NSX-T OPERATION DESIGN GUIDE

Release 3.2

Table of Contents

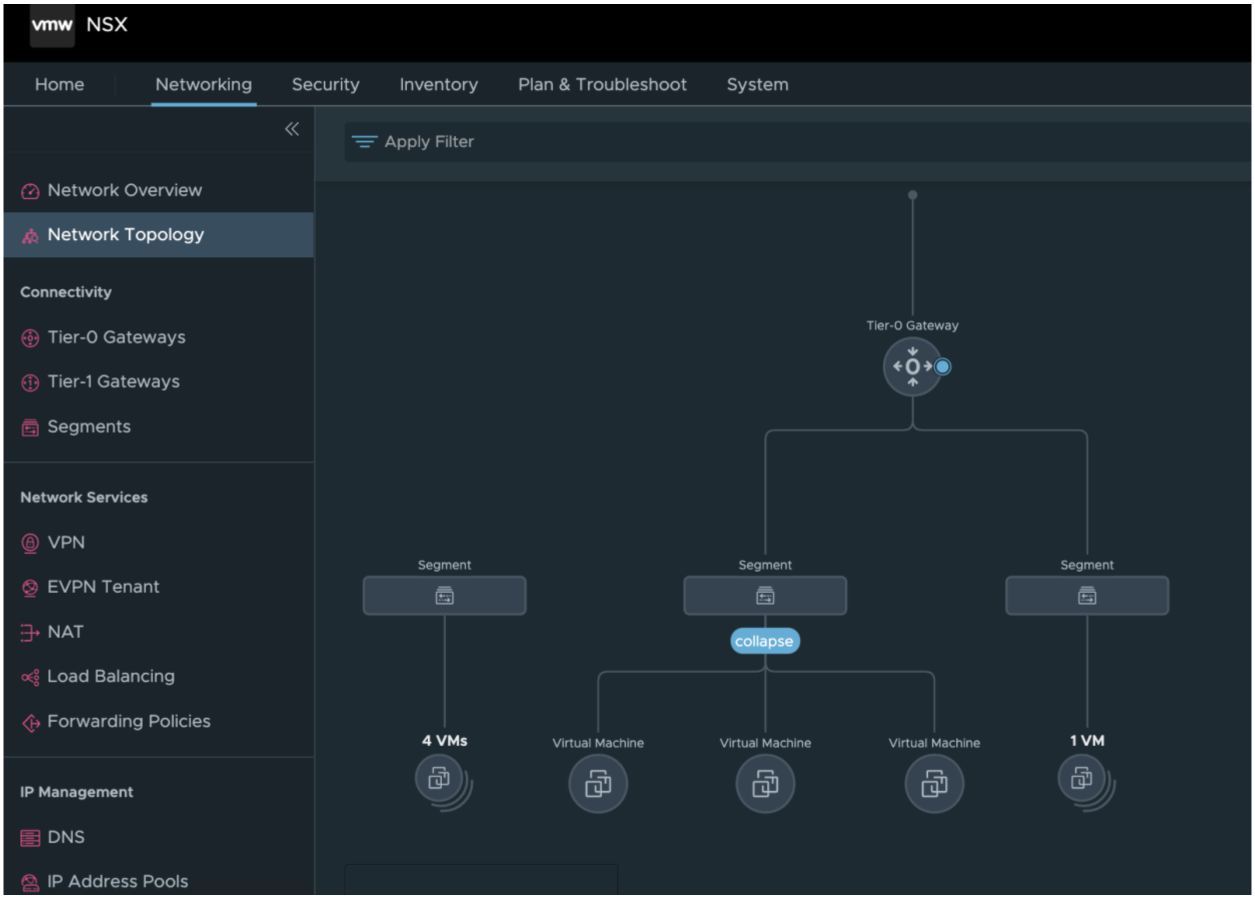
NSX 3.2 Operation Overview

We have made significant improvements to NSX Operation from release 3.0 to 3.2. In this version of the Operation Guide, we will only highlight the new capabilities available in the 3.2 release. A holistic version of 3.2 Operation Guide will be published later.

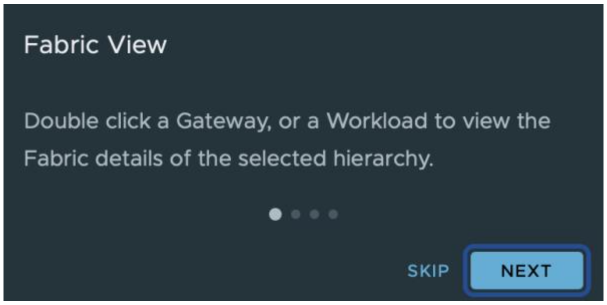
Troubleshooting Tools

Topology View

The network topology feature provides a graphical representation of the network topology which is very helpful when you are verifying your network configuration or troubleshooting errors.

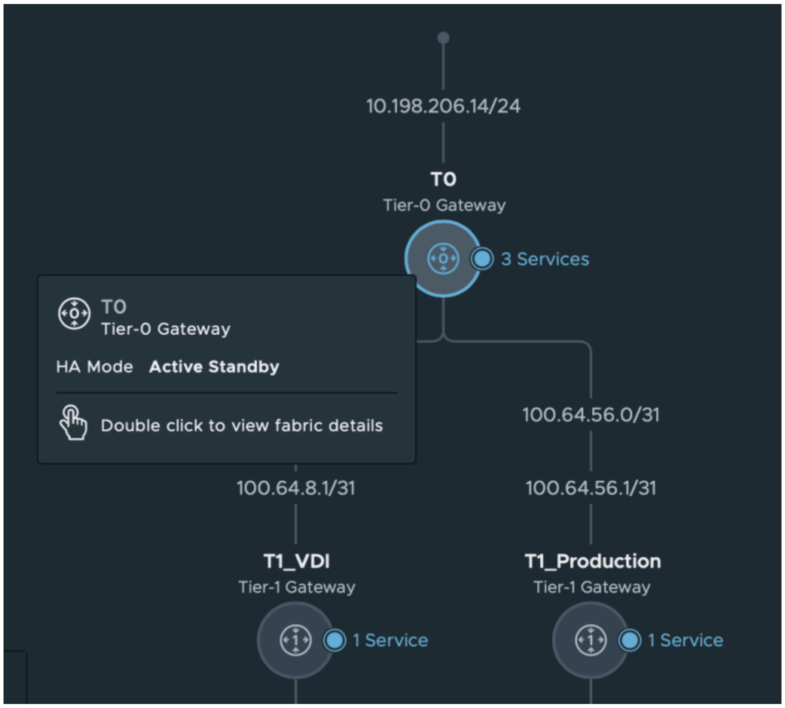


In 3.2, the Network Topology feature is enhanced with Fabric view which provides Fabric details of the NSX constructs and workload.


Fabric View

Double click a Gateway, or a Workload to view the Fabric details of the selected hierarchy.

For the Gateway, with Fabric view, we show the HA Mode, details of all services configured. If there's a VPN, we provide detail visualization of the VPN session.

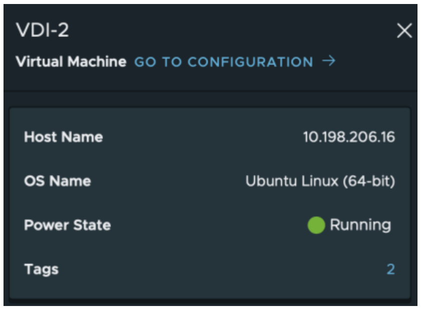


T0
Tier-0 Gateway

HA Mode **Active Standby**

Double click to view fabric details

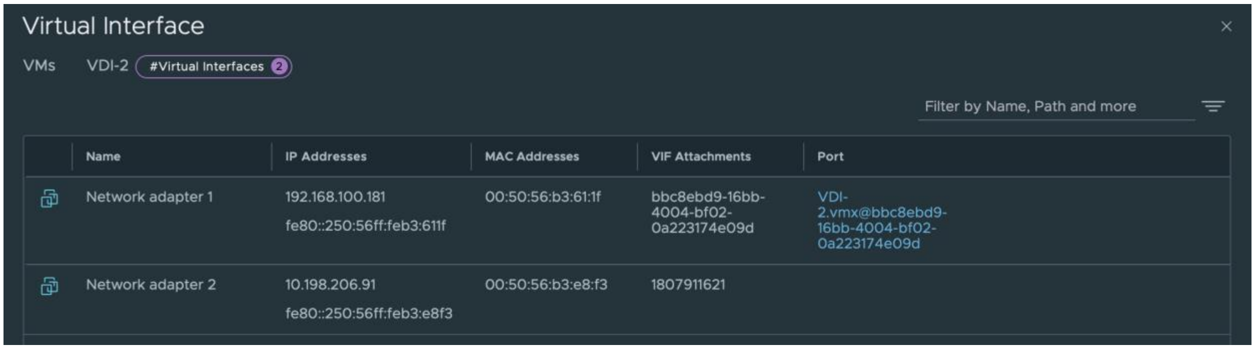
For the workload, with Fabric view we show the hostname, OS type, power state and detail interface information.



VDI-2

Virtual Machine GO TO CONFIGURATION →

Host Name	10.198.206.16
OS Name	Ubuntu Linux (64-bit)
Power State	● Running
Tags	2



Virtual Interface

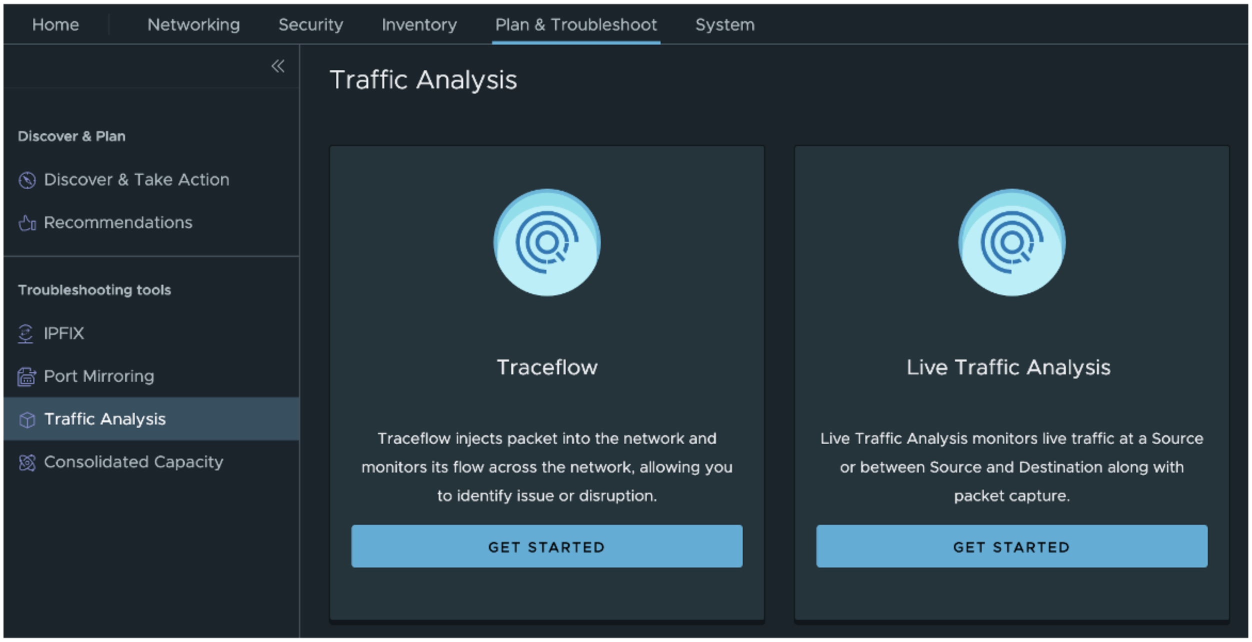
VMs VDI-2 #Virtual Interfaces 2

Filter by Name, Path and more

	Name	IP Addresses	MAC Addresses	VIF Attachments	Port
	Network adapter 1	192.168.100.181 fe80::250:56ff:feb3:611f	00:50:56:b3:61:1f	bbc8ebd9-16bb-4004-bf02-0a223174e09d	VDI-2.vmx@bbc8ebd9-16bb-4004-bf02-0a223174e09d
	Network adapter 2	10.198.206.91 fe80::250:56ff:feb3:e8f3	00:50:56:b3:e8:f3	1807911621	

Live Traffic Analysis (LTA)

Live Traffic Analysis (LTA) is a brand-new feature included in 3.2

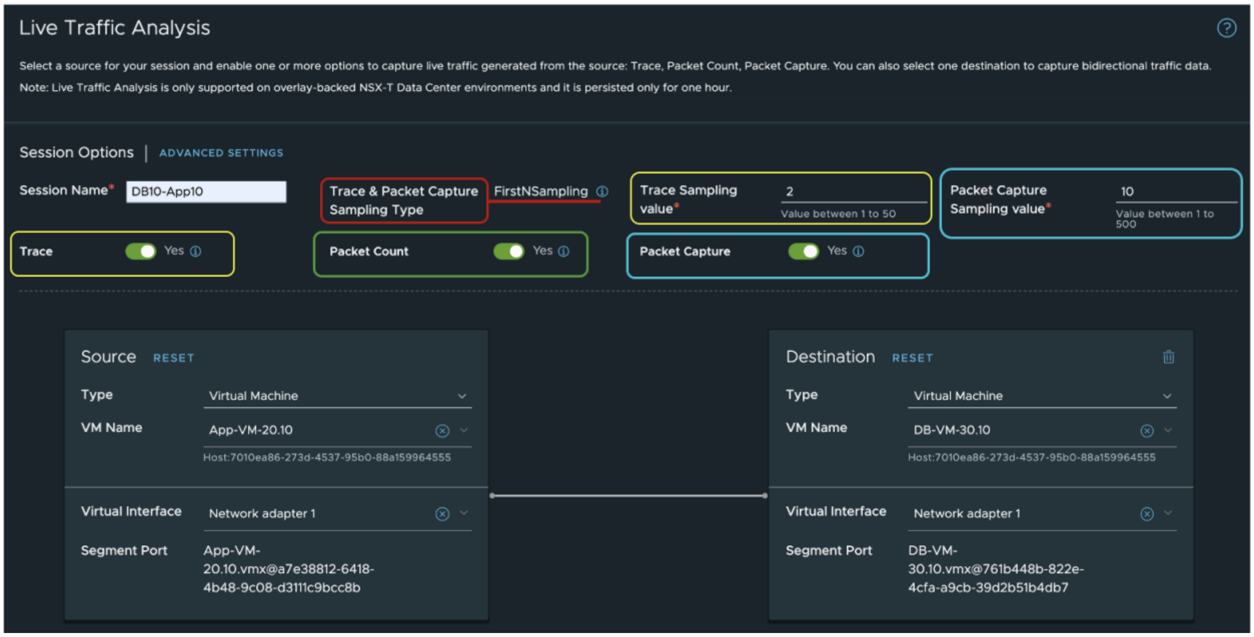


Traceflow injects packet into the network and monitors its flow across the network, allowing you to identify issue or disruption.

Live Traffic Analysis monitors live traffic at a Source or between Source and Destination along with packet capture.

LTA provides more functionalities than Traceflow. But one of the most important differences between Traceflow and LTA is that Traceflow uses a crafted packet to performance the trace and LTA uses the real live packets. So before the workload is deployed, Traceflow can be used to test the expected traffic path.

The LTA feature provides a unified approach of diagnosis. LTA includes 2 independent functionalities, live traffic trace and packet capture. They can be used together or separately.



Live Traffic Analysis

Select a source for your session and enable one or more options to capture live traffic generated from the source: Trace, Packet Count, Packet Capture. You can also select one destination to capture bidirectional traffic data.
Note: Live Traffic Analysis is only supported on overlay-backed NSX-T Data Center environments and it is persisted only for one hour.

Session Options | ADVANCED SETTINGS

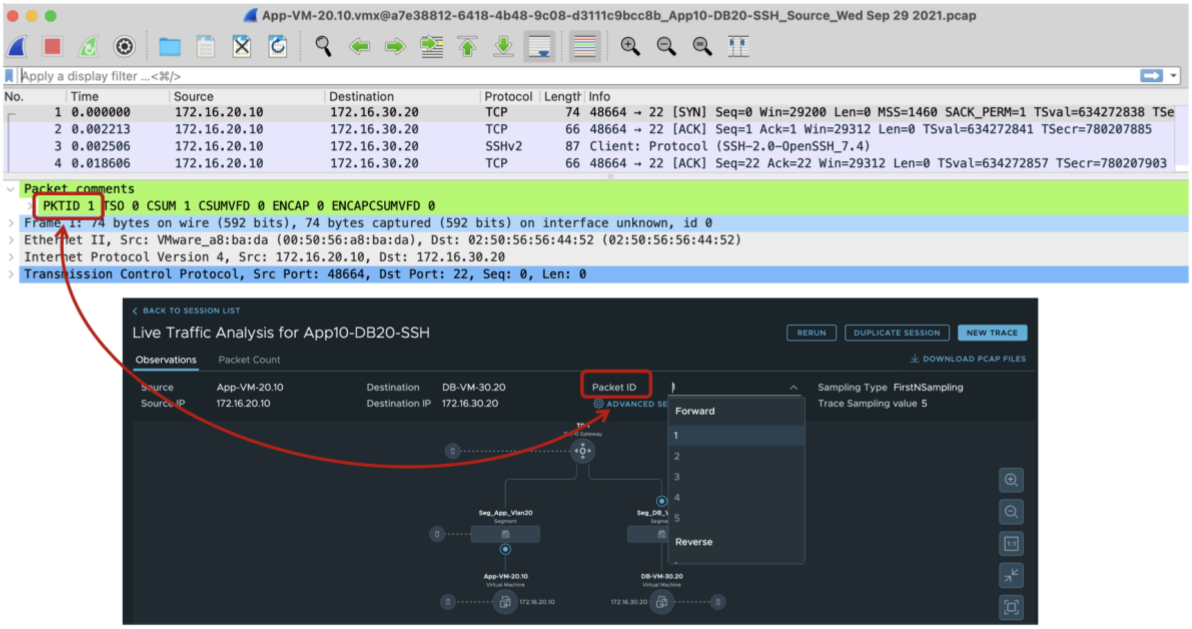
Session Name* DB10-App10
Trace & Packet Capture Sampling Type: FirstNSampling
Trace Sampling value* 2 (Value between 1 to 50)
Packet Capture Sampling value* 10 (Value between 1 to 500)
Trace: Yes
Packet Count: Yes
Packet Capture: Yes

Source RESET
Type: Virtual Machine
VM Name: App-VM-20.10
Host:7010ea86-273d-4537-95b0-88a159964555
Virtual Interface: Network adapter 1
Segment Port: App-VM-20.10.vmx@a7e38812-6418-4b48-9c08-d3111c9bcc8b

Destination RESET
Type: Virtual Machine
VM Name: DB-VM-30.10
Host:7010ea86-273d-4537-95b0-88a159964555
Virtual Interface: Network adapter 1
Segment Port: DB-VM-30.10.vmx@761b448b-822e-4cfa-a9cb-39d2b51b4db7

For how to configure LTA, and how to performance packet trace and packet capture, please check the NSX admin guide for details: https://docs.vmware.com/en/VMware-NSX-T-Data-Center/3.2/administration/GUID-6C76D8E5-0DC4-4365-B7E1-C1A5D276139F.html

As mentioned above, the LTA feature provide 2 independent functionalities, packet trace and packet capture. Trace and capture and be used separately or together. When they are used together, the packet in the capture and trace can be associated together with Packet ID so that observation result of the packet can been found in the trace.



App-VM-20.10.vmx@a7e38812-6418-4b48-9c08-d3111c9bcc8b_App10-DB20-SSH_Source_Wed Sep 29 2021.pcap

Apply a display filter ...<⌘/>

```
No.  Time      Source        Destination   Protocol Length Info
   1 0.000000  172.16.20.10  172.16.30.20  TCP        74 48664 → 22 [SYN] Seq=0 Win=29200 Len=0 MSS=1460 SACK_PERM=1 TSval=634272838 TSe
   2 0.002213  172.16.20.10  172.16.30.20  TCP        66 48664 → 22 [ACK] Seq=1 Ack=1 Win=29312 Len=0 TSval=634272841 TSecr=780207885
   3 0.002506  172.16.20.10  172.16.30.20  SSHv2      87 Client: Protocol (SSH-2.0-OpenSSH_7.4)
   4 0.018606  172.16.20.10  172.16.30.20  TCP        66 48664 → 22 [ACK] Seq=22 Ack=22 Win=29312 Len=0 TSval=634272857 TSecr=780207903

Packet comments
  PKTID 1 TSO 0 CSUM 1 CSUMVFD 0 ENCAP 0 ENCAPCSUMVFD 0
Frame 1: 74 bytes on wire (592 bits), 74 bytes captured (592 bits) on interface unknown, id 0
Ethernet II, Src: VMware_a8:ba:da (00:50:56:a8:ba:da), Dst: 02:50:56:56:44:52 (02:50:56:56:44:52)
Internet Protocol Version 4, Src: 172.16.20.10, Dst: 172.16.30.20
Transmission Control Protocol, Src Port: 48664, Dst Port: 22, Seq: 0, Len: 0
```

< BACK TO SESSION LIST

Live Traffic Analysis for App10-DB20-SSH

Source: App-VM-20.10
Source IP: 172.16.20.10
Destination: DB-VM-30.20
Destination IP: 172.16.30.20
Packet ID: 1
Sampling Type: FirstNSampling
Trace Sampling value: 5

Port Mirroring

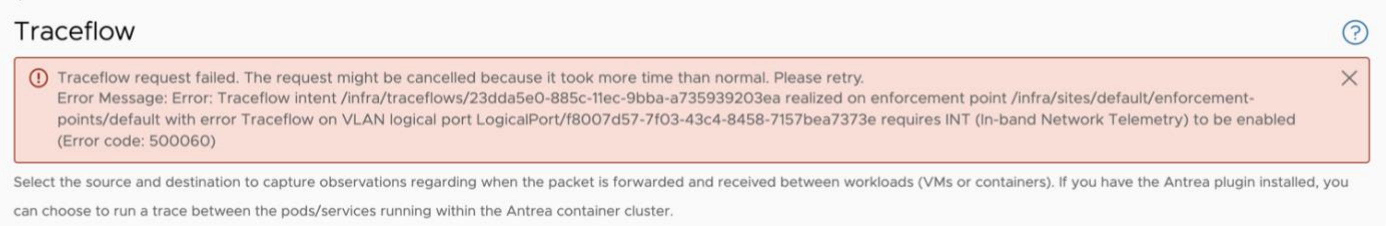
Selective Port Mirroring - Enhanced mirroring with flow-based filtering capability and reduced resource requirements. You can now focus on pertinent flows for effective troubleshooting.

Traceflow

In 3.2, the Traceflow is available for a Vlan based logical network. The Vlan based Traceflow is utilizing INT – In-band Network Telemetry. The In-band Network Telemetry is a framework designed to allow the collection and reporting of network state by the data plane without requiring intervention by the control plane. In the INT architectural model, the INT capable device can add additional header field to the packet which embeds telemetry instructions. These instructions tell other INT-capable devices what state to collect. The network state information may be directly exported by the data plane to the telemetry monitoring system or can be written into the packet as it traverses the network.

So, to perform vlan based Traceflow, INT must be enabled first.

If you see this message when you try to perform Traceflow for a Vlan based logical network,



Traceflow

Traceflow request failed. The request might be cancelled because it took more time than normal. Please retry.
Error Message: Error: Traceflow intent /infra/traceflows/23dda5e0-885c-11ec-9bba-a735939203ea realized on enforcement point /infra/sites/default/enforcement-points/default with error Traceflow on VLAN logical port LogicalPort/f8007d57-7f03-43c4-8458-7157bea7373e requires INT (In-band Network Telemetry) to be enabled (Error code: 500060)

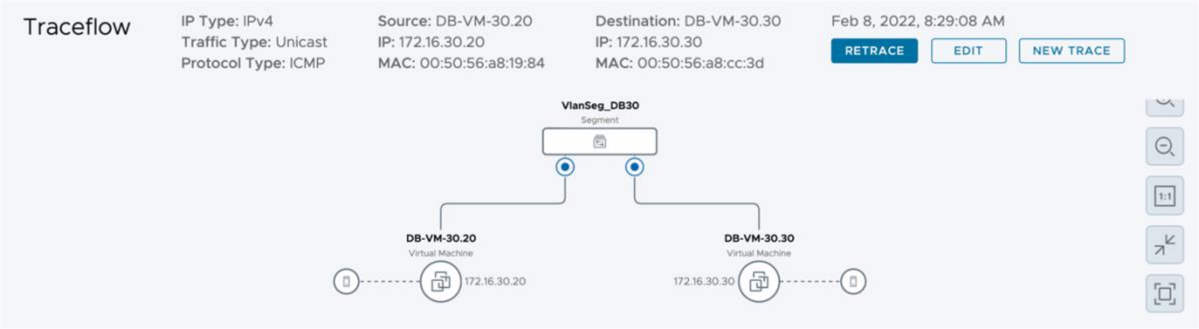
Select the source and destination to capture observations regarding when the packet is forwarded and received between workloads (VMs or containers). If you have the Antrea plugin installed, you can choose to run a trace between the pods/services running within the Antrea container cluster.

You can use the following API to enable the in-band network Telemetry setting,

PUT API /api/v1/infra/ops-global-config

```
{
   "display_name": "ops-global-config",
   "in_band_network_telementry": {
      "dscp_value": 2,
      "indicator_type": "DSCP_VALUE"
   },
   "path": "/infra/ops-global-config",
   "relative_path": "ops-global-config",
   "_revision": 0
}
```

After successfully issue the API, you will be able to perform the Traceflow.



Traceflow

IP Type: IPv4
Traffic Type: Unicast
Protocol Type: ICMP

Source: DB-VM-30.20
IP: 172.16.30.20
MAC: 00:50:56:a8:19:84

Destination: DB-VM-30.30
IP: 172.16.30.30
MAC: 00:50:56:a8:cc:3d

Feb 8, 2022, 8:29:08 AM

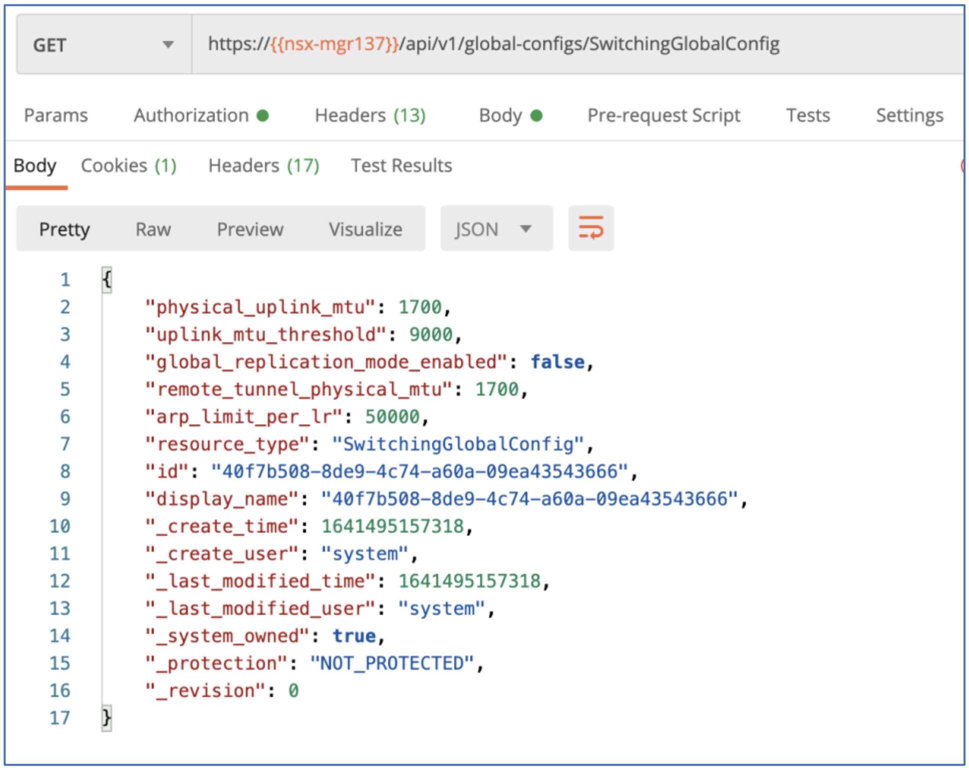
Notes: Vlan based Traceflow is not supported on the Edge node.

Fabric MTU Configuration Check

There are multiple places to configure MTU. First, let's understand the different MTU values in NSX configuration.

Name	What is it for	Default Value
Global MTU	For all the physical uplinks in a NSX domain	1700
Tunnel End MTU	For the tunnel in the same site	1700
Remote Tunnel Endpoint MTU	For Cross-Location communication for Federation	1700
Global Gateway MTU	For all the logical uplinks in a NSX domain	1500

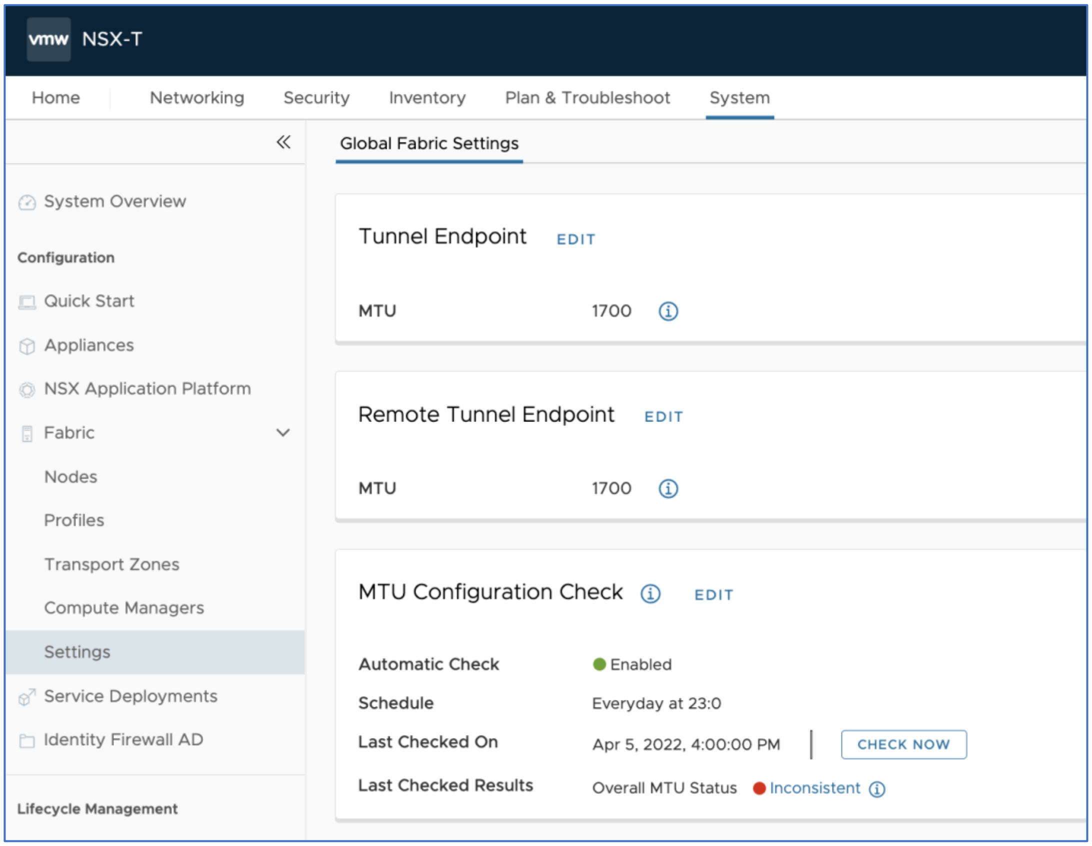
- Global default MTU



```
https://{{nsx-mgr137}}/api/v1/global-configs/SwitchingGlobalConfig
```

```
{
    "physical_uplink_mtu": 1700,
    "uplink_mtu_threshold": 9000,
    "global_replication_mode_enabled": false,
    "remote_tunnel_physical_mtu": 1700,
    "arp_limit_per_lr": 50000,
    "resource_type": "SwitchingGlobalConfig",
    "id": "40f7b508-8de9-4c74-a60a-09ea43543666",
    "display_name": "40f7b508-8de9-4c74-a60a-09ea43543666",
    "_create_time": 1641495157318,
    "_create_user": "system",
    "_last_modified_time": 1641495157318,
    "_last_modified_user": "system",
    "_system_owned": true,
    "_protection": "NOT_PROTECTED",
    "_revision": 0
}
```

The 1700 is the global default MTU for all the physical uplinks in a NSX domain. This is the default value for the optional uplink profile MTU field. If VPN is configured, the global MTU needs to be increased. When the MTU value is not specified in the uplink profile, this global value will be used. This value can be overridden by providing a value for the optional MTU field in the uplink profile. The Whenever this value is updated, the updated value will only be propagated to the uplinks that don't have the MTU value in their uplink profiles. If this value is not set, the default value of 1700 will be used. The node state can be monitored to confirm if the updated MTU value has been realized. Note: the host uplink profile is applicable for VDS7 with NSX-T and N-VDS. The MTU field as part of the uplink profile is not relevant for the VDS7, only when you use the N-VDS. The MTU for VDS7 is configured via vCenter.



Global Fabric Settings

Tunnel Endpoint

MTU 1700

Remote Tunnel Endpoint

MTU 1700

MTU Configuration Check

Automatic Check ● Enabled

Schedule Everyday at 23:0

Last Checked On Apr 5, 2022, 4:00:00 PM

Last Checked Results Overall MTU Status ● Inconsistent

The Tunnel Endpoint MTU setting is used for the tunnel in the same site. The Remote Tunnel Endpoint MTU setting is used for Cross-Location communication for Federation.

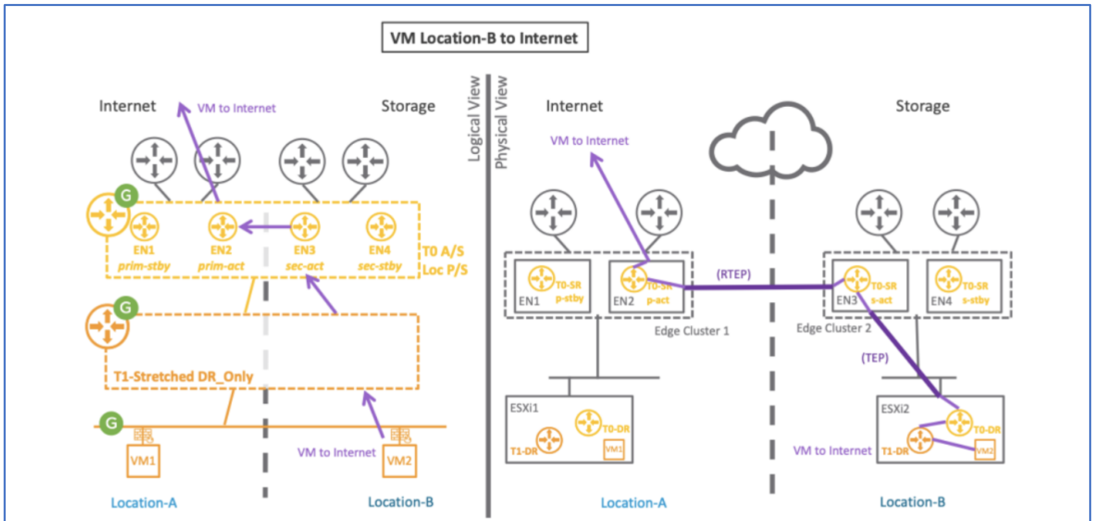
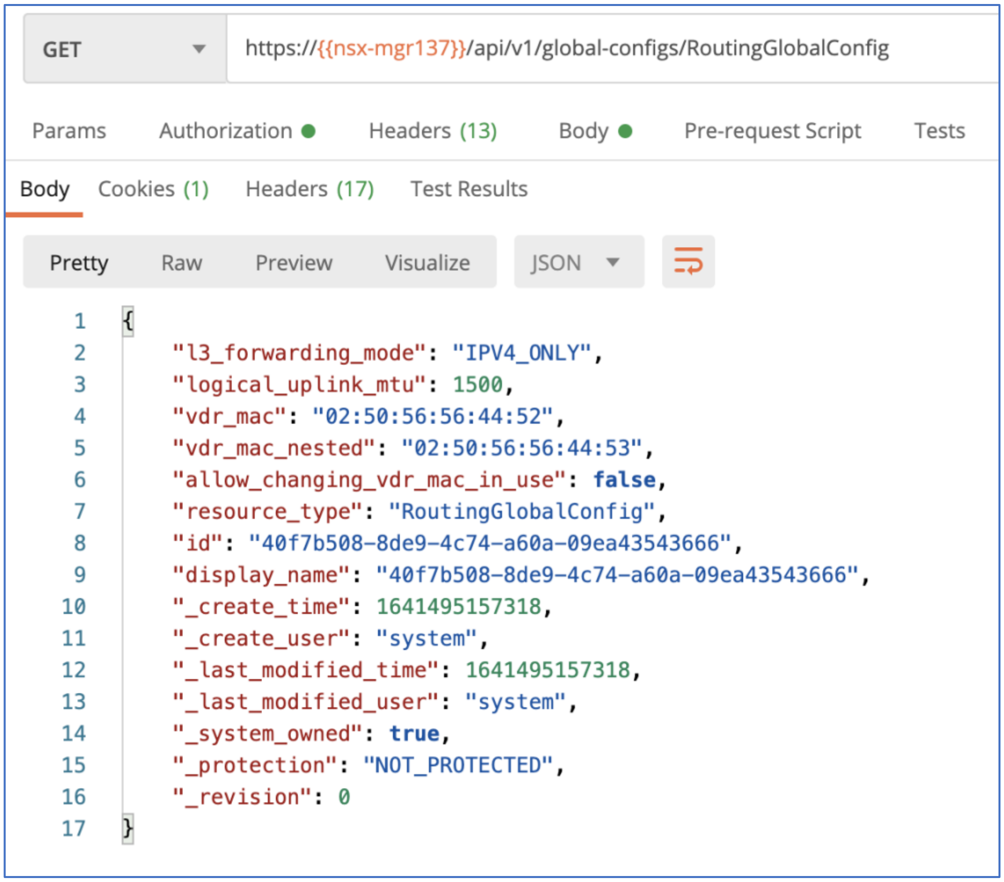


Figure 4-38: NSX-T Federation T0 A/S LocP_S packet walk2

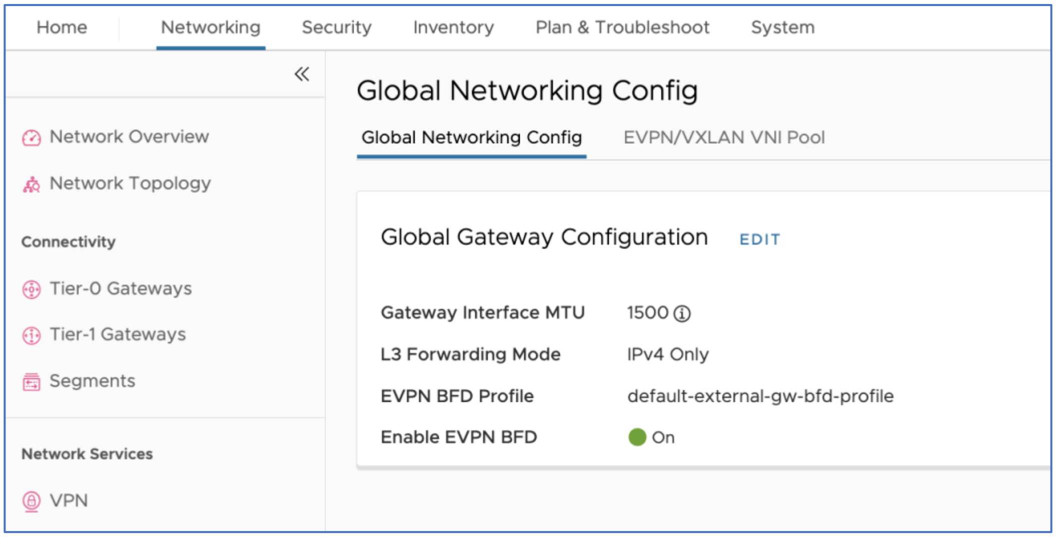
- Global Gateway MTU



GET https://{{nsx-mgr137}}/api/v1/global-configs/RoutingGlobalConfig

```
{
    "l3_forwarding_mode": "IPV4_ONLY",
    "logical_uplink_mtu": 1500,
    "vdr_mac": "02:50:56:56:44:52",
    "vdr_mac_nested": "02:50:56:56:44:53",
    "allow_changing_vdr_mac_in_use": false,
    "resource_type": "RoutingGlobalConfig",
    "id": "40f7b508-8de9-4c74-a60a-09ea43543666",
    "display_name": "40f7b508-8de9-4c74-a60a-09ea43543666",
    "_create_time": 1641495157318,
    "_create_user": "system",
    "_last_modified_time": 1641495157318,
    "_last_modified_user": "system",
    "_system_owned": true,
    "_protection": "NOT_PROTECTED",
    "_revision": 0
}
```

This is the global default MTU for all the logical uplinks in a NSX domain. Currently logical uplink MTU can only be set globally and applies to the entire NSX domain. There is no option to override this value at transport zone level or transport node level. If this value is not set, the default value of 1500 will be used.



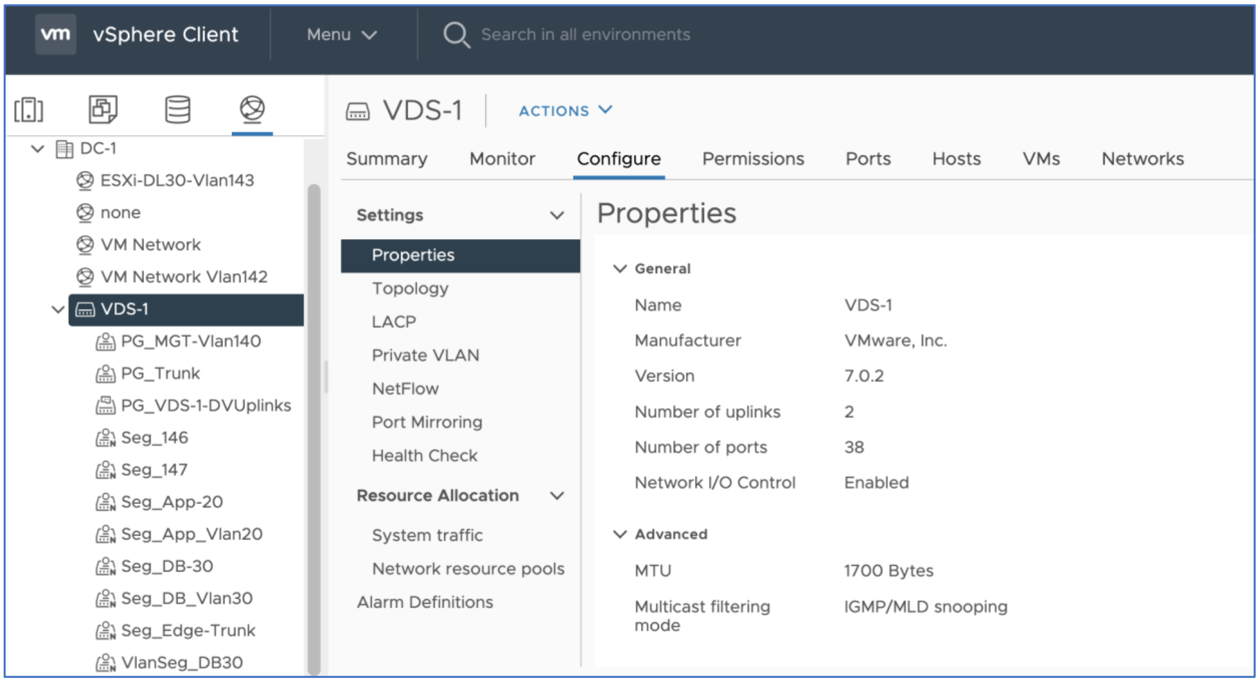
Global Networking Config

Global Gateway Configuration

Gateway Interface MTU 1500 ⓘ
L3 Forwarding Mode IPv4 Only
EVPN BFD Profile default-external-gw-bfd-profile
Enable EVPN BFD ● On

Now let’s understand how the Fabric MTU check feature works.
To avoid potential issues, it’s required that all the hosts in the same NSX domain to have the same uplink MTU setting. The Fabric MTU check feature is to check the MTU value for VDS 7.0 with NSX-T, host N-VDS to see if they are all the same. Note: The Fabric MTU check feature is only to check the configured MTU value for NSX, vmkping with max MTU is needed to verify end to end MTU setting.

MTU for VDS is configured via vCenter:



Properties

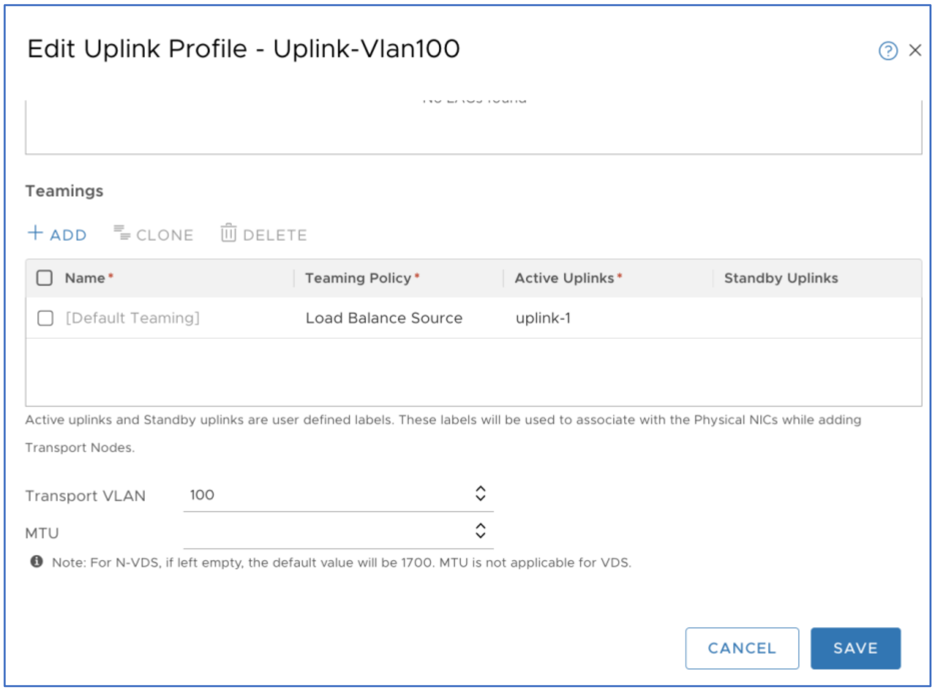
General

Name	VDS-1
Manufacturer	VMware, Inc.
Version	7.0.2
Number of uplinks	2
Number of ports	38
Network I/O Control	Enabled

Advanced

MTU	1700 Bytes
Multicast filtering mode	IGMP/MLD snooping

MTU for N-VDS can be configured via Uplink Profile in NSX:



Edit Uplink Profile - Uplink-Vlan100

No LAGs found

Teamings

□	Name*	Teaming Policy*	Active Uplinks*	Standby Uplinks
□	[Default Teaming]	Load Balance Source	uplink-1	

Active uplinks and Standby uplinks are user defined labels. These labels will be used to associate with the Physical NICs while adding Transport Nodes.

Transport VLAN 100

MTU

Note: For N-VDS, if left empty, the default value will be 1700. MTU is not applicable for VDS.

For N-VDS, if the MTU is not configured, the default Global MTU value will be used.

Monitoring & Alarms

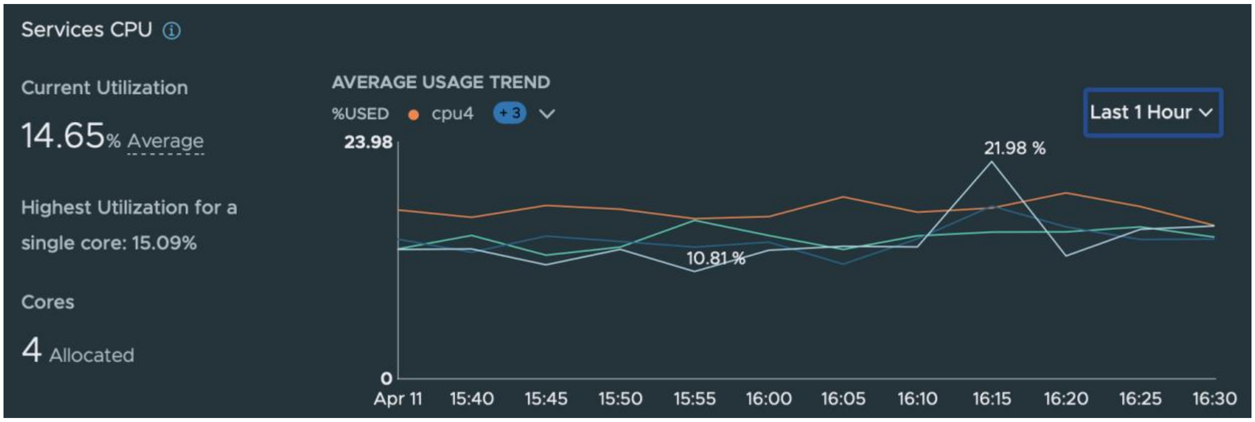
SNMP Monitoring

To download the SNMP MIB files, see Knowledge Base article 1013445: SNMP MIB module file download. Download and use the file named **VMWARE-NSX-MIB.mib**.

Notes: The NSX MIB is a Trap only MIB. SNMP polling is not supported.

Time Series Monitoring - Visibility with NSX Application Platform/NSX Intelligence

Starting from 3.2.1, NSX can provide Timeseries Metrics for the Edge Node resource usage which includes CPU, memory, disk, and interface packet counts.



Services CPU

Current Utilization

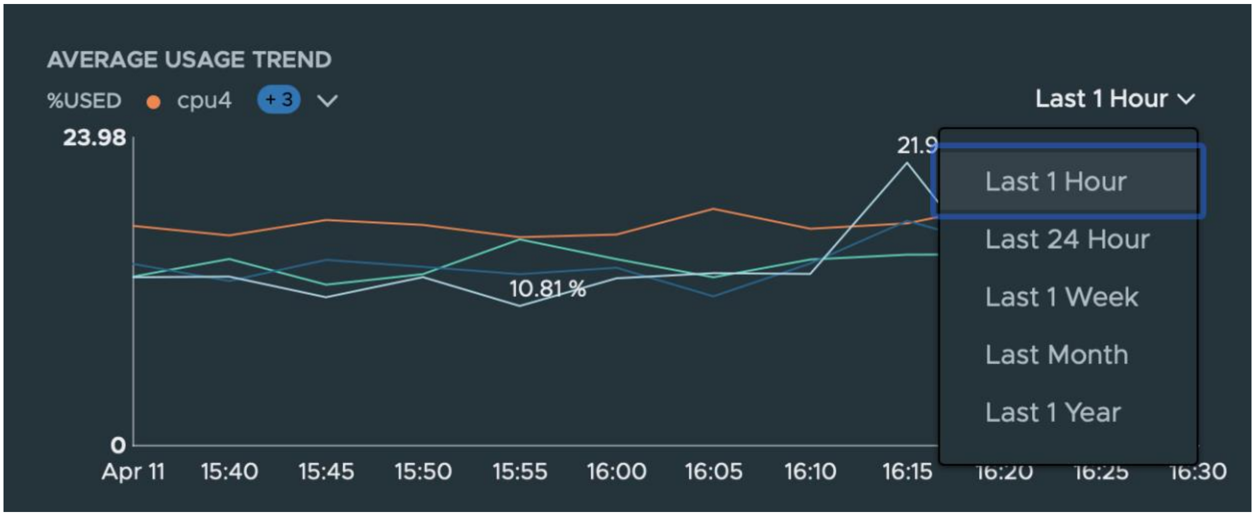
14.65% Average

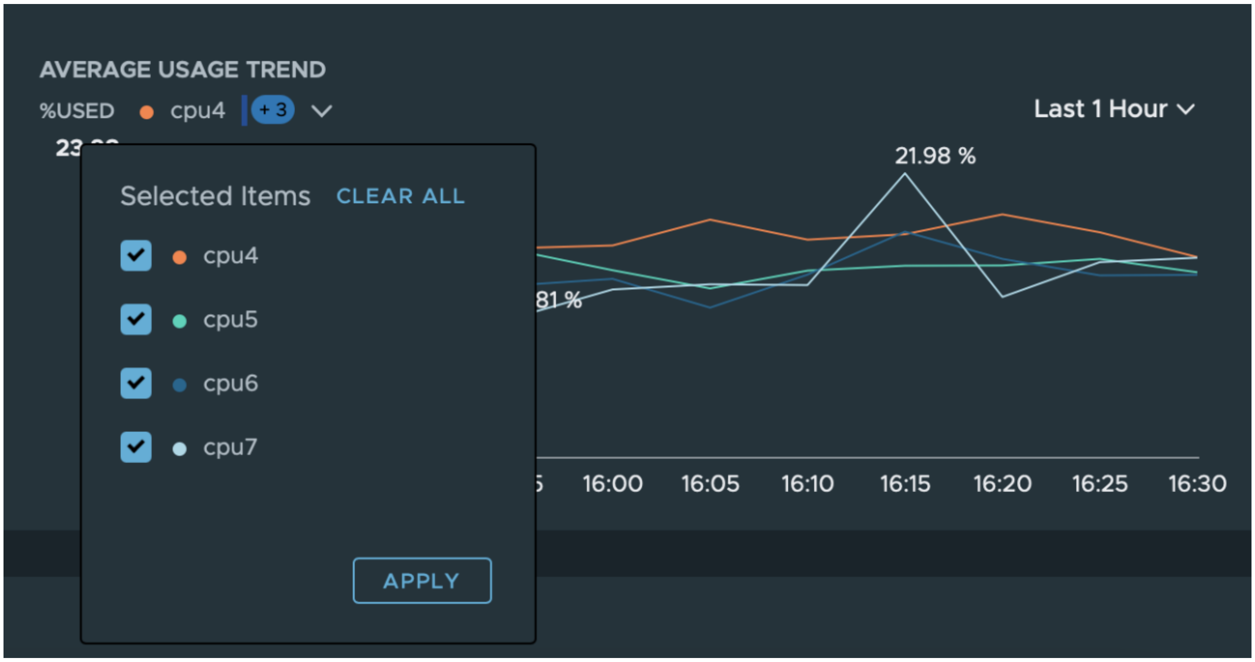
Highest Utilization for a single core: 15.09%

Cores

4 Allocated

All the metrics can be displayed for last hour, 24-hour, week, month, and year with different granularity. For example, for the past hour and the past 24 hours, the granularity will be for 5mins. For the past week, the granularity will be 1 hour. For past month and past year, the granularity will be 1 day.





AVERAGE USAGE TREND

%USED ● cpu4 +3

Last 1 Hour

Selected Items CLEAR ALL

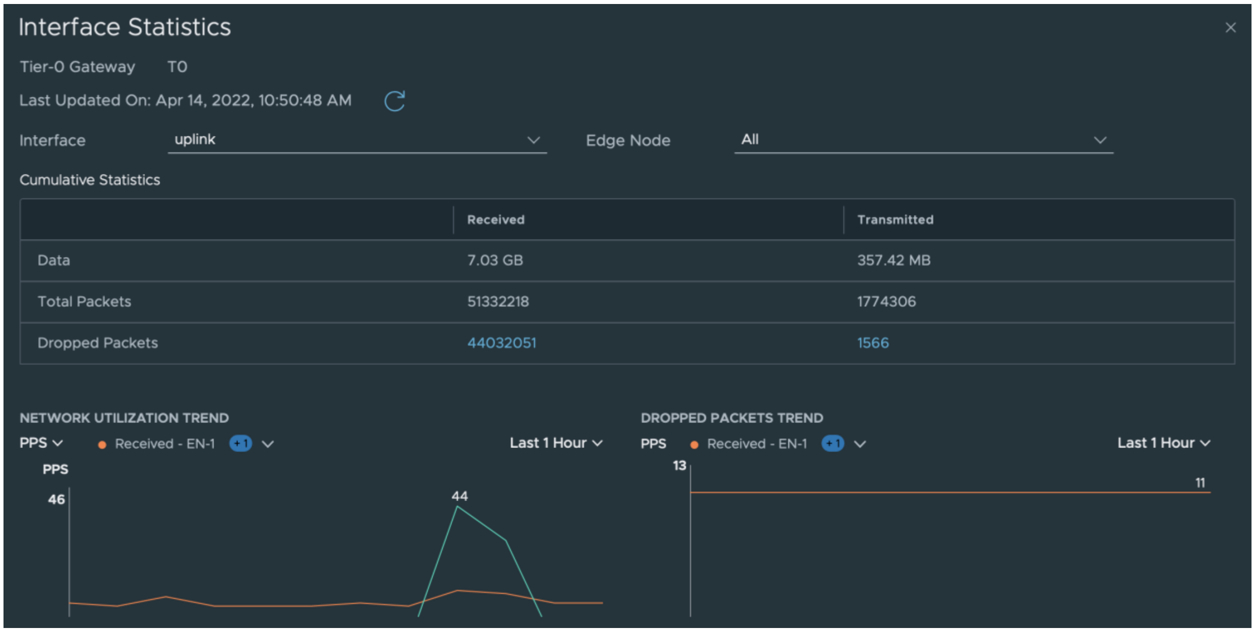
☑ ● cpu4

☑ ● cpu5

☑ ● cpu6

☑ ● cpu7

APPLY



Interface Statistics

Tier-0 Gateway T0

Last Updated On: Apr 14, 2022, 10:50:48 AM

Interface: uplink

Edge Node: All

Cumulative Statistics

	Received	Transmitted
Data	7.03 GB	357.42 MB
Total Packets	51332218	1774306
Dropped Packets	44032051	1566

NETWORK UTILIZATION TREND

PPS ● Received - EN-1 +1 Last 1 Hour

DROPPED PACKETS TREND

PPS ● Received - EN-1 +1 Last 1 Hour

Operation and Maintenance

NSX Upgrade Evaluation Tool

NSX Upgrade Evaluation Tool is a new capability introduced in 3.2.0.1 to help user prepare for upgrading to the latest releases. For details, check the following blogs, https://blogs.vmware.com/networkvirtualization/2022/01/introducing-new-nsx-upgrade-capabilities-for-nsx-t-3-2.html/

Shutdown/Startup Order of NSX

If there is a need to power off the entire NSX environment, the following Shutdown order should be followed.

When you try to power up the NSX environment, the following Startup order should be followed.

Note: We recommend backing up NSX and vCenter before initiating any planned power-off exercises.

The order to shutdown/startup NSX environment,

Shutdown Order:

1. Workload VMs
2. Edge nodes
3. NSX managers
4. vCenter
5. ESXi

Startup Order:

1. ESXi
2. vCenter
3. NSX managers
4. Edge Nodes
5. Workload VMs

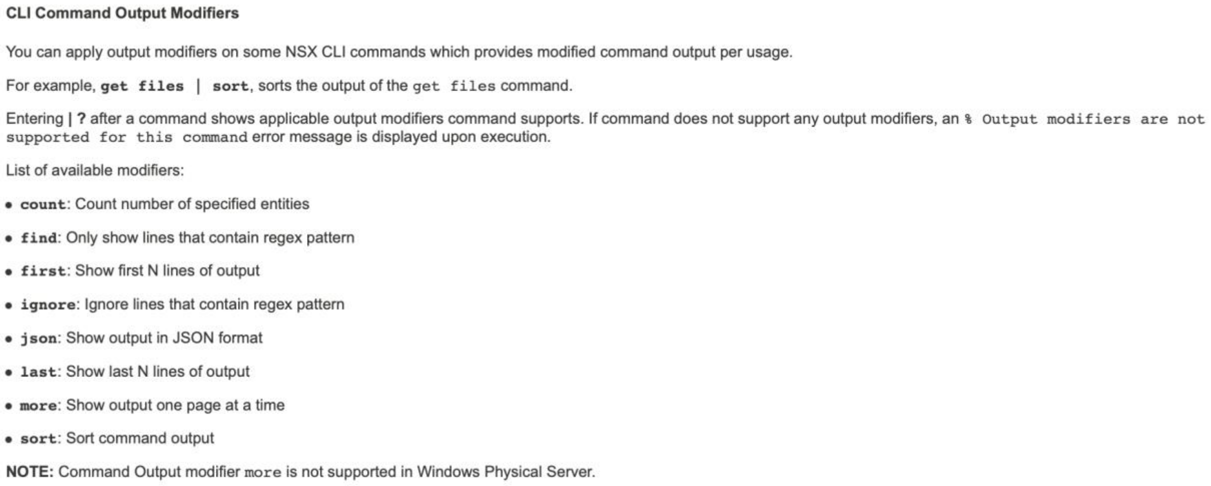
Improved CLI Guide

The 3.2 CLI guide is much more user-friendly. It improves the structure by generalizing common functionalities into defined categories.

https://vdc-download.vmware.com/vmwb-repository/dcr-public/8bc4a9b3-b4fb-447a-a97b-1452c22d6d5d/8537fe7f-36fd-4122-b1a4-fab306cc279d/cli_doc/index.html

Here are the highlights of some of the enhancements of the 3.2 CLI guide,

1. Adding **Modifiers** information:



CLI Command Output Modifiers

You can apply output modifiers on some NSX CLI commands which provides modified command output per usage.

For example, **`get files | sort`**, sorts the output of the `get files` command.

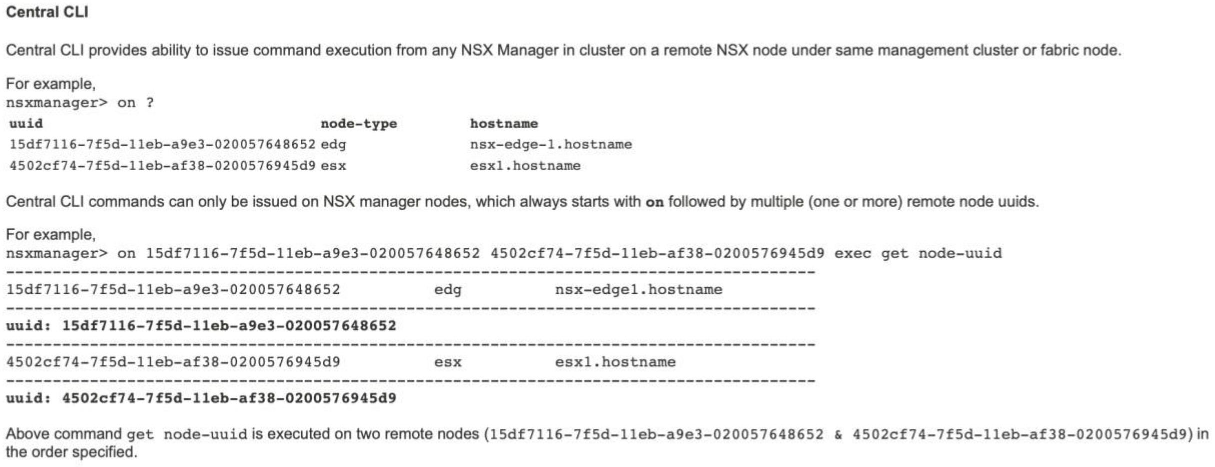
Entering **| ?** after a command shows applicable output modifiers command supports. If command does not support any output modifiers, an `% Output modifiers are not supported for this command` error message is displayed upon execution.

List of available modifiers:

- **`count`**: Count number of specified entities
- **`find`**: Only show lines that contain regex pattern
- **`first`**: Show first N lines of output
- **`ignore`**: Ignore lines that contain regex pattern
- **`json`**: Show output in JSON format
- **`last`**: Show last N lines of output
- **`more`**: Show output one page at a time
- **`sort`**: Sort command output

NOTE: Command Output modifier `more` is not supported in Windows Physical Server.

2. Adding introduction for **Central CLI**,



Central CLI

Central CLI provides ability to issue command execution from any NSX Manager in cluster on a remote NSX node under same management cluster or fabric node.

For example,

```
nsxmanager> on ?
uuid                                  node-type        hostname
15df7116-7f5d-11eb-a9e3-020057648652 edg              nsx-edge-1.hostname
4502cf74-7f5d-11eb-af38-0200576945d9 esx              esx1.hostname
```

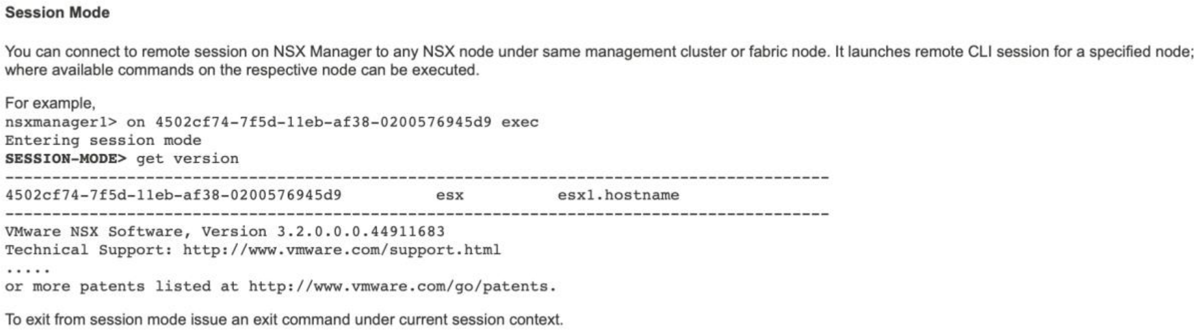
Central CLI commands can only be issued on NSX manager nodes, which always starts with **on** followed by multiple (one or more) remote node uuids.

For example,

```
nsxmanager> on 15df7116-7f5d-11eb-a9e3-020057648652 4502cf74-7f5d-11eb-af38-0200576945d9 exec get node-uuid
-----------------------------------------------------------------------------------------
15df7116-7f5d-11eb-a9e3-020057648652           edg          nsx-edge1.hostname
-----------------------------------------------------------------------------------------
uuid: 15df7116-7f5d-11eb-a9e3-020057648652
-----------------------------------------------------------------------------------------
4502cf74-7f5d-11eb-af38-0200576945d9           esx          esx1.hostname
-----------------------------------------------------------------------------------------
uuid: 4502cf74-7f5d-11eb-af38-0200576945d9
```

Above command `get node-uuid` is executed on two remote nodes (`15df7116-7f5d-11eb-a9e3-020057648652 & 4502cf74-7f5d-11eb-af38-0200576945d9`) in the order specified.

3. Adding introduction for **Session Mode**



Session Mode

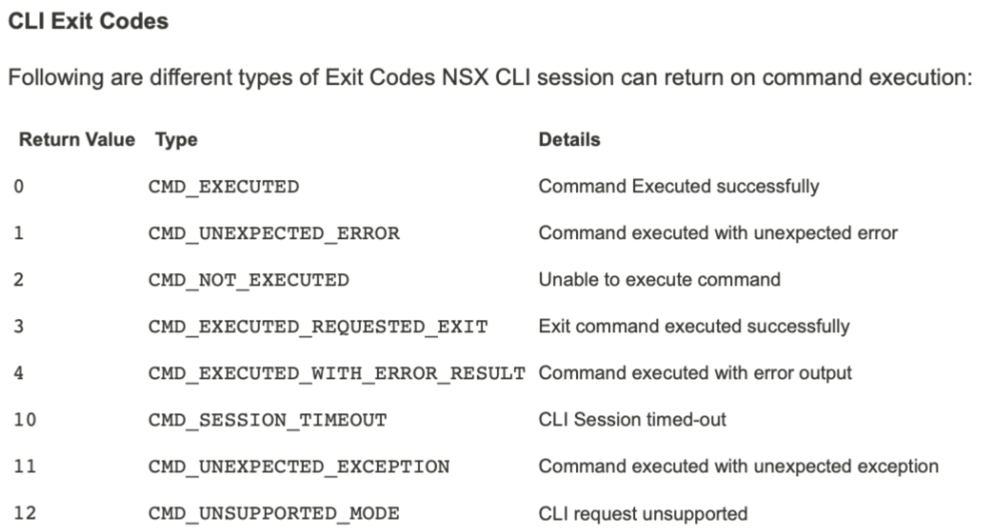
You can connect to remote session on NSX Manager to any NSX node under same management cluster or fabric node. It launches remote CLI session for a specified node; where available commands on the respective node can be executed.

For example,

```
nsxmanager1> on 4502cf74-7f5d-11eb-af38-0200576945d9 exec
Entering session mode
SESSION-MODE> get version
-----------------------------------------------------------------------------------------
4502cf74-7f5d-11eb-af38-0200576945d9           esx          esx1.hostname
-----------------------------------------------------------------------------------------
VMware NSX Software, Version 3.2.0.0.0.44911683
Technical Support: http://www.vmware.com/support.html
.....
or more patents listed at http://www.vmware.com/go/patents.
```

To exit from session mode issue an exit command under current session context.

4. Addition of Exit Codes



CLI Exit Codes

Following are different types of Exit Codes NSX CLI session can return on command execution:

Return Value	Type	Details
0	`CMD_EXECUTED`	Command Executed successfully
1	`CMD_UNEXPECTED_ERROR`	Command executed with unexpected error
2	`CMD_NOT_EXECUTED`	Unable to execute command
3	`CMD_EXECUTED_REQUESTED_EXIT`	Exit command executed successfully
4	`CMD_EXECUTED_WITH_ERROR_RESULT`	Command executed with error output
10	`CMD_SESSION_TIMEOUT`	CLI Session timed-out
11	`CMD_UNEXPECTED_EXCEPTION`	Command executed with unexpected exception
12	`CMD_UNSUPPORTED_MODE`	CLI request unsupported