NSX-T Data Center 故障排除指南

修改日期：2018 年 9 月 19 日

VMware NSX-T Data Center 2.3

vmware®

最新的技术文档可以从 VMware 网站下载：

https://docs.vmware.com/cn/

您如果对本文档有任何意见或建议，请把反馈信息提交至：

docfeedback@vmware.com

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

北京办公室
北京市海淀区科学院南路 2 号
融科资讯中心 C 座南 8 层
www.vmware.com/cn

上海办公室
上海市浦东新区浦东南路 999 号
新梅联合广场 23 楼
www.vmware.com/cn

广州办公室
广州市天河北路 233 号
中信广场 7401 室
www.vmware.com/cn

目录

NSX-T Data Center 故障排除指南 5

1 日志和服务 6
- 日志消息 6
- 对 Syslog 问题进行故障排除 10
- 检查服务 11
- 收集支持包 12

2 对第 2 层连接问题进行故障排除 14
- 检查 NSX Manager 和 NSX Controller 群集状态 14
- 检查逻辑端口 15
- 检查传输节点状态 15
- 检查逻辑交换机状态 16
- 检查逻辑交换机的 CCP 17
- 检查本地控制层面状态 17
- 对配置会话问题进行故障排除 18
- 对 L2 会话问题进行故障排除 19
- 对覆盖网络逻辑交换机的数据层面问题进行故障排除 19
- 对 VLAN 逻辑交换机的数据层面问题进行故障排除 21
- 对覆盖网络逻辑交换机的 ARP 问题进行故障排除 21
- 对 VLAN 逻辑交换机或解析 ARP 时的数据包丢失问题进行故障排除 22

3 安装故障排除 24

4 路由故障排除 28

5 防火墙故障排除 30
- 确定 ESXi 主机上应用的防火墙规则 30
- 确定 KVM 主机上应用的防火墙规则 33
- 防火墙数据包日志 34

6 其他故障排除场景 36
- 无法添加或删除传输节点 36
- 传输节点大约需要 5 分钟才能连接到另一个控制器 37
- NSX Manager 虚拟机降级 37
- NSX 代理与 NSX Manager 通信时超时 38
- 无法添加 ESXi 主机 40
- NSX Controller 状态不正确 40

KVM 虚拟机上的管理 IP 在启用 IPFIX 的情况下无法访问 41

NSX-T Data Center 故障排除指南

*《NSX-T Data Center 故障排除指南》*提供了有关如何在 NSX-T Data Center 环境中对可能出现的问题进行故障排除的信息。

目标读者

本指南适用于 NSX-T Data Center 的系统管理员。假设您熟悉虚拟化、网络和数据中心操作。

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。

日志和服务

1

日志在很多故障排除场景中非常有用。检查服务的状态也十分重要。

本章讨论了以下主题：

- 日志消息
- 对 Syslog 问题进行故障排除
- 检查服务
- 收集支持包

日志消息

来自所有 NSX-T Data Center 组件（包括 ESXi 主机上运行的组件）的日志消息均符合 RFC 5424 中指定的 syslog 格式。来自 KVM 主机的日志消息采用 RFC 3164 格式。日志文件位于 `/var/log` 目录中。

在 NSX-T Data Center 设备上，可以运行以下 NSX-T Data Center CLI 命令以查看日志：

```
get log-file <auth.log | http.log | kern.log | manager.log | node-mgmt.log | syslog> [follow]
```

在管理程序上，可以使用 `tac`、`tail`、`grep` 和 `more` 等 Linux 命令查看日志。也可以在 NSX-T Data Center 设备上使用这些命令。

有关 RFC 5424 的详细信息，请参见 https://tools.ietf.org/html/rfc5424。有关 RFC 3164 的详细信息，请参见 https://tools.ietf.org/html/rfc3164。

RFC 5424 为日志消息定义以下格式：

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

示例日志消息：

```
<187>1 2016-03-15T22:53:00.114Z nsx-manager NSX - SYSTEM [nsx@6876 comp="nsx-manager" errorCode="MP4039"
subcomp="manager"] Connection verification failed for broker '10.160.108.196'. Marking broker unhealthy.
```

每条消息都具有组件 (comp) 和子组件 (subcomp) 信息，可帮助标识消息的来源。

NSX-T Data Center 生成常规日志（程序模块 `local6`，它具有数值 22）和审核日志（程序模块 `local7`，它具有数值 23）。所有 API 调用将触发一个审核日志。

与 API 调用关联的审核日志具有以下信息：

- 实体 ID 参数 `entId`，用于标识 API 的对象。
- 请求 ID 参数 `req-id`，用于标识特定的 API 调用。
- 外部请求 ID 参数 `ereqId`（如果 API 调用包含标头 `X-NSX-EREQID:<string>`）。
- 外部用户参数 `euser`（如果 API 调用包含标头 `X-NSX-EUSER:<string>`）。

RFC 5424 定义以下严重性级别：

严重性级别	说明
0	紧急：系统无法使用
1	警报：必须立即采取措施
2	严重：严重情况
3	错误：错误情况
4	警告：警告情况
5	通知：正常但重大情况
6	信息：信息性消息
7	调试：调试级别消息

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

`MSGID` 字段标识消息的类型。有关消息 ID 列表，请参见日志消息 ID。

配置远程日志记录

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

NSX Manager、NSX Controller、NSX Edge 和管理程序支持远程日志记录。必须分别在每个节点上配置远程日志记录。

在 KVM 主机上，NSX-T Data Center 安装软件包会通过将配置文件置于 `/etc/rsyslog.d` 目录中来自动配置 rsyslog 守护进程。

前提条件

- 配置日志记录服务器以接收日志。

步骤

1 要在 NSX-T Data Center 设备上配置远程日志记录，请执行以下操作：

a 运行以下命令以配置日志服务器和要发送到日志服务器的消息类型。可以将多个设备或消息 ID 指定为逗号分隔列表（不含空格）。

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

有关此命令的详细信息，请参见《*NSX-T CLI 参考*》。您可以多次运行该命令以添加多个日志记录服务器配置。例如：

```
nsx> set logging-server 192.168.110.60 proto udp level info facility syslog messageid
SYSTEM,FABRIC
nsx> set logging-server 192.168.110.60 proto udp level info facility auth,user
```

b 可以使用 `get logging-server` 命令查看日志记录配置。例如，

```
nsx> get logging-servers
192.168.110.60 proto udp level info facility syslog messageid SYSTEM,FABRIC
192.168.110.60 proto udp level info facility auth,user
```

2 要在 ESXi 主机上配置远程日志记录，请执行以下操作：

a 运行以下命令以配置 syslog 并发送测试消息：

```
esxcli network firewall ruleset set -r syslog -e true
esxcli system syslog config set --loghost=udp://<log server IP>:<port>
esxcli system syslog reload
esxcli system syslog mark -s "This is a test message"
```

b 可以运行以下命令以显示配置：

```
esxcli system syslog config get
```

3 要在 KVM 主机上配置远程日志记录，请执行以下操作：

a 针对您的环境编辑文件 `/etc/rsyslog.d/10-vmware-remote-logging.conf`。

b 将以下行添加到该文件中：

```
*.* @<ip>:514;RFC5424fmt
```

c 运行以下命令：

```
service rsyslog restart
```

日志消息 ID

在日志消息中，消息 ID 字段标识消息的类型。您可以在 `set logging-server` 命令中使用 `messageid` 参数筛选将哪些日志消息发送到日志记录服务器。

表 1-1 日志消息 ID

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

表 1-1 日志消息 ID（续）

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

对 Syslog 问题进行故障排除

如果远程日志服务器不接收日志，请执行以下步骤。

- 验证远程日志服务器的 IP 地址。
- 验证 `level` 参数是否正确配置。
- 验证 `facility` 参数是否正确配置。
- 如果协议为 TLS，请将协议设置为 UDP，以查看是否存在证书不匹配问题。
- 如果协议为 TLS，请验证端口 6514 是否在两端都已打开。
- 移除消息 ID 筛选器，并查看服务器是否接收日志。
- 使用命令 `restart service rsyslogd` 重新启动 rsyslog 服务。

rsyslog 配置文件 (`/etc/rsyslog.conf`) 示例：

```
### rsyslog config file. Customized by VMware.
### Do not edit this file by hand. Use the API to make changes.
$PreserveFQDN on
$ModLoad imklog
$ModLoad immark
module(load="imuxsock" sysSock.useSpecialParser="off")
$RepeatedMsgReduction on
$FileOwner syslog
$FileGroup adm
$FileCreateMode 0640
$DirCreateMode 0755
$Umask 0022
$ActionFileDefaultTemplate RSYSLOG_SyslogProtocol23Format
$IncludeConfig /etc/rsyslog.d/*.conf
```

```
$template RFC5424fmt,"<%PRI%>1 %TIMESTAMP:::date-rfc3339% %HOSTNAME% %APP-NAME% %PROCID% %MSGID%
%STRUCTURED-DATA% %msg%\n"
$WorkDirectory /var/spool/rsyslog
$ModLoad imudp
$UDPServerAddress 127.0.0.1
$UDPServerRun 514
$PrivDropToUser syslog
$ActionQueueType LinkedList  # nsx exporter: e7347687-8be7-4519-a8e1-73c5192c9b43
*.info @1.2.3.4:514;RFC5424fmt  # nsx exporter: e7347687-8be7-4519-a8e1-73c5192c9b43
```

检查服务

停止运行或无法启动的服务可能会导致问题。请务必确保所有服务均正常运行。

要检查 NSX Manager 服务的状态，请执行以下操作：

```
nsxmgr> get services
Service name:       cm-inventory
Service state:      stopped

Service name:       http
Service state:      stopped
Session timeout:    1800
Connection timeout: 30
Redirect host:      (not configured)

Service name:       install-upgrade
Service state:      stopped
Enabled:            True

Service name:       liagent
Service state:      stopped

Service name:       manager
Service state:      stopped
Logging level:      info

Service name:       mgmt-plane-bus
Service state:      running

Service name:       node-mgmt
Service state:      running

Service name:       nsx-message-bus
Service state:      running

Service name:       nsx-upgrade-agent
Service state:      running

Service name:       ntp
Service state:      running

Service name:       search
```

```
Service state:      stopped

Service name:       snmp
Service state:      stopped

Start on boot:      False
Service name:       ssh

Service state:      running
Start on boot:      True

Service name:       syslog
Service state:      running
```

在上述示例中，http 服务已停止。您可以使用以下命令启动 http 服务：

```
nsxmgr> start service http
```

SSH 服务

如果部署设备时未启用 SSH 服务，可以管理员身份登录到设备并使用以下命令启用该服务：

```
start service ssh
```

您可以使用以下命令将 SSH 配置为在主机启动时启动：

```
set service ssh start-on-boot
```

要启用 SSH root 登录，可以 root 用户身份登录到设备，编辑文件 /etc/ssh/sshd_config 并替换行

```
PermitRootLogin prohibit-password
```

或者，可通过关闭设备电源并修改其 vApp 属性来启用 SSH 服务并启用 SSH root 访问权限。

为

```
PermitRootLogin yes
```

然后使用以下命令重新启动 sshd 服务器：

```
/etc/init.d/ssh restart
```

收集支持包

您可以在注册的群集和结构层节点上收集支持包，并将这些包下载到您的计算机或上载到文件服务器中。

如果您选择将包下载到您的计算机中，将获得一个存档文件，其中包含每个节点的清单文件和支持包。如果您选择将包上载到文件服务器中，则会将清单文件和各个包单独上载到文件服务器中。

NSX Cloud 说明 如果要收集 CSM 的支持包，请登录到 CSM，转到**系统 > 实用程序 > 支持包**，然后单击**下载**。可以按照以下说明从 NSX Manager 获得 PCG 的支持包。PCG 的支持包还包含所有工作负载虚拟机的日志。

步骤

1 从浏览器中，使用管理员权特权登录到位于 *https://nsx-manager-ip-address* 的 NSX Manager。

2 从导航面板中选择**系统 > 实用程序**。

3 单击**支持包**选项卡。

4 选择目标节点。

可用的节点类型包括“管理节点”、“控制器节点”、“Edge”、“主机”和“公有云网关”。

5 （可选）指定日志期限天数以排除早于指定天数的日志。

6 （可选）切换开关，选择包括或排除核心文件和审核日志。

注意 核心文件和审核日志可能包含敏感信息，例如，密码或加密密钥。

7 （可选）选中相应的复选框以将包上载到文件服务器中。

8 单击**开始收集支持包**以开始收集支持包。

根据存在的日志文件数，每个节点可能需要几分钟的时间。

9 监控收集过程的状态。

状态字段显示完成支持包收集的节点的百分比。

10 如果未设置将包发送到文件服务器的选项，请单击**下载**以下载包。

对第 2 层连接问题进行故障排除 2

如果在连接到同一个逻辑交换机的两个虚拟接口 (virtual interface, VIF) 之间出现通信故障，例如，不能从一个虚拟机 ping 另一个虚拟机，您可以按照本节中的步骤对此故障进行故障排除。

开始之前，请确保没有防火墙规则阻止两个逻辑端口之间的通信。建议您按照本节中的主题顺序对连接问题进行故障排除。

本章讨论了以下主题：

- 检查 NSX Manager 和 NSX Controller 群集状态
- 检查逻辑端口
- 检查传输节点状态
- 检查逻辑交换机状态
- 检查逻辑交换机的 CCP
- 检查本地控制层面状态
- 对配置会话问题进行故障排除
- 对 L2 会话问题进行故障排除
- 对覆盖网络逻辑交换机的数据层面问题进行故障排除
- 对 VLAN 逻辑交换机的数据层面问题进行故障排除
- 对覆盖网络逻辑交换机的 ARP 问题进行故障排除
- 对 VLAN 逻辑交换机或解析 ARP 时的数据包丢失问题进行故障排除

检查 NSX Manager 和 NSX Controller 群集状态

确认 NSX Manager 和 NSX Controller 群集的状态为正常，并且控制器已连接到 NSX Manager。

步骤

1 在 NSX Manager 上运行以下 CLI 命令以确保状态为稳定。

```
NSX-Manager> get management-cluster status
Number of nodes in management cluster: 1
- 192.168.110.47 (UUID 45A8869B-BB90-495D-8A01-69B5FCC56086) Online
```

```
Management cluster status: STABLE

Number of nodes in control cluster: 3
- 192.168.110.201 (UUID 45A8869B-BB90-495D-8A01-69B5FCC56086)
- 192.168.110.202 (UUID 45A8869B-BB90-495D-8A01-69B5FCC56086)
- 192.168.110.203 (UUID 45A8869B-BB90-495D-8A01-69B5FCC56086)
```

2 在 NSX Controller 上运行以下 CLI 命令以确保状态为活动。

```
NSX-Controller1> get control-cluster status
uuid: db4aa77a-4397-4d65-ad33-9fde79ac3c5c
is master: true
in majority: true
uuid                                   address               status
0cfe232e-6c28-4fea-8aa4-b3518baef00d 192.168.110.201       active
bd257108-b94e-4e6d-8b19-7fa6c012961d 192.168.110.202       active
538be554-1240-40e4-8e94-1497e963a2aa 192.168.110.203       active
```

3 在 NSX Controller 上运行以下 CLI 命令以确保它已连接到 NSX Manager。

```
NSX-Controller1> get managers
- 192.168.110.47   Connected
```

检查逻辑端口

检查是否已在同一个逻辑交换机上配置逻辑端口，并且其状态为已连接。

步骤

1 从 NSX Manager GUI 获取逻辑端口 UUID。

2 对每一个逻辑端口执行以下 API 调用，以确保逻辑端口位于同一个逻辑交换机。

```
GET https://<nsx-mgr>/api/v1/logical-ports/<logical-port-uuid>
```

3 对每个逻辑端口执行以下 API 调用，以确保状态为已连接。

```
GET https://<nsx-mgr>/api/v1/logical-ports/<logical-port-uuid>/status
```

检查传输节点状态

检查传输节点的状态。

步骤

- 执行以下 API 调用以获取传输节点的状态。

```
GET https://<nsx-mgr>/api/v1/transport-nodes/<transport-node-ID>/state
```

如果调用返回错误 RPC 超时，请执行以下故障排除步骤：

- 运行 /etc/init.d/nsx-opsAgent status 以查看 opsAgent 是否正在运行。
- 运行 /etc/init.d/nsx-mpa status 以查看 nsx-mpa 是否正在运行。
- 要查看 nsx-mpa 是否已连接到 NSX Manager，请检查 nsx-mpa 检测信号日志。
- 要查看 opsAgent 是否已连接到 NSX Manager，请检查 nsx-opsAgent 日志。如果 opsAgent 已连接到 NSX Manager，您将看到以下消息。

```
Connected to mpa, cookie: ...
```

- 要查看 opsAgent 是否正卡在 HostConfigMsg 的处理上，请检查 nsx-opsAgent 日志。如果是这样，您将看到一条 RMQ 请求消息，但是不会发送回复，或者经过很长时间延迟后才会发送。
- 检查 opsAgent 是否在执行 HostConfigMsg 时已崩溃。
- 要查看 RMQ 消息是否用了很长时间才传递给主机，请比较 NSX Manager 和主机上日志消息的时间戳。

如果调用返回错误 partial_success，则有许多可能的原因。首先查看 nsx-opsAgent 日志。在 ESXi 主机上，检查 hostd.log 和 vmkernel.log。在 KVM 上，syslog 保存着所有日志。

检查逻辑交换机状态

检查逻辑交换机的状态。

步骤

- 执行以下 API 调用以获取逻辑交换机的状态。

```
GET https://<nsx-mgr>/api/v1/logical-switches/<logical-switch-ID>/state
```

如果调用返回错误 partial_success，那么回复将包含 NSX Manager 无法推送逻辑交换机配置的传输节点列表，或者不会获得回复。故障排除步骤类似于用于传输节点的步骤。请检查以下各项：

- 所有必需组件均已安装且正常运行。
- nsx-mpa 已连接到 NSX Manager。
- nsxa 已连接到交换垂直项。
- 通过 grep 查找 nsxa.log 和 nsxaVim.log 中的逻辑交换机 ID，以查看传输节点是否已收到逻辑交换机配置。
- 检查 nsxa 和 nsx-mpa 正常运行时间。通过 grep 查找 syslog 文件中的 nsxa 日志消息，查明 nsxa 何时启动和停止。
- 查明 nsxa 连接到交换垂直项的时间。如果在 nsxa 未连接到交换垂直项时将逻辑交换机配置发送到主机，配置可能不会传递到主机。

在 KVM 上，不会将任何逻辑交换机配置推送到主机。因此，大多数逻辑交换机问题都很可能出在管理层面。

在 ESXi 上，将含糊网络映射到逻辑交换机。要使用逻辑交换机，用户需使用 vCenter Server 或 vSphere API 将虚拟机连接到含糊网络。

检查逻辑交换机的 CCP

确认逻辑交换机处于中央控制层面 (CCP)。

步骤

- 在 NSX Controller 上运行以下 CLI 命令以确保逻辑交换机存在。

```
NSX-Controller1> get logical switches
VNI   UUID                                   Name
52104 feab22ec-94b2-46f4-88f8-f9d44a416272 ls1
```

注意 此 CLI 命令不会列出支持 VLAN 的逻辑交换机。

检查本地控制层面状态

对于覆盖网络逻辑交换机，请检查主机上的 `netcpa` 是否已连接到中央控制层面。

前提条件

查找逻辑交换机所在的控制器。请参见检查逻辑交换机的 CCP。

步骤

1. 通过 SSH 访问逻辑交换机所在的控制器。
2. 运行以下命令，确认控制器显示连接到该 VNI 的管理程序。

```
get logical-switch 5000 connection-table
```

3. 在管理程序上，运行命令 `/bin/nsxcli` 启动 NSX CLI。
4. 运行以下命令以获取 CCP 会话。

```
host1> get ccp-session
Session Index State Controller
Config  0     UP    10.33.74.163
L2      5000  UP    10.33.74.163
```

您应该在 CCP 群集中的其中一个 CCP 节点上看到配置会话。对于每一个覆盖网络逻辑交换机，您应该看到 CCP 群集中其中一个 CCP 节点的 L2 会话。对于 VLAN 逻辑交换机，没有 CCP 连接。

对配置会话问题进行故障排除

如果 CCP 配置会话未启动，请检查 MPA 和 netcpa 的状态。

步骤

1 执行以下 API 调用，以查看 MPA 是否已连接到 NSX Manager。

```
GET https://<nsx-mgr>/api/v1/logical-ports/<logical-port-uuid>
```

2 在管理程序上运行命令 /bin/nsxcli 启动 NSX CLI。

3 运行以下命令以获取 node-uuid。

```
host1> get node-uuid
0c123dd4-8199-11e5-95e2-73cc1cd9b614
```

4 运行以下命令以查看 NSX Manager 是否已将 CCP 信息推送到主机。

```
cat /etc/vmware/nsx/config-by-vsm.xml
```

5 如果 config-by-vsm.xml 有 CCP 信息，请检查是否已在管理程序上配置传输节点。

NSX Manager 在传输节点创建步骤中发送管理程序的主机证书。在接受来自主机的连接之前，CCP 必须具有主机证书。

6 检查 /etc/vmware/nsx/host-cert.pem 中主机证书的有效性。

该证书必须与 NSX Manager 所具有的主机证书相同。

7 运行以下命令以检查 netcpa 的状态。

在 ESXi 上：

```
/etc/init.d/netcpad status
```

在 KVM 上：

```
/etc/init.d/nsx-agent status
```

8 启动或重新启动 netcpa。

在 ESXi 上，启动 netcpa（如果未在运行），或者重新启动它（如果正在运行）。

```
/etc/init.d/netcpad start
```

```
/etc/init.d/netcpad restart
```

在 KVM 上，启动 netcpa（如果未在运行），或者重新启动它（如果正在运行）。

```
/etc/init.d/nsx-agent start
```

```
/etc/init.d/nsx-agent restart
```

9 如果配置会话仍未启动，请收集技术支持包并与 VMware 支持部门联系。

对 L2 会话问题进行故障排除

这仅适用于覆盖网络逻辑交换机。

步骤

1 在管理程序上运行命令 /bin/nsxcli 启动 NSX CLI。

2 运行以下命令以查看在主机上是否存在逻辑交换机。

```
host1> get logical-switches
```

3 检查端口状态是否不是 admin down。

在 ESXi 上，运行 net-dvs 并查看响应。例如，

```
port 63eadf53-ff92-4a0e-9496-4200e99709ff:
com.vmware.port.extraConfig.opaqueNetwork.id = … <- this should match the logical switch UUID
com.vmware.port.opaque.network.id = …. <- this should match the logical switch UUID
com.vmware.port.opaque.network.type = nsx.LogicalSwitch , propType = RUNTIME
com.vmware.common.port.block = false, ... <- Make sure the value is false.
com.vmware.vswitch.port.vxlan = …
com.vmware.common.port.volatile.status = inUse ... <- make sure the value is inUse.
```

如果逻辑端口最终处于被阻止状态，请收集技术支持包并与 VMware 支持部门联系。同时，运行以下命令以获取 DVS 名称：

```
[root@host1:~] net-dvs | grep nsx-switch
com.vmware.common.alias = nsx-switch , propType = CONFIG
```

运行以下命令以取消阻止端口：

```
[root@host1:~] net-dvs -s com.vmware.common.port.block=false <DVS-NAME> -p <logical-port-ID>
```

在 KVM 上，运行 ovs-vsctl list interface 并验证具有相应 VIF UUID 的接口是否存在，并且 admin_state 为已连接。您可以在 external-ids:iface-id 的 OVSDB 中看到 VIF UUID。

对覆盖网络逻辑交换机的数据层面问题进行故障排除

本节中的步骤适用于在配置和运行时状态正常时通过覆盖网络交换机对不同管理程序上虚拟机之间的连接问题进行故障排除。

如果虚拟机位于同一管理程序上，请转到对覆盖网络逻辑交换机的 ARP 问题进行故障排除。

步骤

1 在具有逻辑交换机的控制器上运行以下命令，以查看 CCP 是否具有正确的 VTEP 列表：

```
controller1> get logical-switch 5000 vtep
```

2 在每个管理程序上，运行以下 NSX CLI 命令以查看它是否具有正确的 VTEP 列表：

在 ESXi 上：

```
host1> get logical-switch <logical-switch-UUID> tep-table
```

或者，您可以运行以下 shell 命令查看 VTEP 信息：

```
[root@host1:~] net-vdl2 -M vtep -s vds -n VNI
```

在 KVM 上：

```
host1> get logical-switch <logical-switch-UUID or VNI> tep-table
```

3 检查管理程序上的 VTEP 是否可以相互执行 ping 操作。

在 ESXi shell 提示符下：

```
host1> ping ++netstack=vxlan <remote-VTEP-IP>
```

在 KVM shell 提示符下：

```
host1> ping <remote-VTEP-IP>
```

如果 VTEP 无法相互执行 ping 操作，

a 请确保在创建传输节点时指定的传输 VLAN 与底层网络的期望相匹配。如果在底层网络中使用的是访问端口，那么传输 VLAN 应设置为 0。如果要指定传输 VLAN，那么应将管理程序连接到的底层网络交换机端口配置为在中继模式下接受此 VLAN。

b 检查底层网络连接。

4 检查 VTEP 之间的 BFD 会话是否已启动。

在 ESXi 上，运行 `net-vdl2 -M bfd` 并查看响应。例如，

```
BFD count: 1
==========================
Local IP: 192.168.48.35, Remote IP: 192.168.197.243, Local State: up, Remote State: up, Local
Diag: No Diagnostic, Remote Diag: No Diagnostic, minRx: 1000000, isDisabled: 0
```

在 KVM 上，找到远程 IP 的 GENEVE 接口。

```
ovs-vsctl list interface <GENEVE-interface-name>
```

如果您不知道接口名称，请运行 `ovs-vsctl find Interface type=geneve` 返回所有隧道接口。查找 BFD 信息。

如果找不到远程 VTEP 的 GENEVE 接口，请检查 nsx-agent 是否正在运行以及 OVS 集成网桥是否已连接到 nsx-agent。

```
[root@host1 ~]# ovs-vsctl show
96c9e543-fc68-448a-9882-6e161c313a5b
  Manager "tcp:127.0.0.1:6632"
    is_connected: true
  Bridge nsx-managed
    Controller "tcp:127.0.0.1:6633"
      is_connected: true
    Controller "unix:ovs-l3d.mgmt"
      is_connected: true
    fail_mode: secure
```

对 VLAN 逻辑交换机的数据层面问题进行故障排除

本节中的步骤适用于在配置和运行时状态正常时通过底层网络上配置的 VLAN 对不同管理程序上虚拟机之间的连接问题进行故障排除。

如果虚拟机位于同一管理程序上，且所有配置和运行时状态都正常，请转到对覆盖网络逻辑交换机的 ARP 问题进行故障排除。

步骤

- 检查是否已在中继模式下为逻辑交换机的 VLAN 配置底层网络。

  在 ESXi 上，通过运行 net-dvs 并查找逻辑端口，确认已在逻辑端口上配置 VLAN。例如：

  ```
  port 63eadf53-ff92-4a0e-9496-4200e99709ff:
    com.vmware.common.port.volatile.vlan = VLAN 1000 propType = RUNTIME VOLATILE
  ```

  在 KVM 上，将 VLAN 逻辑交换机配置为集成网桥上的 OpenFlow 规则。换句话说，对于从 VIF 收到的流量，使用 VLAN X 进行标记，然后在修补程序端口上将其转发到 PIF 网桥。运行 `ovs-vsctl list interface` 并确认在 NSX 管理的网桥和 NSX 交换机网桥之间存在修补程序端口。

对覆盖网络逻辑交换机的 ARP 问题进行故障排除

本节中的步骤适用于对覆盖网络交换机的数据包丢失问题进行故障排除。

对于支持 VLAN 的逻辑交换机，请转到对 VLAN 逻辑交换机或解析 ARP 时的数据包丢失问题进行故障排除。

在执行以下故障排除步骤之前，请在每个虚拟机上运行命令 `arp -n`。如果在两个虚拟机上成功解析 ARP，则不需要执行本节中的步骤。而是，转到下一节对 VLAN 逻辑交换机或解析 ARP 时的数据包丢失问题进行故障排除。

步骤

- 如果两个端点都是 ESXi 且已在逻辑交换机上启用 ARP 代理（仅支持覆盖网络逻辑交换机），请在 CCP 和管理程序上检查 ARP 表。

  在 CCP 上：

  ```
  controller1> get logical-switch 5000 arp-table
  ```

  在管理程序上，启动 NSX CLI 并运行以下命令：

  ```
  host1> get logical-switch <logical-switch-UUID> arp-table
  ```

  获取 ARP 表只会告诉我们 ARP 代理状态是否正确。如果未通过代理收到 ARP 响应，或者，如果主机是 KVM 且不支持 ARP 代理，则数据路径应该广播 ARP 请求。可能存在 BUM 流量转发问题。尝试执行以下步骤：

  - 如果逻辑交换机的复制模式是 `MTEP`，请从 NSX Manager GUI 将逻辑交换机的复制模式更改为 `SOURCE`。这可能会修复此问题，使 ping 操作能够正常工作。
  - 添加静态 ARP 条目并查看其余数据路径是否正常工作。

对 VLAN 逻辑交换机或解析 ARP 时的数据包丢失问题进行故障排除

您可以使用自动跟踪流工具，或者手动跟踪数据包对数据包丢失问题进行故障排除。

要运行跟踪流工具，请从 NSX Manager GUI 导航到**工具 > 跟踪流**。有关详细信息，请参见*《NSX-T 管理指南》*。

步骤

- 要手动跟踪数据包，

  在 ESXi 上，运行 `net-stats -l` 以获取 VIF 的交换机端口 ID。如果源和目标 VIF 在同一个管理程序上，请运行以下命令：

  ```
  pktcap-uw --switchport <src-switch-port-ID> --dir=0
  pktcap-uw --switchport <dst-switch-port-ID> --dir=1
  ```

  如果源和目标 VIF 在不同管理程序上，请在托管源 VIF 的管理程序上运行以下命令：

  ```
  pktcap-uw --switchport <src-switch-port-ID> --dir=0
  pktcap-uw --uplink <uplink-name> --dir=1
  ```

在托管目标 VIF 的管理程序上，运行以下命令：

```
pktcap-uw --uplink <uplink-name> --dir=0
pktcap-uw --switchport <dest-switch-port-ID> --dir=1
```

在 KVM 上，如果源和目标 VIF 在同一个管理程序上，请运行以下命令：

```
ovs-dpctl dump-flows
```

安装故障排除

3

本节提供了有关对安装问题进行故障排除的信息。

基本基础架构服务

以下服务必须在设备和管理程序上运行，此外，如果 vCenter Server 用作计算管理器，还必须在 vCenter Server 上运行。

- NTP
- DNS

确保防火墙未阻止 NSX-T 组件和管理程序之间的流量。确保已在组件之间打开所需端口。

要刷新 NSX Manager 上的 DNS 缓存，请以 root 用户身份通过 SSH 访问管理器并运行以下命令：

```
root@nsx-mgr-01:~# /etc/init.d/resolvconf restart
[ ok ] Restarting resolvconf (via systemctl): resolvconf.service.
```

然后，可以检查 DNS 配置文件。

```
root@nsx-mgr-01:~# cat /etc/resolv.conf
# Dynamic resolv.conf(5) file for glibc resolver(3) generated by resolvconf(8)
#     DO NOT EDIT THIS FILE BY HAND -- YOUR CHANGES WILL BE OVERWRITTEN
nameserver 192.168.253.1
search mgt.sg.lab
```

检查主机到控制器和管理器的通信

在 ESXi 主机上使用 NSX-T CLI 命令：

```
esxi-01.corp.local> get managers
- 192.168.110.19   Connected

esxi-01.corp.local> get controllers
 Controller IP    Port     SSL         Status       Is Physical Master   Session State  Controller FQDN
 192.168.110.16   1235   enabled     connected             true               up               NA
```

在 KVM 主机上使用 NSX-T CLI 命令：

```
kvm-01> get managers
- 192.168.110.19   Connected

kvm-01> get controllers
 Controller IP    Port     SSL         Status       Is Physical Master   Session State  Controller FQDN
 192.168.110.16   1235   enabled     connected             true               up               NA
```

在 ESXi 主机上使用主机 CLI 命令：

```
[root@esxi-01:~] esxcli network ip  connection list | grep 1235
tcp         0       0  192.168.110.53:42271                      192.168.110.16:1235   ESTABLISHED
67702  newreno  netcpa
[root@esxi-01:~]
[root@esxi-01:~] esxcli network ip  connection list | grep 5671
tcp         0       0  192.168.110.253:11721             192.168.110.19:5671   ESTABLISHED   2103688
newreno  mpa
tcp         0       0  192.168.110.253:30977             192.168.110.19:5671   ESTABLISHED   2103688
newreno  mpa
```

在 KVM 主机上使用主机 CLI 命令：

```
root@kvm-01:/home/vmware# netstat -nap | grep 1235
tcp        0      0 192.168.110.55:53686    192.168.110.16:1235     ESTABLISHED 2554/netcpa
root@kvm-01:/home/vmware#
root@kvm-01:/home/vmware#
root@kvm-01:/home/vmware# netstat -nap | grep 5671
tcp        0      0 192.168.110.55:50108    192.168.110.19:5671     ESTABLISHED 2870/mpa
tcp        0      0 192.168.110.55:50110    192.168.110.19:5671     ESTABLISHED 2870/mpa

root@kvm-01:/home/vmware# tcpdump -i ens32 port 1235 | grep kvm-01
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on ens32, link-type EN10MB (Ethernet), capture size 262144 bytes
<truncated output>
03:46:27.040461 IP nsxcontroller01.corp.local.1235 > kvm-01.corp.local.38754: Flags [P.], seq
3315301231:3315301275, ack 2671171555, win 323, length 44
03:46:27.040509 IP kvm-01.corp.local.38754 > nsxcontroller01.corp.local.1235: Flags [.], ack 44, win
1002, length 0
^C
<truncated output>
root@kvm-01:/home/vmware#

root@kvm-01:/home/vmware# tcpdump -i ens32 port 5671 | grep kvm-01
tcpdump: verbose output suppressed, use -v or -vv for full protocol decode
listening on ens32, link-type EN10MB (Ethernet), capture size 262144 bytes
03:51:16.802934 IP kvm-01.corp.local.58954 > nsxmgr01.corp.local.amqps: Flags [P.], seq 1153:1222, ack
1790, win 259, length 69
03:51:16.823328 IP nsxmgr01.corp.local.amqps > kvm-01.corp.local.58954: Flags [P.], seq 1790:1891, ack
1222, win 254, length 101
^C
<truncated output>
```

主机注册失败

如果 NSX-T 使用错误的 IP 地址，主机注册将失败。当主机有多个 IP 地址时，可能会发生这种情况。尝试删除传输节点将使其处于“孤立”状态。要解决此问题，请执行以下操作：

- 转到**结构层 > 节点 > 主机**，编辑主机并移除除管理 IP 地址之外的所有其他 IP 地址。
- 单击错误，然后选择**解决**。

KVM 主机问题

KVM 主机问题有时是由磁盘空间不足引起的。`/Boot` 目录可能会快速填满并导致如下错误：

- 无法在主机上安装软件 (Failed to install software on host)
- 设备上没有剩余空间 (No space left on device)

您可以运行命令 **df-h** 检查可用存储。如果 `/boot` 目录达到 100%，可以执行以下操作：

- 运行 `sudo dpkg --list 'linux-image*' | grep ^ii` 以查看安装的所有内核。
- 运行 `uname -r` 以查看当前正在运行的内核。不要移除此内核 (linux-image)。
- 使用 `apt-get purge` 移除不再需要的映像。例如，运行 `sudo apt-get purge linux-image-3.13.0-32-generic linux-image-3.13.0-33-generic`。
- 重新引导主机。
- 在 NSX Manager 中，检查错误，然后选择**解决**。
- 确保虚拟机已打开电源。

部署 Edge 虚拟机时遇到配置错误

部署 Edge 虚拟机之后，NSX Manager 将虚拟机的状态显示为**配置错误**。管理器日志包含类似于以下内容的消息：

```
nsx-manager NSX - FABRIC [nsx@6876 comp="nsx-manager" errorCode="MP16027" subcomp="manager"] Edge
758ad396-0754-11e8-877e-005056abf715 is not ready for configuration error occurred, error detail is NSX
Edge configuration has failed. The host does not support required cpu features: ['aes'].
```

重新启动 Edge 数据路径服务，然后虚拟机应该能够解决该问题。

强制移除传输节点

您可以通过以下 API 调用移除停留在“孤立”状态的传输节点：

```
DELETE https://<NSX Manager>/api/v1/transport-nodes/<TN ID>?force=true
```

NSX Manager 将不对是否有任何活动虚拟机在主机上运行执行任何验证。您负责删除 N-VDS 和 VIB。如果您通过计算管理器添加了节点，请先删除计算管理器，然后再删除该节点。将同时删除传输节点。

路由故障排除

4

NSX-T 内置有用于对路由问题进行故障排除的工具。

跟踪流

您可以使用跟踪流检查数据包流。您可以查看已传送、已丢弃、已接收及已转发的数据包。如果数据包丢弃，将显示原因。例如，数据包可能因防火墙规则而丢弃。

检查路由表

要在服务路由器上查看路由表，请运行以下命令：

```
edge01> get logical-router
Logical Route
UUID                                   VRF    LR-ID  Name               Type                            Ports
736a80e3-23f6-5a2d-81d6-bbefb2786666   0      0                           TUNNEL                          3
c9393d0c-1fcf-4c34-889d-2da1eeee25b8   1      10     SR-t0-router         SERVICE_ROUTER_TIER0            5
9333c94e-5938-46b4-8c7d-5e6ac2c8b7b5   2      8      DR-t1-router01       DISTRIBUTED_ROUTER_TIER1        6
c91eb7c5-0297-4fed-9c22-b96df1c9b80f   3      9      DR-t0-router         DISTRIBUTED_ROUTER_TIER0        4

edge01> vrf 1
edge01(tier0_sr)> get route

Flags: c - connected, s - static, b - BGP, ns - nsx_static
nc - nsx_connected, rl - router_link, t0n: Tier0-NAT, t1n: Tier1-NAT
t1l: Tier1-LB VIP, t1s: Tier1-LB SNAT

Total number of routes: 25

b    10.10.20.0/24        [20/0]        via 192.168.140.1
b    10.10.30.0/24        [20/0]        via 192.168.140.1
b    10.20.20.0/24        [20/0]        via 192.168.140.1
b    10.20.30.0/24        [20/0]        via 192.168.140.1
b    30.0.0.0/8           [20/0]        via 192.168.140.1
rl   100.64.80.0/31       [0/0]         via 169.254.0.1
rl   100.64.80.2/31       [0/0]         via 169.254.0.1
rl   100.64.80.4/31       [0/0]         via 169.254.0.1
<TRUNCATED OUTPUT>
b    192.168.200.0/24     [20/0]        via 192.168.140.1
```

```
b    192.168.210.0/24     [20/0]        via 192.168.140.1
b    192.168.220.0/24     [20/0]        via 192.168.140.1
b    192.168.230.0/24     [20/0]        via 192.168.140.1
b    192.168.240.0/24     [20/0]        via 192.168.140.1
```

要获取接口的 IP 地址，请运行以下命令：

```
edge01(tier0_sr)> get interfaces
Logical Router
UUID                                   VRF  LR-ID  Name           Type
c9393d0c-1fcf-4c34-889d-2da1eeee25b8   1    10     SR-t0-router   SERVICE_ROUTER_TIER0
interfaces
    interface   : 977ac2eb-8ab7-40e9-8abe-782a438c749a
    ifuid       : 285
    name        : uplink01
    mode        : lif
    IP/Mask     : 192.168.140.3/24
    MAC         : 00:50:56:b5:d5:64
    LS port     : 14391f86-efef-4e3d-98c3-f291c17d13f8
    urpf-mode   : STRICT_MODE
    admin       : up
    MTU         : 1600

    interface   : 6af81d72-4d32-5f66-b7ae-403e617290e5
    ifuid       : 270
    mode        : blackhole

    interface   : 015e709d-6079-5c19-9556-8be2e956f775
    ifuid       : 269
    mode        : cpu

    interface   : 3f40f838-eb8a-4f35-854c-ea8bb872dc47
    ifuid       : 272
    name        : bp-sr0-port
    mode        : lif
    IP/Mask     : 169.254.0.2/28
    MAC         : 02:50:56:56:53:00
    VNI         : 25489
    LS port     : 770a208d-27fa-4f8d-afad-a9c41ca6295b
    urpf-mode   : NONE
    admin       : up
    MTU         : 1500

    interface   : 00003300-0000-0000-0000-00000000000a
    ifuid       : 263
    mode        : loopback
    IP/Mask     : 127.0.0.1/8
```

通告 T1 路由

必须通告 T1 路由，以便在 T0 路由器及以上路由器上可见。有不同类型的路由可以通告：NSX 已连接、NAT、静态、LB VIP 和 LB SNAT。

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：

- 确定 ESXi 主机上应用的防火墙规则
- 确定 KVM 主机上应用的防火墙规则
- 防火墙数据包日志

确定 ESXi 主机上应用的防火墙规则

要对 ESXi 主机的防火墙问题进行故障排除，可以查看在主机上应用的防火墙规则。

获取 ESXi 主机上的 dvfilter 列表：

```
[root@esxi-01:~] summarize-dvfilter
<TRUNCATED OUTPUT>
world 70181 vmm0:app-01a vcUuid:'50 35 9c 70 18 8e 99 1d-3c f9 8e cc 6b 27 4c 6f'
 port 50331655 app-01a.eth0
  vNic slot 2
  name: nic-70181-eth0-vmware-sfw.2
 agentName: vmware-sfw
   state: IOChain Attached
   vmState: Detached
   failurePolicy: failClosed
   slowPathID: none
   filter source: Dynamic Filter Creation
world 70179 vmm0:web-02a vcUuid:'50 35 2b f3 4a 4b 10 83-54 72 50 f7 25 10 d8 64'
 port 50331656 web-02a.eth0
  vNic slot 2
  name: nic-70179-eth0-vmware-sfw.2
 agentName: vmware-sfw
   state: IOChain Attached
   vmState: Detached
   failurePolicy: failClosed
   slowPathID: none
   filter source: Dynamic Filter Creation
```

在 dvfilter 中查找特定虚拟机：

```
[root@esxi-01:~] summarize-dvfilter | less -p web

world 70179 vmm0:web-02a vcUuid:'50 35 2b f3 4a 4b 10 83-54 72 50 f7 25 10 d8 64'
 port 50331656 web-02a.eth0
  vNic slot 2
  name: nic-70179-eth0-vmware-sfw.2
 agentName: vmware-sfw
   state: IOChain Attached
   vmState: Detached
   failurePolicy: failClosed
   slowPathID: none
   filter source: Dynamic Filter Creation
.
.
.
```

确定应用于特定 dvfilter 的防火墙规则（在此示例中，`nic-70227-eth0-vmware-sfw.2` 为 dvfilter 名称）：

```
[root@esxi-02:~] vsipioctl getrules -f nic-70227-eth0-vmware-sfw.2
ruleset mainrs {
rule 3072 at 1 inout protocol tcp from any to addrset 48822ec3-2670-497b-82f9-524618c16877 port 443
accept with log;
rule 3072 at 2 inout protocol tcp from any to addrset 48822ec3-2670-497b-82f9-524618c16877 port 80 accept
with log;
rule 3074 at 3 inout protocol tcp from addrset 48822ec3-2670-497b-82f9-524618c16877 to addrset 8b9e75e7-
bc62-4d7f-9a58-a872f393448e port 8443 accept with log;
rule 3074 at 4 inout protocol tcp from addrset 48822ec3-2670-497b-82f9-524618c16877 to addrset 8b9e75e7-
bc62-4d7f-9a58-a872f393448e port 22 accept with log;
rule 3075 at 5 inout protocol tcp from addrset 8b9e75e7-bc62-4d7f-9a58-a872f393448e to addrset
b695c8df-9894-4068-a5e7-5504fe48d459 port 3306 accept with log;
rule 3076 at 6 inout protocol tcp from ip 192.168.110.10 to addrset rdst3076 port 443 accept with log;
rule 3076 at 7 inout protocol icmp typecode 8:0 from ip 192.168.110.10 to addrset rdst3076 accept with
log;
rule 3076 at 8 inout protocol tcp from ip 192.168.110.10 to addrset rdst3076 port 22 accept with log;
rule 3076 at 9 inout protocol tcp from ip 192.168.110.10 to addrset rdst3076 port 80 accept with log;
rule 2 at 10 inout protocol any from any to any accept with log;
}

ruleset mainrs_L2 {
rule 1 at 1 inout ethertype any stateless from any to any accept;
}
```

获取特定 dvfilter 中使用的地址集列表：

```
[root@esxi-02:~]  vsipioctl getaddrsets -f nic-70227-eth0-vmware-sfw.2
addrset 48822ec3-2670-497b-82f9-524618c16877 {
ip 172.16.10.13,
mac 52:54:00:42:4d:38,
}
addrset 8b9e75e7-bc62-4d7f-9a58-a872f393448e {
}
```

```
addrset b695c8df-9894-4068-a5e7-5504fe48d459 {
ip 172.16.30.11,
mac 52:54:00:64:0e:4f,
}
addrset rdst3076 {
ip 172.16.10.13,
ip 172.16.30.11,
mac 52:54:00:42:4d:38,
mac 52:54:00:64:0e:4f,
}
```

检查通过特定 dvfilter 的流量：

```
[root@esxi-02:~] vsipioctl getflows -f nic-75360-eth0-vmware-sfw.2
Count retrieved from kernel active(L3,L4)=20, active(L2)+inactive(L3,L4)=0, drop(L2,L3,L4)=0
a5d914f7a5b85fe5 Active tcp 0800 IN 3076 0 0  192.168.110.10:Unknown(51281) -> 172.16.10.11:ssh(22) 513
FINWAIT2:FINWAIT2  4304 5177 34 33
a5d914f7a5b86001 Active tcp 0800 OUT 2 0 0  172.16.10.11:http(80) -> 100.64.80.1:Unknown(60006) 457
SYNSENT:CLOSED  56 819 1 1
a5d914f7a5b86006 Active igmp 0800 IN 2 0 0  0.0.0.0 -> 224.0.0.1 36 0 1 0
a5d914f7a5b86011 Active tcp 0800 IN 3072 0 0  100.64.80.1:Unknown(60098) -> 172.16.10.11:http(80) 320
FINWAIT2:FINWAIT2  413 5411 9 6
a5d914f7a5b86012 Active tcp 0800 OUT 3074 0 0  172.16.10.11:Unknown(46001) -> 172.16.20.11:Unknown(8443)
815 FINWAIT2:FINWAIT2  7418 1230 10 9
a5d914f7a5b86013 Active udp 0800 OUT 2 0 0  172.16.10.11:Unknown(40080) -> 192.168.110.10:domain(53)
268 140 2 2
a5d914f7a5b86014 Active udp 0800 OUT 2 0 0  172.16.10.11:Unknown(59251) -> 192.168.110.10:domain(53)
268 140 2 2
a5d914f7a5b86015 Active ipv6-icmp 86dd OUT 2 0 0  fe80::250:56ff:feb5:a60e -> ff02::1:ff62:5ed4 135 0 0
72 0 1
a5d914f7a5b86016 Active ipv6-icmp 86dd OUT 2 0 0  fe80::250:56ff:feb5:a60e -> ff02::1:ff62:5ed4 135 0 0
72 0 1
a5d914f7a5b86017 Active tcp 0800 IN 3072 0 0  100.64.80.1:Unknown(60104) -> 172.16.10.11:http(80) 320
FINWAIT2:FINWAIT2  413 5451 9 7
a5d914f7a5b86018 Active tcp 0800 OUT 3074 0 0  172.16.10.11:Unknown(46002) -> 172.16.20.11:Unknown(8443)
815 TIMEWAIT:TIMEWAIT  7314 1230 8 9
a5d914f7a5b86019 Active tcp 0800 IN 3072 0 0  100.64.80.1:Unknown(60110) -> 172.16.10.11:http(80) 320
FINWAIT2:FINWAIT2  373 5451 8 7
a5d914f7a5b8601a Active tcp 0800 OUT 3074 0 0  172.16.10.11:Unknown(46003) -> 172.16.20.11:Unknown(8443)
815 FINWAIT2:FINWAIT2  7418 1230 10 9
a5d914f7a5b8601b Active tcp 0800 IN 3072 0 0  100.64.80.1:Unknown(60114) -> 172.16.10.11:http(80) 328
TIMEWAIT:TIMEWAIT  413 5451 9 7
a5d914f7a5b8601c Active tcp 0800 OUT 3074 0 0  172.16.10.11:Unknown(46004) -> 172.16.20.11:Unknown(8443)
815 TIMEWAIT:TIMEWAIT  7262 1218 7 9
a5d914f7a5b8601d Active tcp 0800 OUT 2 0 0  172.16.10.11:http(80) -> 100.64.80.1:Unknown(60060) 457
SYNSENT:CLOSED  56 819 1 1
a5d914f7a5b8601e Active tcp 0800 IN 3072 0 0  100.64.80.1:Unknown(60120) -> 172.16.10.11:http(80) 320
TIMEWAIT:TIMEWAIT  373 5411 8 6
a5d914f7a5b8601f Active tcp 0800 OUT 3074 0 0  172.16.10.11:Unknown(46005) -> 172.16.20.11:Unknown(8443)
815 FINWAIT2:FINWAIT2  7418 1230 10 9
a5d914f7a5b86020 Active tcp 0800 IN 3072 0 0  100.64.80.1:Unknown(60126) -> 172.16.10.11:http(80) 229
EST:EST  173 5371 3 5
a5d914f7a5b86021 Active tcp 0800 OUT 3074 0 0  172.16.10.11:Unknown(46006) -> 172.16.20.11:Unknown(8443)
815 FINWAIT2:FINWAIT2  7418 1230 10 9
```

确定 KVM 主机上应用的防火墙规则

要对 KVM 主机的防火墙问题进行故障排除，可以查看在主机上应用的防火墙规则。

获取遵守 KVM 主机上防火墙规则的 VIF 列表：

```
# ovs-appctl -t /var/run/openvswitch/nsxa-ctl dfw/vif
Vif ID       : da95fc1e-65fd-461f-814d-d92970029bf0
Port name    : db-01a-eth0
Port number  : 2
```

如果输出为空，请查找节点与控制器之间的连接问题。

获取应用于特定 VIF 的规则列表（在此示例中，da95fc1e-65fd-461f-814d-d92970029bf0 为 VIF ID）：

```
# ovs-appctl -t /var/run/vmware/nsx-agent/nsxa-ctl dfw/rules da95fc1e-65fd-461f-814d-d92970029bf0
Distributed firewall status: enabled

Vif ID       : da95fc1e-65fd-461f-814d-d92970029bf0
ruleset d035308b-cb0d-4e7e-aae5-a428b461db46 {
 rule 3072 inout protocol tcp from any to addrset 48822ec3-2670-497b-82f9-524618c16877 port 443 accept
with log;
 rule 3072 inout protocol tcp from any to addrset 48822ec3-2670-497b-82f9-524618c16877 port 80 accept
with log;
 rule 3074 inout protocol tcp from addrset 48822ec3-2670-497b-82f9-524618c16877 to addrset 8b9e75e7-
bc62-4d7f-9a58-a872f393448e port 8443 accept with log;
 rule 3074 inout protocol tcp from addrset 48822ec3-2670-497b-82f9-524618c16877 to addrset 8b9e75e7-
bc62-4d7f-9a58-a872f393448e port 22 accept with log;
 rule 3075 inout protocol tcp from addrset 8b9e75e7-bc62-4d7f-9a58-a872f393448e to addrset
b695c8df-9894-4068-a5e7-5504fe48d459 port 3306 accept with log;
}

ruleset 3027fed3-60b1-483e-aa17-c28719275704 {
 rule 3076 inout protocol tcp from 192.168.110.10 to addrset b695c8df-9894-4068-a5e7-5504fe48d459 port
443 accept with log;
 rule 3076 inout protocol icmp type 8 code 0 from 192.168.110.10 to addrset b695c8df-9894-4068-
a5e7-5504fe48d459 accept with log;
 rule 3076 inout protocol tcp from 192.168.110.10 to addrset b695c8df-9894-4068-a5e7-5504fe48d459 port
22 accept with log;
 rule 3076 inout protocol tcp from 192.168.110.10 to addrset b695c8df-9894-4068-a5e7-5504fe48d459 port
80 accept with log;
 rule 3076 inout protocol tcp from 192.168.110.10 to addrset 8b9e75e7-bc62-4d7f-9a58-a872f393448e port
443 accept with log;
 rule 3076 inout protocol icmp type 8 code 0 from 192.168.110.10 to addrset 8b9e75e7-bc62-4d7f-9a58-
a872f393448e accept with log;
 rule 3076 inout protocol tcp from 192.168.110.10 to addrset 8b9e75e7-bc62-4d7f-9a58-a872f393448e port
22 accept with log;
 rule 3076 inout protocol tcp from 192.168.110.10 to addrset 8b9e75e7-bc62-4d7f-9a58-a872f393448e port
80 accept with log;
 rule 3076 inout protocol tcp from 192.168.110.10 to addrset 48822ec3-2670-497b-82f9-524618c16877 port
443 accept with log;
 rule 3076 inout protocol icmp type 8 code 0 from 192.168.110.10 to addrset
48822ec3-2670-497b-82f9-524618c16877 accept with log;
```

```
 rule 3076 inout protocol tcp from 192.168.110.10 to addrset 48822ec3-2670-497b-82f9-524618c16877 port
22 accept with log;
 rule 3076 inout protocol tcp from 192.168.110.10 to addrset 48822ec3-2670-497b-82f9-524618c16877 port
80 accept with log;
}

ruleset 5e9bdcb3-adba-4f67-a680-5e6ed5b8f40a {
 rule 2 inout protocol any from any to any accept with log;
}

ruleset ddf93011-4078-4006-b8f8-73f979d7a717 {
 rule 1 inout ethertype any stateless from any to any accept;
}
```

获取特定 VIF 中使用的地址集列表：

```
# ovs-appctl -t /var/run/vmware/nsx-agent/nsxa-ctl dfw/addrsets da95fc1e-65fd-461f-814d-d92970029bf0
48822ec3-2670-497b-82f9-524618c16877 {
 mac 52:54:00:42:4d:38,
 ip 172.16.10.13,
}

8b9e75e7-bc62-4d7f-9a58-a872f393448e {
}

b695c8df-9894-4068-a5e7-5504fe48d459 {
 mac 52:54:00:64:0e:4f,
 ip 172.16.30.11,
}
```

检查通过 Linux Conntrack 模块的连接。在此示例中，我们查找两个特定 IP 地址之间的流量。

```
# ovs-appctl -t ovs-l3d conntrack/show | grep 192.168.110.10 | grep 172.16.10.13
ACTIVE
icmp,orig=(src=192.168.110.10,dst=172.16.10.13,id=1,type=8,code=0),reply=(src=172.16.10.13,dst=192.168.
110.10,id=1,type=0,code=0),start=2018-03-26T04:43:28.325,id=3122159040,zone=23119,status=SEEN_REPLY|
CONFIRMED,timeout=29,mark=3076,labels=0x1f
```

防火墙数据包日志

如果为防火墙规则启用了日志记录，则可以通过查看防火墙数据包日志，对问题进行故障排除。

不管是 ESXi 还是 KVM 主机，该日志文件都是 `/var/log/dfwpktlogs.log`。

```
# tail -f /var/log/dfwpktlogs.log
2018-03-27T10:23:35.196Z INET TERM 3072 IN TCP FIN 100.64.80.1/60688->172.16.10.11/80 8/7 373/5451
2018-03-27T10:23:35.196Z INET TERM 3074 OUT TCP FIN 172.16.10.11/46108->172.16.20.11/8443 8/9 1178/7366
2018-03-27T10:23:35.196Z INET TERM 3072 IN TCP RST 100.64.80.1/60692->172.16.10.11/80 9/6 413/5411
2018-03-27T10:23:35.196Z INET TERM 3074 OUT TCP RST 172.16.10.11/46109->172.16.20.11/8443 9/7 1218/7262
2018-03-27T10:23:37.442Z 71d32787 INET match PASS 3074 IN 60 TCP 172.16.10.12/35770->172.16.20.11/8443 S
2018-03-27T10:23:38.492Z  INET match PASS 2 OUT 1500 TCP 172.16.10.11/80->100.64.80.1/60660 A
2018-03-27T10:23:39.934Z  INET match PASS 3072 IN 52 TCP 100.64.80.1/60720->172.16.10.11/80 S
```

```
2018-03-27T10:23:39.944Z  INET match PASS 3074 OUT 60 TCP 172.16.10.11/46114->172.16.20.11/8443 S
2018-03-27T10:23:39.944Z 71d32787 INET match PASS 3074 IN 60 TCP 172.16.10.11/46114->172.16.20.11/8443 S
2018-03-27T10:23:42.449Z 71d32787 INET match PASS 3074 IN 60 TCP 172.16.10.12/35771->172.16.20.11/8443 S
2018-03-27T10:23:44.712Z INET TERM 3074 IN TCP RST 172.16.10.11/46109->172.16.20.11/8443 9/7 1218/7262
2018-03-27T10:23:44.712Z INET TERM 3074 IN TCP FIN 172.16.10.12/35766->172.16.20.11/8443 9/10 1233/7418
2018-03-27T10:23:44.712Z INET TERM 3074 IN TCP FIN 172.16.10.11/46110->172.16.20.11/8443 9/9 1230/7366
2018-03-27T10:23:44.712Z INET TERM 3074 IN TCP FIN 172.16.10.12/35767->172.16.20.11/8443 9/10 1233/7418
2018-03-27T10:23:44.939Z  INET match PASS 3072 IN 52 TCP 100.64.80.1/60726->172.16.10.11/80 S
2018-03-27T10:23:44.957Z  INET match PASS 3074 OUT 60 TCP 172.16.10.11/46115->172.16.20.11/8443 S
2018-03-27T10:23:44.957Z 71d32787 INET match PASS 3074 IN 60 TCP 172.16.10.11/46115->172.16.20.11/8443 S
2018-03-27T10:23:45.480Z INET TERM 2 OUT TCP TIMEOUT 172.16.10.11/80->100.64.80.1/60528 1/1 1500/56
```

6 其他故障排除场景

本节介绍如何对各种错误场景进行故障排除。

本章讨论了以下主题：

- 无法添加或删除传输节点
- 传输节点大约需要 5 分钟才能连接到另一个控制器
- NSX Manager 虚拟机降级
- NSX 代理与 NSX Manager 通信时超时
- 无法添加 ESXi 主机
- NSX Controller 状态不正确
- KVM 虚拟机上的管理 IP 在启用 IPFIX 的情况下无法访问

无法添加或删除传输节点

您不能删除或添加传输节点。

问题

在以下情况下会出现此错误：

1 ESXi 主机是结构层节点和传输节点。

2 作为传输节点移除主机。但是，传输节点删除失败。传输节点的状态为孤立。

3 立即作为结构层节点移除主机。

4 将主机重新添加为结构层节点。

5 使用新的传输区域和交换机，将主机添加为传输节点。此步骤会导致错误失败/部分成功。

原因

在步骤 2 中，如果等待几分钟时间，传输节点删除将成功，因为 NSX Manager 会重试删除。立即删除结构层节点时，NSX Manager 无法重试，因为已从 NSX-T Data Center 中移除主机。这会导致主机清理不完全，交换机配置仍然存在，从而导致步骤 5 失败。

解决方案

1 从主机上的 vCenter Server 中删除连接到 NSX-T Data Center 交换机的所有 vmknic。

2 使用 `esxcfg-vswitch -l` CLI 命令获取交换机名称。例如：

```
esxcfg-vswitch -l
Switch Name      Num Ports   Used Ports  Configured Ports  MTU     Uplinks
vSwitch0         1536        4           128               1500    vmnic0

  PortGroup Name        VLAN ID  Used Ports  Uplinks
  VM Network            0        0           vmnic0
  Management Network    0        1           vmnic0

Switch Name      Num Ports   Used Ports  Uplinks
nsxvswitch       1536        4
```

3 使用 `esxcfg-vswitch -d <switch-name> --dvswitch` CLI 命令删除交换机名称。例如：

```
esxcfg-vswitch -d nsxvswitch --dvswitch
```

传输节点大约需要 5 分钟才能连接到另一个控制器

在 ESXi 传输节点连接的控制器发生故障时，传输节点大约需要 5 分钟才能连接到另一个控制器。

问题

ESXi 传输节点通常连接到控制器群集中的特定控制器。您可以使用 CLI 命令 `get controllers` 查找已连接的控制器。如果连接的控制器发生故障，传输节点大约需要 5 分钟才能连接到另一个控制器。

原因

传输节点会在一定时间内尝试重新连接到出现故障的控制器，然后才放弃并连接到另一个控制器。整个过程大约需要 5 分钟。这是预期的行为。

NSX Manager 虚拟机降级

部署在 KVM 主机上的 NSX Manager 运行 `get service` 和 `get interface` 等 CLI 命令时返回错误。

问题

CLI 命令 get service 返回错误。例如，

```
nsx-manager-1> get service
% An error occurred while processing the service command
```

其他 CLI 命令也可能返回错误。`get support-bundle` 命令指示 `/tmp` 目录已变为只读。例如，

```
nsx-manager-1> get support-bundle file failed-to-get-service.tgz
% An error occurred while retrieving the support bundle: [Errno 30] Read-only file system:
'/tmp/tmpHzXF1u'
```

`/var/log/messages-<timestamp>` 日志具有如下消息：

```
Nov 17 07:26:48 no kernel: NMI watchdog: BUG: soft lockup - CPU#5 stuck for 23s! [qemu-kvm:4386]
```

原因

NSX Manager 设备上的一个或多个文件系统已损坏。在 https://access.redhat.com/solutions/22621 中记录了一些可能的原因。

要解决此问题，可以修复损坏的文件系统或从备份执行还原。

解决方案

1 选项 1：修复损坏的文件系统。以下步骤专用于 KVM 主机上运行的 NSX Manager。

   a 运行 `virsh destroy` 命令以停止 NSX Manager 虚拟机。

   b 在 qcow2 映像上以写入模式运行 `virt-rescue` 命令。例如，

   ```
   virt-rescue --rw -a nsx-unified-appliance-2.0.0.0.0.6522097.phadniss-p0-DK-to-DGo-on-rhel-
   prod_nsx_manager_1.qcow2
   ```

   c 在 `virt-rescue` 命令提示符下运行 `e2fsck` 命令来修复 `tmp` 文件系统。例如，

   ```
   <rescue> e2fsck /dev/nsx/tmp
   ```

   d 如有必要，再次运行 `e2fsck /dev/nsx/tmp`，直到不再有错误。

   e 使用 `virsh start` 重新启动 NSX Manager。

2 选项 2：从备份执行还原。

   有关说明，请参见*《NSX-T 管理指南》*。

NSX 代理与 NSX Manager 通信时超时

在 ESXi 主机上具有许多传输节点和虚拟机的大型环境中，运行在 ESXi 主机上的 NSX 代理在与 NSX Manager 通信时可能会超时。

问题

某些操作失败，例如当虚拟机 vnic 尝试连接到逻辑交换机时。/var/run/log/nsx-opsagent.log 具有如下消息：

```
level="ERROR" errorCode="MPA41542"] [MP_AddVnicAttachment] RPC call [0e316296-13-14] to NSX management
plane timout
2017-05-15T05:32:13Z nsxa: [nsx@6876 comp="nsx-esx" subcomp="NSXA[VifHandlerThread:-2282640]"
tid="1000017079" level="ERROR" errorCode="MPA42003"] [DoMpVifAttachRpc] MP_AddVnicAttachment() failed:
RPC call to NSX management plane timout
```

原因

在大型环境中，某些操作可能会比平常花费更长的时间，并且可能会因超出默认超时值而失败。

解决方案

1 增加 NSX 代理超时值。

a 在 ESXi 主机上，使用以下命令停止 NSX opsAgent：

```
/etc/init.d/nsx-opsagent stop
```

b 编辑文件 /etc/vmware/nsx-opsagent/nsxa.json，并更改 vifOperationTimeout 的值，例如，从 25 更改为 55。

```
"mp" : {
    /* timeout for VIF operation */
    "vifOperationTimeout" : 25,
```

注意 此超时值必须小于您在步骤 2 中设置的 hostd 超时值。

c 使用以下命令启动 NSX opsAgent：

/etc/init.d/nsx-opsagent start

2 增加 hostd 超时值。

a 在 ESXi 主机上，使用以下命令停止 hostd 代理：

```
/etc/init.d/hostd stop
```

b 编辑文件 /etc/vmware/hostd/config.xml。在 <opaqueNetwork> 下，取消注释 <taskTimeout> 对应的条目，然后将值从 30 更改为 60（示例）。

```
<opaqueNetwork>
    <!-- maximum message size allowed in opaque network manager IPC, in bytes. -->
    <!-- <maxMsgSize> 65536 </maxMsgSize> -->
    <!-- maximum wait time for opaque network response -->
    <!-- <taskTimeout> 30 </taskTimeout> -->
```

c 使用以下命令启动 hostd 代理：

/etc/init.d/hostd start

无法添加 ESXi 主机

您不能将 ESXi 主机添加到 NSX-T Data Center 结构层。

问题

从 NSX Manager GUI 添加 ESXi 主机失败，并显示错误 ...的文件路径由多个非覆盖网络 VIB 声明 (File path of ... is claimed by multiple non-overlay VIBs)。日志文件显示如下消息：

```
Failed to install software on host. Failed to install software on host. 10.172.120.60 :
java.rmi.RemoteException: [DependencyError] File path of '/usr/lib/vmware/vmkmod/nsx-vsip' is claimed by
multiple non-overlay VIBs
```

原因

以前安装的一些 VIB 仍在主机上，可能是因为未彻底卸载。

解决方案

1 从错误消息中获取导致失败的 VIB 的名称。

2 使用 ESXi 命令卸载 VIB。

NSX Controller 状态不正确

NSX Controller 群集中的某些控制器对其中一个控制器报告不正确的状态。

问题

多次关闭和打开某个控制器的电源后，其他控制器报告它处于非活动状态，而它实际上已启动且正常运行。

原因

当关闭后再打开控制器的电源时，有时会出现涉及 ZooKeeper 模块的内部错误，它会导致该控制器与群集中的其他控制器之间出现通信故障。

解决方案

- 从群集中移除报告为非活动状态的控制器节点，从节点中移除群集配置，然后将节点重新加入到群集。有关详细信息，请参见《*NSX 管理指南*》中的“替换 NSX Controller 群集的成员”部分。

KVM 虚拟机上的管理 IP 在启用 IPFIX 的情况下无法访问

在 KVM 主机上的多个虚拟机上启用 IPFIX，同时采样率为 100% 时，某些虚拟机上的管理 IP 可能会间歇性地无法访问。

问题

当您为同一主机上的多个虚拟机启用 IPFIX 且将采样率设置为 100% 时，可能会有大量的 IPFIX 流量。这可能会影响管理流量，从而导致管理 IP 间歇性地无法访问，即使生产流量和管理流量通过不同的 OVS。

原因

主机和虚拟机的工作负载压力过大。

解决方案

- 通过减少启用 IPFIX 的虚拟机数量或降低采样率即可降低主机上的负载。