Contents

NSX-T Data Center Troubleshooting Guide  5

1 Logs and Services  6
   Log Messages  6
   Troubleshooting Syslog Issues  10
   Checking Services  11
   Collecting Support Bundles  13

2 Troubleshooting Layer 2 Connectivity  14
   Check the NSX Manager and NSX Controller Cluster Status  14
   Check the Logical Ports  15
   Check the Transport Node Status  16
   Check the Logical Switch Status  16
   Check the CCP for the Logical Switch  17
   Check the Local Control Plane Status  17
   Troubleshoot Config Session Issues  18
   Troubleshoot L2 Session Issues  19
   Troubleshoot Dataplane Issues for an Overlay logical Switch  20
   Troubleshoot Dataplane Issues for a VLAN logical Switch  21
   Troubleshoot ARP Issues for an Overlay Logical Switch  22
   Troubleshoot Packet Loss for a VLAN logical Switch or When ARP IsResolved  22

3 Troubleshooting Installation  24

4 Troubleshooting Routing  28

5 Troubleshooting the Central Control Plane  30
   Deleting a Controller Node Fails  30
   Removing a Controller from a Cluster Fails  30
   Controller VM Fails to Power On  31
   Controller Fails to Register with NSX Manager  31
   Controller Fails to Come Online after Registration  32
   Controller Fails to Join the Cluster  32
   Controller Clustering Fails  32

6 Troubleshooting Firewall  34
   Determining Firewall Rules that Apply on an ESXi Host  34
   Determining Firewall Rules that Apply on a KVM Host  37
Firewall Packet Logs  38
On an ESXi Host, the get rules Command Shows an Unknown MAC Address  39
Stateful Edge Firewall Not Working  40

7  Other Troubleshooting Scenarios  41
Failure to Add or Delete a Transport Node  41
Transport Node Takes About 5 Minutes to Connect to Another Controller  42
NSX Manager VM Is Degraded  43
NSX Agent Times Out Communicating with NSX Manager  44
Failure to Add an ESXi Host  45
Incorrect NSX Controller Status  45
Management IPs on KVM VMs Not Reachable with IPFIX Enabled  46
Upgrade Fails Due to a Timeout  46
Edge Transport Node Status Degraded if Any Interface is Down  47
NSX-T Data Center Troubleshooting Guide

The *NSX-T Data Center Troubleshooting Guide* provides information on how to troubleshoot issues that might occur in an NSX-T Data Center environment.

**Intended Audience**

This guide is for system administrators of NSX-T Data Center. A familiarity with virtualization, networking, and datacenter operations is assumed.

**VMware Technical Publications Glossary**

VMware Technical Publications provides a glossary of terms that might be unfamiliar to you. For definitions of terms as they are used in VMware technical documentation, go to [http://www.vmware.com/support/pubs](http://www.vmware.com/support/pubs).
Logs and Services

Logs can be helpful in many troubleshooting scenarios. Checking the status of services is also important.

This chapter includes the following topics:
- Log Messages
- Troubleshooting Syslog Issues
- Checking Services
- Collecting Support Bundles

Log Messages

Log messages from all NSX-T Data Center components, including those running on ESXi hosts, conform to the syslog format as specified in RFC 5424. Log messages from KVM hosts are in the RFC 3164 format. The log files are in the directory `/var/log`.

On NSX-T Data Center appliances, you can run the following NSX-T Data Center CLI command to view the logs:

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

On hypervisors, you can use Linux commands such as `tac`, `tail`, `grep`, and `more` to view the logs. You can also use these commands on NSX-T Data Center appliances.


RFC 5424 defines the following format for log messages:

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

A sample log message:

```
```

Every message has the component (comp) and sub-component (subcomp) information to help identify the source of the message.
NSX-T Data Center produces regular logs (facility local6, which has a numerical value of 22) and audit logs (facility local7, which has a numerical value of 23). All API calls trigger an audit log.

An audit log that is associated with an API call has the following information:

- An entity ID parameter entId to identify the object of the API.
- A request ID parameter req-id to identify a specific API call.
- An external request ID parameter ereqId if the API call contains the header X-NSX-EREQID:<string>.
- An external user parameter euser if the API call contains the header X-NSX-EUSER:<string>.

RFC 5424 defines the following severity levels:

<table>
<thead>
<tr>
<th>Severity Level</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Emergency: system is unusable</td>
</tr>
<tr>
<td>1</td>
<td>Alert: action must be taken immediately</td>
</tr>
<tr>
<td>2</td>
<td>Critical: critical conditions</td>
</tr>
<tr>
<td>3</td>
<td>Error: error conditions</td>
</tr>
<tr>
<td>4</td>
<td>Warning: warning conditions</td>
</tr>
<tr>
<td>5</td>
<td>Notice: normal but significant condition</td>
</tr>
<tr>
<td>6</td>
<td>Informational: informational messages</td>
</tr>
<tr>
<td>7</td>
<td>Debug: debug-level messages</td>
</tr>
</tbody>
</table>

All logs with a severity of emergency, alert, critical, or error contain a unique error code in the structured data portion of the log message. The error code consists of a string and a decimal number. The string represents a specific module.

The MSGID field identifies the type of message. For a list of the message IDs, see Log Message IDs.

**Configure Remote Logging**

You can configure NSX-T Data Center appliances and hypervisors to send log messages to a remote logging server.

Remote logging is supported on NSX Manager, NSX Controller, NSX Edge, and hypervisors. You must configure remote logging on each node individually.

On an KVM host, the NSX-T Data Center installation package automatically configures the rsyslog daemon by putting configuration files in the /etc/rsyslog.d directory.

**Prerequisites**

- Configure a logging server to receive the logs.
Procedure

1 To configure remote logging on an NSX-T Data Center appliance:
   a Run the following command to configure a log server and the types of messages to send to the log server. Multiple facilities or message IDs can be specified as a comma delimited list, without spaces.

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

For more information about this command, see the NSX-T CLI Reference. You can run the command multiple times to add multiple logging server configurations. For example:

```
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 You can view the logging configuration with the `get logging-server` command. For example,

```
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 To configure remote logging on an ESXi host:
   a Run the following commands to configure syslog and send a test message:

```
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 You can run the following command to display the configuration:

```
esxcli system syslog config get
```

3 To configure remote logging on a KVM host:
   a Edit the file `/etc/rsyslog.d/10-vmware-remote-logging.conf` for your environment.

   b Add the following line to the file:

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

   c Run the following command:

   ```
   service rsyslog restart
   ```
Log Message IDs

In a log message, the message ID field identifies the type of message. You can use the `messageid` parameter in the `set logging-server` command to filter which log messages are sent to a logging server.

Table 1-1. Log Message IDs

<table>
<thead>
<tr>
<th>Message ID</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>FABRIC</td>
<td>Host node</td>
</tr>
<tr>
<td></td>
<td>Host preparation</td>
</tr>
<tr>
<td></td>
<td>Edge node</td>
</tr>
<tr>
<td></td>
<td>Transport zone</td>
</tr>
<tr>
<td></td>
<td>Transport node</td>
</tr>
<tr>
<td></td>
<td>Uplink profiles</td>
</tr>
<tr>
<td></td>
<td>Cluster profiles</td>
</tr>
<tr>
<td></td>
<td>Edge cluster</td>
</tr>
<tr>
<td></td>
<td>Bridge clusters and endpoints</td>
</tr>
<tr>
<td>SWITCHING</td>
<td>Logical switch</td>
</tr>
<tr>
<td></td>
<td>Logical switch ports</td>
</tr>
<tr>
<td></td>
<td>Switching profiles</td>
</tr>
<tr>
<td></td>
<td>switch security features</td>
</tr>
<tr>
<td>ROUTING</td>
<td>Logical router</td>
</tr>
<tr>
<td></td>
<td>Logical router ports</td>
</tr>
<tr>
<td></td>
<td>Static routing</td>
</tr>
<tr>
<td></td>
<td>Dynamic routing</td>
</tr>
<tr>
<td></td>
<td>NAT</td>
</tr>
<tr>
<td>FIREWALL</td>
<td>Firewall rules</td>
</tr>
<tr>
<td></td>
<td>Firewall rule sections</td>
</tr>
<tr>
<td>FIREWALL-PK TLOG</td>
<td>Firewall connection logs</td>
</tr>
<tr>
<td></td>
<td>Firewall packet logs</td>
</tr>
<tr>
<td>GROUPING</td>
<td>IP sets</td>
</tr>
<tr>
<td></td>
<td>Mac sets</td>
</tr>
<tr>
<td></td>
<td>NSGroups</td>
</tr>
<tr>
<td></td>
<td>NSServices</td>
</tr>
<tr>
<td></td>
<td>NSService groups</td>
</tr>
<tr>
<td></td>
<td>VNI Pool</td>
</tr>
<tr>
<td></td>
<td>IP Pool</td>
</tr>
<tr>
<td>DHCP</td>
<td>DHCP relay</td>
</tr>
</tbody>
</table>
Table 1.1. Log Message IDs (Continued)

<table>
<thead>
<tr>
<th>Message ID</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYSTEM</td>
<td>Appliance management (remote syslog, ntp, etc)</td>
</tr>
<tr>
<td></td>
<td>Cluster management</td>
</tr>
<tr>
<td></td>
<td>Trust management</td>
</tr>
<tr>
<td></td>
<td>Licensing</td>
</tr>
<tr>
<td></td>
<td>User and roles</td>
</tr>
<tr>
<td></td>
<td>Task management</td>
</tr>
<tr>
<td></td>
<td>Install (NSX Manager, NSX Controller)</td>
</tr>
<tr>
<td></td>
<td>Upgrade (NSX Manager, NSX Controller, NSX Edge and host-packages upgrades)</td>
</tr>
<tr>
<td></td>
<td>Realization</td>
</tr>
<tr>
<td></td>
<td>Tags</td>
</tr>
<tr>
<td>MONITORING</td>
<td>SNMP</td>
</tr>
<tr>
<td></td>
<td>Port connection</td>
</tr>
<tr>
<td></td>
<td>Traceflow</td>
</tr>
<tr>
<td></td>
<td>All other log messages.</td>
</tr>
</tbody>
</table>

Troubleshooting Syslog Issues

If logs are not received by the remote log server, perform the following steps.

- Verify the remote log server’s IP address.
- Verify that the level parameter is configured correctly.
- Verify that the facility parameter is configured correctly.
- If the protocol is TLS, set the protocol to UDP to see if there is a certificate mismatch.
- If the protocol is TLS, verify that port 6514 is open on both ends.
- Remove the message ID filter and see if logs are received by the server.
- Restart the rsyslog service with the command `restart service rsyslogd`.

A sample rsyslog configuration file (/etc/rsyslog.conf):

```bash
### 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
```
Checking Services

Services that stop running or fail to start can cause problems. It is important to make sure that all services are running normally.

To check the status of NSX Manager service:

```
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
```
In the example above, the http service is stopped. You can start the http service with the following command:

```
nsxmgr> start service http
```

**SSH Service**

If the SSH service was not enabled when deploying the appliance, you can log in to the appliance as admin and enable it with the following command:

```
start service ssh
```

You can configure SSH to start when the host starts with the following command:

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

To enable SSH root login, you can log in to the appliance as root, edit the file `/etc/ssh/sshd_config` and replace the line

```
PermitRootLogin prohibit-password
```

Alternatively, you can enable the SSH service and enable SSH root access by powering off the appliance and modifying its vApp properties. with

```
PermitRootLogin yes
```

and restart the sshd server with the following command:

```
/etc/init.d/ssh restart
```
Collecting Support Bundles

You can collect support bundles on registered cluster and fabric nodes and download the bundles to your machine or upload them to a file server.

If you choose to download the bundles to your machine, you get a single archive file consisting of a manifest file and support bundles for each node. If you choose to upload the bundles to a file server, the manifest file and the individual bundles are uploaded to the file server separately.

NSX Cloud Note  If you want to collect the support bundle for CSM, log in to CSM, go to System > Utilities > Support Bundle and click on Download. The support bundle for PCG is available from NSX Manager using the following instructions. The support bundle for PCG also contains logs for all the workload VMs.

Procedure

1. From your browser, log in with admin privileges to NSX Manager at https://nsx-manager-ip-address.
2. Select System > Utilities from the navigation panel.
3. Click the Support Bundle tab.
4. Select the target nodes.
   The available types of nodes are Management Nodes, Controller Nodes, Edges, Hosts, and Public Cloud Gateways.
5. (Optional) Specify log age in days to exclude logs that are older than the specified number of days.
6. (Optional) Toggle the switch that indicates whether to include or exclude core files and audit logs.
  _note_  Core files and audit logs might contain sensitive information such as passwords or encryption keys.
7. (Optional) Select a check box to upload the bundles to a file server.
8. Click Start Bundle Collection to start collecting support bundles.
   Depending on how many log files exist, each node might take several minutes.
9. Monitor the status of the collection process.
   The status field shows the percentage of nodes that completed support bundle collection.
10. Click Download to download the bundle if the option to send the bundle to a file server was not set.
Troubleshooting Layer 2 Connectivity

If there is a communication failure between two virtual interfaces (VIFs) that are connected to the same logical switch, for example, you cannot ping one VM from another, you can follow the steps in this section to troubleshoot the failure.

Before you start, make sure that there is no firewall rule blocking traffic between the two logical ports. It is recommended that you follow the order of the topics in this section to troubleshoot the connectivity issue.

This chapter includes the following topics:
- Check the NSX Manager and NSX Controller Cluster Status
- Check the Logical Ports
- Check the Transport Node Status
- Check the Logical Switch Status
- Check the CCP for the Logical Switch
- Check the Local Control Plane Status
- Troubleshoot Config Session Issues
- Troubleshoot L2 Session Issues
- Troubleshoot Dataplane Issues for an Overlay logical Switch
- Troubleshoot Dataplane Issues for a VLAN logical Switch
- Troubleshoot ARP Issues for an Overlay Logical Switch
- Troubleshoot Packet Loss for a VLAN logical Switch or When ARP Is Resolved

Check the NSX Manager and NSX Controller Cluster Status

Verify that the status of NSX Manager and the NSX Controller cluster is normal, and the controllers are connected to the NSX Manager.
Procedure

1 Run the following CLI command on the NSX Manager to make sure the status is stable.

```bash
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 Run the following CLI command on an NSX Controller to make sure the status is active.

```bash
NSX-Controller1> get control-cluster status
uuid: db4aa77a-4397-4d65-ad33-9fde79ac3c5c
is master: true
in majority: true

uuid                     address        status
0cfe232e-6c28-4f0a-8aa4-b3518baef0d 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 Run the following CLI command on an NSX Controller to make sure it is connected to the NSX Manager.

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

Check the Logical Ports

Check that the logical ports are configured on the same logical switch and their status is up.

Procedure

1 From the NSX Manager GUI, get the logical ports UUIDs.

2 Make the following API call for each logical port to make sure the logical ports are on the same logical switch.

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

3 Make the following API call for each logical port to make sure the status is up.

```bash
GET https://<nsx-mgr>/api/v1/logical-ports/<logical-port-uuid>/status
```
Check the Transport Node Status

Check the status of the transport node.

Procedure

- Make the following API call to get the state of the transport node.

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

  If the call returns the error RPC timeout, perform the following troubleshooting steps:
  - Run `/etc/init.d/nsx-opsAgent status` to see if opsAgent is running.
  - Run `/etc/init.d/nsx-mpa status` to see if nsx-mpa is running.
  - To see if nsx-mpa is connected to the NSX Manager, check the nsx-mpa heartbeat logs.
  - To see if opsAgent is connected to the NSX Manager, check the nsx-opsAgent log. You will see the following message if opsAgent is connected to the NSX Manager.
    
    Connected to mpa, cookie: ...
  - To see if opsAgent is stuck processing HostConfigMsg, check the nsx-opsAgent log. If so, you will see an RMQ request message but the reply is not sent or sent after a long delay.
  - Check to see if opsAgent crashed while executing HostConfigMsg.
  - To see if the RMQ messages are taking a long time to be delivered to the host, compare the timestamps of log messages on the NSX Manager and the host.

  If the call returns the error partial_success, there are many possible causes. Start by looking at the nsx-opsAgent logs. On the ESXi host, check hostd.log and vmkernel.log. On KVM, syslog holds all the logs.

Check the Logical Switch Status

Check the status of the logical switch.

Procedure

- Make the following API call to get the state of the logical switch.

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

  If the call returns the error partial_success, the reply will contain a list of transport nodes where the NSX Manager failed to push the logical switch configuration or did not get a reply. The troubleshooting steps are similar to those for the transport node. Check the following:
  - All required components are installed and running.
  - nsx-mpa is connected to the NSX Manager.
- nsxa is connected to the switching vertical.
- Grep the logical switch ID in nsxa.log and nsxaVim.log to see if the logical switch configuration was received by the transport node.
- Check the nsxa and nsx-mpa uptime. Find out when nsxa was started and stopped by grepping nsxa log messages in the syslog file.
- Find out nsxa's connection time to the switching vertical. If the logical switch configuration is sent to the host when nsxa is not connected to the switching vertical, the configuration might not be delivered to the host.

On KVM, no logical switch configuration is pushed to the host. Therefore, most of the logical switch issues are likely to be in the management plane.

On ESXi, an opaque network is mapped to the logical switch. To use the logical switch, users connect VMs to the opaque network using vCenter Server or vSphere API.

Check the CCP for the Logical Switch

Verify that the logical switch is in the central control plane (CCP).

Procedure

- Run the following CLI command on an NSX Controller to make sure that the logical switch is present.

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

  **Note**  This CLI command does not list VLAN-backed logical switches.

Check the Local Control Plane Status

For an overlay logical switch, check that the net:cpa on the host is connected to the central control plane.

Prerequisites

Find the controller that the logical switch is on. See Check the CCP for the Logical Switch.

Procedure

1. SSH to the controller that the logical switch is on.
2. Run the following command and verify that the controller shows the hypervisors that are connected to this VNI.

   ```
   get logical-switch 5000 transport-node-table
   ```

   On the hypervisors, run the command /bin/nsxcli to start NSX CLI.
4 Run the following command to get the CCP sessions.

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

You should see a Config session on one of the CCP nodes in the CCP cluster. For every overlay logical switch, you should see an L2 session to one of the CCP nodes in the CCP cluster. For VLAN logical switches, there are no CCP connections.

## Troubleshoot Config Session Issues

If the CCP config session is not up, check the status of MPA and `netcpa`.

### Procedure

1. Make the following API call to see if MPA is connected to the NSX Manager.

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

2. On the hypervisor, run the command `/bin/nsxcli` to start NSX CLI.

3. Run the following command to get the node-uuid.

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

4. Run the following command to see if the NSX Manager pushed the CCP information to the host.

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

5. If `config-by-vsm.xml` has CCP information, check if a transport node is configured on the hypervisor.

   The NSX Manager sends the host certificate for the hypervisor in the transport node creation step. The CCP must have the host certificate before it accepts connections from the host.

6. Check the validity of the host certificate in `/etc/vmware/nsx/host-cert.pem`.

   The certificate must be the same as the one that the NSX Manager has for the host.

7. Run the following command to check if the status of `netcpa`.

   On ESXi:

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

   On KVM:

   ```bash
   /etc/init.d/nsx-agent status
   ```
8  Start or restart netcpa.

   On ESXi, start netcpa if it is not running, or restart it if it is running.

   /etc/init.d/netcpad start

   /etc/init.d/netcpad restart

   On KVM, start netcpa if it is not running, or restart it if it is running.

   /etc/init.d/nsx-agent start

   /etc/init.d/nsx-agent restart

9  If the config session is still not up, collect the technical support bundles and contact VMware support.

Troubleshoot L2 Session Issues

This applies to overlay logical switches only.

Procedure

1  On the hypervisor, run the command /bin/nsxcli to start NSX CLI.

2  Run the following command to see if the logical switch is present on the host.

   host1> get logical-switches

3  Check that the state of the port is not admin down.

   On ESXi, run net-dvs and look at the response. For example,

   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.

   If the logical port ends up in the blocked state, collect the technical support bundles and contact VMware support. In the meantime, run the following command to get the DVS name:

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

   Run the following command to unblock the port:

   [root@host1:~] net-dvs -s com.vmware.common.port.block=false <DVS-NAME> -p <logical-port-ID>
On KVM, run `ovs-vsctl list interface` and verify that the interface with the corresponding VIF UUID is present and `admin_state` is up. You can see the VIF UUID in OVSDB in `external-ids:iface-id`.

**Troubleshoot Dataplane Issues for an Overlay logical Switch**

The steps in this section are for troubleshooting connectivity issues between VMs on different hypervisors through the overlay switch when the config and runtime states are normal.

If the VMs are on the same hypervisor, go to *Troubleshoot ARP Issues for an Overlay Logical Switch*.

**Procedure**

1. Run the following command on the controller that has the logical switch to see if CCP has the correct VTEP list:

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

2. On each hypervisor, run the following NSX CLI command to see if it has the correct VTEP list:

   On ESXi:
   ```
   host1> get logical-switch <logical-switch-UUID> tep-table
   ```
   Alternatively, you can run the following shell command for the VTEP information:
   ```
   [root@host1:~] net-vdl2 -M vtep -s vds -n VNI
   ```
   On KVM:
   ```
   host1> get logical-switch <logical-switch-UUID or VNI> tep-table
   ```

3. Check to see if the VTEPs on the hypervisors can ping each other.

   At the ESXi shell prompt:
   ```
   host1> ping ++netstack=vxlan <remote-VTEP-IP>
   ```

   At the KVM shell prompt:
   ```
   host1> ping <remote-VTEP-IP>
   ```

   If the VTEPs cannot ping each other,
   a. Make sure the transport VLAN specified when creating the transport node matches what the underlay expects. If you are using access ports in the underlay, the transport VLAN should be set to 0. If you are specifying a transport VLAN, the underlay switch ports that the hypervisors connect to should be configured to accept this VLAN in trunk mode.
b Check underlay connectivity.

4 Check if the BFD sessions between the VTEPs are up.

On ESXi, run `net-vdl2 -M bfd` and look at the response. For example,

```
BFD count: 1
===========================
```

On KVM, find the GENEVE interface to the remote IP.

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

If you don't know the interface name, run `ovs-vsctl find Interface type=geneve` to return all tunnel interfaces. Look for BFD information.

If you cannot find an GENEVE interface to remote VTEP, check if nsx-agent is running and OVS integration bridge is connected to 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
```

**Troubleshoot Dataplane Issues for a VLAN logical Switch**

The steps in this section are for troubleshooting connectivity issues between VMs on different hypervisors through the configured VLAN on the underlay when the config and runtime states are normal.

If the VMs are on the same hypervisor and all the configuration and runtime states are normal, go to *Troubleshoot ARP Issues for an Overlay Logical Switch.*

**Procedure**

- Check that the underlay is configured for the VLAN for the logical switch in trunk mode.

  On ESXi, verify VLAN is configured on the logical port by running `net-dvs` and looking for the logical port. For example:

  ```
  port 63eadf53-ff92-4a0e-9496-4200e99709ff:
    com.vmware.common.port.volatile.vlan = VLAN 1000 propType = RUNTIME VOLATILE
  ```
On KVM, the VLAN logical switch is configured as an openflow rule on integration bridge. In other words, for traffic received from the VIF, tag it with VLAN X and forward it on the patch port to the PIF bridge. Run `ovs-vsctl list interface` and verify the presence of the patch port between the NSX-managed bridge and the NSX-switch bridge.

### Troubleshoot ARP Issues for an Overlay Logical Switch

The steps in this section are for troubleshooting where packets are being lost for an overlay switch.

For a VLAN-backed logical switch, go to Troubleshoot Packet Loss for a VLAN logical Switch or When ARP Is Resolved.

Before performing the following troubleshooting steps, run the command `arp -n` on each VM. If ARP is successfully resolved on both VMs, you do not need to perform the steps in this section. Instead, go to the next section Troubleshoot Packet Loss for a VLAN logical Switch or When ARP Is Resolved.

**Procedure**

- If both endpoints are ESXi and ARP proxy is enabled on the logical switch (only supported for overlay logical switches), check the ARP table on the CCP and the hypervisor.

  On the CCP:

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

  On the hypervisor, start NSX CLI and run the following command:

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

  Fetching the ARP table only tells us whether we have the correct ARP proxy state. If the ARP response is not received via proxy, or if the host is KVM and does not support ARP proxy, the datapath should broadcast the ARP request. There might be a problem with BUM traffic forwarding. Try the following steps:

  - If the replication mode for the logical switch is MTEP, change the replication mode to SOURCE for the logical switch from the NSX Manager GUI. This might fix the issue and ping will start working.
  - Add static ARP entries and see if the rest of the datapath works.

### Troubleshoot Packet Loss for a VLAN logical Switch or When ARP Is Resolved

You can use the automated traceflow tool or manually trace the packets to troubleshoot packet loss.

To run the traceflow tool, from the NSX Manager GUI, navigate to **Tools > Traceflow**. For more information, see the *NSX-T Administration Guide*. 
Procedure

- To manually trace the packets,

  On ESXi, run `net-stats -l` to get the switchport ID of the VIFs. If the source and destination VIFs are on the same hypervisor, run the following commands:

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

  If the source and destination VIFs are on different hypervisors, on the hypervisor hosting the source VIF, run the following commands:

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

  On the hypervisor hosting the destination VIF, run the following commands:

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

  On KVM, if the source and destination VIFs are on the same hypervisor, run the following command:

  ```
  ovs-dpctl dump-flows
  ```
Troubleshooting Installation

This section provides information about troubleshooting installation issues.

**Basic Infrastructure Services**

The following services must be running on the appliances and hypervisors, also on vCenter Server if it is used as a compute manager.

- NTP
- DNS

Make sure that firewall is not blocking traffic between NSX-T components and hypervisors. Make sure that the required ports are open between the components.

To flush the DNS cache on the NSX Manager, SSH as root to the manager and run the following command:

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

You can then check the DNS configuration file.

```
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
```
Checking Communication from Host to Controller and Manager

On an ESXi host using NSX-T CLI commands:

```
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
```

On a KVM host using NSX-T CLI commands:

```
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
```

On an ESXi host using host CLI commands:

```
[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
newreno       mpa
newreno       mpa
```

On a KVM host using host CLI commands:

```
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# 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
```

```
tcpdump -i ens32 port 1235 | grep kvm-01
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
```
Host Registration Failure

If NSX-T uses the wrong IP address, host registration will fail. This can happen when a host has multiple IP addresses. Trying to delete the transport node leaves it in the Orphaned state. To resolve the issue:

- Go to Fabric > Nodes > Hosts, edit the host and remove all IP addresses except the management one.
- Click on the errors and select Resolve.

KVM Host Issues

KVM host issues are sometimes caused by not enough disk space. The /boot directory can fill up quickly and cause errors such as:

- Failed to install software on host
- No space left on device

You can run the command `df -h` to check available storage. If the /boot directory is at 100%, you can do the following:

- Run `sudo dpkg --list 'linux-image*' | grep ^ii` to see all the kernels installed.
- Run `uname -r` to see your currently running kernel. Do not remove this kernel (linux-image).
- Use `apt-get purge` to remove images you don't need anymore. For example, run `sudo apt-get purge linux-image-3.13.0-32-generic linux-image-3.13.0-33-generic`.
- Reboot the host.
- In NSX Manager, check the errors and select Resolve.
- Make sure the VMs are powered on.
Configuration Error when Deploying an Edge VM

After deploying an Edge VM, NSX Manager shows the VM's status as configuration error. The manager log has a message similar to the following:

```
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'].
```

Restarting the edge datapath service and then the VM should resolve the issue.

Force Removing a Transport Node

You can remove a transport node that is stuck in the Orphaned state by making the following API call:

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

NSX Manager will not do any validations as to whether you have any active VMs running on the host. You are responsible for deleting the N-VDS and VIBs. If you have the node added through Compute Manager, delete the Compute Manager first and then delete the node. The transport node will be deleted as well.
**Troubleshooting Routing**

NSX-T has built-in tools to troubleshoot routing issues.

### Traceflow

You can use Traceflow to inspect the flow of packets. You can see delivered, dropped, received, and forwarded packets. If a packet is dropped, a reason is displayed. For example, a packet can be dropped because of a firewall rule.

### Checking Routing Tables

To see the routing table on a service router, run the following commands:

```shell
edge01> get logical-router
Logical Route

<table>
<thead>
<tr>
<th>UUID</th>
<th>VRF</th>
<th>LR-ID</th>
<th>Name</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ports</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>736a80e3-23f6-5a2d-81d6-bbebf278666</td>
<td>0</td>
<td>0</td>
<td>TUNNEL</td>
<td>3</td>
</tr>
<tr>
<td>c9393d8c-1fcf-4c34-889d-2da1eeee25b8</td>
<td>1</td>
<td>10</td>
<td>SR-t0-router</td>
<td>SERVICE_ROUTER_TIER0</td>
</tr>
<tr>
<td>9333c94e-5938-46b4-8c7d-5e6a0c2b87b5</td>
<td>2</td>
<td>8</td>
<td>DR-t1-router01</td>
<td>DISTRIBUTED_ROUTER_TIER1</td>
</tr>
<tr>
<td>c91eb7c5-0297-4fed-9c22-b96df1c9b80f</td>
<td>3</td>
<td>9</td>
<td>DR-t0-router</td>
<td>DISTRIBUTED_ROUTER_TIER0</td>
</tr>
</tbody>
</table>

dedge01> vrf 1
dedge01(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

<table>
<thead>
<tr>
<th>Prefix</th>
<th>Flags</th>
<th>Metric</th>
<th>Next Hop</th>
<th>Interface</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.10.20.0/24</td>
<td>[20/0]</td>
<td>via 192.168.140.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.10.30.0/24</td>
<td>[20/0]</td>
<td>via 192.168.140.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.20.20.0/24</td>
<td>[20/0]</td>
<td>via 192.168.140.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>10.20.30.0/24</td>
<td>[20/0]</td>
<td>via 192.168.140.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>30.0.0.0/8</td>
<td>[20/0]</td>
<td>via 192.168.140.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100.64.80.0/31</td>
<td>[0/0]</td>
<td>via 169.254.0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100.64.80.2/31</td>
<td>[0/0]</td>
<td>via 169.254.0.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>100.64.80.4/31</td>
<td>[0/0]</td>
<td>via 169.254.0.1</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
<TRUNCATED OUTPUT>
To get the IP address of interfaces, run the following command:

```bash
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     : 169.254.0.28
  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
```

**Advertising T1 Routes**

You must advertise T1 routes so that they are visible on T0 router and upwards. There are different types of routes that you can advertise: NSX Connected, NAT, Static, LB VIP, and LB SNAT.
Troubleshooting the Central Control Plane

This section contains information about troubleshooting central control plane (CCP) issues.

This chapter includes the following topics:

- Deleting a Controller Node Fails
- Removing a Controller from a Cluster Fails
- Controller VM Fails to Power On
- Controller Fails to Register with NSX Manager
- Controller Fails to Come Online after Registration
- Controller Fails to Join the Cluster
- Controller Clustering Fails

Deleting a Controller Node Fails

When deleting a controller through the API (POST /api/v1/cluster/nodes/deployments/<node-id>?action=delete) or the NSX Manager UI, you get the error The deletion of cluster node VM has failed.

Cause

- The API call has invalid input, for example, an incorrect node ID.
- There is a connectivity problem between NSX Manager and vCenter Server.

Solution

- Make sure the API parameters are correct.
- Call POST /api/v1/cluster/nodes/deployments/<node-id>?action=delete with the force option to delete the controller. Then manually delete the VM from vCenter Server.

Removing a Controller from a Cluster Fails

When deleting a controller through the API (POST /api/v1/cluster/nodes/deployments/<node-id>?action=delete) or the NSX Manager UI, you get the error Failed to remove the cluster node VM from the cluster.
Cause
- the controller nodes become temporarily unavailable during the process of removing a controller from its cluster.
- The status of the controller cluster is not CONNECTED.

Solution
- From the NSX Manager UI, verify that the status of the controller cluster is healthy. If the controller to be deleted is healthy and the controller cluster is healthy, delete the controller again.
- If either the controller to be deleted or the controller cluster is not healthy, take steps to make sure that both are healthy. Then delete the controller again.

Controller VM Fails to Power On

When deploying a controller through the API (`POST /api/v1/cluster/nodes/deployments`) or the NSX Manager UI, you get the error `The power on of the cluster node VM has failed`.

Cause
- The host where the controller VM is running on might not have enough memory to power on the VM.

Solution
- Log in to vCenter Server and investigate why the controller VM does not power on. If there is insufficient memory, delete the VM, free up memory on the host and redeploy the controller VM. Alternatively, redeploy the VM on a different host.

Controller Fails to Register with NSX Manager

When deploying a controller through the API (`POST /api/v1/cluster/nodes/deployments`) or the NSX Manager UI, the deployment remains in the `Waiting To Register` state indefinitely. The log has the error message `The cluster node VM failed to register itself with the MP within the allotted time`.

Cause
- The controller cannot communicate with the NSX Manager.
- The controller has an internal error.

Solution
- If possible, from the NSX Manager UI, delete the controller.
- If there is no option in the NSX Manager UI to delete the controller, call the API `POST /api/v1/cluster/nodes/deployments/<node-id>?action=delete` to delete the controller.
- Check network connectivity between the controller and the NSX Manager, such as IP address, subset and firewall settings.
- Redeploy the controller.
Controller Fails to Come Online after Registration

When deploying a controller through the API (POST /api/v1/cluster/nodes/deployments) or the NSX Manager UI, you get the error The cluster node VM failed to come online after registration.

Cause
- Internal controller or NSX Manager error.

Solution
- Delete the controller and redeploy.

Controller Fails to Join the Cluster

When deploying a controller through the API (POST /api/v1/cluster/nodes/deployments) or the NSX Manager UI, you get the error Failed to add the cluster node VM to a cluster.

Cause
- The controller cluster becomes unstable or unreachable during the clustering operation.
- The shared secret provided to the controller does not match the shared secret used by the controller cluster.

Solution
- Delete the controller.
- From the NSX Manager UI, check the cluster status of the nodes in the controller cluster.
- Check that the shared secret used by the new controller is the same as the shared secret used by the controller cluster.
- Redeploy the controller.

Controller Clustering Fails

When deploying a controller through the API (POST /api/v1/cluster/nodes/deployments) or the NSX Manager UI, you get the error Failed to add the cluster node VM to a cluster – No stable node or cluster found to join.

Cause
- The controller cluster state is unstable during the clustering operation.

Solution
- Delete the controller.
- From the NSX Manager UI, check that the cluster status of the nodes in the controller cluster is Up.
- Redeploy the controller.
Troubleshooting Firewall

This section provides information about troubleshooting firewall issues.

This chapter includes the following topics:

- Determining Firewall Rules that Apply on an ESXi Host
- Determining Firewall Rules that Apply on a KVM Host
- Firewall Packet Logs
- On an ESXi Host, the getrules Command Shows an Unknown MAC Address
- Stateful Edge Firewall Not Working

Determining Firewall Rules that Apply on an ESXi Host

To troubleshoot firewall issues with an ESXi host, you can look at the firewall rules that apply on the host.

Get the list of dvfilters on the ESXi host:

```
[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
```
Find a dvfilter for a specific VM:

```
[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
```

Determine firewall rules that apply to a specific dvfilter (in this example, nic-70227-eth0-vmware-sfw.2 is the dvfilter name):

```
[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;
}
```

Get the list of address sets used in a specific dvfilter:

```
[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 {
```
Check the flows through a specific dvfilter:

```
[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
a5d914f705b85fe5 Active tcp 0800 IN 3076 0 0 192.168.110.10:Unknown(51281) -> 172.16.10.11:ssh(22)  
      FINWAIT2:FINWAIT2 4304 5177 34 33
a5d914f705b86001 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
a5d914f705b86006 Active igmp 0800 IN 2 0 0 0.0.0.0 -> 224.0.0.1 36 0 1 0
a5d914f705b86011 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
a5d914f705b86012 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
a5d914f705b86013 Active udp 0800 OUT 2 0 0 172.16.10.11:Unknown(40080) -> 192.168.110.10:domain(53)  
      268 140 2 2
a5d914f705b86014 Active udp 0800 OUT 2 0 0 172.16.10.11:Unknown(59251) -> 192.168.110.10:domain(53)  
      268 140 2 2
a5d914f705b86015 Active ipv6-icmp 86dd OUT 2 0 0 fe80::250:6ff:feb5:a60e -> ff02::1:ff62:5ed4 135 0  
      72 0 1
a5d914f705b86016 Active ipv6-icmp 86dd OUT 2 0 0 fe80::250:6ff:feb5:a60e -> ff02::1:ff62:5ed4 135 0  
      72 0 1
a5d914f705b86017 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
a5d914f705b86018 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
a5d914f705b86019 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
a5d914f705b8601a 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
a5d914f705b8601b 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
a5d914f705b8601c 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
a5d914f705b8601d 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
a5d914f705b8601e 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
a5d914f705b8601f 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
```
Determining Firewall Rules that Apply on a KVM Host

To troubleshoot firewall issues with a KVM host, you can look at the firewall rules that apply on the host.

Get the list of VIFs that are subject to firewall rules on the KVM host:

```
# 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
```

If the output is empty, look for connectivity issues between the node and the controllers.

Get the list of rules applied to a specific VIF (in this example, `da95fc1e-65fd-461f-814d-d92970029bf0` is the 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 4882ec3-2670-497b-82f9-524618c16877 port 443 accept with log;
    rule 3072 inout protocol tcp from any to addrset 4882ec3-2670-497b-82f9-524618c16877 port 80 accept with log;
    rule 3074 inout protocol tcp from addrset 4882ec3-2670-497b-82f9-524618c16877 to addrset 8b9e75e7-bc62-4d7f-9a58-a872f393448e port 8443 accept with log;
    rule 3074 inout protocol tcp from addrset 4882ec3-2670-497b-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 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 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 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 8b9e75e7-bc62-4d7f-9a58-a872f393448e port 8443 accept with log;
    rule 3076 inout protocol icmp type 8 code 0 from 192.168.110.10 to addrset 8b9e75e7-bc62-4d7f-9a58-a872f393448e port 3306 accept with log;
}
```
rule 3076 inout protocol tcp from 192.168.110.10 to addrset 8b9e75e7-bc62-4d7f-9a58-a872f393448e port 80 accepts 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;
}

Get the list of address sets used in a specific VIF:

```
# 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 {
}
```

Check connections through the Linux Conntrack module. In this example, we look for flows between two specific IP addresses.

```
# 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
```

Firewall Packet Logs

If logging is enabled for firewall rules, you can look at the firewall packet logs to troubleshoot issues.
The log file is /var/log/dfwpktlogs.log for both ESXi and KVM hosts.

```bash
# 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 3072 IN TCP RST 100.64.80.1/60692->172.16.10.11/80 9/6 413/5411
2018-03-27T10:23:37.442Z  INET TERM 3074 OUT TCP RST 172.16.10.11/46109->172.16.20.11/8443 9/7 1218/7262
```

On an ESXi Host, the `getrules` Command Shows an Unknown MAC Address

On an ESXi host, after configuring a layer-2 firewall rule with one MAC set as source and another MAC set as destination, the `getrules` command on the host shows the destination MAC set having an unknown address.

**Problem**

After configuring a layer-2 firewall rule with one MAC set as source and another MAC set as destination, the `getrules` command on the host shows the destination MAC set as

```
01:00:00:00:00:00/01:00:00:00:00:00.
```

For example,

```
[root@host1:~] vsipioctl getrules -f nic-1000052822-eth1-vmware-sfw.2
ruleset mainrs {
   # generation number: 0
   # realization time : 2018-07-26T12:42:28
   rule 1039 at 1 inout protocol tcp from any to any port 1521 accept as oracle;
   # internal # rule 1039 at 2 inout protocol tcp from any to any port 1521 accept;
   rule 1039 at 3 inout protocol icmp from any to any accept;
   rule 2 at 4 inout protocol any from any to any accept with log;
}
```

```bash
ruleset mainrs_L2 {
   # generation number: 0
   # realization time : 2018-07-26T12:42:28
   rule 1039 at 1 inout protocol tcp from any to any port 1521 accept as oracle;
   # internal # rule 1039 at 2 inout protocol tcp from any to any port 1521 accept;
   rule 1039 at 3 inout protocol icmp from any to any accept;
   rule 2 at 4 inout protocol any from any to any accept with log;
}
```

VMware, Inc.
The internal OUT rule with the address 01:00:00:00:00:00/01:00:00:00:00:00 is created by design to handle outbound broadcasting packets and does not indicate a problem.

**Solution**

None required. The firewall rule will work as configured.

**Stateful Edge Firewall Not Working**

Stateful Edge firewall does not work between Tier-0 uplinks.

**Problem**

A loopback or hairpin is created when a Tier-0 router has multiple uplinks, and traffic ingresses on one of the uplinks and egresses on another uplink. When this occurs, firewall rules and NAT are only processed while the packet ingresses on the original uplink. This causes the reply returning on the second uplink to not match the original session, and the packet may be dropped.

**Cause**

Services are processed once during the hairpinning process, and not on both interfaces. This causes the reply to be considered another flow, rather than part of the original flow, because the direction of the packet for both the initial and the reply is IN.

**Solution**

- If no destination NAT rules are present on the SR, add one. A destination NAT rule will cause the reply to be tried to be matched against the original session, rather than being treated as a new session, and the packet will not be dropped. See Configure Source and Destination NAT on a Tier-0 Router in the *NSX-T Data Center Administration Guide*.  

```c
# realization time : 2018-07-26T12:42:28
rule 1040 at 1 inout ethertype any stateless from addrset d83a1523-0d07-4b18-8a5b-77a634540b57 to addrset 9ad9c6ef-c7dd-4682-833d-57097b415e41 accept;
# internal # rule 1040 at 2 in ethertype any stateless from addrset d83a1523-0d07-4b18-8a5b-77a634540b57 to addrset 9ad9c6ef-c7dd-4682-833d-57097b415e41 accept;
# internal # rule 1040 at 3 out ethertype any stateless from addrset d83a1523-0d07-4b18-8a5b-77a634540b57 to mac 01:00:00:00:00:00/01:00:00:00:00:00 accept;
rule 1 at 4 inout ethertype any stateless from any to any accept;
}
```
Other Troubleshooting Scenarios

This section describes how to troubleshoot various error scenarios.

This chapter includes the following topics:

- Failure to Add or Delete a Transport Node
- Transport Node Takes About 5 Minutes to Connect to Another Controller
- NSX Manager VM Is Degrade
- NSX Agent Times Out Communicating with NSX Manager
- Failure to Add an ESXi Host
- Incorrect NSX Controller Status
- Management IPs on KVM VMs Not Reachable with IPFIX Enabled
- Upgrade Fails Due to a Timeout
- Edge Transport Node Status Degraded if Any Interface is Down

Failure to Add or Delete a Transport Node

You cannot delete or add a transport node.

Problem

The error occurs in the following scenario:

1. An ESXi host is a fabric node and a transport node.
2. The host is removed as a transport node. However, transport node deletion fails. The state of the transport node is Orphaned.
3. The host is removed as a fabric node immediately.
4. The host is added as a fabric node again.
5. The host is added as a transport node with a new transport zone and switch. This step results in the error Failed/Partial Success.
Cause

In step 2, if you wait for a few minutes, the transport node deletion will succeed because NSX Manager will retry the deletion. When you delete the fabric node immediately, NSX Manager cannot retry because the host is removed from NSX-T Data Center. This results in incomplete cleanup of the host, with the switch configuration still present, which causes step 5 to fail.

Solution

1. Delete all vmknics from vCenter Server on the host that are connected to the NSX-T Data Center switch.

2. Get the switch name using the `esxcfg-vswitch -l` CLI command. For example:

   ```
   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. Delete the switch name using the `esxcfg-vswitch -d <switch-name> --dvswitch` CLI command. For example:

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

Transport Node Takes About 5 Minutes to Connect to Another Controller

When an ESXi transport node's connected controller goes down, it takes about 5 minutes for the transport node to connect to another controller.

Problem

An ESXi transport node is normally connected to a specific controller in a controller cluster. You can find the connected controller with the CLI command `get controllers`. If the connected controller goes down, it takes about 5 minutes for the transport node to be connected to another controller.

Cause

The transport node attempts to re-connect to the controller that is down for a certain amount of time before giving up and connecting to another controller. The whole process takes about 5 minutes. This is expected behavior.
NSX Manager VM Is Degraded

NSX Manager that is deployed on a KVM host returns an error when running CLI commands such as `get service` and `get interface`.

Problem

The CLI command `get service` returns an error. For example,

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

Other CLI commands might also return an error. The `get support-bundle` command indicates that the `/tmp` directory has become read-only. For example,

```
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'
```

The `/var/log/messages-<timestamp>` log has the a message such as the following:

```
```

Cause

One or more file systems on the NSX Manager appliance were corrupted. Some possible causes are documented in https://access.redhat.com/solutions/22621.

To resolve the issue, you can repair the corrupt file systems or perform a restore from a backup.

Solution

1. Option 1: Repair the corrupt file systems. The following steps are specifically for NSX Manager running on a KVM host.
   a. Run the `virsh destroy` command to stop the NSX Manager VM.
   b. Run the `virt-rescue` command in write mode on the qcow2 image. For example,
      ```
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. In the `virt-rescue` command prompt run the `e2fsck` command to fix the `tmp` file system. For example,
      ```
      <rescue> e2fsck /dev/nsx/tmp
      ```
   d. If necessary, run the `e2fsck /dev/nsx/tmp` again until there are no more errors.
   e. Restart NSX Manager with the `virsh start`.
Option 2: Perform a restore from a backup.

For instructions, see the NSX-T Administration Guide.

**NSX Agent Times Out Communicating with NSX Manager**

In a large-scale environment with many transport nodes and VMs on ESXi hosts, NSX agents, which run on ESXi hosts, might time out when communicating with NSX Manager.

**Problem**

Some operations, such as when a VM vnic tries to attach to a logical switch, fail. The /var/run/log/nsx-opsagent.log has messages such as:

```
level="ERROR" errorCode="MPA41542"] [MP_AddVnicAttachment] RPC call [0e316296-13-14] to NSX management plane timeout
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 timeout
```

**Cause**

In a large-scale environment, some operations might take longer than usual and fail because the default timeout values are exceeded.

**Solution**

1. Increase the NSX agent timeout value.
   a. On the ESXi host, stop the NSX opsAgent with the following command:

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

   b. Edit the file `/etc/vmware/nsx-opsagent/nsxa.json` and change the `vifOperationTimeout` value from 25 to, for example, 55.

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

   **Note** This timeout value must be less than the hostd timeout value that you set in step 2.

   c. Start the NSX opsAgent with the following command:

   `/etc/init.d/nsx-opsagent start`
2 Increase the hostd timeout value.
   a On the ESXi host, stop the hostd agent with the following command:

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

   b Edit the file `/etc/vmware/hostd/config.xml`. Under `<opaqueNetwork>`, uncomment the entry for `<taskTimeout>` and change the value from 30 to, for example, 60.

   ```xml
   <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> -->
   </opaqueNetwork>
   ```

   c Start the hostd agent with the following command:

   ```bash
   /etc/init.d/hostd start
   ```

### Failure to Add an ESXi Host

You are not able to add an ESXi host to the NSX-T Data Center fabric.

#### Problem

From the NSX Manager GUI, adding an ESXi hosts fails with the error `File path of ... is claimed by multiple non-overlay VIBs`. The log file shows messages such as the following:

```log
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
```

#### Cause

Some VIBs from a previous install are still on the host, probably because a clean uninstall did not occur.

#### Solution

1. From the error message, get the names of VIBs that are causing the failure.
2. Use ESXi commands to uninstall the VIBs.

### Incorrect NSX Controller Status

Some controllers in an NSX Controller cluster report incorrect status for one of the controllers.

#### Problem

After a controller is powered off and on a number of times, the other controllers report that it is inactive when it is up and running.
Cause
An internal error involving the ZooKeeper module sometimes occurs when a controller is powered off and on and causes a communication failure between this controller and the other controllers in the cluster.

Solution
◆ Remove the controller node that is reported to be inactive from the cluster, remove the cluster configuration from the node and rejoin the node to the cluster. For more information, see the section "Replace a Member of the NSX Controller Cluster" in the *NSX-T Administration Guide*.

Management IPs on KVM VMs Not Reachable with IPFIX Enabled
When IPFIX is enabled on multiple VMs on a KVM host and the sampling rate is 100%, the management IPs on some of the VMs might intermittently be unreachable.

Problem
When you enable IPFIX for multiple VMs on the same host and you set the sampling rate to be 100%, there can be a large amount of IPFIX traffic. This can impact management traffic, causing the management IPs to be intermittently unreachable, even if the production traffic and management traffic go through different OVSes.

Cause
The workload is too stressful for the host and the VMs.

Solution
◆ Reduce the load on the host by reducing the number of VMs with IPFIX enabled or reducing the sampling rate.

Upgrade Fails Due to a Timeout
An event during the upgrade process fails and the message from the Upgrade Coordinator indicates a timeout error.

Problem
During the upgrade process, the following events might fail because they do not complete within a specific period of time. The Upgrade Coordinator reports a timeout error for the event and the upgrade fails.

<table>
<thead>
<tr>
<th>Event</th>
<th>Timeout Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Putting a host into maintenance mode</td>
<td>4 hours</td>
</tr>
<tr>
<td>Waiting for a host to reboot</td>
<td>32 minutes</td>
</tr>
<tr>
<td>Waiting for the NSX service to be running</td>
<td>13 minutes</td>
</tr>
</tbody>
</table>
Solution

- For the maintenance mode issue, log in to vCenter Server and check the status of tasks related to the host. Take actions to resolve any issues.
- For the host reboot issue, check the host to see why it failed to reboot.
- For the NSX service issue, log in to the NSX Manager UI, go to the Fabric > Nodes page and see if the host has an installation error. If so, you can resolve it from the NSX Manager UI. If the error cannot be resolved, you can refer to the upgrade logs to determine the cause of the failure.

**Edge Transport Node Status Degraded if Any Interface is Down**

All interfaces are counted in an edge transport node status, and should be connected and Up.

**Problem**

An NSX Edge transport node has a status that is degraded, but the data plane is functioning normally. The Edge transport node degraded state also triggers the corresponding transport zone status to be degraded. The Edge transport node has one or more network interfaces in a Down state.

**Cause**

The NSX-T Data Center management plane declares an Edge transport node to be in a degraded state if any interface is down, regardless of whether that interface is used or configured. If the Edge is a virtual machine, the vNIC may be disconnected. If you have a bare metal Edge, the NIC port may not be connected, or may have a link state down.

If the interface that is Down is not used, then there is no functional impact to the Edge.

**Solution**

- To avoid a degraded state, ensure all Edge interfaces are connected and Up, regardless of whether they are in use or not.