

vRealize Network Insight FAQs

VMware vRealize Network Insight 4.1



vmware®

You can find the most up-to-date technical documentation on the VMware website at:

<https://docs.vmware.com/>

If you have comments about this documentation, submit your feedback to

docfeedback@vmware.com

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

Copyright © 2019 VMware, Inc. All rights reserved. [Copyright and trademark information.](#)

Contents

1	About vRealize Network Insight FAQs Guide	4
2	General	5
3	Installation and Configuration	7
4	Adding or Configuring vCenter Servers as Data Source	13
5	Micro-Segmentation and Flows	15
6	Clustering	17
	Clustering - General	17
	Clustering - Installing and Configuring	19
	Clustering - Scaling	20
	Clustering - Deployment	22
7	Data Management and Processing	24
8	IPFIX	26

About vRealize Network Insight FAQs Guide

1

The vRealize Network Insight FAQs guide provides the user with frequently asked questions about vRealize Network Insight.

Intended Audience

This information is intended for users who work with vRealize Network Insight.

How do I Create a Support Bundle?

Refer to the support-bundle section in the *vRealize Network Insight Command Line Reference Guide*.

How to Create read only Admin Role in Palo Alto Networks Panorama for XML API access?

To add an **Admin** role for XML API access:

- 1 Select **Panorama** → **Admin Roles**
- 2 Click **Add** to add a new Admin Role to open Admin Role Profile dialog box
- 3 In Admin Role Profile dialog box
 - a Give a name to the role (for example, api-only-admin)
 - b Select the **Role** as **Panorama**
 - c Disable all entries in the Web UI tab
 - d Enable all entries except **Commit** in the XML API tab
 - e Click **OK** to close the dialog box, a new **Admin Role** appears in the list with the name
 - f Click **Commit** to commit the changes to Panorama
- 4 Assign this **Admin** role to an administrative account.

When is a service considered shared?

The following ports are configured as shared:

Protocol	Port
DNS	53
Bootpc	68
Kerberos	110
sunrpc	111

Protocol	Port
NTP	123
map	143
Imap3	220
SMTP	25
LDAP	389
IGMPv3Lite	465
syslog	514
Submission	587
syslog-conn	601
LDAPS	636
IMAPS	993
POP3S	995
NFS	2049
MSFT-GC	3268
MSFT-GC-SSL	3269

I see an event/error on Data Source as "The data source identity information such as a certificate or the key has changed." What does it mean?

vRealize Network Insight has received a new certificate from a data source which is not the same as stored in the product. vRealize Network Insight auto accepts the certificate presented by data sources. During the process, you get an event on the data sources where you can download old and new certificates.

Installation and Configuration

What are the resource requirements for vRealize Network Insight?

Refer to the vRealize Network Insight Installation Guide for Resource Requirements.

What happens if I enter the incorrect key during vRealize Network Insight Proxy OVA Deployment?

The secret key is not validated during vRealize Network Insight Proxy OVA deployment. The deployment completes even with incorrect secret key. However, pairing can fail and vRealize Network Insight Proxy does not show up as detected on vRealize Network Insight UI.

To correct the shared secret, log in to vRealize Network Insight Proxy CLI and run the `set-proxy-shared-secret` command to set the correct secret key. This command replaces the old key with the new one, and therefore, vRealize Network Insight Platform detects vRealize Network Insight Proxy and pairs up.

How do I configure DNS after vRealize Network Insight Proxy OVA is deployed?

Log in to vRealize Network Insight Proxy CLI, and run the `change-network-settings` command. This interactive command will provide the user an option to add or modify DNS after which the vRealize Network Insight Proxy will be reconfigured with the new DNS.

If any of the network parameters is not configured correctly, use the `change-network-settings` command to modify the network configuration parameters.

How do I find out vRealize Network Insight Proxy VM IP from the UI?

Go to Settings page and select vRealize Network Insight Infrastructure menu option. The IP address of both, vRealize Network Insight Platform and vRealize Network Insight Proxy VMs is displayed.

What should I do if vRealize Network Insight Proxy is not detected in 5 minutes after deploying vRealize Network Insight Proxy OVA?

Log in to vRealize Network Insight Proxy using `consoleuser` (refer to the vRealize Network Insight Command-Line Reference Guide) and verify the following:

- Verify vRealize Network Insight Platform pairing status with vRealize Network Insight Proxy using the CLI `show-connectivity-status`.
- If the pairing status is shown **Passed**, open the Platform UI in a new browser window and login to check status.
- If the pairing status is showing **Failed**, then the shared secret key specified during the vRealize Network Insight Proxy OVA deployment could be wrong. To fix this problem, use the `set-proxy-shared-secret` command to set the correct secret key. This command replaces the old key with the new one, and therefore, vRealize Network Insight Platform can detect vRealize Network Insight Proxy.
- If the `show-connectivity-status` shows network reachability to vRealize Network Insight Platform as **Failed**, then verify whether vRealize Network Insight Platform is reachable from vRealize Network Insight Proxy VM using the `ping` command.
- If it is not reachable, then verify if NTP, DNS, gateway, and other network parameters are configured correctly using `show-config` command.
- If any of the network parameters is not configured correctly, use the `setup` command to modify the network configuration parameters.

What should I do if I forget my login credentials?

If you are the UI local user: Contact vRealize Network Insight UI administrator to reset the credentials for you.

If you are the administrator: From vRealize Network Insight 3.4, the UI credentials can be changed by using CLI `modify-password`. Refer the CLI guide for details. If you are working on vRealize Network Insight versions previous to 3.4, contact support.

How do I change the login password?

To change the login password:

- 1 Go to **Administrator > Settings**, and then click **My Profile** on the left pane.
- 2 On the **Change Password** page, fill in the required information and click **Save**.

What do I do if I get the login screen before detecting the vRealize Network Insight Proxy VM?

- This behavior is expected when the browser is refreshed or URL is opened in a new window before detecting the proxy.
- Log in by using the credentials set during license activation for the `admin@local` username.

Does vRealize Network Insight support multiple vCenter Server/NSX Manager?

Yes, vRealize Network Insight supports multiple vCenter Servers and NSX Manager.

Which services of vRealize Network Insight need Internet access and why?

vRealize Network Insight supports remote home calling feature that requires Internet access. This feature or services allow the vRealize Network Insight team to gain a better understanding of customer environments and proactively troubleshoot or repair issues. The following services need Internet access:

- 1 Upgrade Service (`svc.ni.vmware.com:443`): vRealize Network Insight uses this service to contact the remote upgrade host and pull in newly released bits as they become available. Certain metrics related to key services and performance of vRealize Network Insight are periodically gathered and uploaded using upgrade host for the vRealize Network Insight Support team to monitor and identify any anomaly in the environment so that they can act before it impacts critical services.
- 2 Metric service (`svc.ni.vmware.com:443`): Certain metrics related to key services and performance of vRealize Network Insight are periodically gathered and uploaded for the vRealize Network Insight Support team to monitor and identify any anomaly in the environment so that they can act before it impacts critical services. It can be enabled/disabled while deploying the vApp or through "telemetry" CLI later.
- 3 Support Service (`support2.ni.vmware.com:443`): This service establishes remote secured tunnels to the vRealize Network Insight support host that allow authorized personnel to remotely access and work on deployments. It is disabled by default and can be enabled/disabled through UI as well as "support-tunnel" CLI.

- 4 Registration Service (reg.ni.vmware.com:443): For registering the appliance with all external services. It will enable trusted communication between above mentioned services. When setup has access to internet, registration happens automatically. In an isolated environment it can be done using "offline-registration" CLI (Please refer to CLI guide for more details). It is required for enabling Support Tunnel.

Note If the vRealize Network Insight platform is behind an Internet proxy, whitelist the following domain names and ports:

Table 3-1.

Service	URL	Port
Upgrade Service/Metric Service	svc.ni.vmware.com	443
Support Tunnel Service	support2.ni.vmware.com	443
Registration Service	reg.ni.vmware.com	443

What is port aggregation and what is the mechanism to do it?

Port aggregation is built in to aggregate the ephemeral port flows – like dynamic FTP, Oracle, MS-RPC etc. This helps in reducing the number of flows in system and provide an aggregated view for large number of flows that are essentially for the same service.

The mechanism to do it is as follows:

- For first three days of noticing a destination_ip, we will aggregate destination ports on that particular IP in buckets of 10K and start building a port-profile for that IP (build a port-profile per destination IP).
- After three days, once we have built a profile, we will start aggregating port ranges where the port density is high (reflect ephemeral port opening pattern). The ranges themselves will be dynamic in size such as 100, 1,000, 10,000, and will be created depending on how many ports are being opened and how widespread they are in the given range of aggregation.

Note This decision happens independently for each server IP address.

- This will enable high-port flows to be reported with no aggregation where there is no bulk port open activity happening and also let dynamic aggregation to be applied where such activity is happening.
- The profile is continually updated in a time-decayed manner to account for new ports opening up or older ones being not used any more.

How do I change the IP Address/Gateway/Netmask after vRealize Network Insight OVA is deployed?

To change vRealize Network Insight Platform/Proxy network settings, log in to CLI and run the `change-network-settings` command. This interactive command will provide the user an option to modify the IP address, gateway, netmask, and so forth after which the vRealize Network Insight appliance is reconfigured with new details.

Note

- This task must be done using VM console session as the appliance reboots in the end.
- If vRNI Platform IP is modified and it is paired with proxies then on each proxy VM run this CLI command:

```
vrni-proxy set-platform --ip-or-fqdn <New_Platform_IP>
```

How do I change from an Evaluation License to a Perpetual License?

Refer to the Change License section in the vRealize Network Insight User Guide.

How are licenses characterized in vRealize Network Insight?

Table 3-2.

License Name	License Type	Features
Enterprise	Full/Production: It can be perpetual or time bound.	The following features are enabled: <ul style="list-style-type: none"> ■ AWS as data provider ■ Tunable data retention policies ■ Infoblox DNS data source ■ Physical IP and DNS Mapping ■ Analytics
Advanced	Full/Production: It can be perpetual or time bound.	NA

Note ALL licenses are capacitated per CPU socket and CCU (Concurrent Users). The evaluation licenses can be renewed or converted to **Production** with the updated key through **UI -> Settings -> About**. Refer to the user guide for more details.

How to take a backup of the VMs in vRealize Network Insight ?

Refer to *VMware Best Practices* to take the backup of VMs such as VMware VADP/VDP API. It is recommended that you take a backup before creating or expanding clusters.

Adding or Configuring vCenter Servers as Data Source

4

What if I am getting a "Request timed out" message while adding vCenter Server using IP address?

- Verify that the vCenter Server IP address is reachable from the vRealize Network Insight Proxy VM.
- Log in to vRealize Network Insight Proxy CLI and use the ping to ensure that IP is reachable and telnet to ensure that the vCenter Server is reachable on port 443.
- If vCenter Server is reachable, then retry adding.
- If IP address is not reachable, then verify whether the gateway is correctly configured from vRealize Network Insight Proxy VM using command `show-config`.
- If gateway is incorrect, then correct it using the `setup` command.

What if I am getting a "IP/FQDN is invalid" message while adding vCenter Server?

- Verify whether provided IP/FQDN for vCenter Server is correct.
- Verify whether FQDN is reachable from vRealize Network Insight Proxy VM using `ping` command.
- If it is not reachable, then verify if the DNS is configured correctly on vRealize Network Insight Proxy VM using `nslookup FQDN` and `show-config` command.
- If DNS is incorrect, then correct it using the `setup` command

What privileges does the vRealize Network Insight Security and Operations Platform require?

vRealize Network Insight requires the VMware vCenter Server credentials with the following privileges:

- Distributed Switch: Modify
- dvPort group: Modify

What if I am getting error "User does not have required privileges." while enabling IPFIX on vCenter Server Data source page?

vRealize Network Insight requires the VMware vCenter Server credentials with the following privileges to enable IPFIX:

- Distributed Switch: Modify
- dvPort group: Modify

Please make sure that provided VMware vCenter Server user have permission on vCenter Server's root folder and all of its child entities e.g all folders and all datacenters.

How frequently is the data fetched from environment?

vRealize Network Insight Proxy fetches data every 10 minutes from environment.

How soon the analysis of data will start after adding the vCenter Server?

Analysis of data starts right away after adding a vCenter Server. Product UI will show partial picture of data within few minutes which can take two hours to get complete.

Note Flow traffic data changes continuously and include at least 24 hours of data in its analysis.

How do I clean up IPFIX settings in vCenter Server if I have deleted vRealize Network Insight OVAs?

- Using VMware vSphere Web Client: Go to **Home > Networking > VDS (Name) > Netflow** Settings. Remove vRealize Network Insight Proxy IP from Collector settings.
- Using VMware vSphere Windows Client: Go to **Home > Inventory > Networking > VDS (Name) > Edit** Settings. Remove vRealize Network Insight Proxy IP from Collector settings in Netflow tab. This step is required to be done for each VDS for which IPFIX is enabled.

How do I clean up IPFIX configuration in vRealize Network Insight?

In vRealize Network Insight UI, goto **settings > Data Sources**, delete the vCenter server. This removes IPFIX configuration done by vRealize Network Insight.

Micro-Segmentation and Flows

What do the numbers in the Traffic Distribution Pin represent?

The percentage gives an overview of the traffic distribution based on flow analysis.

Table 5-1.

Traffic	Description
East-West (EW)	East-West traffic as the percentage of the traffic of the total group
Switched (% of EW)	Switched traffic as the percentage of East-West traffic
Routed (% of EW)	Routed traffic as the percentage (%) of East-West traffic
Within Host (% of VM-VM)	Traffic with source and destination on same host as percentage of virtual machine to virtual machine traffic
VM to VM (% of EW)	Virtual machine to virtual machine traffic as percentage of East-West traffic
Internet	Internet traffic as percentage of the traffic of the total group

How are ports aggregated in flows?

Port aggregation is built in to aggregate the ephemeral port flows - like dynamic FTP, Oracle, MS-RPC etc. This helps in reducing the number of flows in system and provide an aggregated view for large number of flows that are essentially for the same service. The mechanism to do this is as follows:

- For first three days of noticing a destination_ip , we will aggregate dst ports on that IP in buckets of 10K and start building a port-profile for that IP.
- Once three days are over - and we have built a profile that can be used with confidence - we will start aggregating port ranges where the port density is high (in other words - reflect ephemeral port opening pattern). The ranges themselves will be dynamic in size - 100, 1,000, 10,000 and will be created depend on how many ports are being opened and how widespread they are in the given range of aggregation.
- This will enable high-port flows to be reported with no aggregation where there is no bulk port open activity happening; and also let dynamic aggregation to be applied where such activity is happening.

- The profile is continually updated in a time-decayed manner to account for new ports opening up or older ones being not used any more.

What does the 240.240.240.240 IP address signify in vRealize Network Insight?

240.240.240.240 is a place holder IP address in vRealize Network Insight. This IP address is used if there are very large number of IP addresses (> 5000) hitting some particular IP. All further incoming Internet IPs (5001th onwards) with this placeholder IP 240.240.240.240 can be replaced for that service end point.

This is to limit the number of flows in the system, as publicly exposed service that log each internet client individually could result in very large number of flows - which would result in increased system load.

For all the flows that have been replaced with this placeholder IP, all the metrics are aggregated on the corresponding flow with this IP address, so there is no loss of statistics at an aggregate level.

All the destination IP for the flows reported in the flows view are shown as originating from 240.240.240.240 are actually being hit by large count of internet IP (> 5000).

Clustering

This chapter includes the following topics:

- [Clustering - General](#)
- [Clustering - Installing and Configuring](#)
- [Clustering - Scaling](#)
- [Clustering - Deployment](#)

Clustering - General

Can a proxy or a collector VM be clustered?

No. Clustering for collector/proxy VMs is not supported.

Does vRealize Network Insight need a load balancer like vRealize Log Insight?

vRealize Network Insight clustering is a scale-out solution and not an HA solution. If the primary platform VM/master node fails, then the whole service becomes unavailable.

What happens if connection between remote proxy & platform goes down?

In case connection between Platform and Proxy VM is down, Proxy VM will store data locally (depending on the disk space) and will send it whenever it is connected again.

Is vRealize Log Insight integrated with vRealize Network Insight?

Yes, vRealize Log Insight has been integrated with vRealize Network Insight 3.4. The alerts are sent to syslog which can be vRealize Log Insight.

What happens if a node reboots?

If a node reboots, it automatically joins the cluster and continues to be operational. If it is the primary node, then there is a complete loss of service during the time it was down.

How to change IP of any platform node or proxy in a cluster?

In a N node cluster you can change IP of any platform node using the `change-network-settings` CLI. You need to reflect new IP on all other platform nodes in a cluster using `update-IP-change` CLI. It must be done using VM console session as the appliance reboots in the end.

Note If platform1 (primary platform from where clustering action was triggered) IP is changed you also need to reflect new IP on all proxies using `vrni-proxy` CLI.

For example consider a 3-Node cluster:

- Use case 1: Only the platform2 IP is changed
 - a Run `change-network-settings` on platform2 to change its IP.
 - b Run `update-IP-change` on platform1 and platform3 to reflect new platform2 IP
- Use case 2: Only the platform1 IP is changed
 - a Run `change-network-settings` on platform1 to change its IP.
 - b Run `update-IP-change` on platform2 and platform3 to reflect new platform1 IP
 - c Run `vrni-proxy` on all proxies to reflect new platform1 IP
- Use case 3: The platform1 and proxyX IP are changed
 - a Run all steps mentioned in Use case 2.
 - b Run `change-network-settings` on proxyX to change its IP. Refer to the CLI guide for more details on each command.

How much disk space is needed on platform1?

Platform1 requires more disk space compared to other nodes in cluster as some of the configuration data is stored on Platform1 only.

What happens if any of the node ran out of disk space?

The UI starts showing error messages when disk space on any particular platform node reaches a certain threshold. Add more disk space to the platform node by logging in to

vCenter

How many times data is replicated in cluster?

The data replication mechanism is depending on the components present in the platform node.

Clustering - Installing and Configuring

Do all platform VMs have to be on the same L2/L3 segment?

No. However, it is best to keep all platform nodes on a common network with low latencies between nodes. This is because many of the distributed components replicate data among the nodes and high latencies can cause system performance and stability issues.

Can a cluster be upgraded using in-product upgrade feature?

Online upgrades are not supported for cluster till 3.7. From 3.8 and the succeeding releases, a cluster can be upgraded using the online upgrade method.

What happens if there is a failure during the cluster creation process?

It is a best practice to snapshot primary platform and proxies before starting the cluster creation process. If there is a failure, delete the secondary platform nodes and recover primary platform and proxy VMs from the snapshots.

What happens to the existing data and configuration when I expand the single node deployment to a cluster?

All data and configuration is maintained without any change. The data will be accessible after cluster creation.

Can you have platform VM in different regions?

No, we require the Platform nodes to co-located be in the same site. The proxy servers can be geo-distributed.

Can platform hosted on vSAN Stretch clusters (2 Datacenters ...)?

Yes, vSAN clusters within same or across datacenters would still ensure certain IO performance like local storage.

Can we host cluster nodes on different vSAN Clusters?

Yes, Different nodes of a Platform cluster could be hosted on different underlying datastores.

Do you need to backup platform nodes?

Yes, backups should be done using VMware recommended snapshot/backup technologies (as of now). vRNI cluster High-availability feature is in road map which will remove the need for external backups.

How to estimate the bandwidth between the cluster proxy VM on a region and the platform VM cluster on another region?

In some large deployments, we have seen this number ranging from 1 mbps to 20 mbps. There is much of deduplication or compression that happens in Proxy VM before data is sent to Platform VM.

How much network traffic will be between cluster node?

Traffic usually depends on size of cluster & type of datacenter environment.

For installations with 30-50k VMs:

- Between clusters: 50-400Mbps approx.
- Between proxy & platform: 100Kbps-15Mbps approx.

What is the maximum admissible latency between nodes in a cluster?

The platform nodes have to be co-located in the same site. In such cases, the latency is minimal. If the platform nodes are hosted on vSAN stretch clusters (two data centers), the vSAN clusters within or across the clusters ensure certain IO performance like local storage. The applications running on data centers such as vRealize Network Insight work fine. You can host different nodes of a platform cluster on different underlying datastores. But you need to ensure that all the platform VMs in a cluster are co-located within the same site.

What is the maximum admissible latency between the proxy VMs on a region and the platform VM cluster on another region?

You can have geo-distributed proxies in your setup. There is an HTTPS connection from Proxy VM to Platform VM so it can tolerate high latencies, to order of few seconds. vRealize Network Insight supports maximum of 10 nodes in a cluster (30,000 VMs w/ flows Or 50,000 VMs without flows).

What should be size of proxy/platform VM?

Use large brick configuration: refer installation guide.

Clustering - Scaling

Can I extend an already created cluster?

Yes, extending a cluster is supported up to 10 nodes.

What happens if a non-primary platform VM becomes unavailable?

Internal services have limited resiliency to non-primary node failures. In general, due to a node failure NI loses compute power.

What kind of load balancing is supported?

Mapping of proxy to platform is fixed. Once data from any Proxy VM reaches any Platform VM, its processing is load-balanced internally among all the Platform VMs.

Will creating a platform cluster increase bandwidth consumption?

The proxy or collector VMs continue to talk only to primary or platform VM. The bandwidth requirement for platform VM clustering communication is minimal. Therefore there is no significant increase in bandwidth consumption.

What is the frequency of data transmission between proxy VM to platform VM?

The proxy VM sends the deduplicated or compressed data continuously to the platform VM.

Does any optimization of data take place in the proxy VM?

Various deduplications, compressions, reduction, or batching steps happen in the proxy VM. When the connection between the platform VM and the proxy VM is down, the proxy VM stores data locally (depending on the disk space) and sends it whenever the connection is restored.

Is there any optimization done for network bandwidth?

Yes, various deduplication/compressions/reduction/batching steps happen on the proxy VM.

Is clustering possible on the proxy servers?

No. Clustering is not possible on the proxy servers.

How does vCenter send traffic to the proxy server?

vCenters do not send traffic to the proxy server. The proxy servers actually connect to their designated vCenter to fetch the information.

When deploying a cluster , how does vCenter send traffic to the various proxy servers?

Actually proxies get connected to vCenter to fetch the information. Respective proxy will connect to designated vCenter & fetch the information. There is no clustering available on proxy

Clustering - Deployment

How do I access UI after scaling out the cluster?

The UI access is restricted from Platform1 only.

What is Platform1 and why do I need to remember this node?

The platform node from which cluster creation process is initiated is treated as **Platform1**. The UI should be accessed only from this node out of the n nodes in cluster.

How is data retrieved from the other nodes in a cluster if the UI access is restricted to platform1?

The data of the datacenter is distributed across all nodes in a cluster. And when the UI layer requests data on platform1, the platform1 node gets the data stored on all nodes and sends a response to the UI.

Can I use a platform node which is deployed in different data center for creating clusters?

All nodes in a cluster exchange data between them. So, to avoid latency issues, it is recommended to use the platform nodes deployed in the same data center to create a cluster

What happens to data on existing platform when I scale out the platform node?

The data on an existing platform node is preserved and distributed across all nodes in a cluster.

Does the number of proxy VMs matter in determining how many platform bricks I need?

No. Only the total number of VMs across all vCenters and status of the flows (enabled or disabled) have impact on number of bricks needed. Refer to the brick model table in the *vRealize Network Insight Installation Guide*.

Does the number of vCenters or the number of physical devices (like routers) or any other type of data sources have impact on the number of platform bricks I need?

No. Only the total number of VMs across all vCenters and status of the flows (enabled or disabled) have impact on the number of bricks needed. Refer to the brick model table in the *vRealize Network Insight Installation Guide*.

Does vRNI support platform cluster distributed across 2 datacenters for HA reasons?

No. The platform Cluster doesn't support splitting across datacenters. All Platform Cluster VMs should be in same site. Platform Cluster doesn't support HA today. It is on the roadmap. The customers can use SRM for HA against DR across 2 sites.

Does vRNI support single vCenter with more than 6000 VMs and flows enabled?

Up to Release 3.5, the vRNI proxies don't support collecting data from a single large vCenter with more than 6000 VMs with flows. This is on the roadmap.

How much disk space is needed on Platform1?

Platform1 requires more disk space compared to other nodes in cluster as some of the configuration data is stored on Platform1 only.

What happens if any of the node ran out of disk space?

The UI starts showing error messages when disk space on any particular platform node reaches a certain threshold. Add more disk space to the platform node by logging in to vCenter.

How many times data is replicated in cluster?

The data replication mechanism depends on the components present in the platform node.

How do the clusters work?

- All proxies in a deployment connect to one platform (Platform1). The connectivity between platform and proxy is through https on port 443. So only port 443 is visible to proxies from Platform1.
- Upon receiving the requests from proxy, Platform1 node load balances requests to other platform nodes in cluster in round robin fashion.
- The platform node normalizes the data and put them in messaging queue for processing by computation engine.
- The computation engine distributes the data across all nodes in cluster by using data replication mechanism. That way there won't be any data loss if any of the node (except Platform1) goes down in cluster.
- Some of the configuration data is stored explicitly on Platform1 node that is not replicated. That's the reason why the high availability solution is not supported.

Data Management and Processing

7

How does data processing pipeline behave in boundary conditions such as when platform-proxy server communication breaks?

- What is the default retention period?

30 days. It can be increased from UI with Enterprise License. Note: When increasing make sure to follow disk guidelines.

- How data is handled on Proxy?

All data on proxy is converted to SDM (Self Describing Message) before sending it to platform including flow data. It includes all config, inventory and metric data from any data source. If platform is not reachable or SDM upload to Kafka queue fails then they are written on disk on proxy vm (under /var/BLOB_STORE).

- When will data start to purge on Proxy?

For non-flow data: There is 10GB space allocated to store SDMs on disk (BLOB_STORE). When this store fills, collector starts deleting older SDMs and adds new SDMs to the disk. It depends on the size of data gathered from all data sources how quickly this limit is breached.

For flow data: There is 15 GB space allocated to store raw flows (under /var/flow/vds/nfcapd). As soon as this space is consumed flow processor starts deleting older flow files. At incoming raw flows rate of ~2M/min it would take ~10hrs till rotation start to occur.

- What is the purge logic?

Oldest SDMs get deleted first.

- When will new data stop being processed in Proxy?

Never, as long as services are running properly.

- Assuming disconnect between Platform and Proxy and No purge condition met, would all data be reconciled on Platform on re-connect?

All data stored on disk will be sent to platform. It should be reconciled completely except if data loss conditions exist on platform (more info below).

- What are the conditions when data loss can occur on Platform?

Platform starts to drop SDMs that are on Kafka queue for more than 6hrs (18hrs if it is a 3-node cluster). Another possibility is if the queue is saturated. It can happen when there is Lag built up in system and incoming data rate is high.

- Will latest SDM be published first or earliest one in that order?

Oldest SDMs are sent first. There is one known issue until v3.9 which will result in some data loss. Contact GSS for more information.

- Is data stored on disk in Proxy and then pushed to Platform when there is no communication problem?

If there is no communication issue then SDMs are not stored on disk. It is sent to platform from memory itself. Whenever proxy receives that there was a problem in sending SDM then only it is stored on disk.

- In event of any issue how proxy learn which was last processed flow file?

Flow-processor maintains bookmark in DB on which was last processed nfcapd file.

- What is the max size of SDM that can be processed without any issue? How can user learn about this breach?

There is 15MB limit on SDM size. Starting v3.9 an event is raised whenever platform drops large SDM.

What is IPFIX?

IPFIX is an IETF protocol for exporting flow information. A flow is defined as a set of packets transmitted in a specific timeslot, and sharing 5-tuple values - source IP address, source port, destination IP address, destination port, and protocol. The flow information may include properties such as timestamps, packets/bytes count, Input/output interfaces, TCP Flags, VXLAN ID, Encapsulated flow information and so on. This is often referred to as Netflow. However, IPFIX is the standard IETF protocol.

What flow information is exported by the VDS?

A VDS in vSphere environment can be configured to export flow information using IPFIX. Enable flow monitoring on all the port groups attached to the VDS. If packets arrive on port X of a VDS and exit from port Y, a corresponding flow record is emitted if flow monitoring is enabled on port Y. The direction of every flow record is set as "Egress".

How does vRealize Network Insight use IPFIX?

vRealize Network Insight uses VMware VDS IPFIX to collect network traffic data. Every session has two paths. For example: Session A↔C has A→C packets and C→A packets. To analyze the complete information of any session, IPFIX data about packets in both the directions is required. Refer following diagram where VM-A is connected to DVPG-A and is talking to VM-C. Here DVPG-A will only provide data about the C→A packets, and DVPG-Uplink will provide data about A→C packets. To get the complete information of A's traffic, IPFIX should be enabled on DVPG-A, DVPG-uplink.

How do I troubleshoot vRealize Network Insight Flow Collection?

- 1 Please ensure that the specific VDS and its DVPGs and Uplink properties has Netflow monitoring **Enabled** and the collector IP address is that of vRealize Network Insight Collector.

- 2 IPFIX Netflow packets getting dropped in between by a firewall (NSX, Virtual or Physical). Please ensure that the Netflow packets destined for UDP port 2055 on vRealize Network Insight Collector IP is allowed by any firewall that may be present in the route between ESXi Host and the vRealize Network Insight Collector.
- 3 The ESXi host has ceased to send IPFIX Netflow packets. The ESXi host backs off sending the Netflow packets after some time if UDP port 2055 is not reachable. This may happen due to firewall dropping the packets.
- 4 The vRealize Network Insight Collector is not reachable by ESXi Host due to network routing problem. Please ensure that the proper route exist between ESXi Host and the vRealize Network Insight Collector.

Which VMware KB articles should I be aware of, related to IPFIX?

VMware ESXi 6.0 Update 1:

[2135956](#)