行以下过程：

步骤

- 1 登录到 vSphere Web Client。
- 2 单击**网络和安全 (Networking & Security)**，然后单击**安装 (Installation)**。

3 在**管理 (Management)**下面，选择要删除的控制器。

4 单击**删除 (x) (Delete (x))** 图标。

5 选择**删除 (Delete)**或**强制删除 (Forcefully Delete)**。

- ◆ 在选择**强制删除 (Forcefully Delete)**选项时，将强制而不是正常删除控制器。该选项忽略任何失败并从数据库中清除该数据。您应该确认已手动处理任何可能的失败。您必须确认已成功删除控制器虚拟机。如果失败，您必须通过 **vCenter Server** 删除控制器虚拟机。

注 如果删除群集中的最后一个控制器，您必须选择**强制删除 (Forcefully Delete)**选项以移除最后一个控制器节点。当系统中不存在任何控制器时，主机将在所谓的“无头”模式下工作。新虚拟机或已执行 vMotion 操作的虚拟机将遇到网络问题，直至部署了新的控制器并且同步已完成为止。

- ◆ 如果未选择该选项，将正常删除控制器。

6 单击**是 (Yes)**。正常控制器删除执行以下步骤：

- 关闭节点电源。
- 检查群集运行状况。
- 如果群集未正常运行，则打开控制器电源并放弃移除请求。
- 如果群集正常运行，则移除控制器虚拟机并释放节点的 IP 地址。
- 从群集中移除控制器虚拟机的标识。

将删除选定的控制器。

7 单击**操作 > 更新控制器状态 (Actions > Update Controller State)**以重新同步控制器状态。

重新部署 NSX Controller

删除已损坏的控制器节点后，部署新的控制器节点。

步骤

- 1 登录到 vSphere Web Client。
- 2 从**网络和安全 (Networking & Security)** 中，单击**安装 > 管理 (Installation > Management)**。
- 3 在 **NSX Controller** 节点部分中，单击受影响的控制器。请截屏或记下 **NSX Controller 详细信息** 屏幕中的配置信息，以供将来参考。

例如：



- 4 单击**添加节点 (+) (Add Node (+))** 图标以部署新的 NSX Controller 节点。
- 5 在“添加控制器”对话框中，选择要添加节点的数据中心，然后配置控制器设置。
 - a 选择适当的群集。
 - b 在群集和存储中选择一个主机。
 - c 选择分布式端口组。
 - d 选择要将其中的 IP 地址分配给节点的 IP 池。
 - e 单击**确定 (OK)**，等待安装完成，并确保该节点的状态为**正常**。有关添加控制器节点的详细信息，请参见 NSX 安装指南 中的“部署 NSX Controller 群集”。
- 6 重新同步控制器状态，方法是单击**操作 > 更新控制器状态**。

“更新控制器状态”将当前 VXLAN 和分布式逻辑路由器配置（包括跨 vCenter NSX 部署中的通用对象）从 NSX Manager 推送到控制器群集。

方法 2：重新部署 NSX Controller 群集

在此方法中，删除所有三个控制器节点，然后添加新的控制器节点，以便维护完全正常运行的三节点群集。

如果存在以下任一情况，VMware 建议删除 NSX Controller 群集：

- 一个或多个控制器节点面临灾难性或不可恢复的错误。
- 控制器虚拟机无法访问且无法修复。

在这种情况下，即使某些控制器节点看似正常运行，但最好还是删除所有控制器节点。

重新部署新的控制器群集，然后更新 NSX Manager 上的控制器状态。更新控制器状态会导致重新同步 VXLAN 并重新部署分布式逻辑路由器。

步骤

- 1 登录到 vSphere Web Client。
- 2 导航到**网络和安全 > 安装 > 管理**。
- 3 在 **NSX Controller 节点** 部分中，将三个控制器节点全部删除。一次选择一个节点，然后单击**删除** (✖) 图标。

当系统中不存在任何控制器时，主机将在“无头 (headless)”模式下运行。新虚拟机或已迁移的虚拟机将遇到网络问题，直至部署了新的控制器并且完成同步为止。

- 4 部署三个新控制器节点以创建完全正常运行的 NSX Controller 群集。

有关添加控制器群集的详细信息，请参见 NSX 安装指南 中的“部署 NSX Controller 群集”。
- 5 重新同步控制器状态，方法是单击**操作 > 更新控制器状态**。

幻影控制器

幻影控制器可能是实时控制器虚拟机 (VM) 或不存在的虚拟机，它们可能会加入群集，也可能不会加入群集。NSX Manager 同步 vCenter Server 清单中的所有虚拟机的列表。如果 vCenter Server 或主机删除控制器虚拟机而没有来自 NSX Manager 的请求，或者 vCenter Server 清单更改控制器虚拟机的引用 MOID，则会创建一个幻影控制器。

如果从 NSX 中创建控制器，配置信息将存储在 NSX Manager 中。NSX Manager 通过 vCenter Server 部署新的控制器虚拟机。

NSX 管理员为 NSX Manager 提供配置（包括 IP 地址池）以创建一个控制器。NSX Manager 从池中移除一个 IP 地址，并将该 IP 以及其余控制器配置作为虚拟机创建请求推送到 vCenter Server。NSX Manager 等待 vCenter Server 确认该请求的状态。

- **The controller creation process was successful:** 如果成功创建控制器虚拟机，vCenter Server 将启动控制器虚拟机。NSX Manager 存储虚拟机的受管对象 ID (MOID) 以及其余控制器配置信息。MOID（或 MO-REF）是 vCenter 为其清单中的每个对象分配的唯一标识符。如果虚拟机仍然是 vCenter Server 清单的一部分，vCenter Server 还使用该 MOID 跟踪虚拟机。
- **The controller creation process was not successful:** 如果 IP 和网络连接配置不正确，则 NSX Manager 可能无法连接到 vCenter Server。NSX Manager 等待预设的时间以创建一个单节点控制器群集（对于第一个群集），或者将新控制器加入活动群集。如果定时器到期，NSX Manager 将请求 vCenter Server 删除虚拟机。IP 地址将返回到池中，并且 NSX 声明控制器创建失败。

幻影控制器如何生成

在 NSX Manager 请求删除控制器时，vCenter Server 使用 MOID 查找要删除的控制器虚拟机。

不过，如果任何 vCenter 活动导致从 vCenter Server 清单中移除控制器虚拟机，vCenter 将从其数据库中移除该 MOID。请注意，即使从 vCenter 清单中移除后，控制器虚拟机可能仍会在 NSX Manager 上处于活动状态。但对于 vCenter Server 来说，控制器虚拟机不再存在。即使 vCenter Server 已从其清单中移除虚拟机，可能也不会删除该虚拟机。如果虚拟机仍处于活动状态，则它仍在加入或尝试加入 NSX 控制器群集。

下面是幻影控制器如何生成的最常见示例：

- vCenter Server 管理员从清单中移除包含控制器虚拟机的主机。然后，重新添加主机。在移除主机时，vCenter Server 删除与主机以及其中的虚拟机关联的所有 MOID。在随后重新添加主机时，vCenter Server 为主机和虚拟机分配全新的 MOID。对于 NSX 用户来说，主机和虚拟机仍然是相同的，但从 vCenter Server 的角度看，主机和虚拟机是全新的对象。但实际上，主机和虚拟机仍然是相同的。在主机和虚拟机中运行的应用程序不会发生变化。
- vCenter Server 管理员通过 vCenter Server 或使用主机管理删除控制器虚拟机。这种删除不是由 NSX Manager 发起的请求。
- 此处的“删除”还包括导致虚拟机丢失的任何主机/存储故障。在这种情况下，虚拟机对于 vCenter Server 以及群集和 NSX Manager 为缺失状态。但由于删除不是由 NSX Manager 发起的请求，因此，NSX Manager 和控制器群集认为控制器仍然有效。返回到 NSX Manager 的控制器状态指示该控制器节点已关闭并且不属于群集，而不会显示在 UI 上。NSX Manager 还具有一些日志以指示无法再访问控制器。

发现幻影控制器时执行的操作

- 1 按照 [NSX Controller 已断开连接](#) 中所述，同步控制器。
- 2 请参见日志条目。如果控制器虚拟机已意外删除或损坏，您必须使用**强制删除 (Forcefully Delete)**选项从 NSX Manager 数据库中清除该条目。有关详细信息，请参阅[删除 NSX Controller](#)。
- 3 在删除控制器后，请确认：
 - 实际删除了控制器虚拟机。
 - `show controller-cluster startup-nodes` 命令仅显示有效的控制器。
 - NSX Manager 的 syslog 条目不再显示额外的控制器。

从 NSX 6.2.7 或更高版本开始，NSX Manager 根据原始 MOID 验证 vCenter 清单，以确保控制器虚拟机在清单中仍然存在。如果 NSX Manager 在清单中找不到控制器虚拟机，则 NSX Manager 使用虚拟机的实例 UUID 搜索虚拟机。实例 UUID 存储在虚拟机中，因此，即使将虚拟机重新添加到 vCenter 清单，该 ID 也不会发生变化。如果 NSX Manager 可以使用实例 UUID 找到虚拟机，则 NSX Manager 使用新的 MOID 更新其数据库。

不过，如果您克隆控制器虚拟机，则克隆的虚拟机具有与原始虚拟机相同的属性以及新的实例 UUID。NSX Manager 无法检测克隆的虚拟机的 MOID。

幻影控制器的日志条目

在检测到幻影控制器时，将显示以下错误级别日志条目：

- 2017-07-31 22:15:05.844 UTC ERROR NVPStatusCheck ControllerServiceImpl:2146 – Controller <#> does not exist, might be deleted already. Skip saving its connectivity info.
- 2017-07-31 22:15:05.769 UTC ERROR NVPStatusCheck ControllerServiceImpl:2580 – the node is created by this NSX Manager <#>, but database has no record and delete might be in progress.

NSX Controller 已断开连接

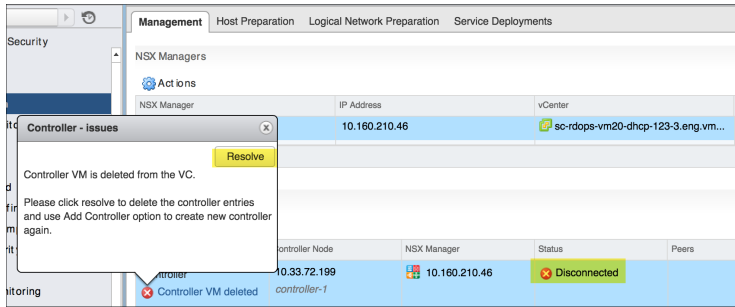
如果从 vCenter Server 中关闭 NSX Controller 虚拟机电源或从 vCenter Server 中删除控制器虚拟机，**安装 (Installation) > 管理 (Management)** 页面的**状态 (Status)**列将显示不同步 (**Out of sync**) 状态。

前提条件

已关闭控制器虚拟机电源或从 vCenter Server 中删除控制器虚拟机。

步骤

- 1 在 vSphere Web Client 中，导航到**网络和安全 (Networking & Security) > 安装 (Installation) > 管理 (Management)**。



- 2 单击**错误 (Error)**链接以查看这种不同步状态的详细原因。
- 3 单击**解决 (Resolve)**按钮以解决该问题。

结果

如果关闭了控制器虚拟机电源，管理层面将为控制器触发 **power on** 命令。

如果删除了控制器虚拟机，则会从管理层面中删除控制器条目，并且管理层面将控制器删除信息传送到中央控制层面。

后续步骤

使用**添加节点 (Add Node)**选项创建新的控制器。有关详细信息，请参阅《NSX 管理指南》。

控制层面代理 (netcpa) 问题

在 NSX for vSphere 上，控制层面 (netcpa) 用作本地代理守护进程，从而与 NSX Manager 和控制器群集进行通信。**通信通道运行状况 (Communication Channel Health)**功能是一项主动运行状况检查，以定期向 NSX Manager 报告中央控制层面到本地控制层面状态并显示在 NSX Manager UI 中。该报告还作为检测信号以检测 NSX Manager 到 ESXi 主机 netcpa 通道的运行状态。它在通信故障期间提供错误详细信息，在通道进入错误状态时生成事件，以及生成从 NSX Manager 到主机的检测信号消息。

问题

在控制层面代理和控制器之间出现连接问题。

原因

如果缺少任何连接，控制层面代理可能无法正常工作。

解决方案

- 1 使用以下命令验证在通道进入错误状态时的连接状态：

```
GET https://<NSX_Manager_IP>/api/2.0/vdn/inventory/host/{hostId}/connection/status
```

下面是返回值示例：

```
<?xml version="1.0" encoding="UTF-8"?>
<hostConnStatus>
<hostName>10.161.246.20</hostName>
<hostId>host-21</hostId>
<nsxMgrToFirewallAgentConn>UP</nsxMgrToFirewallAgentConn>
<nsxMgrToControlPlaneAgentConn>UP</nsxMgrToControlPlaneAgentConn>
<hostToControllerConn>DOWN</hostToControllerConn>
<fullSyncCount>-1</fullSyncCount>
<hostToControllerConnectionErrors>
<hostToControllerConnectionError>
<controllerIp>10.160.203.236</controllerIp>
<errorCode>1255604</errorCode>
<errorMessage>Connection Refused</errorMessage>
</hostToControllerConnectionError>
<hostToControllerConnectionError>
<controllerIp>10.160.203.237</controllerIp>
<errorCode>1255603</errorCode>
<errorMessage>SSL Handshake Failure</errorMessage>
</hostToControllerConnectionError>
</hostToControllerConnectionErrors>
</hostConnStatus>
```

支持以下错误代码：

1255602：控制器证书不完整；1255603：SSL 握手失败；1255604：已拒绝连接；1255605：保持活动状态超时；
1255606：SSL 异常；1255607：消息错误；1255620：未知错误

2 按照以下方式确定控制层面代理关闭的原因：

- a 在 ESXi 主机上运行 `/etc/init.d/netcpad status` 命令以检查主机上的控制层面代理状态。

```
[root@esx-01a:~] /etc/init.d/netcpad status
netCP agent service is running
```

- b 使用 `more /etc/vmware/netcpa/config-by-vsm.xml` 命令检查控制层面代理配置。将会列出 NSX Controller 的 IP 地址。

```
[root@esx-01a:~] more /etc/vmware/netcpa/config-by-vsm.xml
<config>
  <connectionList>
    <connection id="0000">
      <port>1234</port>
      <server>192.168.110.31</server>
      <sslEnabled>true</sslEnabled>
      <thumbprint>A5:C6:A2:B2:57:97:36:F0:7C:13:DB:64:9B:86:E6:EF:1A:7E:5C:36</thumbprint>
    </connection>
    <connection id="0001">
      <port>1234</port>
      <server>192.168.110.32</server>
      <sslEnabled>true</sslEnabled>
```



```

    <thumbprint>12:E0:25:B2:E0:35:D7:84:90:71:CF:C7:53:97:FD:96:EE:ED:7C:DD</thumbprint>
  </connection>
  <connection id="0002">
    <port>1234</port>
    <server>192.168.110.33</server>
    <sslEnabled>true</sslEnabled>
    <thumbprint>BD:DB:BA:B0:DC:61:AD:94:C6:0F:7E:F5:80:19:44:51:BA:90:2C:8D</thumbprint>
  </connection>
</connectionList>
...

```

- 3 使用以下命令验证从控制层面代理到控制器的连接。输出是每个控制器一个连接。

```

>[root@esx-01a:~] esxcli network ip connection list | grep 1234
tcp      0    0 192.168.110.51:16594    192.168.110.31:1234    ESTABLISHED    36754    newreno
netcpa-worker
tcp      0    0 192.168.110.51:46917    192.168.110.33:1234    ESTABLISHED    36754    newreno
netcpa-worker
tcp      0    0 192.168.110.51:47891    192.168.110.32:1234    ESTABLISHED    36752    newreno
netcpa-worker

```

- 4 运行以下命令，以验证从控制层面代理到控制器的连接是否显示 CLOSED 或 CLOSE_WAIT 状态：

```

esxcli network ip
    connection list |grep "1234.*netcpa*" | egrep "CLOSED|CLOSE_WAIT"

```

- 5 如果控制层面代理已关闭很长时间，则连接可能根本不存在。要验证这种情况，请运行以下命令。输出是每个控制器一个连接。

```

esxcli network ip
    connection list |grep "1234.*netcpa*" |grep ESTABLISHED

```

- 6 控制层面代理 (netcpa) 自动恢复机制：**自动控制层面代理监控过程检测处于错误状态的控制层面代理。在控制层面代理处于错误状态时，将停止响应并随后自动尝试进行恢复。

- a** 在控制层面代理停止响应时，将会生成实时核心文件。您可以按以下方式查找核心文件：

```
ls /var/core  
netcpa-worker-zdump.000
```

- b** 在 *vmkwarning.log* 文件中报告 **syslog** 错误。

```
cat /var/run/log/vmkwarning.log | grep NETCPA  
2017-08-11T06:32:17.994Z cpu1:1000044539)ALERT: Critical - NETCPA is hanged  
Taking live-dump & restarting netcpa process!
```

注 如果控制层面代理监控器由于延迟的状态检查响应而遇到临时故障，则可能会在 **VMKernel** 日志中报告类似于以下内容的警告消息。

```
Warning - NETCPA getting netcpa status failed!
```

您可以忽略该警告。

- 7** 如果未自动恢复该问题，请按以下方式重新启动控制层面代理：

- a** 以 **root** 身份通过 **SSH** 或控制台登录到 **ESXi** 主机。
- b** 运行 `/etc/init.d/netcpad restart` 命令以在 **ESXi** 主机上重新启动控制层面代理。

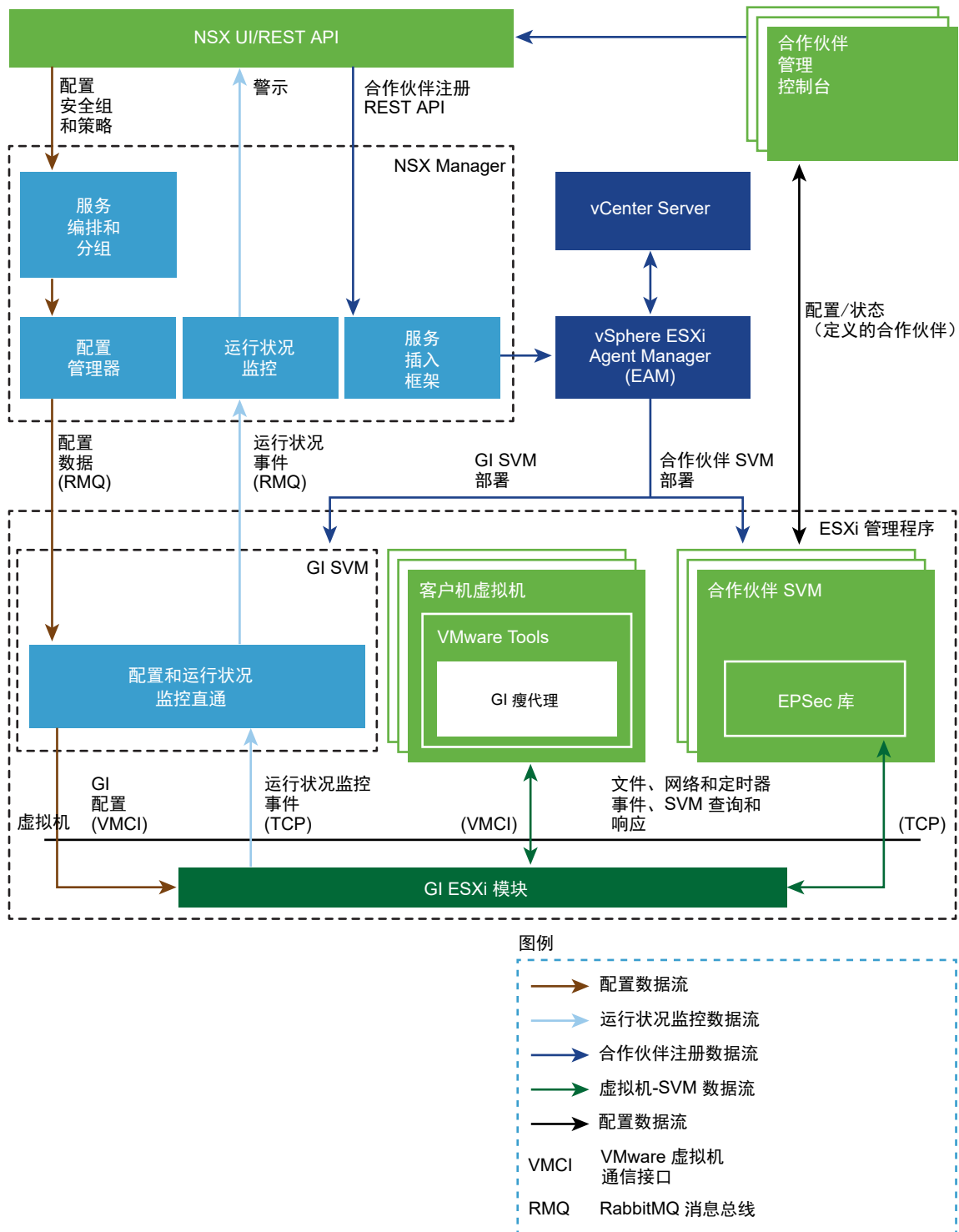
Guest Introspection 故障排除

9

本章讨论了以下主题：

- [Guest Introspection 架构](#)
- [Guest Introspection 日志](#)
- [收集 Guest Introspection 环境和工作详细信息](#)
- [Linux 或 Windows 上的瘦代理故障排除](#)
- [ESX GI 模块 \(MUX\) 故障排除](#)
- [EPSecLib 故障排除](#)

Guest Introspection 架构



Guest Introspection 日志

您可以捕获几种不同的日志以在排除 Guest Introspection 故障时使用。

ESX GI 模块 (MUX) 日志

如果 ESXi 主机上的虚拟机未使用 Guest Introspection，或者在主机上出现有关到 SVA 的通信的警报，则可能是 ESXi 主机上的 ESX GI 模块出现问题。

日志路径和示例消息

MUX 日志路径

/var/log/syslog

var/run/syslog.log

ESX GI 模块 (MUX) 消息采用以下格式: <timestamp>EPSecMUX<[ThreadID]>: <message>

例如:

```
2017-07-16T05:44:49Z EPSecMux[38340669]: [ERROR] (EPSEC) [38340669]
Attempted to recv 4 bytes from sd 49, errno = 104 (Connection reset by peer)
```

在上面的示例中

- [ERROR] 是消息类型。其他类型可能是 [DEBUG]、[INFO]
- (EPSEC) 表示消息是端点安全特定的消息

启用和查看日志文件

要查看在主机上安装的 ESX GI 模块 VIB 版本，请运行 `#esxcli software vib list | grep epsec-mux` 命令。

要启用完整日志记录，请在 ESXi 主机命令 `shell` 上执行以下步骤:

- 1 运行 `ps -c | grep Mux` 命令以查找当前运行的 ESX GI 模块进程。

例如:

```
~ # ps -c | grep Mux
192223 192223 sh /bin/sh /sbin/watchdog.sh -s vShield-Endpoint-Mux -q 100 -t 1000000 /usr/lib/
vmware/vShield-Endpoint-Mux 900 -c 910
192233 192233 vShield-Endpoint-Mux /usr/lib/vmware/vShield-Endpoint-Mux 900 -c 910
192236 192233 vShield-Endpoint-Mux /usr/lib/vmware/vShield-Endpoint-Mux 900 -c 910
```

- 2 如果该服务未运行，您可以使用以下命令重新启动该服务: `/etc/init.d/vShield-Endpoint-Mux start` 或 `/etc//init.d/vShield-Endpoint-Mux restart`。
- 3 要停止运行的 ESX GI 模块进程（包括 `watchdog.sh` 进程），请运行 `~ # kill -9 192223 192233 192236` 命令。

请注意，生成了两个 ESX GI 模块进程。

- 4 使用新的 `-d` 选项启动 ESX GI 模块。请注意，对于 `epsec-mux` 内部版本 5.1.0-01255202 和 5.1.0-01814505，无法使用 `-d` 选项：`~ # /usr/lib/vmware/vShield-Endpoint-Mux -d 900 -c 910`
- 5 查看 ESXi 主机上的 `/var/log/syslog.log` 文件中的 ESX GI 模块日志消息。请检查是否正确指定了与全局解决方案、解决方案 ID 和端口号对应的条目。

示例： 示例 `muxconfig.xml` 文件

```
<?xml version="1.0" encoding="UTF-8"?>

<EndpointConfig>

  <InstalledSolutions>

    <Solution>

      <id>100</id>

      <ipAddress>xxx.xxx.xxx.xxx</ipAddress>

      <listenOn>ip</listenOn>

      <port>48655</port>

      <uuid>42383371-3630-47b0-8796-f1d9c52ab1d0</uuid>

      <vmxPath>/vmfs/volumes/7adf9e00-609186d9/EndpointService (216)/EndpointService (216).vmx</vmxPath>

    </Solution>

    <Solution>

      <id>102</id>

      <ipAddress>xxx.xxx.xxx.xxx</ipAddress>

      <listenOn>ip</listenOn>

      <port>48651</port>

      <uuid>423839c4-c7d6-e92e-b552-79870da05291</uuid>

      <vmxPath>/vmfs/volumes/7adf9e00-609186d9/apoon/EndpointSVM-alpha-01/EndpointSVM-alpha-01.vmx</vmxPath>

    </Solution>

    <Solution>

      <id>6341068275337723904</id>

      <ipAddress>xxx.xxx.xxx.xxx</ipAddress>
```

```
<listenOn>ip</listenOn>

<port>48655</port>

<uuid>42388025-314f-829f-2770-a143b9cbd1ee</uuid>

<vmxPath>/vmfs/volumes/7adf9e00-609186d9/DlpService (1)/DlpService (1).vmx</vmxPath>

</Solution>

</InstalledSolutions>

<DefaultSolutions/>

<GlobalSolutions>

  <solution>

    <id>100</id>

    <tag></tag>

    <order>0</order>

  </solution>

  <solution>

    <id>102</id>

    <tag></tag>

    <order>10000</order>

  </solution>

  <solution>

    <id>6341068275337723904</id>

    <tag></tag>

    <order>10001</order>

  </solution>

</GlobalSolutions>

</EndpointConfig>
```

GI 瘦代理日志

瘦代理安装在虚拟机客户机操作系统上，可以检测用户登录详细信息。

日志路径和示例消息

瘦代理由 GI 驱动程序组成 - vsepflt.sys、vnetflt.sys、vnetwfp.sys（Windows 10 和更高版本）。

瘦代理日志位于 ESXi 主机上，它是 vCenter 日志包的一部分。日志路径是 /vmfs/volumes/<datastore>/<vmname>/vmware.log。例如：/vmfs/volumes/5978d759-56c31014-53b6-1866abaace386/Windows10-(64-bit)/vmware.log

瘦代理消息采用以下格式：<timestamp> <VM Name><Process Name><[PID]>: <message>。

在下面的日志示例中，Guest: vnet or Guest:vsep 指示与相应的 GI 驱动程序相关的日志消息，后跟调试消息。

例如：

```
2017-10-17T14:25:19.877Z| vcpu-0| I125: Guest: vnet: AUDIT: DriverEntry :
vnetFilter build-4325502 loaded
2017-10-17T14:25:20.282Z| vcpu-0| I125: Guest: vsep:
AUDIT: VFileSocketMgrConnectHelper : Mux is connected
2017-10-17T14:25:20.375Z| vcpu-0| I125:
Guest: vsep: AUDIT: DriverEntry : vfileFilter build-4286645 loaded

2017-10-17T18:22:35.924Z| vcpu-0| I125: Guest: vsep: AUDIT:
VFileSocketMgrConnectHelper : Mux is connected
2017-10-17T18:24:05.258Z| vcpu-0| I125: Guest: vsep: AUDIT:
VFileFltPostOpCreate : File (\Windows\System32\Tasks\Microsoft\Windows\
SoftwareProtectionPlatform\SvcRestartTask) in a transaction, ignore
```

示例：启用 vShield Guest Introspection 瘦代理驱动程序日志记录

由于调试设置可能会导致 vmware.log 文件填满而对其进行限制，因此，我们建议您在收集所需的所有信息后立即禁用调试模式。

该过程要求您修改 Windows 注册表。在修改注册表之前，请确保创建注册表的备份。有关备份和还原注册表的详细信息，请参见 Microsoft 知识库文章 [136393](#)。

要为瘦代理驱动程序启用调试日志记录，请执行以下操作：

- 1 单击**开始 > 运行 (Start > Run)**。输入 regedit，然后单击**确定 (OK)**。将打开“注册表编辑器”窗口。有关详细信息，请参见 Microsoft 知识库文章 [256986](#)。
- 2 使用注册表编辑器创建以下项：HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\services\vsepflt\parameters。
- 3 在新创建的参数项下面，创建以下 DWORD。在输入这些值时，请确保选择十六进制值：

```
Name: log_dest
Type: DWORD
Value: 0x2
```



```
Name: log_level
Type: DWORD
Value: 0x10
```

log_level 参数项的其他值：

```
Audit 0x1
Error 0x2
Warn 0x4
Info 0x8
Debug 0x10
```

- 4 以管理员身份打开命令提示符。运行以下命令以卸载并重新加载 vShield Endpoint 文件系统小型驱动程序：

- fltmc unload vsepflt
- fltmc load vsepflt

您可以在位于虚拟机的 vmware.log 文件中找到这些日志条目。

启用 vShield GI 网络自检驱动程序日志记录

由于调试设置可能会导致 vmware.log 文件填满而对其进行限制，因此，我们建议您在收集所需的所有信息后立即禁用调试模式。

该过程要求您修改 Windows 注册表。在修改注册表之前，请确保创建注册表的备份。有关备份和还原注册表的详细信息，请参见 Microsoft 知识库文章 [136393](#)。

- 1 单击 **开始 > 运行 (Start > Run)**。输入 regedit，然后单击 **确定 (OK)**。将打开“注册表编辑器”窗口。有关详细信息，请参见 Microsoft 知识库文章 [256986](#)。
- 2 编辑注册表：

```
Windows Registry Editor Version 5.0
[HKEY_LOCAL_MACHINE\SYSTEM\CurrentControlSet\Services\vnetflt\Parameters]
"log_level" = DWORD: 0x0000001F
"log_dest" = DWORD: 0x00000001
```

- 3 重新引导虚拟机。

vsepflt.sys 和 vnetflt.sys 日志文件位置

在使用 log_dest 注册表设置 DWORD:0x00000001 时，Endpoint 瘦代理驱动程序将日志记录到调试程序中。运行调试程序（SysInternals 中的 DbgView 或 windbg）以捕获调试输出。

或者，您也可以将 log_dest 注册表设置设为 DWORD:0x00000002，在这种情况下，驱动程序日志将输出到 vmware.log 文件，该文件位于 ESXi 主机上的相应虚拟机文件夹中。

启用 UMC 日志记录

Guest Introspection 用户模式组件 (UMC) 在受保护的虚拟机上的 VMware Tools 服务中运行。

- 1 在 Windows XP 和 Windows Server 2003 上，创建一个 `tools.config` 文件（如果在以下路径中不存在）：`C:\Documents and Settings\All Users\Application Data\VMware\VMware Tools\tools.conf`。
- 2 在 Windows Vista、Windows 7 和 Windows Server 2008 上，创建一个 `tools.config` 文件（如果在以下路径中不存在）：`C:\ProgramData\VMware\VMware Tools\tools.conf`。
- 3 在 `tools.conf` 文件中添加以下行以启用 UMC 组件日志记录。

```
[logging]
log = true
vsep.level = debug
vsep.handler = vmx
```

在使用 `vsep.handler = vmx` 设置时，UMC 组件将日志记录到 `vmware.log` 文件中，该文件位于 ESXi 主机上的相应虚拟机文件夹中。

在使用以下设置时，将在指定的日志文件中输出 UMC 组件日志。

```
vsep.handler = file
vsep.data = c:/path/to/vsep.log
```

GI EPSecLib 和 SVM 日志

EPSecLib 从 ESXi 主机 ESX GI 模块 (MUX) 中接收事件。

日志路径和示例消息

EPSecLib 日志路径

`/var/log/syslog`

`var/run/syslog`

EPSecLib 消息采用以下格式：`<timestamp> <VM Name><Process Name><[PID]>: <message>`

在下面的示例中，`[ERROR]` 是消息类型，`(EPSEC)` 表示 Guest Introspection 特定的消息。

例如：

```
Oct 17 14:26:00 endpoint-virtual-machine EPSecTester[7203]: [NOTICE] (EPSEC)
[7203] Initializing EPSec library build: build-00000

Oct 17 14:37:41 endpoint-virtual-machine EPSecSample: [ERROR] (EPSEC) [7533] Event
terminated reading file. Ex: VFileGuestEventTerminated@tid=7533: Event id: 3554.
```

收集日志

要为 EPSec 库（GI SVM 中的组件）启用调试日志记录，请执行以下操作：

- 1 从 NSX Manager 中获取控制台密码以登录到 GI SVM。
- 2 创建 `/etc/epseclib.conf` 文件并添加：


```
ENABLE_DEBUG=TRUE
ENABLE_SUPPORT=TRUE
```
- 3 运行 `chmod 644 /etc/epseclib.conf` 命令以更改权限。
- 4 运行 `/usr/local/sbin/rcusvm restart` 命令以重新启动 GI-SVM 进程。

这会为 GI SVM 上的 EPSecLib 启用调试日志记录，可以在 `/var/log/messages` 中找到适用于 NSX for vSphere 6.2.x 和 6.3.x 的调试日志。由于调试设置可能会导致 `vmware.log` 文件填满而对其进行限制，因此，我们建议您在收集所需的所有信息后立即禁用调试模式。

GI SVM 日志

在捕获日志之前，请确定主机 ID 或主机 MOID：

- 在 NSX Manager 中运行 `show cluster all` 和 `show cluster <cluster ID>` 命令。

例如：

```
nsxmgr-01a> show cluster all
```

No.	Cluster Name	Cluster Id	Datacenter Name	Firewall Status
1	RegionA01-COMP01	domain-c26	RegionA01	Enabled
2	RegionA01-MGMT01	domain-c71	RegionA01	Enabled

```
nsxmgr-01a> show cluster domain-c26
```

```
Datacenter: RegionA01
Cluster: RegionA01-COMP01
```

No.	Host Name	Host Id	Installation Status
1	esx-01a.corp.local	host-29	Ready
2	esx-02a.corp.local	host-31	Ready

- 1 要确定当前日志记录状态，请运行以下命令：

```
GET https://nsxmanager/api/1.0/usvmlogging/host-##/com.vmware.vshield.usvm
```

```
GET https://nsxmanager/api/1.0/usvmlogging/host-##/root
```

- 2 要更改当前日志记录状态，请运行以下命令：

```
POST https://nsxmanager/api/1.0/usvmlogging/host-##/changelevel
```

```
## Example to change root logger ##

<?xml version="1.0" encoding="UTF-8" ?>
<logginglevel>
<loggerName>root</loggerName>
```

```
<level>DEBUG</level>
</logginglevel>

## Example to change com.vmware.vshield.usvm ##

<?xml version="1.0" encoding="UTF-8" ?>
<logginglevel>
<loggerName>com.vmware.vshield.usvm</loggerName>
<level>DEBUG</level>
</logginglevel>
```

- 3 要生成日志，请运行以下命令：

```
GET https://NSXMGR_IP/api/1.0/hosts/host.###/techsupportlogs
```

选择 Send 和 Download。

请注意，该命令生成 GI SVM 日志，并将该文件另存为 `techsupportlogs.log.gz` 文件。由于调试设置可能会导致 `vmware.log` 文件填满而对其进行限制，因此，我们建议您在收集所需的所有信息后立即禁用调试模式。

收集 Guest Introspection 环境和工作详细信息

在检查组件兼容性时，收集环境详细信息是非常有用的。

- 1 确定是否在客户环境中使用 NSX Guest Introspection。如果未使用，请移除虚拟机的 Guest Introspection 服务并确认已解决该问题。
- 2 收集环境详细信息：
 - a ESXi 内部版本 - 在 ESXi 主机上运行 `uname -a` 命令，或者在 vSphere Web Client 中单击一个主机，然后在右侧窗格顶部查找内部版本号。
 - b Linux 产品版本和内部版本号
 - c `/usr/sbin/vsep -v` 将提供生产版本

```
Build number
-----
Ubuntu
dpkg -l | grep vmware-nsx-gi-file
SLES12 and RHEL7
rpm -qa | grep vmware-nsx-gi-file
```

- 3 VMware NSX® for vSphere® 版本和以下信息：
 - 合作伙伴解决方案名称和版本号
 - 合作伙伴解决方案使用的 EPSec 库版本号：登录到 GI SVM 并运行 `#strings path to EPSec library/libEPSec.so | grep BUILD`
 - 虚拟机中的客户机操作系统
 - 任何其他第三方应用程序或文件系统驱动程序

- 4 ESX GI 模块 (MUX) 版本 - 运行 `esxcli software vib list | grep epsec-mux` 命令。
- 5 收集工作负载详细信息，如服务器类型。
- 6 收集 ESXi 主机日志。有关详细信息，请参见[收集 VMware ESX/ESXi 的诊断信息 \(653\)](#)。
- 7 从合作伙伴解决方案中收集服务虚拟机 (GI SVM) 日志。与您的合作伙伴联系以了解 GI SVM 日志收集的更多详细信息。
- 8 在出现问题时，收集挂起状态文件，请参见[挂起 ESX/ESXi 上的虚拟机 \(2005831\)](#) 以收集诊断信息。
- 9 在收集数据后，请比较 vSphere 组件的兼容性。有关详细信息，请参见 [VMware 产品互操作性列表](#)。

Linux 或 Windows 上的瘦代理故障排除

可以使用 VMware Tools™ 在每个客户机虚拟机上安装 Guest Introspection 瘦代理。

Linux 上的瘦代理故障排除

如果虚拟机的读取和写入操作以及文件解压缩或保存速度变慢，则瘦代理可能出现問題。

- 1 检查涉及的所有组件的兼容性。兼容性是 Endpoint 的主要问题之一。您需要了解 ESXi、vCenter Server、NSX Manager 以及选择的任何安全解决方案（Trend Micro、McAfee、Kaspersky、Symantec 等）的内部版本号。在收集该数据后，比较 vSphere 组件的兼容性。有关详细信息，请参见 [VMware 产品互操作性列表](#)。
- 2 确保在系统上安装了文件自检。
- 3 使用 `service vsep status` 命令验证瘦代理是否正在运行。在执行该命令后，将会看到 vsep 服务处于运行状态。
- 4 如果您认为系统性能问题是由瘦代理造成的，请运行 `service vsep stop` 命令以停止该服务。
- 5 接下来，执行测试以获取基准。然后，您可以运行 `service vsep start` 命令以启动 vsep 服务并再次执行测试。
- 6 为 Linux 瘦代理启用调试：
 - a 打开 `/etc/vsep/vsep.conf` 文件
 - b 将所有日志的 `DEBUG_LEVEL=4` 更改为 `DEBUG_LEVEL=7`
 - c 对于中等大小的日志，可以将其设置为 `DEBUG_LEVEL=6`
 - d 默认日志目标 (`DEBUG_DEST=2`) 是 `vmware.log`（在主机上）；要将其更改为客户机（即 `/var/log/message` 或 `/var/log/syslog`），请设置 `DEBUG_DEST=1`

注 启用完整日志记录可能会产生较高的日志活动以填充 `vmware.log` 文件，这可能会导致该文件变得非常大。请尽快禁用完整日志记录。

Windows 上的瘦代理故障排除

- 1 检查涉及的所有组件的兼容性。您需要了解 ESXi、vCenter Server、NSX Manager 以及选择的任何安全解决方案（Trend Micro、McAfee、Kaspersky、Symantec 等）的内部版本号。在收集所有这些数据后，您可以比较 vSphere 组件的兼容性。有关详细信息，请参见 [VMware 产品互操作性列表](#)。
- 2 确保 VMware Tools™ 是最新版本。如果您发现仅特定虚拟机受到影响，请参见在 [vSphere 中安装和升级 VMware Tools \(2004754\)](#)。
- 3 运行 Powershell 命令 `fltmc` 以验证是否加载了瘦代理。

在执行该命令后，将会在驱动程序列表中看到名称 `vsepfilt`。如果未加载该驱动程序，您应该可以使用 `fltmc load vsepfilt` 命令加载该驱动程序。

- 4 如果系统性能问题是由瘦代理造成的，请使用以下命令卸载该驱动程序：`fltmc unload vsepfilt`。接下来，执行测试以获取基准。然后，您可以运行以下命令以加载该驱动程序并再次执行测试：

```
fltmc load vsepfilt.
```

如果确认瘦代理存在性能问题，请参见在 [NSX 和 vCloud Networking and Security 中升级 VMware tools 后虚拟机速度缓慢 \(2144236\)](#)。

- 5 如果未使用网络自检，请移除或禁用该驱动程序。
也可以通过修改 VMware Tools 安装程序移除网络自检：

- a 挂载 VMware Tools 安装程序。
- b 导航到 **控制面板 > 程序和功能 (Control Panel > Programs and Features)**。
- c 右键单击 **VMware Tools > 修改 (VMware Tools > Modify)**。
- d 选择 **完整安装 (Complete install)**。
- e 找到 **NSX 文件自检**。网络自检应具有一个专用子文件夹。
- f 禁用 **网络自检 (Network Introspection)**。
- g 重新引导虚拟机以完成驱动程序卸载。

- 6 为瘦代理启用调试日志记录。有关详细信息，请参见 [Guest Introspection 日志](#)。所有调试信息配置为记录到该虚拟机的 `vmware.log` 文件中。
- 7 查看 `procmon` 日志以查看瘦代理的文件扫描。有关详细信息，请参见 [解决使用防病毒软件的 vShield Endpoint 的性能问题 \(2094239\)](#)。

收集环境和工作负载详细信息

- 1 确定是否在客户环境中使用 NSX Guest Introspection。如果未使用，请移除虚拟机的 Guest Introspection 服务并确认已解决该问题。只有在需要使用 Guest Introspection 时，才需要对 Guest Introspection 问题进行故障排除。

2 收集环境详细信息：

- a ESXi 内部版本 - 在 ESXi 主机上运行 `uname -a` 命令，或者在 vSphere Web Client 中单击一个主机，然后在右侧窗格顶部查找内部版本号。
- b Linux 产品版本和内部版本号
- c `/usr/sbin/vsep -v` 将提供生产版本

```
Build number
-----
Ubuntu
dpkg -l | grep vmware-nsx-gi-file
SLES12 and RHEL7
rpm -qa | grep vmware-nsx-gi-file
```

3 VMware NSX® for vSphere® 版本和以下信息：

- 合作伙伴解决方案名称和版本号
- 合作伙伴解决方案使用的 EPSec 库版本号：登录到 SVM 并运行 `#strings path to EPSec library/libEPSec.so | grep BUILD`
- 虚拟机中的客户机操作系统
- 任何其他第三方应用程序或文件系统驱动程序

4 ESX GI 模块 (MUX) 版本 - 运行 `esxcli software vib list | grep epsec-mux` 命令。

5 收集工作负载详细信息，如服务器类型。

6 收集 ESXi 主机日志。有关详细信息，请参见[收集 VMware ESX/ESXi 的诊断信息 \(653\)](#)。

7 从合作伙伴解决方案中收集服务虚拟机 (SVM) 日志。与您的合作伙伴联系以了解 SVM 日志收集的更多详细信息。

8 在出现问题时，收集挂起状态文件，请参见[挂起 ESX/ESXi 上的虚拟机 \(2005831\)](#) 以收集诊断信息。

瘦代理崩溃故障排除

如果瘦代理崩溃，则会在 `/directory` 中生成核心文件。从 `location/directory` 中收集核心转储文件（核心）。可以使用 `file` 命令检查核心是不是由 `vsep` 生成的。例如：

```
# file core
core: ELF 64-bit LSB core file x86-64, version 1 (SYSV), SVR4-style, from '/usr/sbin/vsep'
```

虚拟机挂起或冻结

收集处于挂起状态的虚拟机的 VMware vmss 文件，请参见[在 ESX/ESXi 上挂起虚拟机以收集诊断信息 \(2005831\)](#)，或者使虚拟机崩溃并收集完整内存转储文件。VMware 提供了一个实用程序以将 ESXi vmss 文件转换为核心转储文件。有关详细信息，请参见 [Vmss2core fling](#)。

ESX GI 模块 (MUX) 故障排除

ESX GI 模块 (MUX)

如果 ESXi 主机上的所有虚拟机未正常使用 Guest Introspection，或者在特定主机上出现有关到 GI SVA 的通信的警报，则可能是 ESXi 主机上的 ESX GI 模块出现问题。

- 1 运行 `# /etc/init.d/vShield-Endpoint-Mux status` 命令，以检查是否正在 ESXi 主机上运行该服务：

例如：

```
# /etc/init.d/vShield-Endpoint-Mux status
vShield-Endpoint-Mux is running
```

- 2 如果发现未运行该服务，请重新启动该服务，或者使用以下命令启动该服务：

```
/etc/init.d/vShield-Endpoint-Mux start
```

或

```
/etc/init.d/vShield-Endpoint-Mux restart
```

请注意，可以在生产时间安全地重新启动该服务，因为它不会造成任何重大影响并在几秒钟内重新启动。

- 3 要更好地了解 ESX GI 模块执行的操作或检查通信状态，您可以检查 ESXi 主机上的日志。ESX GI 模块日志将写入到主机 `/var/log/syslog` 文件中。它还包含在 ESXi 主机支持日志中。

有关详细信息，请参见[使用 vSphere Web Client 收集 ESX/ESXi 主机和 vCenter Server 的诊断信息 \(2032892\)](#)。

- 4 ESX GI 模块的默认日志记录选项为“信息”，并且可以将其升级到“调试”以收集更多信息：

有关详细信息，请参见[Guest Introspection 日志](#)。

- 5 重新安装 ESX GI 模块也可以修复很多问题，尤其是安装了不正确的版本，或者将 ESXi 主机添加到以前安装了 Endpoint 的环境中。需要将其移除并重新安装。

要移除 VIB，请运行以下命令：`esxcli software vib remove -n epsec-mux`

- 6 如果您遇到 VIB 安装问题，请检查 ESXi 主机上的 `/var/log/esxupdate.log` 文件。该日志显示有关为什么未成功安装该驱动程序的最明确的信息。这是 ESX GI 模块安装的一个常见问题。有关详细信息，请参见在[VMware NSX for vSphere 6.x](#) 中，在 ESXi 主机上安装 NSX Guest Introspection 服务（ESX GI 模块 VIB）失败 (2135278)。

- 7 要检查损坏的 ESXi 映像，请查找类似于以下内容的消息：

```
esxupdate: esxupdate: ERROR: Installation Error:
(None, 'No image profile is found on the host or image profile is empty.
An image profile is required to install or remove VIBs. To install an image profile,
use the esxcli image profile install command.')
```


8 要确认映像已损坏，请在 ESXi 主机上运行 `cd /vmfs/volumes` 命令。

a 运行以下命令以搜索 `imgdb.tgz` 文件：`find * | grep imgdb.tgz`。

该命令通常会找到两个匹配项。例如：

`0ca01e7f-cc1ea1af-bda0-1fe646c5ceea/imgdb.tgz` 或 `edbf587b-da2add08-3185-3113649d5262/imgdb.tgz`

b 为每个匹配项运行以下命令：`ls -l match_result`

例如：

```
> ls -l 0ca01e7f-cc1ea1af-bda0-1fe646c5ceea/imgdb.tgz -rwx-----
1 root root 26393 Jul 20 19:28 0ca01e7f-cc1ea1af-bda0-1fe646c5ceea/imgdb.tgz
> ls -l edbf587b-da2add08-3185-3113649d5262/imgdb.tgz -rwx-----
1 root root 93 Jul 19 17:32 edbf587b-da2add08-3185-3113649d5262/imgdb.tgz
```

`imgdb.tgz` 文件的默认大小远大于另一个文件，或者，如果一个文件仅具有几个字节，则表示该文件已损坏。要解决该问题，唯一支持的方法是为该特定 ESXi 主机重新安装 ESXi。

EPSecLib 故障排除

NSX Manager 处理该虚拟机的部署问题。

EPSecLib

过去（使用 vShield），第三方 SVA 解决方案处理部署问题。现在，该解决方案连接到 NSX Manager。NSX Manager 处理该 SVA 的部署问题。如果环境中的 SVA 存在警报，请通过 NSX Manager 重新部署这些 SVA。

- 所有配置将会丢失，因为它们均存储在 NSX Manager 中。
- 最好重新部署 SVA 虚拟机，而不是重新引导这些虚拟机。
- NSX 依靠 EAM 部署和监控主机上的 VIB 和 SVM，如 SVA。
- EAM 是用于确定安装状态的事实来源。
- NSX 用户界面 (UI) 中的安装状态只能确定是否安装了 VIB 或是否打开了 SVM 电源。
- NSX UI 中的服务状态指示虚拟机中的功能是否正常工作。

SVA 部署以及 NSX 和 vCenter Server 进程之间的关系

- 1 在选择为 Endpoint 准备群集时，将在 EAM 上创建一个代理机构以部署 SVA。
- 2 然后，EAM 使用它创建的代理机构信息将 `ovf` 部署到 ESXi 主机中。
- 3 NSX Manager 验证 EAM 是否部署了 `ovf`。
- 4 NSX Manager 验证 EAM 是否打开了虚拟机电源。
- 5 NSX Manager 通知合作伙伴 SVA 解决方案管理器已打开虚拟机电源并注册了虚拟机。
- 6 EAM 向 NSX 发送事件以指示安装已完成。

- 7 合作伙伴 SVA 解决方案管理器向 NSX 发送事件，以指示 SVA 虚拟机中的服务已启动并正在运行。
- 8 如果 SVA 出现问题，您可以在两个地方查看日志。您可以检查 EAM 日志，因为 EAM 处理这些虚拟机的部署问题。有关详细信息，请参见[收集 VMware vCenter Server 4.x、5.x 和 6.0 的诊断信息 \(1011641\)](#)。或者，查看 SVA 日志。

有关详细信息，请参见 [Guest Introspection 日志](#)。

- 9 如果出现 SVA 部署问题，则可能是 EAM 以及与 NSX Manager 的通信出现问题。您可以检查 EAM 日志，最简单的办法是重新启动 EAM 服务。有关详细信息，请参见[主机准备](#)。
- 10 如果所有上述问题一切正常，但您希望测试 Endpoint 功能，您可以使用 Eicar 测试文件测试该问题：
 - 使用任何标签创建一个文本文件。例如：eicar.test。
 - 该文件的内容应仅为以下字符串：
`X50!P%@AP[4\PZX54(P^)7CC)7}$EICAR-STANDARD-ANTIVIRUS-TEST-FILE! $H+H*`
 - 保存该文件。在保存后，将会看到已删除该文件。这确认了 Endpoint 解决方案正常工作。