Editing AI Routed Fabric Settings, Release 4.3.1

Table of Contents

New and changed information . . . 1
Editing AI Routed fabric settings . . . 2
General . . . 3
Fabric Management . . . 4
  General Parameters . . . 4
  vPC . . . 7
  Protocols . . . 8
  Security . . . 9
  Advanced . . . 10
  Freeform . . . 15
  Manageability . . . 15
  Bootstrap . . . 16
  Configuration Backup . . . 17
  Flow Monitor . . . 18
Telemetry . . . 20
  Edit configuration settings . . . 21
  Edit NAS settings . . . 21
  Disable Telemetry . . . 22
  Perform force disable telemetry on your fabric . . . 22
  NAS . . . 22
  Guidelines and limitations for network attached storage . . . 23
  Add network attached storage to export flow records . . . 23
    Add NAS to Nexus Dashboard . . . 23
    Add the onboarded NAS to Nexus Dashboard . . . 24
  Flow collection . . . 26
  Understanding flow telemetry . . . 26
    Flow telemetry guidelines and limitations . . . 26
    Configure flows . . . 28
    Monitor the subnet for flow telemetry . . . 31
  Understanding Netflow . . . 32
    Understanding Netflow types . . . 32
    Netflow guidelines and limitations . . . 33
    Configure Netflow . . . 34
  Understanding sFlow . . . 35
    Guidelines and limitations for sFlow . . . 35
    Configure sFlow telemetry . . . 35
External streaming . . . 37
  Configure external streaming settings . . . 37
  Guidelines and limitations . . . 38
    Guidelines and limitations in NX-OS fabrics . . . 38
  Email . . . 38

Message bus . . . 40
Add Kafka broker configuration . . . 40
Configure Kafka exports in fabric settings . . . 42
Anomalies . . . 44
Advisories . . . 45
Statistics . . . 45
Faults . . . 46
Audit Logs . . . 46
Syslog . . . 46
Guidelines and limitations for syslog . . . 47
Add syslog server configuration . . . 47
Configure syslog to enable exporting anomalies data to a syslog server . . . 47
Additional settings . . . 49
Retrieving the authentication key . . . 49
Retrieve the 3DES encrypted OSPF authentication key . . . 49
Retrieve the encrypted IS-IS authentication key . . . 49
Retrieve the 3DES encrypted BGP authentication key . . . 50
Retrieve the encrypted BFD authentication key . . . 50
Guidelines for VXLAN Fabric With eBGP Underlay . . . 51
Adding Switches . . . 52
Support for Cisco N9164E-NS4-O and Cisco N9348Y12C-SE1 switches . . . 52
Assigning Switch Roles . . . 52
Understanding switches running SONiC . . . 52
Switches running SONiC: On AI Routed and Routed fabrics . . . 53
Switches running SONiC: External fabrics . . . 55
Switches running SONiC: General information . . . 55
Troubleshooting issues . . . 55
AI QoS classification and queuing policies . . . 56
Understanding AI QoS classification and queuing policies . . . 56
Guidelines and limitations for AI QoS classification and queuing policies . . . 56
Configure AI QoS classification and queuing policies . . . 57
Create a policy using the custom QoS templates . . . 58
Border Gateway Support . . . 59
Border Gateway Route Filtering . . . 59
Guidelines and Limitations . . . 61
Layer 3 VNI without VLAN . . . 61
Guidelines and limitations for Layer 3 VNI without VLAN . . . 61
MACsec support in Data Center VXLAN EVPN and BGP fabrics . . . 62
Guidelines . . . 62
Enable MACsec . . . 63
Disable MACsec . . . 68
vPC fabric peering . . . 68
Guidelines and limitations . . . 68

QoS for fabric vPC-peering . . . 69
Create a virtual peer link . . . 71
Convert a physical peer link to a virtual peer link . . . 72
Convert a virtual peer link to a physical peer link . . . 73
Deploying Fabric Underlay eBGP Policies . . . 74
Deploying Fabric Overlay eBGP Policies . . . 74
Deploying Spine Switch Overlay Policies . . . 74
Deploying Leaf Switch Overlay Policies . . . 75
Adding a Super Spine Switch to an Existing VXLAN BGP EVPN Fabric . . . 76
Understanding enhanced analytics for AI fabrics . . . 77
AI endpoint and topology discovery . . . 78
Prerequisites . . . 78
Guidelines and limitations: Analytics for AI fabrics . . . 79
Add Slurm as an integration . . . 80
DCGM Exporter . . . 80
Configure DOCA Telemetry Service and Fluent Bit . . . 82
GPU/NIC Topology Fleet Collector . . . 94
Collect data in secured mode . . . 98
Associate a fabric with an AI cluster . . . 99
Associate a switch with a scalable unit . . . 100
View information on AI resources . . . 102
View AI job information . . . 110
Copyright . . . 128

New and changed information

The following table provides an overview of the significant changes up to this current release. The table does not provide an exhaustive list of all changes or of the new features up to this release.

Release Version	Feature	Description
Nexus Dashboard 4.3.1	Support for switches running SONiC images	Support is now available for certain switches running SONiC (Software for Open Networking in Cloud) images in these Nexus Dashboard fabric types: • AI Routed (this article) • Routed • External For more information, refer to Understanding switches running SONiC.
Nexus Dashboard 4.3.1	AI job monitoring with server NIC statistics	Beginning with Nexus Dashboard 4.3.1, AI job monitoring includes server network interface card (NIC) statistics and supports external AI fabrics. You can correlate AI jobs with GPU, switch-interface, and server-interface data, including CRC errors, discards, bandwidth, Remote Direct Memory Access (RDMA) negative acknowledgment (NACK), congestion notification packet (CNP), operational status, and link speed. For more information, refer to Understanding enhanced analytics for AI fabrics.
Nexus Dashboard 4.3.1	Enhanced border node functions and routing isolation for multi-tier VXLAN fabrics	Beginning with Nexus Dashboard 4.3.1, you can use unique ASNs on applicable border switches in AI routed fabrics when the fabric uses eBGP and is configured in Same-Tier-AS mode. For more information, refer to General Parameters.
Nexus Dashboard 4.3.1	Support for Cisco N9164E-NS4-O and N9348Y12C-SE1 switches	Starting with Nexus Dashboard 4.3.1, Nexus Dashboard adds support for Cisco N9164E-NS4-O and N9348Y12C-SE1 switches. For more information, refer to Support for Cisco N9164E-NS4-O and Cisco N9348Y12C-SE1 switches.
Nexus Dashboard 4.3.1	Cisco Nexus 9000 Series name change	Cisco Nexus 9000 Series is now the Cisco N9000 Series.

Editing AI Routed fabric settings

An **AI Routed** fabric is a type of fabric that provides automated provisioning of AI ready BGP Layer 3 networks on Cisco Nexus (NX-OS) devices.

When you first create an AI Routed fabric using the procedures provided in Creating Fabrics and Fabric Groups, the standard workflow allows you to create a fabric using the bare minimum settings so that you are able to create a fabric quickly and easily. Use the procedures in this article to make more detailed configurations for your routed fabric.

1. Navigate to the main Fabrics window:

   Manage > Fabrics

2. Locate the fabric that you want to edit.
3. Click the circle next to the routed fabric that you want to edit to select that fabric, then click **Actions > Edit Fabric Settings**.

   The **Edit *fabric_name* Settings** window appears.

4. Click the appropriate tab to edit these settings for the fabric:
   - General
   - Fabric Management
   - Telemetry (if the **Telemetry** feature is enabled for the fabric)

General

Use the information in this section to edit the settings in the **General** window for your routed fabric.

Change the general parameters that you configured previously for the routed fabric, if necessary, or click another tab to leave these settings unchanged.

Fabric type	Description
Name	The name for the fabric. This field is not editable.
Type	The fabric type for this fabric. This field is not editable.
Location	Choose the location for the fabric.
BGP ASN for Spines	Enter the BGP autonomous system number (ASN) for the fabric's spine switches.
Network operating system	A non-editable field that displays whether you chose **NX-OS** or **SONiC** as the network operating system when you created this fabric.
AI QOS & Queuing policy	Choose the queuing policy from the drop-down list based on the predominant fabric link speed for certain switches in the fabric. For more information, see AI QoS classification and queuing policies. Options are: • **400G**: Enable QoS queuing policies for an interface speed of 400 Gb. This is the default value. • **100G**: Enable QoS queuing policies for an interface speed of 100 Gb. • **25G**: Enable QoS queuing policies for an interface speed of 25 Gb.
License tier	Choose the licensing tier for the fabric: • **Essentials** • **Advantage** • **Premier** Click on the information icon (i) next to License tier to see what functionality is enabled for each license tier.
Enabled features	Check the **Telemetry** check box to enable telemetry for the fabric. This is the equivalent of enabling the Nexus Dashboard Insights service in previous releases.
Telemetry collection	This option becomes available if you choose to enable **Telemetry** in the **Enabled features** field above. Choose either **Out-of-band** or **In-band** for telemetry collection.
Telemetry streaming	This option becomes available if you choose to enable **Telemetry** in the **Enabled features** field above. Choose either **IPv4** or **IPv6** for telemetry streaming.
Security domain	Choose the security domain for the fabric.

Fabric Management

Use the information in this section to edit the settings in the **Fabric Management** window for your routed fabric. The tabs and their fields in the screen are explained in the following sections. The fabric-level parameters are included in these tabs.

If you chose **SONiC** as the network operating system for this fabric, then these tabs are not applicable and are therefore not shown:

- vPC
- Authentication
- Security
- Freeform
- Configuration Backup
- Flow Monitor

In addition, for tabs that are shown for SONiC configurations, certain fields in some tabs might not be shown. For example, even though the **General Parameters** tab is shown for SONiC configurations, the **Super Spine ASN** field under **General Parameters** will not be shown since the super spine switch role is not supported in a fabric with a SONiC operating system.

- General Parameters
- vPC
- Protocols
- Security
- Advanced
- Freeform
- Manageability
- Bootstrap
- Configuration Backup
- Flow Monitor

General Parameters

The **General Parameters** tab is displayed, by default. The fields in this tab are described in the following table.

Field	Description
First Hop Redundancy Protocol	This field is available if you did not select the **Enable EVPN VXLAN Overlay** option above. This field is only applicable to a routed fabric. Specify the First Hop Redundancy Protocol that you want to use: - **hsrp**: Hot Standby Router Protocol - **vrrp**: Virtual Router Redundancy Protocol After a network has been created, you cannot change this fabric setting. You should delete all networks, and then change the First Hop Redundancy Protocol setting.
BGP ASN for Super Spines	Enter the ASN used for super spine and border super spines, if the fabric contains any super spine or border super spine switches.
BGP ASN for Leafs	Enter the ASN used for leaf switches, if the fabric contains any leaf switches.
BGP ASN for Border and Border Gateway switches	Enter the ASN used for border and border gateway switches, if the fabric contains any border and border gateway switches. If the **Allow Unique ASN on Border** option is enabled, this field is unavailable because the ASN for each applicable switch is configured individually by using the leaf_bgp_asn policy.
BGP AS Mode	Choose **Multi-AS** or **Same-Tier-AS**. - In a **Multi-AS** fabric, a unique AS number per leaf or border switch is used. - In a **Same-Tier-AS** fabric, all leaf nodes share one unique AS and all border nodes share another unique AS by default. In both **Multi-AS** and **Same-Tier-AS**, all spine switches in a fabric share one unique AS number. The fabric is identified by the spine switch ASN. If the **Allow Unique ASN on Border** option is enabled, you can use different ASNs for applicable border switches and border gateways even when the fabric is configured in **Same-Tier-AS** mode.
ASN Auto Allocation	Automatically allocates and tracks BGP ASN for leafs, borders and border gateways in Multi-AS mode.
BGP ASN range	Enter the BGP ASN range for auto-allocation. If not specified and ASN auto-allocation is enabled, Nexus Dashboard will automatically generate the ASN range (min: 1 or 1.0, max: 4294967295 or 65535.65535).
Allow Same ASN On Leafs	Uses the same ASN on all the leaf nodes even when you have configured Multi-AS mode.

Field	Description
Allow Unique ASN on Border	Check the check box to allow applicable border switches and border gateways to use different ASNs when the fabric is configured in **Same-Tier-AS** mode. When this option is enabled, the **BGP ASN for Border and Border Gateway switches** field is unavailable. You must configure the ASN for each applicable switch individually by using the leaf_bgp_asn policy. When the **Allow Unique ASN on Border** option is enabled, you must manually apply the leaf_bgp_asn policy to each applicable switch. When this option is enabled, the system also allows border switches and border gateway switches to use the same ASN without generating a validation error. If this option is later disabled, no validation error is generated for existing switches that already use different ASNs. Validation is enforced only when the leaf_bgp_asn policy is created or updated.
Enable IPv6 routed fabric or VXLAN with IPv6 underlay	Enables IPv6 routed fabric or IPv6 underlay. With the checkbox cleared, the system configures IPv4 routed fabric or IPv4 underlay. To configure IPv6 underlay, you must also configure the VXLAN overlay parameters in the **EVPN** tab.
Underlay Subnet IP Mask	Specifies the subnet mask for the fabric interface IP addresses.
Manual Underlay IP Address Allocation	Check the check box to disable dynamic underlay IP address allocations.
Underlay Routing Loopback IP Range or Router ID Range	Specifies the loopback IPv4 addresses for protocol peering.
Assign IPv4 address to routing loopback	In an IPv6 routed fabric or VXLAN EVPN fabric with IPv6 underlay, assign IPv4 address used for BGP router ID to the routing loopback interface.
Underlay Subnet IP Range	Specifies the IPv4 addresses for underlay P2P routing traffic between interfaces.
Underlay Routing Loopback IPv6 Range	Specifies the loopback IPv6 addresses for protocol peering.
Disable Route-Map Tag	Disables subnet redistribution.
Route-Map Tag	Configures a route tag for redistributing subnets. By default, the tag value of 12345 is configured, when enabled.
Subinterface Dot1q Range	Specifies the subinterface range when Layer 3 sub interfaces are used.

Field	Description
Enable Performance Monitoring	Check the check box to enable performance monitoring. Ensure that you do not clear interface counters from the command-line interface of the switches. Clearing interface counters can cause the **Performance Monitor** to display incorrect data for traffic utilization. If you must clear the counters and the switch has both clear counters and clear counters snmp commands (not all switches have the clear counters snmp command), ensure that you run both the main and the SNMP commands simultaneously. For example, you must run the clear counters interface ethernet slot/port command followed by the clear counters interface ethernet slot/port snmp command. This can lead to a one-time spike.  Performance monitoring is supported on switches with NX-OS Release 9.3.6 and later.
Network VLAN Range	VLAN ranges for the Layer 3 VRF and overlay network.

vPC

Field	Description
vPC Peer Link VLAN	VLAN used for the vPC peer link SVI.
Make vPC Peer Link VLAN as Native VLAN	Enables vPC peer link VLAN as Native VLAN.
vPC Peer Keep Alive option	From the drop-down list, select **management** or **loopback**. To use IP addresses assigned to the management port and the management VRF, select **management**. To use IP addresses assigned to loopback interfaces (in non-management VRF), select **loopback**. If you use IPv6 addresses, you must use loopback IDs.
vPC Auto Recovery Time	Specifies the vPC auto recovery time-out period in seconds.
vPC Delay Restore Time	Specifies the vPC delay restore period in seconds.
vPC Peer Link Port Channel Number	Specifies the Port Channel ID for a vPC Peer Link. By default, the value in this field is 500.
vPC IPv6 ND Synchronize	Enables IPv6 Neighbor Discovery synchronization between vPC switches. The check box is enabled by default. Uncheck the check box to disable the function.
Fabric wide vPC Domain Id	Enables the usage of same vPC Domain Id on all vPC pairs in the fabric. When you select this field, the **vPC Domain Id** field is editable.
vPC Domain Id	Specifies the vPC domain ID to be used on all vPC pairs. Otherwise unique vPC domain IDs are used (in increment of 1) for each vPC pair.

Field	Description
Enable Qos for Fabric vPC-Peering	Enables QoS on spines for guaranteed delivery of vPC Fabric Peering communication.  QoS for vPC fabric peering and queuing policies options in fabric settings are mutually exclusive.
Qos Policy Name	Specifies QoS policy name that should be same on all spines.

Protocols

Field	Description
Routing Loopback Id	The loopback interface ID is populated as 0 by default. It is used as the BGP router ID.
VTEP Loopback Id	The loopback interface ID is populated as 1 and it is used for VTEP peering purposes.
BGP Maximum Paths	Specifies maximum number for BGP routes to be installed for same prefix on the switches for ECMP.
Enable BGP Authentication	Check the check box to enable BGP authentication. If you enable this field, the **BGP Authentication Key Encryption Type** and **BGP Authentication Key** fields are enabled.
BGP Authentication Key Encryption Type	Choose the three for 3DES encryption type, or seven for Cisco encryption type.
BGP Authentication Key	Enter the encrypted key based on the encryption type.  Plain-text passwords are not supported. Log on to the switch, retrieve the encrypted key. Enter the key in the **BGP Authentication Key** field. For more information, see Retrieve the encrypted BFD authentication key.
Enable PIM Hello Authentication	Enables the PIM hello authentication.
PIM Hello Authentication Key	Specifies the PIM hello authentication key.

Field	Description

Enable BFD	Check the **Enable BFD** check box to enable **feature bfd** on all switches in the fabric. This feature is valid only on IPv4 underlay and the scope is within a fabric.
	
	Nexus Dashboard supports BFD within a fabric. The BFD feature is disabled by default in the Fabric Settings. If enabled, BFD is enabled for the underlay protocols with the default settings. Any custom BFD configurations requires configurations to be deployed via the per switch freeform or per interface freeform policies.
	
	The following configuration is pushed after you enable BFD.
	
	'feature bfd'
	
	For Nexus Dashboard with BFD-enabled, the following configurations are pushed on all the P2P fabric interfaces:
	
	`no ip redirects`
	`no ipv6 redirects`
	
	For information about BFD feature compatibility, refer your respective platform documentation and for information about the supported software versions, see *Cisco Nexus Dashboard Fabric Controller Compatibility Matrix*.

Enable BFD for BGP	Check the check box to enable BFD for the BGP neighbor. This option is disabled by default.
Enable BFD Authentication	Check the check box to enable BFD authentication. If you enable this field, the **BFD Authentication Key ID** and **BFD Authentication Key** fields are enabled.
BFD Authentication Key ID	Specifies the BFD authentication key ID for the interface authentication.
BFD Authentication Key	Specifies the BFD authentication key. For information about how to retrieve the BFD authentication parameters, see Retrieve the encrypted BFD authentication key.

Security

Field	Description
Enable MACsec	Check the check box to enable MACsec in the fabric. MACsec configuration is not generated until MACsec is enabled on an intra-fabric link. Perform a **Recalculate and deploy** operation to generate the MACsec configuration and deploy the configuration on the switch.

Field	Description
MACsec Cipher Suite	Choose one of the following MACsec cipher suites for the MACsec policy: · **GCM-AES-128** · **GCM-AES-256** · **GCM-AES-XPN-128** · **GCM-AES-XPN-256** The default value is **GCM-AES-XPN-256**.
MACsec Primary Key String	Specify a Cisco Type 7 encrypted octet string that is used for establishing the primary MACsec session. For **AES_256_CMAC**, the key string length must be 130 and for **AES_128_CMAC**, the key string length must be 66. If these values are not specified correctly, an error is displayed when you save the fabric.  The default key lifetime is infinite.
MACsec Primary Cryptographic Algorithm	Choose the cryptographic algorithm used for the primary key string. It can be **AES_128_CMAC** or **AES_256_CMAC**. The default value is **AES_128_CMAC**. You can configure a fallback key on the device to initiate a backup session if the primary session fails.
MACsec Fallback Key String	Specify a Cisco Type 7 encrypted octet string that is used for establishing a fallback MACsec session. For **AES_256_CMAC**, the key string length must be 130 and for **AES_128_CMAC**, the key string length must be 66. If these values are not specified correctly, an error is displayed when you save the fabric.
MACsec Fallback Cryptographic Algorithm	Choose the cryptographic algorithm used for the fallback key string. It can be **AES_128_CMAC** or **AES_256_CMAC**. The default value is **AES_128_CMAC**.
MACsec Status Report Timer	Specify the MACsec operational status periodic report timer in minutes.

Advanced

Field	Description
Intra Fabric Interface MTU	Specifies the MTU for the intra fabric interface. This value must be an even number.
Layer 2 Host Interface MTU	Specifies the MTU for the Layer 2 host interface. This value must be an even number.
Power Supply Mode	Choose the appropriate power supply mode.
CoPP Profile	From the drop-down list, select the appropriate Control Plane Policing (CoPP) profile policy for the fabric. By default, the **strict** is selected.
VTEP HoldDown Time	Specifies the NVE source interface hold down time.

Field	Description
VRF Lite Subnet IP Range	These fields are prefilled with the DCI subnet details. Update the fields as needed. The values shown on the page are automatically generated. If you want to update the IP address ranges, VXLAN Layer 2/Layer 3 network ID ranges or the VRF/network VLAN ranges, ensure that each fabric has its own unique range and is distinct from any underlay range to avoid possible duplication. You should only update one range of values at a time.
VRF Lite Subnet Mask	If you want to update more than one range of values, do it in separate instances. For example, if you want to update Layer 2 and Layer 3 ranges, you should do the following. 1. Update the Layer 2 range and click **Save**. 2. Click the **Edit Fabric** option again, update the Layer 3 range, and click **Save**.
Enable CDP for Bootstrapped Switch	Check the check box to enable CDP for switches discovered using Bootstrap.
Enable NX-API	Check the check box to enable NX-API on HTTPS. This check box is checked by default.
Enable NX-API on HTTP	Specifies enabling of NX-API on HTTP. Check **Enable NX-API on HTTP** and **Enable NX-API** check boxes to use HTTP. This check box is checked by default. If you uncheck this check box, the applications that use NX-API and supported by Nexus Dashboard, such as Endpoint Locator (EPL), Layer 4-Layer 7 services (L4-L7 services), VXLAN OAM, and so on, start using the HTTPS instead of HTTP. If you check both **Enable NX-API** and **Enable NX-API on HTTP** check boxes, applications use HTTP.
Enable Strict Config Compliance	Enable the Strict Configuration Compliance feature by selecting this check box. For more information, see the section "Strict Configuration Compliance" in Configuration Compliance.
Enable AAA IP Authorization	Enables AAA IP authorization (make sure IP Authorization is enabled in the AAA Server).
Enable Nexus Dashboard as Trap Host	Check the check box to enable Nexus Dashboard as a trap host.
Enable TCAM Allocation	TCAM commands are automatically generated for VXLAN and vPC Fabric Peering when enabled.
Greenfield Cleanup Option	Enable the switch cleanup option for greenfield switches without performing a switch reload. This option is typically recommended only for the Data centers with the Cisco N9000v Switches.

Field	Description
Enable Default Queuing Policies	Check the check box to apply QoS policies on all the switches in this fabric. To remove the QoS policies that you applied on all the switches, uncheck this check box, update all the configurations to remove the references to the policies, and deploy the configuration. Pre-defined QoS configurations are included that can be used for various Cisco N9000 Series Switches. When you check this check box, the appropriate QoS configurations are pushed to the switches in the fabric. The system queuing is updated when configurations are deployed to the switches. You can perform the interface marking with defined queuing policies, if required, by adding the necessary configuration to the peer interface freeform block. Review the actual queuing policies by opening the policy file in the template editor. From the Nexus Dashboard Web UI, choose **Operations > Template**. Search for the queuing policies by the policy file name, for example, **queuing_policy_default_8q_cloudscale**. Choose the file and click the **Modify/View template** icon to edit the policy. See the *Cisco N9000 Series NX-OS Quality of Service Configuration Guide* for platform specific details.
N9K Cloud Scale Platform Queuing Policy	Choose the queuing policy from the drop-down list to be applied to all Cisco N9200 Series Switches and the Cisco N9000 Series Switches that ends with EX, FX, and FX2 in the fabric. The valid values are **queuing_policy_default_4q_cloudscale** and **queuing_policy_default_8q_cloudscale**. Use the **queuing_policy_default_4q_cloudscale** policy for FEXs. You can change from the **queuing_policy_default_4q_cloudscale** policy to the **queuing_policy_default_8q_cloudscale** policy only when FEXs are offline.
N9K R-Series Platform Queuing Policy	Select the queuing policy from the drop-down list to be applied to all Cisco Nexus switches that ends with R in the fabric. The valid value is **queuing_policy_default_r_series**.
Other N9K Platform Queuing Policy	Choose the queuing policy from the drop-down list to be applied to all other switches in the fabric other than the switches mentioned in the above two options. The valid value is **queuing_policy_default_other**.
Advanced QoS Configuration	Check the box to enable AI QoS and queuing policies in the BGP fabric. For more information, see AI QoS classification and queuing policies.  This option is not available if you also enabled either of the following options: · **Enable Qos for Fabric vPC-Peering** option in the **vPC** tab · **Enable Default Queuing Policies** option in the **Advanced** tab

Field	Description
AI QoS & Queuing Policy	This field is available if you checked the **Advanced QoS Configuration** option above. Choose the queuing policy from the drop-down list based on the predominant fabric link speed for certain switches in the fabric. For more information, see AI QoS classification and queuing policies. Options are: · **800G**: Enable QoS queuing policies for an interface speed of 400 Gb. · **400G**: Enable QoS queuing policies for an interface speed of 400 Gb. · **100G**: Enable QoS queuing policies for an interface speed of 100 Gb. · **25G**: Enable QoS queuing policies for an interface speed of 25 Gb. · **User-defined**: Provides customization options for parameters that are fixed in port speed templates.
 The **ROCEv2 DSCP** through **Bandwidth remaining %** fields are editable if you choose **User-defined** in the **AI QoS & queuing policy** field.	
ROCEv2 DSCP	Enter the DSCP value (0–63) for RDMA traffic. The default value is 26.
CNP DSCP	Enter the DSCP value (0–63) for congestion notification. The default value is 48.
WRED minimum threshold (in kbytes)	Enter the minimum queue size for WRED. The default value is 950.
WRED maximum threshold (in kbytes)	Enter the maximum queue size for WRED. The default value is 3000.
Drop probability %	Enter the value to control WRED. The default value is 7.
WRED weight	Enter the value to control WRED, which influences queue changes. The default value is 0.
Bandwidth remaining %	Enter the percentage of remaining bandwidth allocated to AI traffic queue. The default value is 50.
Enable Dynamic load balancing	Click to **Enable** the options. Choose the **DLB Mode** from the drop-down list and enter the aging period in the **Flowlet aging timer** field. For more information, see Add a Dynamic Load Balancing (DLB) policy template.  · Flowlet aging is applicable when you choose the **flowlet**, **policy-driven-flowlet**, or **policy-driven-mixed-mode** mode option from the **DLB Mode** drop-down list when you add a Dynamic Load Balancing (DLB) policy template. It is not effective when you choose the **per-packet** or **policy-driven-per-packet** mode option from the **DLB Mode** drop-down list. · The NX-OS CLI does accept a flowlet aging value, regardless of the mode; however, it may have a no operation result based on the mode choice.

Field	Description
Enable IP Load Sharing	Click to enable IP load sharing using source and destination addresses for AI workloads.
Enable MACsec	Check the box to enable MACsec in the BGP fabric. For more information, see Enable MACsec.
MACsec Primary Key String	Specify a Cisco Type 7 encrypted octet string that is used for establishing the primary MACsec session. For AES_256_CMAC, the key string length must be 130 and for AES_128_CMAC, the key string length must be 66. If these values are not specified correctly, an error is displayed when you save the fabric. The default key lifetime is infinite.
MACsec Primary Cryptographic Algorithm	Choose the cryptographic algorithm used for the primary key string. It can be **AES_128_CMAC** or **AES_256_CMAC**. The default value is **AES_128_CMAC**. You can configure a fallback key on the device to initiate a backup session if the primary session fails.
MACsec Fallback Key String	Specify a Cisco Type 7 encrypted octet string that is used for establishing a fallback MACsec session. For AES_256_CMAC, the key string length must be 130 and for AES_128_CMAC, the key string length must be 66. If these values are not specified correctly, an error is displayed when you save the fabric.
MACsec Fallback Cryptographic Algorithm	Choose the cryptographic algorithm used for the fallback key string. It can be **AES_128_CMAC** or **AES_256_CMAC**. The default value is **AES_128_CMAC**.
MACsec Cipher Suite	Choose one of the following MACsec cipher suites for the MACsec policy: • GCM-AES-128 • GCM-AES-256 • GCM-AES-XPN-128 • GCM-AES-XPN-256 The default value is **GCM-AES-XPN-256**. The MACsec configuration is not deployed on the switches after the fabric deployment is complete. You need to enable MACsec on intra-fabric links to deploy the MACsec configuration on the switch.
MACsec Status Report Timer	Specifies the MACsec operational status periodic report timer in minutes.
Leaf Freeform Config	Add CLIs that should be added to switches that have Leaf, Border, and Border Gateway roles.
Spine Freeform Config	Add CLIs that should be added to switches with Spine, Border Spine, and Border Gateway Spine roles.

Field	Description
Intra-fabric Links Additional Config	Add CLIs that should be added to the intra-fabric links.

Freeform

The fields in this tab are shown below. For more information, see "Enable freeform configurations on fabric switches" in Working with Inventory in Your Nexus Dashboard LAN or IPFM Fabrics.

Field	Description
Leaf Pre-Interfaces Freeform Config	Enter additional CLIs, added before interface configurations, for all Leafs and Tier2 Leafs as captured from Show Running Configuration.
Spine Pre-Interfaces Freeform Config	Enter additional CLIs, added before interface configurations, for all Spines as captured from Show Running Configuration.
Leaf Post-Interfaces Freeform Config	Enter additional CLIs, added after interface configurations, for all Leafs and Tier2 Leafs as captured from Show Running Configuration.
Spine Post-Interfaces Freeform Config	Enter additional CLIs, added after interface configurations, for all Spines as captured from Show Running Configuration.
Intra-fabric Links Additional Config	Add CLIs that should be added to the intra-fabric links.

Manageability

Field	Description
DNS Server IPs	Specifies the comma-separated list of IP addresses (v4/v6) of the DNS servers.
DNS Server VRFs	Specifies one VRF for all DNS servers or a comma separated list of VRFs, one per DNS server.
NTP Server IPs	Specifies the comma-separated list of IP addresses (v4/v6) of the NTP server.
NTP Server VRFs	Specifies one VRF for all NTP servers or a comma-separated list of VRFs, one per NTP server.
Syslog Server IPs	Specifies the comma-separated list of IP addresses (v4/v6) IP address of the syslog servers, if used.
Syslog Server Severity	Specifies the comma-separated list of syslog severity values, one per syslog server. The minimum value is 0 and the maximum value is 7. To specify a higher severity, enter a higher number.
Syslog Server VRFs	Specifies one VRF for all syslog servers or a comma-separated list of VRFs, one per syslog server.

Field	Description
AAA Freeform Config	Specifies the AAA freeform configs. If AAA configs are specified in the fabric settings, **switch_freeform** PTI with source as **UNDERLAY_AAA** and description as **AAA Configurations** will be created.

Bootstrap

The **Bootstrap** page is a group of settings that are mandatory if you chose SONiC as your network operating system when you created or edited this fabric.

Field	Description
Enable Bootstrap	Check the **Enable Bootstrap** check box to enable the bootstrap feature. After you enable bootstrap, you can enable the DHCP server for automatic IP address assignment using one of the following methods: · **External DHCP Server** - Enter information about the external DHCP server in the **Switch Mgmt Default Gateway** and **Switch Mgmt IP Subnet Prefix** fields. · **Local DHCP Server** - Check the **Local DHCP Server** check box and enter details for the remaining mandatory fields.
Enable Local DHCP Server	Check the **Enable Local DHCP Server** check box to enable DHCP service on Nexus Dashboard and initiate automatic IP address assignment. When you check this check box, the **DHCP Scope Start Address** and **DHCP Scope End Address** fields become editable. If you do not check this check box, Nexus Dashboard uses the remote or external DHCP server for automatic IP address assignment.
DHCP Version	Select **DHCPv4** or **DHCPv6** from this drop-down list. When you select DHCPv4, the **Switch Mgmt IPv6 Subnet Prefix** field is disabled. If you select DHCPv6, the **Switch Mgmt IP Subnet Prefix** is disabled. Nexus Dashboard IPv6 POAP is not supported with Cisco Nexus 7000 Series Switches. Cisco N9000 and Cisco Nexus 3000 Series Switches support IPv6 POAP only when switches are either Layer2 adjacent (eth1 or out-of-band subnet must be a /64) or Layer3 adjacent residing in some IPv6 /64 subnet. Subnet prefixes other than /64 are not supported.
DHCP Scope Start Address	Specifies the first and last IP addresses of the IP address range. IPs from this scope are allocated to the switches during the POAP bootstrap process.
DHCP Scope End Address	

Field	Description
Switch Mgmt Default Gateway	Specifies the default gateway for the DHCP scope.
Switch Mgmt IP Subnet Prefix	Specifies the prefix length for DHCP scope.
DHCP scope and management default gateway IP address specification	If you specify the management default gateway IP address 10.0.1.1 and subnet mask 24, ensure that the DHCP scope is within the specified subnet, between 10.0.1.2 and 10.0.1.254.
Switch Mgmt IPv6 Subnet Prefix	Specifies the IPv6 prefix for the Mgmt0 interface on the switch. The prefix should be between 112 and 126. This field is editable if you enable IPv6 for DHCP.
Enable AAA Config	Check the check box to include AAA configs from the **Manageability** tab during device bootup.
Bootstrap Freeform Config	Enter additional commands as needed. For example, if you are using AAA or remote authentication related configurations, you need to add these configurations in this field to save the intent. After the devices boot up, they contain the intent defined in the **Bootstrap Freeform Config** field. Copy-paste the running-configuration to a **freeform config** field with correct indentation, as seen in the running configuration on the NX-OS switches. The freeform config must match the running config. For more information, see the section "Resolve freeform config errors in switches" in Working with Inventory in Your Nexus Dashboard LAN or IPFM Fabrics.
DHCPv4/DHCPv6 Multi Subnet Scope	Specifies the field to enter one subnet scope per line. This field is editable after you check the **Enable Local DHCP Server** check box. The format of the scope should be defined as: **DHCP Scope Start Address, DHCP Scope End Address, Switch Management Default Gateway, Switch Management Subnet Prefix** For example: 10.6.0.2, 10.6.0.9, 10.6.0.1, 24

Configuration Backup

Field	Description
Hourly Fabric Backup	Check the **Hourly Fabric Backup** check box to enable an hourly backup of fabric configurations and the intent. You can enable an hourly backup for fresh fabric configurations and the intent. If there is a configuration push in the previous hour, Nexus Dashboard takes a backup. Intent refers to configurations that are saved in Nexus Dashboard but yet to be provisioned on the switches.

Field	Description
Scheduled Fabric Backup	Check the check box to enable a daily backup. This backup tracks changes in running configurations on the fabric devices that are not tracked by configuration compliance.
Scheduled Time	Specifies the scheduled backup time in a 24-hour format. This field is enabled if you check the **Scheduled Fabric Backup** check box. 1. Select both the check boxes to enable both back up processes. The backup process is initiated after you click **Save**.  Hourly and scheduled backup processes happen only during the next periodic configuration compliance activity, and there can be a delay of up to an hour. 2. To trigger an immediate backup, do the following: a. Choose **Overview > Topology**. b. Click within the specific fabric box. The fabric topology screen comes up. c. From the **Actions** pane at the left part of the screen, click **Re-Sync Fabric**. You can also initiate the fabric backup in the fabric topology window. Click **Backup Now** in the **Actions** pane. 3. Click **Save** after filling and updating relevant information.

Flow Monitor

Field	Description
Enable Netflow	Check the **Enable Netflow** check box to enable Netflow on VTEPs for this Fabric. By default, Netflow is disabled.  When Netflow is enabled on the fabric, you can choose not to have netflow on a particular switch by having a fake **no_netflow** PTI. If netflow is not enabled at the fabric level, an error message is generated when you enable netflow at the interface, network, or VRF level. For information about Netflow support for Nexus Dashboard, see the "Configuring Netflow support" section in Creating Fabrics and Fabric Groups.

Field	Description
Netflow Exporter	To add Netflow exporters for receiving netflow data: 1. In the **Netflow Exporter** area, choose **Actions > Add**. The **Add Item** page appears. 2. In the **Exporter Name** field, enter a name of the exporter. 3. In the **IP** field, enter the IP address of the exporter. 4. In the **VRF** field, specify the VRF over which the exporter is routed. 5. In the **Source Interface** field, enter the source interface name. 6. In the **UDP Port** field, enter the UDP port number over which the netflow data is exported. 7. Click **Save** to configure the exporter.
Netflow Record	To add Netflow records: 1. In the **Netflow Record** area, choose **Actions > Add** to add one or more Netflow records. 2. In the **Record Name** field, enter a name for the record. 3. In the **Record Template** field, select the required templates. From Release 12.0.2, the following two record templates are available for use. You can create custom netflow record templates. Custom record templates saved in the template library are available for use here. ○ **netflow_ipv4_record** – to use the IPv4 record template. ○ **netflow_l2_record** – to use the Layer 2 record template. 4. Check the **Is Layer2 Record** check box if the record is for Layer2 netflow. 5. Click **Save** to configure the report.
Netflow Monitor	To add Netflow monitors: 1. In the **Netflow Monitor** area, choose **Actions > Add**. 2. In the **Monitor Name** field, enter a name for the monitor. 3. In the **Record Name** field, enter a name for the record. 4. In the **Exporter1 Name** field, enter the name of the exporter for the netflow monitor. 5. (Optional) In the **Exporter2 Name** field, enter the name of the secondary exporter for the netflow monitor. The record name and exporters referred to in each netflow monitor must be defined in the **Netflow Record** and **Netflow Exporter** configuration sections in the **Flow Monitor** tab. . Click **Save** to configure the flow monitor.

Telemetry

For a Nexus Dashboard fabric in Monitored mode, Nexus Dashboard will not deploy the telemetry configuration to the switches in the fabric. In order to activate, pause, or deactivate telemetry, you must copy and paste the expected configuration on the switches.

1. Navigate to **Admin > System Status > Telemetry > Switches**
2. Click the checkbox next to a switch, then click **Actions > Expected configuration**.
3. View and copy the configuration information in the **Software Telemetry** and **Flow Telemetry** area,.
4. Using the command line, log in to the switch.
5. Enter this command.

```
switch(config)# copy running-config startup-config
```

The telemetry feature in Nexus Dashboard allows you to collect, manage, and monitor real-time telemetry data from your Nexus Dashboard. This data provides valuable insights into the performance and health of your network infrastructure, enabling you to troubleshoot proactively and optimize operations. When you enable telemetry, you gain enhanced visibility into network operations and efficiently manage your fabrics.

Follow these steps to enable telemetry for a specific fabric.

1. Navigate to the **Fabrics** page.

   Go to **Manage > Fabrics**.

2. Choose the fabric for which you want to enable telemetry.
3. From the **Actions** drop-down list, choose **Edit fabric settings**.

   The **Edit *fabric-name* settings** page displays.

   You can also access the **Edit *fabric-name* settings** page for a fabric from the **Fabric Overview** page. In the **Fabric Overview** page, click the **Actions** drop-down list and choose **Edit fabric settings**.

4. In the **Edit *fabric-name* settings** page, click the **General** tab.
5. Under the **Enabled features** section, check the **Telemetry** check box.
6. Click **Save**.

   Navigate back to the **Edit *fabric-name* settings** page. The **Telemetry** tab displays.

   The **Telemetry** tab appears only when you enable the **Telemetry** option under the **General** tab in the **Edit *fabric-name* settings** page.

The **Telemetry** tab includes these options.

- **Configuration**—allows you to manage telemetry settings and parameters.
- **NAS**—provides Network Analytics Service (NAS) features for advanced insights.

Edit configuration settings

The **Configuration** tab includes these settings.

- **General**—allows you to enable analysis.

  You can enable these settings.

  - **Enable assurance analysis**—enables you to collect of telemetry data from devices to ensure network reliability and performance.
  - **Enable Microburst sensitivity** - allows you to monitor traffic to detect unexpected data bursts within a very small time window (microseconds). Choose the sensitivity type from the **Microburst Sensitivity Level** drop-down list. The options are **High sensitivity**, **Medium sensitivity**, and **Low sensitivity**.

    The **Enable Microburst sensitivity** option is available only for ACI fabrics.

- **Flow collection modes**—allows you to choose the mode for telemetry data collection. Modes include **NetFlow**, **sFlow**, and **Flow Telemetry**.

  For more information see: Flow collection and Configure flows.

- **Flow collection rules**—allows you to define rules for monitoring specific subnets or endpoints. These rules are pushed to the relevant devices, enabling detailed telemetry data collection.

  For more information, see Flow collection.

Edit NAS settings

Nexus Dashboard allows you to export captured flow records to a remote NAS device using the Network File System (NFS) protocol. Nexus Dashboard defines the directory structure on NAS where the flow records are exported.

You can choose between the two export modes.

- **Full**—exports the complete data for each flow record.
- **Base**—exports only the essential 5-tuple data for each flow record.

Nexus Dashboard needs both read and write permissions on the NAS to perform the export successfully. If Nexus Dashboard cannot write to the NAS, it will generate an alert to notify you of the issue.

Disable Telemetry

You can uncheck the **Telemetry** check box on your fabric's **Edit Fabric Settings > General >** page to disable the telemetry feature for your fabric. Disabling telemetry puts the telemetry feature in a transition phase and eventually the telemetry feature is disabled.

In certain situations, the disable telemetry workflow can fail, and you may see the **Force disable telemetry** option on your fabric's **Edit Fabric Settings** page.

If you disable the telemetry option using the instructions provided in Perform force disable telemetry on your fabric, Nexus Dashboard acknowledges the user intent to disable telemetry feature for your fabric, ignoring any failures.

The Nexus Dashboard **Force disable telemetry** allows you to perform a force disable action for the telemetry configuration on your fabric. This action is recommended when the telemetry disable workflow has failed and you need to disable the telemetry feature on your fabric.

Using the **Force disable telemetry** feature may leave switches in your fabric with stale telemetry configurations. You must manually clean up these stale configurations on the switches before re-enabling telemetry on your fabric.

Perform force disable telemetry on your fabric

Follow these steps to perform a force disable telemetry on your fabric.

1. (Optional) Before triggering a force disable of telemetry configuration, resolve any telemetry configuration anomalies flagged on the fabric.
2. On the **Edit Fabric Settings** page of your fabric, a banner appears to alert you that telemetry cannot be disabled gracefully, and a **Force Disable** option is provided with the alert message.
3. Disable telemetry from the Nexus Dashboard UI using one of these options.
    a. Click the **Force disable** option in the banner that appears at the top of your fabric's **Edit Fabric Settings** page to disable telemetry for your fabric gracefully.
    b. Navigate to your fabric's **Overview** page and click the **Actions** drop-down list to choose **Telemetry > Force disable telemetry** option.

    Once the force disable action is executed, the **Telemetry** configuration appears as disabled in **Edit Fabric Settings > General > Enabled features > Telemetry** area, that is, the **Telemetry** check box is unchecked.
4. Clean up any stale telemetry configurations from the fabric before re-enabling telemetry on Nexus Dashboard.

NAS

You can export flow records captured by Nexus Dashboard on a remote Network Attached Storage (NAS) with NFS.

Nexus Dashboard defines the directory structure on NAS where the flow records are exported.

You can export the flow records in Base or Full mode. In Base mode, only 5-tuple data for the flow record is exported. In Full mode the entire data for the flow record is exported.

Nexus Dashboard requires read and write permission to NAS in order to export the flow record. A system issue is raised if Nexus Dashboard fails to write to NAS.

Guidelines and limitations for network attached storage

- In order for Nexus Dashboard to export the flow records to an external storage, the Network Attached Storage added to Nexus Dashboard must be exclusive for Nexus Dashboard.
- Network Attached Storage with Network File System (NFS) version 3 must be added to Nexus Dashboard.
- Flow Telemetry and Netflow records can be exported.
- Export of FTE is not supported.
- Average Network Attached Storage requirements for 2 years of data storage at 20k flows per sec:
  - Base Mode: 500 TB data
  - Full Mode: 2.8 PB data
- If there is not enough disk space, new records will not be exported and an anomaly is generated.

Add network attached storage to export flow records

The workflow to add Network Attached Storage (NAS) to export flow records includes the following steps:

1. Add NAS to Nexus Dashboard.
2. Add the onboarded NAS to Nexus Dashboard to enable export of flow records.

Add NAS to Nexus Dashboard

Follow these steps to add NAS to Nexus Dashboard.

1. Navigate to **Admin** > **System Settings** > **General**.
2. In the **Remote storage** area, click **Edit**.
3. Click **Add Remote Storage Locations**.
4. Complete the following fields to add NAS to Nexus Dashboard.
   a. Enter the name of the Network Attached Storage and a description, if desired.
   b. In the **Remote storage location type** field, click **NAS Storage**.
   c. In the **Type** field, choose **Read Write**.

      Nexus Dashboard requires read and write permission to export the flow record to NAS. A system issue is raised if Nexus Dashboard fails to write to NAS.

   d. In the **Hostname** field, enter the IP address of the Network Attached Storage.
   e. In the **Port** field, enter the port number of the Network Attached Storage.

f. In the **Export path** field, enter the export path.

   Using the export path, Nexus Dashboard creates the directory structure in NAS for exporting the flow records.

g. In the **Alert threshold** field, enter the alert threshold time.

   Alert threshold is used to send an alert when the NAS is used beyond a certain limit.

h. In the **Limit (Mi/Gi)** field, enter the storage limit in Mi/Gi.

i. Click **Save**.

In a three-node Nexus Dashboard cluster deployment, NAS backup operations require that all cluster nodes are in an operational and healthy state. If any node is down, unreachable, or not in a fully functional state, NAS backup operations do not proceed.

This ensures cluster consistency and data integrity during backup operations. Verify cluster health and confirm that all three nodes are up and synchronized before initiating or scheduling NAS backups.

Add the onboarded NAS to Nexus Dashboard

Follow these steps to add the onboarded NAS to Nexus Dashboard.

1. Navigate to the Fabrics page:

   Manage > Fabrics

2. Choose the fabric with the telemetry feature enabled.
3. Choose **Actions > Edit Fabric Settings**.
4. Click **Telemetry**.
5. Click the **NAS** tab in the **Telemetry** page.
6. Make the necessary configurations in the **General settings** area.
   a. Enter the name in the **Name** field.
   b. In the **NAS server** field, choose the NAS server added to Nexus Dashboard from the drop-down list.
7. In the **Collection settings** area, choose the flow from the **Flows** drop-down list.
   - In Base mode, only 5-tuple data for the flow record is exported.
   - In Full mode, the entire data for the flow record is exported.
8. Click **Save**.

The traffic from the flows displayed in the **Flows** page is exported as a JSON file to the external NAS in the following directory hierarchy.

Beginning with Nexus Dashboard 4.3.1, flows are exported to a path with the unified Nexus Dashboard version in the directory name instead of the NIR version.

```
└── NDI-<VERSION>-FLOW-JSON/
    └── fabricName=<fabricName>/
        └── year=2022/
            ├── month=01/
            │   └── date=01/
            │       ├── hour=01/
            │       │   ├── 52170795-0b94-481c-800a-c47f0fa41fac.json
            │       │   └── fa92c70c-96fc-4e32-ac76-324bdd5139d4.json
            │       └── hour=23/
            │           ├── 737f4292-bf29-4630-bdd9-ccb80885ddc1.json
            │           └── 68b434d9-0957-4fe4-be01-e0688cb4336d.json
            └── month=02/
                └── date=20/
                    ├── hour=10/
                    │   ├── e05ce8fb-88af-45db-8c52-4b00e1841b16.json
                    │   └── 6fd2b652-dfe1-430e-905a-020abd399e3e.json
                    └── hour=23/
                        ├── eeb6784a-33a0-4ae3-b13e-db4db93fe48b.json
                        └── b289c75e-a709-4284-a018-b38ab101d90f.json
```

Navigate to **Analyze** > **Flows** to view the flows that will be exported.

Each flow record is written as a line delimited JSON.

JSON output file format for a flow record in base mode

{" fabricName" :" myapic" ," terminalTs" :1688537547433," originTs" :1688537530376," srcIp" :" 2000:201:1:1::1" ," dstIp" :" 2000:201:1:1::3" ," srcPort" :1231," dstPort" :1232," ingressVrf" :" vrf1" ," egressVrf" :" vrf1" ," ingressTenant" :" FSV1" ," egressTenant" :" FSV1" ," protocol" :" UDP" }

{" fabricName" :" myapic" ," terminalTs" :1688537547378," originTs" :1688537530377," srcIp" :" 201.1.1.127" ," dstIp" :" 201.1.1.1" ," srcPort" :0," dstPort" :0," ingressVrf" :" vrf1" ," egressVrf" :" " ," ingressTenant" :" FSV2" ," egressTenant" :" " ," protocol" :" ANY-HOST" }

JSON output file format for a flow record in full mode

{" fabricName" :" myapic" ," terminalTs" :1688538023562," originTs" :1688538010527," srcIp" :" 201.1.1.121" ," dstIp" :" 201.1.1.127" ," srcPort" :0," dstPort" :0," ingressVrf" :" vrf1" ," egressVrf" :" vrf1" ," ingressTenant" :" FSV2" ," egressTenant" :" FSV2" ," protocol" :" ANY-HOST" ," srcEpg" :" ext-epg" ," dstEpg" :" ext-epg1" ," latencyMax" :0," ingressVif" :" eth1/15" ," ingressVni" :0," latency" :0," ingressNodes" :" Leaf1-2" ," ingressVlan" :0," ingressByteCount" :104681600," ingressPktCount" :817825," ingressBurst" :0," ingressBurstMax" :34768," egressNodes" :" Leaf1-2" ," egressVif" :" po4" ,

" egressVni" :0," egressVlan" :0," egressByteCount" :104681600," egressPktCount" :817825," egressBurst" :0," egressBurstMax" :34768," dropPktCount" :0," dropByteCount" :0," dropCode " :" " ," dropScore" :0," moveScore" :0," latencyScore" :0," burstScore" :0," anomalyScore" :0," hashCollision" :false," dropNodes" :" []" ," nodeNames" :" [\" Leaf1-2\"]" ," nodeIngressVifs" :" [\" Leaf1-2,eth1/15\"]" ," nodeEgressVifs" :" [\" Leaf1-2,po4\"]“ ," srcMoveCount" :0," dstMoveCount" :0," moveCount" :0," prexmit" :0," rtoOutside" :false," events" :" [[\\\" 1688538010527,Leaf1-2,0,3,1,no,no,eth1/15,,po4,po4,,,,,0,64,0,,,,,,,,\\\"]]" }

Flow collection

Understanding flow telemetry

Flow telemetry allows users to see the path taken by different flows in detail. It also allows you to identify the EPG and VRF instance of the source and destination. You can see the switches in the flow with the help of flow table exports from the nodes. The flow path is generated by stitching together all the exports in order of the flow.

You can configure the Flow Telemetry rule for the following interface types:

- VRF instances
- Physical interfaces
- Port channel interfaces
- Routed sub-interfaces (Cisco ACI fabric)
- SVIs (Cisco ACI fabric)

In a Cisco ACI fabric, if you want to configure routed sub-interfaces from the UI, select L3 Out.

In an NX-OS fabric, physical or port channel flow rules are supported only on routed interfaces.

Flow telemetry monitors the flow for each fabric separately, as there is no stitching across the fabrics in a fabric group. Therefore, flow telemetry is for individual flows. For example, if there are two fabrics (fabric A and fabric B) within a fabric group, and traffic is flowing between the two fabrics, they will be displayed as two separate flows. One flow will originate from Fabric A and display where the flow exits. And the other flow from Fabric B will display where it enters and where it exits.

Flow telemetry guidelines and limitations

- All flows are monitored as a consolidated view in a unified pipeline for Cisco ACI and NX-OS fabrics, and the flows are aggregated under the same umbrella.
- Even if a particular node (for example, a third-party switch) is not supported for Flow Telemetry, Nexus Dashboard will use LLDP information from the previous and next nodes in the path to identify the switch name and the ingress and egress interfaces.
- You can enable Flow Telemetry at the fabric level in Nexus Dashboard, even if only a subset of switches in the fabric supports the feature. Enabling Flow Telemetry at the fabric level does not

require all switches to support it; provided at least one switch supports Flow Telemetry, you can enable the feature for the entire fabric.

- Nexus Dashboard supports Kafka export for Flow anomalies. However, Kafka export is not currently supported for Flow Event anomalies.

Flow telemetry guidelines and limitations for NX-OS fabrics

- Ensure that you have configured NTP and enabled PTP in Nexus Dashboard. See Cisco Nexus Dashboard Deployment Guide and Precision Time Protocol (PTP) for Cisco Nexus Dashboard Insights for more information. You are responsible for configuring the switches with external NTP servers.
- The **Fabric Overview** page displays the correct PTP status only when **Flow telemetry** or **Traffic analytics** is enabled under **Flow collection**. Enabling PTP in the fabric settings alone does not enable PTP monitoring.
- In the **Edit Flow** page, you can enable all three telemetry types. sFlow is most restrictive, Netflow has some more capability, and Flow Telemetry has the most capability. We recommend that you enable Flow Telemetry if it is available for your configuration. If Flow Telemetry is not available, then use Netflow. If Netflow is not available, use sFlow.
- If there are multiple Nexus Dashboard clusters onboarded to Nexus Dashboard, partial paths will be generated for each fabric.
- For inter-fabric flows that traverse an external fabric and return to the source fabric (a U-turn path), Nexus Dashboard displays the egress interface as N/A when the external fabric contains Cisco N9K-C9364C switches. Nexus Dashboard cannot capture egress port metadata for these flows. The remote port value remains valid.
- If you manually configure the fabric to use with Nexus Dashboard and Flow Telemetry support, the Flows Exporter port changes from 30000 to 5640. To prevent a breakage of the Flows Exporter, adjust the automation.
- Flow telemetry is supported in -FX3 platform switches for the following NX-OS versions:
  - 9.3(7) and later
  - 10.1(2) and later
  - Flow telemetry is not supported in -FX3 platform switches for NX-OS version 10.1(1).
- Interface based Flow Telemetry is only supported on modular chassis with -FX land -GX line cards on physical ports and port-channels rules.
- If interface-based Flow Telemetry is pushed from Nexus Dashboard for **Classic LAN** and **External Connectivity Network** fabrics, perform the following steps:
  - Choose the fabric.
  - Choose **Policies** > **Action** > **Add policy** > **Select all** > **Choose template** > host_port_resync and click **Save**.
  - In the Fabric Overview page, choose **Actions** > **Recalculate and deploy**.
- For VXLAN fabrics, interface-based Flow Telemetry is not supported on switch links between spine switch and leaf switch.
- If you want to use the default VRF instance for flow telemetry, you must create the VRF instance with a name of " default" in lowercase. Do not enter the name with any capital letters.

- Flow telemetry is not supported in classic LAN topologies with 2-level VPC access layers.
- If you want to enable Flow Telemetry, ensure that there are no pre-existing Netflow configurations on the switches. If there are any pre-existing configurations, the switch configuration may fail.

  To enable Flow Telemetry without configuration issues, follow these steps:

  - Ensure that there are no pre-existing Netflow configurations on the switches. If such configurations exist, enabling Flow Telemetry might result in a system anomaly with an error message stating invalid command match IP source address.
  - If you encounter the error, disable Flow Telemetry.
  - Remove any existing Netflow configurations from the switches.
  - Re-enable Flow Telemetry.
  - For some flows, latency information is not available, which could happen due to latency issues. In these cases, latency information will be reported as 0.

Flow telemetry rules guidelines and limitations for NX-OS fabrics

- If you configure an interface rule (physical or port channel) on a subnet, it can monitor only incoming traffic. It cannot monitor outgoing traffic on the configured interface rule.
- If a configured port channel that contains two physical ports, only the port channel rule is applicable. Even if you configure physical interface rules on the port, only port channel rule takes precedence.
- For NX-OS release 10.3(2) and earlier, if a flow rule are configured on an interface, then global flow rules are not matched.
- For NX-OS release 10.3(3) and later, a flow rule configured on an interface is matched first and then the global flow rules are matched.

Configure flows

Configure flow collection modes

Follow these steps to configure flow collection modes.

1. Navigate to **Admin** > **System Settings** > **Flow collection**.
2. In the **Flow collection mode** area, choose **Flow telemetry**.

Enabling Flow Telemetry automatically activates Flow Telemetry Events. Whenever a compatible event takes place, an anomaly will be generated, and the What's the impact? section in the **Anomaly** page will display the associated flows. You can manually configure a Flow Telemetry rule to acquire comprehensive end-to-end information about the troublesome flow.

Configure flow collection rules in an NX-OS fabric

Follow these steps to configure flow collection rules in an NX-OS fabric.

1. Navigate to the **Telemetry** page for your fabric.
   a. Navigate to the main **Fabrics** page:

Manage > Fabrics

b. In the table showing all of the Nexus Dashboard fabrics that you have already created, locate the LAN or IPFM fabric where you want to configure telemetry settings.

c. Single-click on that fabric.

The **Overview** page for that fabric appears.

d. Click **Actions > Edit Fabric Settings**.

The **Edit *fabric_name* Settings** page appears.

e. Verify that the **Telemetry** option is enabled in the **Enabled features** area.

The Telemetry tab doesn't become available unless the **Telemetry** option is enabled in the **Enabled features** area.

f. Click the **Telemetry** tab to access the telemetry settings for this fabric.

2. Click the **Flow collection** tab in the **Telemetry** page.
3. In the **Mode** area, click **Flow telemetry**.
4. In the **Flow collections rules** area, determine what sort of flow collection rule that you want to add.
   - VRF
   - Physical interface
   - Port channel

VRF

Follow these steps to add a VRF rule.

1. Click the **VRF** tab.

   A table with already-configured VRF flow collection rules is displayed.

   For any VRF flow collection rule in this table, click the ellipsis (...), then click **Edit rule** to edit that rule or **Delete rule** to delete it.

2. Add a new rule by clicking **Create flow collection rule**.
   a. In the **General** area, complete the following:
      i. Enter the name of the rule in the **Rule Name** field.
      ii. The VRF field is disabled. The flow rule applies to all the VRF instances.
      iii. In the **Flow Properties** area, select the protocol for which you intend to monitor the flow traffic.
      iv. Enter the source and destination IP addresses. Enter the source and destination port.
      v. Click **Save**.

Physical interface

Follow these steps to add a physical interface rule.

1. Click the **Physical interface** tab.

   A table with already-configured physical interface flow collection rules is displayed.

   For any physical interface flow collection rule in this table, click the ellipsis (...), then click **Edit rule** to edit that rule or **Delete rule** to delete it.

2. Add a new rule by clicking **Create flow collection rule**.
   a. In the **General** area, complete the following:
      i. Enter the name of the rule in the **Rule Name** field.
      ii. Check the **Enabled** check box to enable the status. If you enable the status, the rule will take effect. Otherwise, the rule will be removed from the switches.
      iii. In the **Flow Properties** area, select the protocol for which you intend to monitor the flow traffic.
      iv. Enter the source and destination IP addresses. Enter the source and destination port.
      v. In the **Interface List** area, click **Select a Node**. Use the search box to select a node.
      vi. From the drop-down list, select an interface. You can add more than one row (node+interface combination) by clicking **Add Interfaces**. However, within the rule, a node can appear only once. Configuration is rejected if more than one node is added.
      vii. Click **Save**.

Port channel

Follow these steps to add a port channel rule.

1. Click the **Port channel** tab.

   A table with already-configured port channel flow collection rules is displayed.

   For any port channel flow collection rule in this table, click the ellipsis (...), then click **Edit rule** to edit that rule or **Delete rule** to delete it.

2. Add a new rule by clicking **Create flow collection rule**.
   a. In the **General** area, enter the name of the rule in the **Rule Name** field.
      i. Select the **Enabled** check box to enable the status. If you enable the status, the rule will take effect. Otherwise, the rule will be removed from the switches.
      ii. In the **Flow Properties** area, select the protocol for which you intend to monitor the flow traffic.
      iii. Enter the source and destination IP addresses. Enter the source and destination port.
      iv. From the drop-down list, select an interface. You can add more than one row (node+interface combination) by clicking **Add Interfaces**. However, within the rule, a node can appear only once. Configuration is rejected if more than one node is added.
      v. Click **Save**.

3. Click **Done**.

Monitor the subnet for flow telemetry

In the following example, the configured rule for a flow monitors the specific subnet provided. The rule is pushed to the fabric which pushes it to the switches. So, when the switch sees traffic coming from a source IP or the destination IP, and if it matches the subnet, the information is captured in the TCAM and exported to the Nexus Dashboard service. If there are 4 nodes (A, B, C, D), and the traffic moves from A > B > C > D, the rules are enabled on all 4 nodes and the information is captured by all the 4 nodes. Nexus Dashboard stitches the flows together. Data such as the number of drops and the number of packets, anomalies in the flow, and the flow path are aggregated for the 4 nodes.

Follow these steps to monitor the subnet for flow telemetry.

1. Navigate to **Manage** > **Fabric**.
2. Choose a fabric.
3. Verify that your **Fabrics** and the **Snapshot** values are appropriate. The default snapshot value is 15 minutes. Your choice will monitor all the flows in the chosen fabric or snapshot fabric.
4. Navigate to **Connectivity** > **Flows** to view a summary of all the flows that are being captured based on the snapshot that you chose.

The related anomaly score, record time, the nodes sending the flow telemetry, flow type, ingress and egress nodes, and additional details are displayed in a table format. If you click a specific flow in the table, specific details are displayed in the sidebar for the particular flow telemetry. In the sidebar, if you click the Details icon, the details are displayed in a larger page. In this page, in addition to other details, the **Path Summary** is also displayed with specifics related to source and destination. If there are flows in the reverse direction, that will also be visible in this location.

For a bi-directional flow, there is an option to choose to reverse the flow and see the path summary displayed. If there are any packet drops that generate a flow event, they can be viewed in the Anomaly dashboard.

Understanding Netflow

Netflow is an industry standard where Cisco routers monitor and collect network traffic on an interface. Netflow version 9 is supported.

Netflow enables the network administrator to determine information such as source, destination, class of service, and causes of congestion. Netflow is configured on the interface to monitor every packet on the interface and provide telemetry data. You cannot filter on Netflow.

Netflow in Nexus series switches is based on intercepting the packet processing pipeline to capture summary information of network traffic.

The components of a flow monitoring setup are as follows:

- Exporter: Aggregates packets into flows and exports flow records towards one or more collectors
- Collector: Reception, storage, and pre-processing of flow data received from a flow exporter
- Analysis: Used for traffic profiling or network intrusion
- The following interfaces are supported for Netflow:

Supported interfaces for Netflow

Interfaces	5 Tuple	Nodes	Ingress	Egress	Path	Comments
Routed Interface/Port Channel	Yes	Yes	Yes	No	Yes	Ingress node is shown in path
Sub Interface/Logical (Switch Virtual Interface)	Yes	Yes	No	No	No	No

In an NX-OS fabric, port channel support is available if you monitor only the host-facing interfaces.

Understanding Netflow types

You can use these Netflow types.

- Full Netflow
- Sampled Netflow

Full Netflow

With Full Netflow, all packets on the configured interfaces are captured into flow records in a flow table. Flows are sent to the supervisor module. Records are aggregated over configurable intervals and exported to the collector. Except in the case of aliasing (multiple flows hashing to the same entry in the flow table), all flows can be monitored regardless of their packet rate.

Cisco N9000 Series switches with the Fabric Controller type as well as switches in a Cisco ACI fabric

support Full Netflow.

Sampled Netflow

With Sampled Netflow, packets on configured interfaces are time sampled. Flows are sent to the supervisor or a network processor for aggregation. Aggregated flow records are exported at configured intervals. The probability of a record for a flow being captured depends on the sampling frequency and packet rate of the flow relative to other flows on the same interface.

Nexus 7000 and Nexus 7700 Series switches with F/M line cards and the Fabric Controller type, support Sampled Netflow.

Netflow guidelines and limitations

- In Cisco N9000 series switches, Netflow supports a small subset of the published export fields in the RFC.
- Netflow is captured only on the ingress port of a flow as only the ingress switch exports the flow. Netflow cannot be captured on fabric ports.

Netflow guidelines and limitations for Cisco ACI fabrics

- We recommend that you enable Flow Telemetry. If that is not available for your configuration, use Netflow. However, you can determine which mode of flow to use based upon your fabric configuration.
- Enabling both Flow Telemetry and Netflow is not supported.
- After you enable Netflow, you must obtain the Netflow collector IP address and configure Cisco APIC with the collector IP address. See Cisco APIC and NetFlow.

  To obtain the Netflow collector IP address, navigate to **Admin > System Settings > General**, then locate the **External Pools** area. Click **View all** at the bottom left area of the **External Pools** tile; the **Telemetry-collector** persistent IP addresses listed in the table are used for the Netflow collector IP address.

- The Netflow and sFlow flow collection modes do not support any anomaly.

Netflow guidelines and limitations for NX-OS fabrics

- In the **Edit Flow** page, you can enable all three modes. Choose the best possible mode for a product. sFlow is the most restrictive, Netflow has more capabilities, and Flow Telemetry has the most capabilities. We recommend that you enable Flow Telemetry if it is available for your configuration. If Flow Telemetry is not available, then use Netflow. If Netflow is not available, use sFlow.
- In Cisco Nexus 7000 and Cisco N9000 Series switches, only the ingress host-facing interface configured for Netflow are supported (either in VXLAN or Classic LAN).
- The Netflow supported fabrics are Classic and VXLAN. VXLAN is not supported on fabric ports.
- Netflow configurations will not be pushed. However, if a fabric is managed, the software sensors will be pushed.
- If you manually configure the fabric to use with Nexus Dashboard and Netflow support, the Flows Exporter port changes from 30000 to 5640. To prevent a breakage of the Flows Exporter, adjust

the automation.

- To configure Netflow on fabric switches, see the **Configuring Netflow** section in the Cisco N9000 Series NX-OS System Management Configuration Guide.

Configure Netflow

Follow these steps to configure Netflow.

1. Navigate to the **Telemetry** page for your LAN or IPFM fabric.
2. Click the **Flow collection** tab on the **Telemetry** page.
3. In the **Mode** area, make the following choices:
    - Choose **Netflow**.
    - Choose **Flow Telemetry**.
4. Click **Save**.

Understanding sFlow

sFlow is an industry standard technology traffic in data networks containing switches and routers. Nexus Dashboard supports sFlow version 5 on Cisco Nexus 3000 series switches.

sFlow provides the visibility to enable performance optimization, an accounting and billing for usage, and defense against security threats.

The following interfaces are supported for sFlow:

Supported interfaces for sFlow

Interfaces	5 Tuple	Nodes	Ingress	Egress	Path	Comments
Routed Interface	Yes	Yes	Yes	Yes	Yes	Ingress node is shown in path

Guidelines and limitations for sFlow

- Nexus Dashboard supports sFlow with Cisco Nexus 3000 series switches.
- It is recommended to enable Flow Telemetry if it is available for your configuration. If it is not available for your configuration, use Netflow. If Netflow, is not available for your configuration, then use sFlow.
- For sFlow, Nexus Dashboard requires the configuration of persistent IPs under cluster configuration, and 6 IPs in the same subnet as the data network are required.
- sFlow configurations will not be pushed. However, if a fabric is managed, the software sensors will be pushed.
- If you manually configure the fabric to use with Nexus Dashboard and sFlow support, the Flows Exporter port changes from 30000 to 5640. To prevent a breakage of the Flows Exporter, adjust the automation.
- Nexus Dashboard does not support sFlow in the following Cisco Nexus 3000 Series switches:
  - Cisco Nexus 3600-R Platform Switch (N3K-C3636C-R)
  - Cisco Nexus 3600-R Platform Switch (N3K-C36180YC-R)
  - Cisco Nexus 3100 Platform Switch (N3K-C3132C-Z)
- Nexus Dashboard does not support sFlow in the following Cisco N9000 Series fabric modules:
  - Cisco N9508-R fabric module (N9K-C9508-FM-R)
  - Cisco N9504-R fabric module (N9K-C9504-FM-R)
- To configure sFlow on fabric switches, see the **Configuring sFlow** section in the Cisco N9000 Series NX-OS System Management Configuration Guide.

Configure sFlow telemetry

Prerequisites

Follow these steps to configure sFlow telemetry.

1. Navigate to the **Telemetry** page for your LAN or IPFM fabric.
2. Click the **Flow collection** tab on the **Telemetry** page.
3. In the **Mode** area, make the following choices:
    - Choose **sFlow**.
    - Choose **Flow Telemetry**.
4. Click **Save**.

External streaming

The **External streaming** tab in Nexus Dashboard allows you export data that Nexus Dashboard collects over Kafka, email, and syslog. Nexus Dashboard generates data such as advisories, anomalies, audit logs, faults, statistical data, and risk and conformance reports. When you configure a Kafka broker, Nexus Dashboard writes all data to a topic. By default, the Nexus Dashboard collects export data every 30 seconds or at a less frequent interval.

For ACI fabrics, you can also collect data for specific resources (CPU, memory, and interface utilization) every 10 seconds from the leaf and spine switches using a separate data pipeline. To export this data, select the **Usage** option under **Collection Type** in the **Message bus** export settings. Additionally, CPU and memory data is collected for the controllers.

Nexus Dashboard does not store the collected data in Elasticsearch; instead, it exports the data directly to your repository or data lake using a Kafka broker for consumption. By using the Kafka export functionality, you can then export this data to your Kafka broker and push it into your data lake for further use.

You can configure an email scheduler to define the type of data and the frequency at which you want to receive information via email. You can also export anomaly records to an external syslog server. To do this, select the **Syslog** option under the **External Streaming** tab.

Configure external streaming settings

Follow these steps to configure external streaming settings.

1. Navigate to the **Fabrics** page.

   Go to **Manage > Fabrics**.

2. Choose the fabric for which you configure streaming settings.
3. From the **Actions** drop-down list, choose **Edit fabric settings**.

   The **Edit *fabric-name* settings** page displays.

   You can also access the **Edit *fabric-name* settings** page for a fabric from the **Fabric Overview** page. In the **Fabric Overview** page, click the **Actions** drop-down list and choose **Edit fabric settings**.

4. In the **Edit *fabric-name* settings** page, click the **External streaming** tab.

   You can view these options.

   - Email
   - Message bus
   - Syslog

Guidelines and limitations

- Intersight connectivity is required to receive the reports by email.
- You can configure up to five emails per day for periodic job configurations.
- A maximum of six exporters is supported for export across all types of exporters including email, message bus, and syslog. You must provide unique names for each export.
- The scale for Kafka exports is increased to support up to 20 exporters per cluster. However, statistics selection is limited to any six exporters.
- Before configuring your Kafka export, you must add the external Kafka IP address as a known route in your Nexus Dashboard cluster configuration and verify that Nexus Dashboard can reach the external Kafka IP address over the network.
- The anomalies in Kafka and email messages include categories such as Resources, Environmental, Statistics, Endpoints, Flows, and Bugs.
- Export data is not supported for snapshot fabrics.
- You must provide unique names for each exporter, and they may not be repeated between Kafka export for **Alerts and Events** and Kafka export for **Usage**.
- Nexus Dashboard supports Kafka export for flow anomalies. However, Kafka export is not currently supported for flow Event anomalies.

Guidelines and limitations in NX-OS fabrics

- Remove all configurations in the *Message Bus Configuration* and *Email* page before you disable Software Telemetry on any fabric and remove the fabric from Nexus Dashboard.
- If you use interface groups while enabling PTP in an Enhanced Classic LAN fabric, you must add the following commands to the freeform section of the interface group.
  - ttag
  - ttag-strip

    If you do not add these commands, endpoints connected behind these interfaces will experience traffic loss.

Email

The email scheduler feature in Nexus Dashboard automates the distribution of summarized data collected from Nexus Dashboard. It allows customization of selection of email recipients, choice of email format, scheduling frequency settings, and configuring the types of alerts and reports.

Follow these steps to configure an email scheduler.

1. Navigate to the **Fabrics** page.

   Go to **Manage > Fabrics**.

2. Choose the fabric for which you configure streaming settings.
3. From the **Actions** drop-down list, choose **Edit fabric settings**.

The **Edit *fabric-name* settings** page displays.

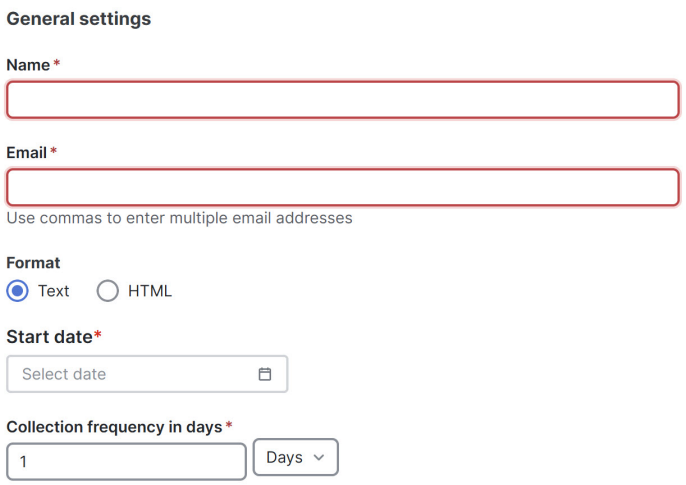
4. In the **Edit *fabric-name* settings** page, click the **External streaming** tab.
5. Click the **Email** tab.
6. Review the information provided in the **Email** tab for already-configured email configurations.

   The following details display under **Email** tab.

Field	Description
Name	The name of the email configuration.
Email	The email addresses used in the email configuration.
Start time	The start date used in the email configuration.
Frequency	The frequency in days or weeks set in the email configuration.
Anomalies	The severity level for anomalies and advisories set in the email configuration.
Advisories	
Risk and conformance reports	The status of the overall inventory for a fabric, including software release, hardware platform, and a combination of software and hardware conformance.

To add a new email configuration, click **Add email** in the **Email** page.

1. Follow these steps to configure **General** Settings.
   a. In the **Name** field, enter the name of the email scheduler.
   b. In the **Email** field, enter one or more email addresses separated by commas.
   c. In the **Format** field, choose **Text** or **HTML** email format.
   d. In the **Start date** field, choose the start date when the scheduler should begin sending emails.
   e. In the **Collection frequency in days** field, specify how often the summary is sent, you can choose days or weeks.



2. Follow these steps to configure **Collection Settings**.
   a. In the **Mode** field, choose one of the following modes.

- **Basic**—displays the severity levels for anomalies and advisories.
- **Advanced**—displays the categories and severity levels for anomalies and advisories.

b. Check the **Only include active alerts in email** check box, to include only active anomaly alerts.

c. Under **Anomalies** choose the categories and severity levels for the anomalies.

d. Under **Advisories** choose the categories and severity levels for the advisories.

e. Under **Risk and Conformance Reports**, choose from the following options.

- Software
- Hardware

3. Click **Save**.

   The **Email** area displays the configured email schedulers.

You will receive an email about the scheduled job on the provided *Start Date* and at the time provided in the **Collection frequency in days** field. The subsequent emails follow after *Collect Every* frequency expires. If the provided time is in the past, Nexus Dashboard will send you an email immediately and trigger the next email after the duration from the provided start time expires.

Message bus

Add Kafka broker configuration

Follow these steps to configure the message bus and add kafka broker.

1. Configure the message bus at the **System Settings** level.
   a. Navigate to **Admin** > **System Settings** > **General**.
   b. In the **Message bus configuration** area, click **Edit**.

   The **Message bus configuration** dialog box opens.

c. Click **Add message bus configuration**.

The **Add message bus configuration** dialog box opens.

d. In the **Name** field, enter a name for the configuration.

e. In the **Hostname/IP address** and **Port** fields, enter the IP address of the message bus consumer and the port that is listening on the message bus consumer.

f. In the **Topic name** field, enter the name of the Kafka topic to which Nexus Dashboard must send the messages.

g. In the **Mode** field, choose the security mode.

The supported modes are **Unsecured**, **Secured SSL** and **SASLPLAIN**. The default value is **Unsecured**.

- For **Unsecured**, no other configurations are needed.
- For **Secured SSL**, fill out the following field:

  Client certification name—The System Certificate name configured at the Certificate Management level. The CA certificate and System Certificate (which includes Client certificate and Client key) are added at the Certificate Management level.

  Refer to Step 2 for step-by-step instructions on managing certificates. Navigate to **Admin** > **Certificate Management** to manage the following certificates:

  - **CA Certificate**—The CA certificate used for signing consumer certificate, which will be stored in the trust-store so that Nexus Dashboard can trust the consumer.
  - **Client Certificate**—The CA signed certificate for Nexus Dashboard. The certificate is signed by the same CA, and the same CA certificate will be in the truststore of the consumer. This will be stored in Nexus Dashboard's Kafka keystore that is used for exporting.
  - **Client Key**—A private key for the Kafka producer, which is Nexus Dashboard in this case. This will be stored in Nexus Dashboard's Kafka keystore that is used for exporting.

- For **SASLPLAIN**, fill out these fields:

  - **Username**—The username for the SASL/PLAIN authentication.
  - **Password**—The password for the SASL/PLAIN authentication.

h. Click **Save**

2. Add CA certificates and System certificates at the **Certificate Management level**.

a. Navigate to **Admin** > **Certificate Management**.

b. In the **Certificate management** page, click the **CA Certificates** tab, then click **Add CA certificate**.

The fields in the **CA Certificates** tab are described in the following table.

Field	Description
Certificate name	The name of the CA certificate.
Certificate details	The details of the CA certificate.
Attached to	The CA signed certificate attached to Nexus Dashboard.
Expires on	The Expiry date and time of the CA certificate.
Last updated time	The last updated time of the CA certificate.

c. In the **Certificate management** page, click the **System certificates** tab, then click **Add system certificate** to add Client Certificate and Client key. Note that the Client certificate and Client key should have same names except extensions as .cer/.crt/.pem for Client certificate and .key for Client key.

You must add a valid CA Certificate before adding the corresponding System Certificate.

The fields in the **System Certificates** tab are described in the following table.

Field	Description
Certificate name	The name of the Client certificate.
Certificate details	The details of the Client certificate.
Attached to	The feature to which the system certificate is attached to, in this case, the message bus.
Expires on	The Expiry date and time of the CA certificate.
Last updated time	The last updated time of the CA certificate.

To configure message bus, the System Certificate should be attached to message bus feature.

To attach a System Certificate to the message bus feature:

a. Choose the System Certificate that you want to use and click the ellipses (...) on that row.

b. Choose **Manage Feature Attachments** from the drop-down list.

The **Manage Feature Attachments** dialog box opens.

c. In the **Features** field, choose **messageBus**.

d. Click **Save**.

For more information on CA certificates, see Managing Certificates in your Nexus Dashboard.

Configure Kafka exports in fabric settings

1. Navigate to the **External streaming** page for your fabric.

   a. Navigate to the **Fabrics** page.

   Go to **Manage > Fabrics**.

b. Choose the fabric for which you configure streaming settings.

c. From the **Actions** drop-down list, choose **Edit fabric settings**.

The **Edit *fabric-name* settings** page displays.

d. In the **Edit *fabric-name* settings** page, click the **External streaming** tab.

e. Click the **Message bus** tab.

2. Review the information provided in the **Message bus** tab for already-configured message bus configurations, or click **Add message bus** to add a new message bus configuration.

Skip to Step 3 if you are adding a message bus.

The fields in the **Message bus** tab are described in the following table.

Field	Description
Message bus stream	The name of the message bus stream configuration.
Collection type	The collection type used by the message bus stream.
Mode	The mode used by the message bus stream.
Anomalies	The severity level for anomalies and advisories set in the message bus stream configuration.
Advisories	
Statistics	The statistics that were configured for the message bus stream.
Faults	The severity level for faults set in the message bus stream configuration.
Audit Logs	The audit logs that were configured for the message bus stream.

3. To configure a new message bus stream, in the **Message bus** page, click **Add message bus**.
4. In the **Message bus stream** field, choose the message bus stream that you want to edit.
5. In the **Collection Type** area, choose the appropriate collection type.

Depending on the **Collection Type** that you choose, the options displayed in this area will change.

- **Alerts and events**: This is the default setting. Continue to Step 7, if you choose **Alerts and events**.
- **Usage**: In the **Collection settings** area, under **Data**, the Resources, and Statistics for the collection settings are displayed. By default, the data for CPU, Memory, and Interface Utilization are collected and exported. You cannot choose to export a subset of these resources.

Usage is applicable only for ACI Fabrics. This option is disabled for other fabrics.

6. Click **Save**. The configured message bus streams are displayed in the **Message bus** area. This configuration now sends immediate notification when the selected anomalies or advisories occur.
7. If you choose **Alerts and events** as the **Collection Type**, in the **Mode** area, choose either **Basic** or **Advanced**.

The configurations that are available in each collection settings section might vary, depending on

the mode that you set.

8. Determine which area you want to configure for the message bus stream.

   The following areas appear in the page:

   - Anomalies
   - Advisories
   - Statistics
   - Faults
   - Audit Logs

   After you complete the configurations on this page, click **Save**. Nexus Dashboard displays the configured message bus streams in the **Message bus** area. This configuration now sends immediate notification when the selected anomalies or advisories occur.

Anomalies

- If you chose **Basic** in the **Mode** area, choose one or more of the following severity levels for anomaly statistics that you want to configure for the message bus stream:
  - Critical
  - Major
  - Warning
  - Minor

  Or click **Select all** to select all available statistics for the message bus stream.

- If you chose **Advanced** in the **Mode** area:
  - Choose one or more of the following categories for anomaly statistics that you want to configure for the message bus stream:
    - Active Bugs
    - Capacity
    - Compliance
    - Configuration
    - Connectivity
    - Hardware
    - Integrations
    - System
  - Choose one or more of the following severity levels for anomaly statistics that you want to configure for the message bus stream:
    - Critical
    - Major
    - Warning

        - Minor

    Or click **Select all** to select all available categories and statistics for the message bus stream.

For more information on anomaly levels, see Detecting Anomalies in Your Nexus Dashboard.

Advisories

- If you chose **Basic** in the **Mode** area, choose one or more of the following severity levels for advisory statistics that you want to configure for the message bus stream:
    - Critical
    - Major
    - Warning
    - Minor

    Or click **Select all** to select all available statistics for the message bus stream.

- If you chose **Advanced** in the **Mode** area:
    - Choose one or more of the following categories for advisory statistics that you want to configure for the message bus stream:
        - Best Practices
        - Field Notices
        - HW end-of-life
        - SW end-of-life
        - PSIRT
    - Choose one or more of the following severity levels for advisory statistics that you want to configure for the message bus stream:
        - Critical
        - Major
        - Warning
        - Minor

    Or click **Select all** to select all available categories and statistics for the message bus stream.

For more information on advisory levels, see Identifying Advisories in Your Nexus Dashboard.

Statistics

There are no differences in the settings in the **Statistics** area when you choose **Basic** or **Advanced** in the **Mode** area.

Choose one or more of the following categories to configure the statistics you want to stream over the message bus.

- Interfaces—streams statistics related to network interfaces, such as traffic volume, error rates, and interface status.
- Protocol—streams protocol-specific statistics, including packet counts, protocol errors, and handshake success rates.

- Resource Allocation—streams data about system resources, such as CPU usage, memory consumption, and bandwidth allocation.
- Environmental—streams environmental metrics such as temperature, humidity, and power supply status.
- Endpoints—streams statistics about connected endpoints, including connection status, data throughput, and session durations.

Faults

There are no differences in the settings in the **Faults** area when you choose **Basic** or **Advanced** in the **Mode** area.

Choose one or more of the following severity levels for fault statistics that you want to configure for the message bus stream:

- Critical
- Major
- Minor
- Warning
- Info

Audit Logs

There are no differences in the settings in the **Audit Logs** area when you choose **Basic** or **Advanced** in the **Mode** area.

Choose one or more of the following categories for audit logs that you want to configure for the message bus stream:

- Creation
- Deletion
- Modification

Syslog

Nexus Dashboard supports the export of anomalies in syslog format. You can use the syslog configuration feature to develop network monitoring and analytics applications on top of Nexus Dashboard, integrate with the syslog server to get alerts, and build customized dashboards and visualizations.

After you choose the fabric where you want to configure the syslog exporter and set up the syslog export configuration, Nexus Dashboard establishes a connection with the syslog server and sends data to the syslog server.

Nexus Dashboard exports anomaly records to the syslog server. With syslog support, you can export anomalies to your third-party tools even if you do not use Kafka.

Guidelines and limitations for syslog

If the syslog server is not operational at a certain time, messages generated during that downtime will not be received by a server after the server becomes operational.

Add syslog server configuration

Follow these steps to add syslog server configuration.

1. Navigate to **Admin > System Settings > General**.
2. In the **Remote streaming servers** area, click **Edit**.

   The **Remote streaming servers** page displays.
3. Click **Add server**.

   The **Add server** page displays.
4. Choose the **Service** as **Syslog**.
5. Choose the **Protocol**.

   You have these options.

   - TCP
   - UDP
6. In the **Name** field, provide the name for the syslog server.
7. In the **Hostname/IP address** field, provide the hostname or IP address of the syslog server.
8. In the **Port** field, specify the port number used by the syslog server.
9. If you want to enable secure communication, check the **TLS** check box.

Before you enable **TLS** you must upload the CA certificate for the syslog destination host to Nexus Dashboard. For more information see, Upload a CA certificate.

Configure syslog to enable exporting anomalies data to a syslog server

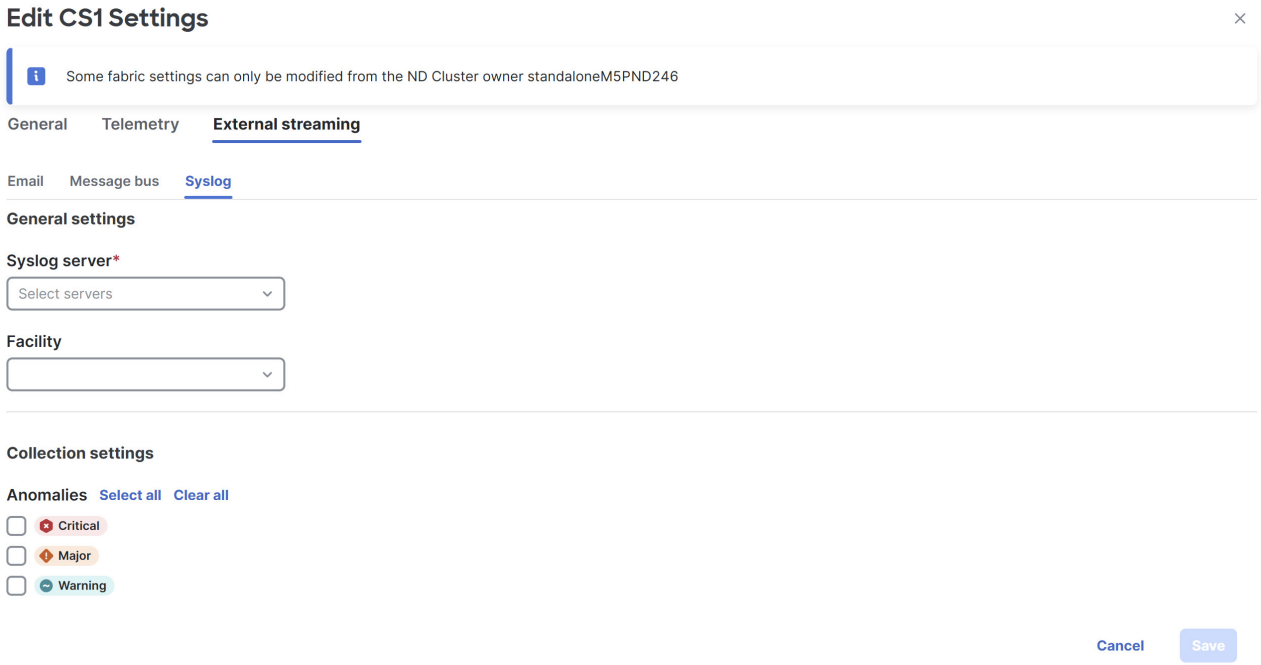
Follow these steps to configure syslog to enable exporting anomalies data to a syslog server.

1. Navigate to the **Fabrics** page.

   Go to **Manage > Fabrics**.
2. Choose the fabric for which you configure streaming settings.
3. From the **Actions** drop-down list, choose **Edit fabric settings**.

   The **Edit *fabric-name* settings** page displays.
4. In the **Edit *fabric-name* settings** page, click the **External streaming** tab.
5. Click the **Syslog** tab.

The following details display under **Syslog** tab.



6. Make the necessary configurations in the **General settings** area.
   a. In the **Syslog server** drop down list, choose a syslog server.

      The **Syslog server** drop down list displays the syslog servers that you added in the **System Settings** level. For more information, see Add syslog server configuration.

   b. In the **Facility** field, from the drop-down list, choose the appropriate facility string.

      A facility code is used to specify the type of system that is logging the message. For this feature, the **local0-local7** keywords for locally used facility are supported.

7. In the **Collection settings** area, choose the desired severity options.

   The options available are **Critical**, **Major**, and **Warning**.

8. Click **Save**.

Upload a CA certificate

Follow these steps to upload a CA certificate for syslog server **TLS**.

1. Navigate to **Admin** > **Certificate Management**.
2. In the **Certificate management** page, click the **CA certificates** tab, then click **Add CA certificate**.

   You can upload multiple files at a single instance.

3. Browse your local directory and choose the certificate-key pair to upload.

   You can upload certificates with the .pem/.cer/.crt/ file extensions.

4. Click **Save** to upload the selected files to Nexus Dashboard.

   A successful upload message appears. The uploaded certificates are listed in the table.

Additional settings

The following sections provide information for additional settings that might be necessary when editing the settings for an AI Routed fabric.

Retrieving the authentication key

Retrieve the 3DES encrypted OSPF authentication key

1. SSH into the switch.
2. On an unused switch interface, enable the following:

```
config terminal
   feature ospf
   interface Ethernet1/1
     no switchport
     ip ospf message-digest-key 127 md5 ospfAuth
```

   In the example, **ospfAuth** is the unencrypted password.

   This Step 2 is needed when you want to configure a new key.

3. Enter the **show run interface Ethernet1/1** command to retrieve the password.

```
Switch # show run interface Ethernet1/1
  interface Ethernet1/1
    no switchport
    ip ospf message-digest key 127 md5 3 sd8478f4fsw4f4w34sd8478fsdfw
    no shutdown
```

   The sequence of characters after **md5 3** is the encrypted password.

4. Update the encrypted password into the **OSPF Authentication Key** field.

Retrieve the encrypted IS-IS authentication key

To get the key, you must have access to the switch.

1. SSH into the switch.
2. Create a temporary keychain.

```
config terminal
  key chain isis
  key 127
```

```
  key-string isisAuth
```

In the example, **isisAuth** is the plaintext password. This will get converted to a Cisco type 7 password after the CLI is accepted.

3. Enter the **show run | section "key chain"** command to retrieve the password.

```
key chain isis
  key 127
    key-string 7 071b245f5a
```

The sequence of characters after key-string 7 is the encrypted password. Save it.

4. Update the encrypted password into the ISIS Authentication Key field.
5. Remove any unwanted configuration made in Step 2.

Retrieve the 3DES encrypted BGP authentication key

1. SSH into the switch and enable BGP configuration for a non-existent neighbor.

Non-existent neighbor configuration is a temporary BGP neighbor configuration for retrieving the password.

```
router bgp
  neighbor 10.2.0.2 remote-as 65000
  password bgpAuth
```

In the example, **bgpAuth** is the unencrypted password.

2. Enter the show run bgp command to retrieve the password. A sample output:

```
neighbor 10.2.0.2
  remote-as 65000
  password 3 sd8478fswerdfw3434fsw4f4w34sdsd8478fswerdfw3434fsw4f4w3
```

The sequence of characters after password 3 is the encrypted password.

3. Update the encrypted password into the **BGP Authentication Key** field.
4. Remove the BGP neighbor configuration.

Retrieve the encrypted BFD authentication key

1. SSH into the switch.
2. On an unused switch interface, enable the following:

```
switch# config terminal
switch(config)# int e1/1
switch(config-if)# bfd authentication keyed-SHA1 key-id 100 key passwd
```

In the example, **passwd** is the unencrypted password and the key ID is **100**.

This Step 2 is needed when you want to configure a new key.

3. Enter the **show running-config interface** command to retrieve the key.

```
switch# show running-config interface Ethernet1/1

interface Ethernet1/1
description connected-to- switch-Ethernet1/1
no switchport
mtu 9216
bfd authentication Keyed-SHA1 key-id 100 hex-key 636973636F313233
no ip redirects
ip address 10.4.0.6/30
no ipv6 redirects
ip ospf network point-to-point
ip router ospf 100 area 0.0.0.0
no shutdown
```

The BFD key ID is **100** and the encrypted key is **636973636F313233**.

4. Update the key ID and key in the **BFD Authentication Key ID** and **BFD Authentication Key** fields.

Guidelines for VXLAN Fabric With eBGP Underlay

- Brownfield migration is not supported for eBGP fabrics.
- You cannot change the leaf switch Autonomous System (AS) number after it is created and the configuration is deployed. You must delete the **leaf_bgp_asn** policy and perform **Recalculate & Deploy** to remove BGP configuration related to this AS. Then, add the **leaf_bgp_asn policy** with the new AS number.
- To switch between Multi-AS and Same-Tier-AS modes, remove all manually added BGP policies (including **leaf_bgp_asn** on the leaf switch and the EBGP overlay policies), and perform the **Recalculate & Deploy** operation before the mode change.
- You cannot change or delete the leaf switch **leaf_bgp_asn** policy if there are ebgp overlay policies present on the device. You need to delete the eBGP overlay policy first, and then delete the **leaf_bgp_asn** policy.
- Intra-fabric links only support IPv6 link local addresses when the underlay is IPv6.
- On a border device, VRF-Lite is supported with manual mode.

- VXLAN eBGP fabrics are supported as child fabrics in VXLAN Multi-Site fabrics.
- TRM (Tenant Routed Multicast) is supported with an eBGP fabric with an IPv4 or an IPv6 underlay.
- VXLAN with an IPv6 underlay does not support the following features:
  - Bidirectional Forwarding Detection (BFD)
  - MACSec
  - Flexible Netflow
  - BGP authentication

Adding Switches

Switches can be added to a single fabric at any point in time. To add switches to a fabric and discover existing or new switches, refer to the section "Add switches to a fabric" in Working with Inventory in Your Nexus Dashboard LAN or IPFM Fabrics.

Support for Cisco N9164E-NS4-O and Cisco N9348Y12C-SE1 switches

Beginning with this release, Nexus Dashboard adds support for Cisco N9348Y12C-SE1 switch running NX-OS release 10.6(3) and Cisco N9164E-NS4-O switch running NX-OS release 10.6.2(n) and later.

For more information on Cisco N9164E-NS4-O switch specific details, see the Cisco N9000 Series NX-OS Release Notes.

The Cisco N9164E-NS4-O switch is limited to Layer-3 operations only. The following features are not supported on this switch:

- Layer-2: SVI, vPC (including pairing), and Port-channels.
- Traffic Management: Multicast, PACL, VACL, and VXLAN.
- Advanced Analysis: Connectivity Analysis, NetFlow, and sFlow (resulting in no support for Traffic Analysis/Flow Telemetry).

Assigning Switch Roles

To assign roles to switches on Nexus Dashboard Fabric Controller refer to the "Assign switch roles" in Working with Inventory in Your Nexus Dashboard LAN or IPFM Fabrics.

Understanding switches running SONiC

Support is now available for certain switches running SONiC (Software for Open Networking in Cloud) images in these Nexus Dashboard fabric types:

- AI Routed
- Routed
- External

Support is available for these switches running SONiC images:

- Cisco N9164E-NS4-O
- Cisco N93108TC-FX3 (N9K-C93108TC-FX3)

These sections provide more detailed information based on the fabric type.

- Switches running SONiC: On AI Routed and Routed fabrics
- Switches running SONiC: External fabrics
- Switches running SONiC: General information

Switches running SONiC: On AI Routed and Routed fabrics

For AI Routed and Routed fabrics, you will configure SONiC as the network operating system for the switches as part of the process of creating the fabric. See Creating Fabrics and Fabric Groups for more information.

When you choose the SONiC Network operating system in the **Advanced Configuration** mode when creating a fabric and you add the relevant switches to the fabric, the switches are bootstrapped through ZTP (Zero Touch Provisioning). During ZTP, the switch discovers onboarding parameters through DHCP and establishes an initial communication with Nexus Dashboard, which delivers the appropriate SONiC image and startup configuration.

New SONiC-specific fabric, interface, and network templates are also available as part of this feature.

Since the switches running the SONiC image are fully managed by Nexus Dashboard, when you add a configuration policy for the switch, only the leaf_bgp_asn policy is available in the **Add policy** page for you to choose. You would specify the leaf switch's ASN in the **Leaf BGP AS #** field in this page in the case where the **ASN auto allocation** option is disabled under **General Parameters** in fabric settings. See Working with Inventory in Your Nexus Dashboard LAN or IPFM Fabrics for more information.

The show commands, which are allowed to be run on SONiC switches, are listed in the sonic_allowed_show_clis template. The contents of that template are validated against when you choose **Actions > Maintenance > Show commands** and you choose sonic_show from the **Commands** pull-down list on the **Switch** show commands page.

Guidelines and limitations: Switches running SONiC on AI Routed and Routed fabrics

- You must open the Nexus Dashboard port 30022 for switches running SONiC images to function in Routed and AI Routed fabrics.
- Telemetry is always enabled on Routed and AI Routed fabrics with switches running SONiC images and cannot be disabled through the Nexus Dashboard GUI. It can be paused, which stops streaming but retains the configuration, and then resumed, which sends a full baseline then streams on-change updates. Telemetry can only be fully disabled by deleting the fabric.
- Per-packet dynamic load-balancing is supported on Cisco N9164E-NS4-O switches. Per-packet dynamic load-balancing is not supported on Cisco N93108TC-FX3 switches.
- AI job monitoring is not supported.

These are the additional guidelines and limitations for switches running SONiC.

- Deployment constraints: Switches running SONiC on AI Routed and Routed fabrics

- Unsupported features: Switches running SONiC on AI Routed and Routed fabrics

Deployment constraints: Switches running SONiC on AI Routed and Routed fabrics

- **Greenfield only**—Only greenfield deployments with IPv4 underlay and OOB management are supported.
- **Management access**—Devices can only be managed via OOB (eth0). Out-of-band configuration changes are not supported.
- **Multi-AS only**—Only Multi-AS mode is supported.
- **Port breakout**—Must be done through the Nexus Dashboard GUI workflow.
- **Default VRF**—Consistent with existing routed fabrics, only default VRF and external peering out of border switches in default VRF are supported.

Unsupported features: Switches running SONiC on AI Routed and Routed fabrics

The following features are **not** supported in this release:

Roles

In Routed and AI Routed fabrics with switches running SONiC images, only switch roles of spine, leaf and border are supported; all other switch roles are not supported.

Networking

- Port-channels, vPC, PVLAN
- HSRP/VRRP, BGP Authentication, BFD Authentication, MACsec

On the multi-attach of network screen, if a network is shown to be attached to short-formed interface eth0 for the SONiC switch, that interface actually refers to Ethernet0 rather than eth0 of the SONiC switch management interface. In other words:

- eth0 is the management port
- Ethernet0 to Ethernet47 are the front panel ports 1-48
- Ethernet48 is the front panel port 49. If this port is configured as the breakout port, then:
  - Ethernet48 to Ethernet51 are configured as the breakout port, and
  - Ethernet52 is the front panel port 50, and so on

Configuration

- Freeform configuration
- Banner and AAA configurations

Telemetry and Analytics

NetFlow, Flow Telemetry, Traffic Analytics, sFlow, Flow Rules (unavailable at the fabric level), Policy CAM Analysis, Delta Analysis, Traffic Analytics, Conformance, Connectivity Analysis, Energy Management, Log Collector, Bug Scan

Switches running SONiC: External fabrics

For External fabrics, you will configure SONiC as the network operating system for the switches as part of the process of adding switches to the External fabric and discovering those switches. See Editing External Fabric Settings for more information.

These are the additional guidelines and limitations for switches running SONiC.

- Deployment constraints: Switches running SONiC on External fabrics

Deployment constraints: Switches running SONiC on External fabrics

- **Config compliance and freeform**—For External fabrics, config compliance and freeform configs are not supported for switches running on the SONiC image. The **Replay config** option is provided instead for the user to upload the JSON format configurations to merge or replace the running configurations on the SONiC switches.

Switches running SONiC: General information

- General guidelines and limitations: Switches running SONiC
- General deployment constraints: Switches running SONiC
- General deployment model: Switches running SONiC
- Troubleshooting issues

General guidelines and limitations: Switches running SONiC

- A fabric that contains switches running SONiC images can only contain those switch types. You cannot have a mix of switches running SONiC images and switches with other types of images in the same fabric.
- For switches running on SONiC, you can only update the image on those switches using a SONiC image. In addition, you must choose **Disruptive** as the update type for SONiC images. See Managing Your Fabric Software for more information.
- Having the same subnet scope range between NX-OS fabrics and fabrics with switches running SONiC images is not supported in this release.

General deployment constraints: Switches running SONiC

- **Topology discovery**—LLDP-based only.

General deployment model: Switches running SONiC

Telemetry is only supported in a co-hosted deployment (the colocation model is not supported). This means fabrics with switches running SONiC images are not supported on virtual cluster (app 500G) cluster types or colocated setups.

Troubleshooting issues

To help facilitate troubleshooting, Nexus Dashboard provides **Begin** and **End** troubleshooting actions to create or update the credential of the user rescue-user with corresponding permissions on switches running SONiC images. You can then access the switches with the username rescue-user

and execute commands for troubleshooting. Once you are finished with troubleshooting, you can choose the **End troubleshooting** action to remove rescue-user from the switches.

AI QoS classification and queuing policies

These sections provide information about the AI QoS classification and queuing policies.

- Understanding AI QoS classification and queuing policies
- Guidelines and limitations for AI QoS classification and queuing policies
- Configure AI QoS classification and queuing policies
- Create a policy using the custom QoS templates

Understanding AI QoS classification and queuing policies

Support is available for configuring a low latency, high throughput, and lossless fabric configuration that can be used for artificial intelligence (AI) fabric based traffic.

The AI QoS feature allows you to:

- Easily configure a network with homogeneous interface speeds, where most or all of the links run at 400Gb, 100Gb, or 25Gb speeds.
- Provide customizations to override the predominate queuing policy for a host interface.

When you apply the AI QoS policy, Nexus Dashboard will automatically pre-configure any inter-fabric links with QoS and system queuing policies, and will also enable Priority Flow Control (PFC). If you enable the AI QoS feature on a VXLAN EVPN fabric, then the Network Virtual (NVE) interface will have the attached AI QoS policies.

You can enable this feature and set queuing policy parameters based on interface speed using new fields available during BGP fabric configuration.

+ Policies defined with these custom Classification and Queuing templates can be used in various host interface polices. For more information, see Create a policy using the custom QoS templates.

When enabling the AI feature, priority-flow-control watchdog-interval on is enabled on all of your configured devices, intra-fabric links, and all your host interfaces where Priority Flow Control (PFC) is also enabled. The PFC watchdog interval is for detecting whether packets in a no-drop queue are being drained within a specified time period. This release also adds the **Priority flow control watch-dog interval** field on the **Advanced tab**. When you create or edit a Data Center VXLAN EVPN fabric or other fabrics and AI is enabled, you can set the **Priority flow control watch-dog interval** field to a non-system default value (the default is 100 milliseconds). For more information on the PFC watchdog interval for Cisco NX-OS, see Configuring a priority flow control watchdog Interval in the *Cisco N9000 Series NX-OS Quality of Service Configuration Guide*.

If you perform an upgrade from an earlier release, and then do a **Recalculate and deploy**, you may see additional **priority-flow-control watchdog-interval on** configurations.

Guidelines and limitations for AI QoS classification and queuing policies

Following are the guidelines and limitations for the AI QoS and queuing policy feature:

- Apply AI QoS policies at the fabric level rather than on individual switches to ensure consistent traffic classification and uniform policy enforcement.
- On Cisco N9K-C9808 and N9K-C9804 series switches, the command **priority-flow-control watch-dog-interval** is not supported in either global or interface configuration modes and the command **hardware qos nodrop-queue-thresholds queue-green** is not supported in global configuration mode.
- Cisco N9K-C9808 and N9K-C9804 series switches only support AI fabric type from NX-OS version 10.5(1) and later.
- This feature does not automate any per-interface speed settings.
- This feature is supported only on Nexus devices with Cisco Cloud Scale technology, such as the Cisco N9300-FX2, Cisco N9300-FX3, Cisco N9300-GX, and Cisco N9300-GX2 series switches.
- This feature is not supported in fabrics with devices that are assigned with a ToR role.

Configure AI QoS classification and queuing policies

Follow these steps to configure AI QoS and queuing policies:

1. Enable AI QoS and queuing policies at the fabric level.
   a. Create a fabric as you normally would.
   b. In the **Advanced** tab in those instructions, make the necessary selections to configure AI QoS and queuing policies at the fabric level.
   c. Configure any remaining fabric-level settings as necessary in the remaining tabs.
   d. When you have completed all the necessary fabric-level configurations, click **Save**, then click **Recalculate and deploy**.

      At this point in the process, the network QoS and queuing policies are configured on each device, the classification policy is configured on NVE interfaces (if applicable), and priority flow control and classification policy is configured on all intra-fabric link interfaces.

2. For host interfaces, selectively enable priority flow control, QoS, and queuing by editing the policy associated with that host interface.

   See Working with Connectivity for LAN Fabrics for more information.

   a. Within a fabric where you enabled AI QoS and queuing policies in the previous step, click the **Interfaces** tab.

      The configured interfaces within this fabric are displayed.

   b. Locate the host interface where you want to enable AI QoS and queuing policies, then click the box next to that host interface to select it and click **Actions > Edit**.

      The **Edit Interfaces** page is displayed.

   c. In the **Policy** field, verify that the policy that is associated with this interface contains the necessary fields that will allow you to enable AI QoS and queuing policies on this host interface.

      For example, these policy templates contain the necessary AI QoS and queuing policies fields:

- int_access_host
- int_dot1q_tunnel_host
- int_pvlan_host
- int_routed_host
- int_trunk_host

d. Locate the **Enable priority flow control** field and click the box next to this field to enable Priority Flow Control for this host interface.

e. In the **Enable QoS Configuration** field, click the box next to this field to enable AI QoS for this host interface.

   This enables the QoS classification on this interface if AI queuing is enabled at the fabric level.

f. If you checked the box next to the **Enable QoS Configuration** field in the previous step and you created a custom QoS policy using the procedures provided in Create a policy using the custom QoS templates, enter that custom QoS classification policy in the **Custom QoS Policy for this interface** field to associate that custom QoS policy with this host interface, if necessary.

   If this field is left blank, then Nexus Dashboard will use the default QOS_CLASSIFICATION policy, if available.

g. If you created a custom queuing policy using the procedures provided in Create a policy using the custom QoS templates, enter that custom queuing policy in the **Custom Queuing Policy for this interface** field to associate that custom queuing policy with this host interface, if desired.

h. Click **Save** when you have completed the AI QoS and queuing policy configurations for this host interface.

Create a policy using the custom QoS templates

Follow these procedures to use the custom QoS templates to create a policy, if desired. See Managing Your Template Library for general information on templates.

1. Within a fabric where you enabled AI QoS and queuing policies, click **Inventory > Switches**, then double-click the switch that has the host interface where you enabled AI QoS and queuing policies.

   The **Switch overview** page for that switch appears.

2. Choose **Configuration Policies > Policies**.
3. Click **Actions > Add policy**.

   The **Create Policy** page appears.

4. Set the priority and enter a description for the new policy.

   Note that the priority for this policy must be lower (must come before) the priority that was set for the host interface.

5. In the **Select Template** field, click the **No Policy Selected** text.

The **Select Policy Template** page appears.

1. Make the necessary QoS classification or queuing configurations in the template that you selected, then click **Save**.

Any custom QoS policy created using these procedures are now available to use when you configure QoS and queuing policies for the host interface.

Border Gateway Support

The following border gateway roles are now supported in an eBGP fabric:

- Border gateway
- Border gateway spine
- Border gateway super spine

Support is also available for eBGP fabrics as child fabrics within a VXLAN fabric group. Several new fields are available when creating a BGP fabric that allow you to enter an ASN and to advertise the Anycast Border Gateway PIP as a Virtual Tunnel End Point (VTEP) for the BGP fabric in the VXLAN fabric group.

The way that the BGP ASN is allocated for the border gateway will match the settings for the leaf switches and border switches. You also might have to make additional configurations, based on the device role that you assign to the switch:

- If you set the device role as **border gateway**, then you must manually add the **leaf_bgp_asn** policy to the device to specify the BGP ASN that will be used on the switch, following the same Multi-AS mode or Same-Tier-AS mode rule that you have set at the fabric level.
- If you set the device role as **border gateway spine**, then the device ASN is the same value as the value provided in the **BGP ASN for Spines** option at the fabric level.
- If you set the device role as **border gateway super spine**, then the device ASN is the same value as the value provided in the **BGP ASN for Super Spines** option at the fabric level.

Nexus Dashboard Fabric Controller will automatically generate the BGP underlay configuration after you click **Recalculate & Deploy** at the end of the configuration process. You will then need to manually add the eBGP overlay policies for the border gateways using the procedures provided in the section " Deploying Fabric Overlay eBGP Policies" in

Border Gateway Route Filtering

Two prefix-list policies are created for each border gateway when you perform a **Recalculate & Deploy** at the VXLAN EVPN Multi-Site fabric level:

- **fabric-internal**: This policy is applied to an eBGP session toward fabric internal, and is used to avoid routes that are learned from DCI from being propagated to other nodes within the fabric.

  For this policy, permit only the following:

  - Routing loopback (typically loopback0) subnet
  - VTEP loopback (typically loopback1) subnet

- Anycast RP loopback (typically loopback254) subnet
- Multisite loopback IP address

Following is a sample configuration for this policy:

```
ip prefix-list ebgp-fabric-to-internal seq 10 permit 8.2.0.0/22 eq 32
ip prefix-list ebgp-fabric-to-internal seq 20 permit 8.3.0.0/22 eq 32
ip prefix-list ebgp-fabric-to-internal seq 30 permit 8.254.254.0/24 eq 32
ip prefix-list ebgp-fabric-to-internal seq 40 permit 10.10.0.2/32
```

```
route-map ebgp-rmap-filter-to-internal permit 10
 match ip address prefix-list ebgp-fabric-to-internal
route-map ebgp-rmap-filter-to-internal deny 20
```

```
router bgp 65028
 neighbor 8.4.0.1
   address-family ipv4 unicast
     route-map ebgp-rmap-filter-to-internal out
```

- **fabric-to-dci**: This policy is applied to the VXLAN EVPN Multi-Site underlay eBGP session, and is used to avoid VXLAN fabric underlay routes from being propagated to external fabrics.

  For this policy, permit only the BGW and the local loopback interface IP address of its neighbor BGW (loopback0/1/VPC VIP/multisite loopback).

Following is a sample configuration for this policy:

```
ip prefix-list ebgp-fabric-to-dci seq 10 permit 8.2.0.2/32
ip prefix-list ebgp-fabric-to-dci seq 20 permit 8.3.0.1/32
ip prefix-list ebgp-fabric-to-dci seq 30 permit 10.10.0.2/32
```

```
route-map ebgp-rmap-filter-to-dci permit 10
 match ip address prefix-list ebgp-fabric-to-dci
route-map ebgp-rmap-filter-to-dci deny 20
```

```
router bgp 65028
 neighbor 10.10.1.2
   address-family ipv4 unicast
     route-map ebgp-rmap-filter-to-dci out
     next-hop-self
```

You can edit either of these two prefix-list policies using the template ipv4_prefix_list_internal to add the custom prefix-list entries. For more information on adding or editing a policy, see the section "Add a policy" in Working with Configuration Policies for Your Nexus Dashboard LAN or IPFM Fabrics.

Guidelines and Limitations

Following are the guidelines and limitations with the border gateway support in BGP fabrics:

- Border gateway support for an eBGP fabric is only available when EVPN is enabled, and when the underlay is IPv4.
- Tenant Routed Multicast (TRM) is also supported with border gateway roles in a BGP fabric.

Layer 3 VNI without VLAN

Following is the upper-level process to enable the Layer 3 VNI without VLAN feature in a fabric:

1. (Optional) When configuring a new fabric, check the **Enable L3VNI w/o VLAN** field to enable the Layer 3 VNI without VLAN feature at the fabric level. The setting at this fabric-level field affects the related field at the VRF level, as described below.
2. When creating or editing a VRF, check the **Enable L3VNI w/o VLAN** field to enable the Layer 3 VNI without VLAN feature at the VRF level. The default setting for this field varies depending on the following factors:
    - For existing VRFs, the default setting is disabled (the **Enable L3VNI w/o VLAN** box is unchecked).
    - For newly-created VRFs, the default setting is inherited from the fabric settings, as described above.
    - This field is a per-VXLAN fabric variable. For VRFs that are created from a VXLAN EVPN Multi-Site fabric, the value of this field is inherited from the fabric setting in the child fabric. You can edit the VRF in the child fabric to change the value, if desired.

      See the "Create a VRF" section in Working with Segmentation and Security for Your Nexus Dashboard VXLAN Fabric for more information.

The VRF attachment (new or edited) then uses the new Layer 3 VNI without VLAN mode if the following conditions are met:

- The **Enable L3VNI w/o VLAN** is enabled at the VRF level
- The switch supports this feature and the switch is running on the correct release (see Guidelines and limitations for Layer 3 VNI without VLAN)

The VLAN is ignored in the VRF attachment when these conditions are met.

Guidelines and limitations for Layer 3 VNI without VLAN

Following are the guidelines and limitations for the Layer 3 without VLAN feature:

- The Layer 3 VNI without VLAN feature is supported on the -EX, -FX, and -GX versions of the Cisco N9000 switches. When you enable this feature at the VRF level, the feature setting on the

VRF will be ignored on switch models that do not support this feature.

- When used in a Campus VXLAN EVPN fabric, this feature is only supported on Cisco N9000 series switches in that type of fabric. This feature is not supported on Cisco Catalyst 9000 series switches in the Campus VXLAN EVPN fabric; those switches require VLANs for Layer 3 VNI configurations.
- This feature is supported on switches running on NX-OS release 10.3.1 or later. If you enable this feature at the VRF level, the feature setting on the VRF is ignored on switches running an NX-OS image earlier than 10.3.1.
- When you perform a brownfield import in a Data Center VXLAN EVPN fabric, if one switch configuration is set with the **Enable L3VNI w/o VLAN** configuration at the VRF level, then you should also configure this same setting for the rest of the switches in the same fabric that are associated with this VRF, if the switch models and images support this feature.
- For VXLAN fabric or fabric group deployments with Campus switches, use a VNI range within 4096 to 1,677,721 to align with the best practices for VNI allocation in Campus environments.

MACsec support in Data Center VXLAN EVPN and BGP fabrics

MACsec is supported in Data Center VXLAN EVPN and BGP fabrics for intra-fabric links. You should enable MACsec on the fabric and on each required intra-fabric link to configure MACsec. Unlike CloudSec, auto-configuration of MACsec is not supported.

Support is available for MACsec for inter-fabric links, DCI MACsec, with a QKD server to generate keys or using preshared keys. For more information, see the section "About connecting two NX-OS fabrics with MACsec using QKD" in Creating Fabrics and Fabric Groups.

MACsec is supported on switches with minimum Cisco NX-OS releases 7.0(3), I7(8), and 9.3(5).

Guidelines

- If MACsec cannot be configured on the physical interfaces of the link, an error is displayed when you click **Save**. MACsec cannot be configured on the device and link due to the following reasons:
  - The minimum NX-OS version is not met.
  - The interface is not MACsec capable.
- MACsec global parameters in the fabric settings can be changed at any time.
- MACsec and CloudSec can coexist on a BGW device.
- MACsec status of a link with MACsec enabled is displayed on the **Links** page.
- Brownfield migration of devices with MACsec configured is supported using switch and interface freeform configs.

  For more information about MACsec configuration, which includes supported platforms and releases, see the Configuring MACsec chapter in *Cisco N9000 Series NX-OS Security Configuration Guide*.

The following sections show how to enable and disable MACsec in Nexus Dashboard.

Enable MACsec

The process described in this section is for configuring MACsec for intra-fabric links and inter-fabric links.

For information on configuring data center interconnect (DCI) MACsec for inter-fabric links and using a quantum key distribution (QKD) server, see the section "About connecting two NX-OS fabrics with MACsec using QKD" in Creating Fabrics and Fabric Groups.

1. Navigate to **Manage > Fabrics**.
2. Click **Create Fabric** to create a new fabric or click **Actions > Edit Fabric Settings** on an existing Data Center VXLAN EVPN fabric.
3. Click **Fabric Management > Security** and specify the MACsec details.

Field	Description
Enable MACsec	Check the check box to enable MACsec in the fabric. MACsec configuration is not generated until MACsec is enabled on an intra-fabric link. Perform a **Recalculate and deploy** operation to generate the MACsec configuration and deploy the configuration on the switch.
MACsec Cipher Suite	Choose one of the following MACsec cipher suites for the MACsec policy: · **GCM-AES-128** · **GCM-AES-256** · **GCM-AES-XPN-128** · **GCM-AES-XPN-256** The default value is **GCM-AES-XPN-256**.
MACsec Primary Key String	Specify a Cisco Type 7 encrypted octet string that is used for establishing the primary MACsec session. For **AES_256_CMAC**, the key string length must be 130 and for **AES_128_CMAC**, the key string length must be 66. If these values are not specified correctly, an error is displayed when you save the fabric. The default key lifetime is infinite.
MACsec Primary Cryptographic Algorithm	Choose the cryptographic algorithm used for the primary key string. It can be **AES_128_CMAC** or **AES_256_CMAC**. The default value is **AES_128_CMAC**. You can configure a fallback key on the device to initiate a backup session if the primary session fails.
MACsec Fallback Key String	Specify a Cisco Type 7 encrypted octet string that is used for establishing a fallback MACsec session. For **AES_256_CMAC**, the key string length must be 130 and for **AES_128_CMAC**, the key string length must be 66. If these values are not specified correctly, an error is displayed when you save the fabric.

Field	Description
MACsec Fallback Cryptographic Algorithm	Choose the cryptographic algorithm used for the fallback key string. It can be **AES_128_CMAC** or **AES_256_CMAC**. The default value is **AES_128_CMAC**.
Enable DCI MACsec	Check the check box to enable DCI MACsec on DCI links.  If you enable the **Enable DCI MACsec** option and disable the **Use Link MACsec Setting** option, Nexus Dashboard uses the fabric settings for configuring DCI MACsec on the DCI links.
Enable QKD	Check the check box to enable the QKD server for generating quantum keys for encryption.  If you choose to not enable the **Enable QKD** option, Nexus Dashboard generates preshared keys instead of using the QKD server to generate the keys. If you disable the **Enable QKD** option, all the fields pertaining to QKD are grayed out.
DCI MACsec Cipher Suite	Choose one of the following DCI MACsec cipher suites for the DCI MACsec policy: · **GCM-AES-128** · **GCM-AES-256** · **GCM-AES-XPN-128** · **GCM-AES-XPN-256** The default value is **GCM-AES-XPN-256**.
DCI MACsec Primary Key String	Specify a Cisco Type 7 encrypted octet string that is used for establishing the primary DCI MACsec session. For **AES_256_CMAC**, the key string length must be 130 and for **AES_128_CMAC**, the key string length must be 66. If these values are not specified correctly, an error is displayed when you save the fabric.  The default key lifetime is infinite.
DCI MACsec Primary Cryptographic Algorithm	Choose the cryptographic algorithm used for the primary key string. It can be **AES_128_CMAC** or **AES_256_CMAC**. The default value is **AES_128_CMAC**. You can configure a fallback key on the device to initiate a backup session if the primary session fails.

Field	Description
DCI MACsec Fallback Key String	Specify a Cisco Type 7 encrypted octet string that is used for establishing a fallback MACsec session. For **AES_256_CMAC**, the key string length must be 130 and for **AES_128_CMAC**, the key string length must be 66. If these values are not specified correctly, an error is displayed when you save the fabric.  If you disabled the **Enable QKD** option, you need to specify the **DCI MACsec Fallback Key String** option.
DCI MACsec Fallback Cryptographic Algorithm	Choose the cryptographic algorithm used for the fallback key string. It can be **AES_128_CMAC** or **AES_256_CMAC**. The default value is **AES_128_CMAC**.
QKD Profile Name	Specify the crypto profile name. The maximum size is 63.
KME Server IP	Specify the IPv4 address for the Key Management Entity (KME) server.
KME Server Port Number	Specify the port number for the KME server.
Trustpoint Label	Specify the authentication type trustpoint label. The maximum size is 64.
Ignore Certificate	Enable this check box to skip verification of incoming certificates.
MACsec Status Report Timer	Specify the MACsec operational status periodic report timer in minutes.

4. Click **Save**.
5. Click a fabric to view the **Summary** in the side panel.
6. Click the **Launch** icon to open the **Fabric Overview** page.
7. Click the **Fabric Overview > Links** tab.
8. Choose an inter-fabric connection (IFC) link on which you want to enable MACsec and click **Actions > Edit**. For more information on creating a VRF-Lite inter-fabric link, see the section "Establishing inter-fabric connectivity using VRF Lite" Working with Connectivity in Your Nexus Dashboard LAN Fabrics.

   The **Link Management - Edit Link** page displays.
9. From the **Link Management - Edit Link** page, navigate to the **Security** tab and enter the required parameters.

   For an inter-fabric link, the **Enable MACsec** option is in Security.

Field	Description
Enable MACsec	Check the check box to enable MACsec on the VRF-Lite inter-fabric connection (IFC) link. If you enable the **Enable MACsec** option and you disable the **Use Link MACsec Setting** option, Nexus Dashboard uses the fabric settings for configuring MACsec on the VRF-Lite IFC. When MACsec is configured on the link, Nexus Dashboard generates the following configurations: - Switch-level MACsec configurations if this is the first link that enables MACsec. - MACsec configurations for the link.
Source MACsec Policy/Key-Chain Name Prefix	Specify the prefix for the policy and key-chain names for the MACsec configuration at the source. The default value is **DCI**, and you can change the value.
Destination MACsec Policy/Key-Chain Name Prefix	Specify the prefix for the policy and key-chain names for the MACsec configuration at the destination. The default value is **DCI**, and you can change the value.
Enable QKD	Check the check box to enable the QKD server for generating quantum keys for encryption. If you choose to not enable the **Enable QKD** option, Nexus Dashboard uses preshared keys provided by the user instead of using the QKD server to generate the keys. If you disable the **Enable QKD** option, all the fields pertaining to QKD are grayed out.
Use Link MACsec Setting	Check this check box as the override option for using the link settings instead of using the fabric settings.
MACsec Cipher Suite	Choose one of the following MACsec cipher suites for the MACsec policy: - **GCM-AES-128** - **GCM-AES-256** - **GCM-AES-XPN-128** - **GCM-AES-XPN-256** The default value is **GCM-AES-XPN-256**.

Field	Description
MACsec Primary Key String	Specify a Cisco Type 7 encrypted octet string that is used for establishing the primary DCI MACsec session. For **AES_256_CMAC**, the key string length must be 130 and for **AES_128_CMAC**, the key string length must be 66. If these values are not specified correctly, an error is displayed when you save the fabric.  The default key lifetime is infinite.
MACsec Primary Cryptographic Algorithm	Choose the cryptographic algorithm used for the primary key string. It can be **AES_128_CMAC** or **AES_256_CMAC**. The default value is **AES_128_CMAC**. You can configure a fallback key on the device to initiate a backup session if the primary session fails.
MACsec Fallback Key String	Specify a Cisco Type 7 encrypted octet string that is used for establishing a fallback MACsec session. For **AES_256_CMAC**, the key string length must be 130 and for **AES_128_CMAC**, the key string length must be 66. If these values are not specified correctly, an error is displayed when you save the fabric.  This parameter is mandatory if **Enable QKD** is not selected.
MACsec Fallback Cryptographic Algorithm	Choose the cryptographic algorithm used for the fallback key string. It can be **AES_128_CMAC** or **AES_256_CMAC**. The default value is **AES_128_CMAC**.
Source QKD Profile Name	Specify the source crypto profile name. The maximum size is 63.
Source KME Server IP	Specify the source IPv4 address for the Key Management Entity (KME) server.
Source KME Server Port Number	Specify the source port number for the KMEserver.
Source Trustpoint Label	Specify the source authentication type trustpoint label. The maximum size is 64.
Destination QKD Profile Name	Specify the destination crypto profile name.
Destination KME Server IP	Specify the destination IPv4 address for the KME server.
Destination KME Server Port Number	Specify the destination port number for the KME server.
Destination Trustpoint Label	Specify the destination authentication type trustpoint label. The maximum size is 64.
Ignore Certificate	Specify if you want to skip verification of incoming certificates.

10. Click **Save**.
11. From the **Fabric Overview > Switches** tab, select **Actions > Recalculate and deploy** to deploy the MACsec configuration on the switch.

Disable MACsec

The process described in this section is for disabling MACsec on an inter-fabric link for Data Center VXLAN EVPN and BGP fabrics.

For information on disabling MACsec for an inter-fabric link with a quantum key distribution (QKD) server, see the section "About connecting two NX-OS fabrics with MACsec using QKD" in Creating Fabrics and Fabric Groups].

Nexus Dashboard performs the following operations when you disable MACsec:

- Disables MACsec on an inter-fabric link using QKD or a preshared key.
- If this is the last link where MACsec is enabled, Nexus Dashboard deletes the switch-level MACsec configuration from the switch.
    1. Navigate to the **Fabric Overview > Links** tab.
    2. Choose the link on which you want to disable MACsec on the inter-fabric link.
    3. From the **Link Management - Edit Link** page, navigate to the appropriate tab and unselect the **Enable MACsec** option.

       For an intra-fabric link, the **Enable MACsec** option is in the Advanced tab.

       For an inter-fabric link, the **Enable MACsec** option is in the Security tab.
    4. Click **Save**.
    5. From the **Fabric Overview > Switches** tab, select **Actions > Deploy** to remove the MACsec configuration from the switch.

vPC fabric peering

vPC Fabric Peering provides an enhanced dual-homing access solution without the overhead of wasting physical ports for vPC Peer Link. This feature preserves all the characteristics of a traditional vPC. For more information, see *Information about vPC Fabric Peering* section in *Cisco N9000 Series NX-OS VXLAN Configuration Guide*.

You can create a virtual peer link for two switches or change the existing physical peer link to a virtual peer link. Nexus Dashboard supports vPC fabric peering in both greenfield as well as brownfield deployments. This feature is applicable for **Data Center VXLAN EVPN** and **BGP Fabric** fabric templates.

The **BGP Fabric** fabric does not support brownfield import.

Guidelines and limitations

The following are the guidelines and limitations for vPC fabric pairing.

- vPC fabric peering is supported from Cisco NX-OS Release 9.2(3).
- Only the Cisco N9K-C9332C Switch, Cisco N9K-C9364C Switch, Cisco N9K-C9348GC-FXP Switch, and Cisco N9000 Series Switches that end with FX or FX2 support vPC fabric peering.
- Cisco N9K-C93180YC-FX3S and N9K-C93108TC-FX3P platform switches support vPC fabric peering.
- Cisco N9300-EX, and 9300-FX/FXP/FX2/FX3/GX/GX2 platform switches support vPC Fabric Peering. Cisco N9200 and Cisco N9500 platform switches do not support vPC Fabric Peering. For more information, see *Guidelines and Limitations for vPC Fabric Peering* section in *Cisco N9000 Series NX-OS VXLAN Configuration Guide*.
- If you use other Cisco N9000 Series Switches, a warning will appear during **Recalculate & Deploy**. A warning appears in this case because these switches will be supported in future releases.
- If you try pairing switches that do not support vPC fabric peering, using the **Use Virtual Peerlink** option, a warning will appear when you deploy the fabric.
- You can convert a physical peer link to a virtual peer link and vice-versa with or without overlays.
- Switches with border gateway leaf roles do not support vPC fabric peering.
- vPC fabric peering is not supported for Cisco N9000 Series Modular Chassis and FEXs. An error appears during **Recalculate & Deploy** if you try to pair any of these.
- Brownfield deployments and greenfield deployments support vPC fabric peering in Cisco Nexus Dashboard.
- However, you can import switches that are connected using physical peer links and convert the physical peer links to virtual peer links after **Recalculate & Deploy**. To update a TCAM region during the feature configuration, use the hardware access-list tcam ingress-flow redirect512 command in the configuration terminal.

QoS for fabric vPC-peering

In the **Data Center VXLAN EVPN** fabric settings, you can enable QoS on spines for guaranteed delivery of vPC Fabric Peering communication. Additionally, you can specify the QoS policy name.

Note the following guidelines for a greenfield deployment:

- If QoS is enabled and the fabric is newly created:
  - If spines or super spines neighbor is a virtual vPC, make sure neighbor is not honored from invalid links, for example, super spine to leaf or borders to spine when super spine is present.
  - Based on the Cisco N9000 Series Switch model, create the recommended global QoS config using the **switch_freeform** policy template.
  - Enable QoS on fabric links from spine to the correct neighbor.
- If the QoS policy name is edited, make sure policy name change is honored everywhere, that is, global and links.
- If QoS is disabled, delete all configuration related to QoS fabric vPC peering.
- If there is no change, then honor the existing PTI.

For more information about a greenfield deployment, see Creating Fabrics and Fabric Groups.

Note the following guidelines for a brownfield deployment:

Brownfield Scenario 1:

- If QoS is enabled and the policy name is specified:

  You need to enable only when the policy name for the global QoS and neighbor link service policy is same for all the fabric vPC peering connected spines.

  - Capture the QoS configuration from switch based on the policy name and filter it from unaccounted configuration based on the policy name and put the configuration in the **switch_freeform** with PTI description.
  - Create service policy configuration for the fabric interfaces as well.
  - Greenfield configuration should make sure to honor the brownfield configuration.
- If the QoS policy name is edited, delete the existing policies and brownfield extra configuration as well, and follow the greenfield flow with the recommended configuration.
- If QoS is disabled, delete all the configuration related to QoS fabric vPC peering.

  No cross check for possible or error mismatch user configuration, and user might see the diff.

Brownfield Scenario 2:

- If QoS is enabled and the policy name is not specified, QoS configuration is part of the unaccounted switch freeform config.
- If QoS is enabled from fabric settings after **Recalculate & Deploy** for brownfield, QoS configuration overlaps and you will see the diff if fabric vPC peering config is already present.

For more information about a brownfield deployment, see Creating Fabrics and Fabric Groups.

To view the vPC pairing page of a switch, from the fabric topology page, right-click the switch and choose **vPC Pairing**. The vPC pairing page for a switch has the following fields:

Field	Description
Use Virtual Peerlink	Allows you to enable or disable the virtual peer linking between switches.
Switch name	Specifies all the peer switches in a fabric.NOTE: When you have not paired any peer switches, you can see all the switches in a fabric. After you pair a peer switch, you can see only the peer switch in the vPC pairing page.
Recommended	Specifies if the peer switch can be paired with the selected switch. Valid values are **true** and **false**. Recommended peer switches will be set to **true**.

Field	Description
Reason	Specifies why the vPC pairing between the selected switch and the peer switches is possible or not possible.
Serial Number	Specifies the serial number of the peer switches.

You can perform the following with the **vPC Pairing** option:

Create a virtual peer link

Follow these steps to create a virtual peer link.

1. Choose **Home > Topology**.
2. Choose a fabric with the **Data Center VXLAN EVPN** or **BGP Fabric** fabric type.
3. Right-click a switch and choose **vPC Pairing** from the drop-down list.

   The page to choose the peer appears.

   Alternatively, you can also navigate to the fabric **Overview** page. Choose a switch in the **Inventory > Switches** tab and click on **Actions** > **vPC Pairing** to create, edit, or unpair a vPC pair. However, you can use this option only when you choose a Cisco Nexus switch.

   You will get the following error when you choose a switch with the border gateway leaf role. <switch-name> has a Network/VRF attached. Please detach the Network/VRF before vPC Pairing/Unpairing

4. Check the **Use Virtual Peerlink** check box.
5. Choose a peer switch and check the **Recommended** column to see if pairing is possible.

   If the value is **true**, pairing is possible. You can pair switches even if the recommendation is **false**. However, you will get a warning or error during **Recalculate & Deploy**.

6. Click **Save**.
7. In the **Topology** page, choose **Recalculate & Deploy**.

   The **Deploy Configuration** page appears.

8. Click the field against the switch in the **Preview Config** column.

   The **Config Preview** page appears for the switch.

9. View the vPC link details in the pending configuration and side-by-side configuration.
10. Close the page.
11. Click the pending errors icon next to **Recalculate & Deploy** icon to view errors and warnings, if any.

    If you see any warnings that are related to TCAM, click the **Resolve** icon. A confirmation dialog box about reloading switches appears. Click **OK**. You can also reload the switches from the

topology page. For more information, see *Guidelines and Limitations for vPC Fabric Peering* and *Migrating from vPC to vPC Fabric Peering* sections in *Cisco N9000 Series NX-OS VXLAN Configuration Guide*.

The switches that are connected through vPC fabric peering, are enclosed in a gray cloud.

Convert a physical peer link to a virtual peer link

Before you begin

- Perform the conversion from physical peer link to virtual peer link during the maintenance window of switches.
- Ensure the switches support vPC fabric peering. Only the following switches support vPC fabric peering:
  - Cisco N9K-C9332C Switch, Cisco N9K-C9364C Switch, and Cisco N9K-C9348GC-FXP Switch.
  - Cisco N9000 Series Switches that ends with FX, FX2, and FX2-Z.
  - Cisco N9300-EX, and Cisco N9300-FX/FXP/FX2/FX3/GX/GX2 platform switches. For more information, see *Guidelines and Limitations for vPC Fabric Peering* section in *Cisco N9000 Series NX-OS VXLAN Configuration Guide*.

Follow these steps to convert a physical peer link to a virtual peer link.

1. Choose **Home > Topology**.
2. Choose a fabric with the **Data Center VXLAN EVPN** or **BGP Fabric** fabric type.
3. Right-click the switch that is connected using the physical peer link and choose **vPC Pairing** from the drop-down list.

   The page to choose the peer appears.

   Alternatively, you can also navigate to the fabric **Overview** page. Choose a switch in the **Inventory > Switches** tab and click on **Actions** > **vPC Pairing** to create, edit, or unpair a vPC pair. However, you can use this option only when you choose a Cisco Nexus switch.

   You will get the following error when you choose a switch with the border gateway leaf role. <switch-name> has a Network/VRF attached. Please detach the Network/VRF before vPC Pairing/Unpairing

4. Check the **Recommended** column to see if pairing is possible.

   If the value is **true**, pairing is possible. You can pair switches even if the recommendation is **false**. However, you will get a warning or error during **Recalculate & Deploy**.

5. Check the **Use Virtual Peerlink** check box.

   The **Unpair** icon changes to **Save**.

6. Click **Save**.

After you click **Save**, the physical vPC peer link is automatically deleted between the switches even without deployment.

7. In the **Topology** page, choose **Recalculate & Deploy**.

   The **Deploy Configuration** page appears.

8. Click the field against the switch in the **Preview Config** column.

   The **Config Preview** page appears for the switch.

9. View the vPC link details in the pending configuration and the side-by-side configuration.
10. Close the page.
11. Click the pending errors icon next to the **Recalculate & Deploy** icon to view errors and warnings, if any.

    If you see any warnings that are related to TCAM, click the **Resolve** icon. A confirmation dialog box about reloading switches appears. Click **OK**. You can also reload the switches from the fabric topology page.

    The physical peer link between the peer switches turns red. Delete this link. The switches are connected only through a virtual peer link and are enclosed in a gray cloud.

Convert a virtual peer link to a physical peer link

Before you begin

Connect the switches using a physical peer link before disabling the vPC fabric peering.

Follow these steps to convert a virtual peer link to a physical peer link.

1. Choose **Home > Topology**.
2. Choose a fabric with the **Data Center VXLAN EVPN** or **BGP Fabric** fabric type.
3. Right-click the switch that is connected through a virtual peer link and choose **vPC Pairing** from the drop-down list.

   The page to choose the peer appears.

Alternatively, you can also navigate to the fabric **Overview** page. Choose a switch in the **Inventory > Switches** tab and click on **Actions** > **vPC Pairing** to create, edit, or unpair a vPC pair. However, you can use this option only when you choose a Cisco Nexus switch.

4. Uncheck the **Use Virtual Peerlink** check box.

   The **Unpair** icon changes to **Save**.

5. Click **Save**.
6. In the **Topology** page, choose **Recalculate & Deploy**.

   The **Deploy Configuration** page appears.

7. Click the field against the switch in the **Preview Config** column.

   The **Config Preview** page appears for the switch.

8. View the vPC peer link details in the pending configuration and the side-by-side configuration.
9. Close the page.
10. Click the pending errors icon next to the **Recalculate & Deploy** icon to view errors and warnings, if any.

    If you see any warnings that are related to TCAM, click the **Resolve** icon. The confirmation dialog box about reloading switches appears. Click **OK**. You can also reload the switches from the fabric topology page.

    The virtual peer link, represented by a gray cloud, disappears and the peer switches are connected through a physical peer link.

Deploying Fabric Underlay eBGP Policies

To deploy fabric underlay eBGP policy, you must manually add the **leaf_bgp_asn** policy on each leaf switch to specify the BGP AS number used on the switch. Implementing the **Recalculate & Deploy** operation afterwards will generate eBGP peering over the physical interface between the leaf and spine switches to exchange underlay reachability information. If **Same-Tier-AS mode** is used, you can deploy the **leaf_bgp_asn** policy on all leafs at the same time as they share the same BGP ASN.

In a Multi-AS fabric, add the the **leaf_bgp_asn** policy on each leaf node and the fabric. In a vPC switch pair, they share the same AS number.

To add a policy to the required switch, see the section "Add a policy" in Working with Configuration Policies for Your Nexus Dashboard LAN or IPFM Fabrics.

Deploying Fabric Overlay eBGP Policies

You must manually add the eBGP overlay policy for overlay peering. Nexus Dashboard provides the built-in eBGP leaf and spine overlay peering policy templates that you must manually add to the eBGP leaf and spine switches to form the EVPN overlay peering.

Deploying Spine Switch Overlay Policies

Add the **ebgp_overlay_spine_all_neighbor** policy on the spine or super spine switches. This policy can be deployed on all the spine switches at once, since they share the same field values. If the network contains spine switches and super spine switches, you must deploy the policy only on the super spine switches.

1. Navigate to **Manage > Fabrics** and double-click on the fabric.

   The **Fabric Overview** window appears.

2. On the **Policies** tab, choose **Actions > Add policy**.
3. Select all the spine switches to which you want to add the **ebgp_overlay_spine_all_neighbor** policy and click **Next**.

The **Create Policy** window appears.

4. Click **Choose Template** and select the **ebgp_overlay_spine_all_neighbor** policy.
5. Enter the necessary field values for the following fields, as required and click **Save**.

Field	Description
Leaf IP List	Specifies the IPv6 or the IPv4 address of the connected leaf switch routing loopback interface. If you have enabled IPv6 underlay, ensure you enter the IPv6 address in this field.
Leaf BGP ASN	The BGP AS numbers of the leaf switches.
BGP Update-Source Interface	This is the source interface for BGP updates. Underlay routing loopback (**loopback0**) is used by default.

1. At the top-right of the **Fabric Overview** window, choose **Actions > Recalculate and deploy**.
2. After the configuration deployment is complete, click **Close**.

   You can use the **Edit policy** option to edit the policy and click **Push Configuration** to deploy the configuration.

Deploying Leaf Switch Overlay Policies

Add the **ebgp_overlay_leaf_all_neighbor** policy on all the leaf switches, to establish eBGP overlay peering towards the spine switch. This policy can be deployed on all leaf switches at once, since they share the same field values.

1. Navigate to **Manage > Fabrics** and double-click on the fabric.

   The **Fabric Overview** window appears.

2. On the **Policies** tab, choose **Actions > Add policy**.
3. Select all the spine switches to which you want to add the **ebgp_overlay_leaf_all_neighbor** policy and click **Next**.

   The **Create Policy** window appears.

4. Click **Choose Template** and select the **ebgp_overlay_leaf_all_neighbor** policy.
5. Enter the necessary field values for the following fields, as required and click **Save**.

Field	Description
Spine/Super Spine IPv4/IPv6 List	Specifies the IPv4 or IPv6 addresses of the routing loopback interfaces of spine or super spine switches for BGP peering. If you have enabled IPv6 underlay, enter the IPv6 address in this field. If the fabric has any super spine or border super spine, provide the IP addresses of the super spines or border super spines.
BGP Update-Source Interface	This is the source interface for BGP updates. Underlay routing loopback (**loopback0**) is used by default.

6. At the top-right of the **Fabric Overview** window, choose **Actions > Recalculate and deploy**.
7. After the configuration deployment is complete, click **Close**.

   You can use the **Edit policy** option to edit the policy and click **Push Configuration** to deploy the configuration.

Adding a Super Spine Switch to an Existing VXLAN BGP EVPN Fabric

If your fabric contains both spine and super spine switches, you must reconfigure the fabric to use the super spine for deploying the overlay between the leaf and the border devices. This topic describes steps to integrate a super spine switch to your existing fabric which has a leaf switch and a spine switch with an overlay between them.

1. To add the **ebgp_overlay_spine_all_neighbor** policy to super spine switches that are newly added to an existing VXLAN BPG EVPN fabric:
   a. Navigate to the **Fabric Overview** window for your fabric and click the **Policies** tab.
   b. Select the super spine switches to which you want to add the **ebgp_overlay_spine_all_neighbor** policy and click **Next**.

      The **Create Policy** window appears.

   c. Click **Choose Template** and select the **ebgp_overlay_spine_all_neighbor** policy.
   d. Enter the IPv4 or IPv6 addresses of the leaf switches in the **Leaf IP List** field.
   e. Enter the AS numbers for the leaf switches in the **Leaf BGP ASN** field and click **Save**.
2. To modify the existing **ebgp_overlay_leaf_all_neighbor** policy on each leaf node:
   a. Find the existing policy by filtering based on **ebgp_overlay_leaf_all_neighbor** template name.

      Ensure you modify only one policy at a time.

   b. Select a policy and choose **Action > Edit policy**.
   c. Enter the IP addresses of the super spine routing loopback interfaces in the **Spine/Super Spine IP List** field and click **Save**.
3. Select the leaf and the super spine switches that you have added in step 2 and choose **Actions > Deploy**.
4. On the **Links** tab, click on Protocol View and verify that eBGP peering between the super spine and leaf switches are established.
5. Remove the existing overlay between the leaf and the spine switches as follows:
   a. On the spine switch, select the **ebgp_overlay_spine_all_neighbor** policy and choose **Actions > Delete policy**.
   b. On the leaf switch, select the **ebgp_overlay_leaf_all_neighbor** policy and choose **Actions > Edit policy**.
   c. Remove the IP address of the spine switch in the **Spine/Super Spine IPv4/IPv6 List** field and click **Save**.
6. To deploy the updated configuration on spine and leaf switches , choose **Actions > Recalculate and deploy** at the top-right of the **Fabric Overview** window.

or

Select the leaf and the spine switches and choose **Actions > Deploy**.

Understanding enhanced analytics for AI fabrics

Enhanced analytics is available for workloads in AI routed and VXLAN fabrics within Nexus Dashboard. This enhancement provides end-to-end visibility and actionable insights for AI infrastructures by integrating job completion and GPU statistics with network statistics, providing a detailed overview of network topologies along with GPUs.

In Nexus Dashboard 4.3.1, enhanced analytics for AI fabrics adds server network interface card (NIC) statistics to AI job monitoring. Before this release, AI job monitoring correlated Slurm job data, GPU telemetry, and switch-side network telemetry. With this release, you can correlate job data with server NIC telemetry, including error, discard, bandwidth, RDMA NACK, CNP, operational status, and link speed counters. You can use this correlation to determine whether job degradation is related to server-side NIC saturation or errors, switch-side network conditions, or both.

You can also use AI job monitoring with an existing AI fabric that is onboarded in Nexus Dashboard as an AI Observability Fabric. For onboarded AI fabrics, Nexus Dashboard can show AI jobs, server NIC statistics, correlated anomalies, and topology information when the required Slurm, GPU, server NIC, and topology telemetry is available.

For server NIC visibility, Nexus Dashboard collects server NIC counters every 30 seconds and retains the data for 15 days. The collected data includes CRC errors, transmit and receive discards, bandwidth Rx and Tx, NACK Rx and Tx, CNP Rx and Tx, operational status, and link speed. Server-interface optics counters are not available from the NVIDIA DOCA Telemetry Service (DTS) in this release.

These enhanced analytics for AI fabrics include:

- An AI resources page, which provides enhanced analytics data on these AI resources:
  - AI clusters
  - Scalable units
  - Servers
  - Live jobs
  - GPUs

  For more information on AI resources, see View information on AI resources.

- An AI job dashboard that correlates job performance with network health (such as optics, congestion, and errors), server NIC health (such as CRC errors, discards, bandwidth, NACK, CNP, operational status, and link speed), and GPU health (such as utilization, memory, temperature, and power). This enables proactive anomaly detection, efficient troubleshooting of job degradation, and optimization of network link utilization. For more information on AI jobs, see View AI job information.

  In order to view job details for AI fabrics, you must first integrate Slurm (Simple Linux Utility for Resource Management), which is an open-source job scheduler used to manage resources on high-performance computing clusters. For more information, see Working with Integrations in

Your Nexus Dashboard.

AI endpoint and topology discovery

Nexus Dashboard release 4.2.1 introduces enhanced discovery capabilities to provide unparalleled visibility into AI workloads. This feature extends traditional network device discovery in Nexus Dashboard to include detailed host-level information, offering a foundational understanding of your AI infrastructure, from network fabric to individual GPU servers.

Nexus Dashboard automatically identifies and collects the following rich data from GPU servers connected to your fabric.

- Host-level resources—Information about network interface cards (NICs) and GPU health.
- Intelligent correlation—Multiple network interfaces (NICs/IP addresses) from a single physical server are intelligently correlated and presented as a unified host entity within Nexus Dashboard. This simplified management provides a holistic view of each server, even with multi-NIC configurations.
- Logical grouping—Nexus Dashboard organizes discovered hosts and their connected leaf switches into logical constructs called scalable units and AI Clusters, which are essential for structured management and analysis.

The core of this enhanced discovery relies on LLDP. When enabled for AI fabric types, Nexus Dashboard actively queries connected devices to gather LLDP details. This protocol provides key host attributes such as system name and description, local and remote port IDs, MAC and IP addresses, and management addresses.

For AI fabric types, Nexus Dashboard enables LLDP endpoint discovery on the switches by default. You must also enable LLDP on the hosts. Nexus Dashboard then automatically filters switches and routers based on endpoint descriptions, such as vendor names in the system description.

Nexus Dashboard introduces specific logical groupings, to effectively manage and visualize AI environments.

- Scalable unit—A logical grouping of leaf switches and servers. Nexus Dashboard calculates the operational status of a scalable unit based on the health of its constituent leaf switches and servers. Individual link statuses (e.g., between a leaf and a server/GPU) do not directly participate in this high-level scalable unit status calculation.
- AI cluster—An AI cluster is essentially a fabric group that provides a dedicated view and management space specifically for your AI infrastructure. You can associate a fabric with an AI cluster if you see the **Associate with AI cluster** option in that fabric, as described in Associate a fabric with an AI cluster.

For more information, see View information on AI resources.

Prerequisites

- The enhanced discovery is specifically available for AI fabric types (AI Routed and VXLAN EVPN fabrics) and the LLDP endpoint discovery is enabled by default for these fabric types. You must also enable LLDP on the hosts as well.
- To organize resources into scalable units, you must assign switches with the leaf role to a scalable

unit through user-driven actions within AI fabrics. See Associate a switch with a scalable unit for more information.

- To view job details for AI fabrics, you must integrate Simple Linux Utility for Resource Management (Slurm) with Nexus Dashboard. For more information, see Working with Integrations in Your Nexus Dashboard.
- To collect GPU statistics, DCGM (Nvidia Data Center GPU Manager) Exporter should be running on the host. See DCGM Exporter for more information.
- To collect server NIC statistics, NVIDIA DOCA Telemetry Service (DTS) and Fluent Bit must be configured on the GPU servers to send server NIC metrics to Nexus Dashboard.
- GPU servers must be reachable from the Nexus Dashboard data network.
- The hostname reported by the Data Center GPU Manager (DCGM) Exporter, Data Center Infrastructure on a Chip Architecture (DOCA) Telemetry Service, Slurm, and LLDP advertisements must match the server hostname.
- You must also run the script to collect GPU-to-NIC mapping information for AI fabric analytics. See GPU/NIC Topology Fleet Collector for more information.

Guidelines and limitations: Analytics for AI fabrics

These are the guidelines and limitations for the analytics for AI fabrics feature:

- Only Config-Only backup and restore operations are supported with this feature. Full backup and restore operations are not supported.
- This functionality is only available for co-hosted clusters.
- To view server NIC statistics, configure the GPU servers to expose NVIDIA DOCA Telemetry Service (DTS) metrics and use Fluent Bit to send the server NIC metrics to Nexus Dashboard.
- Server NIC counters are collected every 30 seconds and retained for 15 days.
- Server-interface optics counters are not available from DTS in this release.
- For external AI fabrics, AI job monitoring depends on successful fabric onboarding, Slurm integration, GPU and NIC topology mapping, and telemetry availability. If switch telemetry is not available, switch-side interface metrics are not displayed.
- Run the GPU/NIC Topology Fleet Collector scripts:
  - During the initial setup,
  - If you change the hostname of the server at any point after you have set up your AI fabrics, or
  - If you have any topology change, interface rename or NIC replacement.

  Upload the generated zip file into the **Update version** box in the **Job monitoring telemetry** area under **Edit <fabric> Settings > AI settings**. See GPU/NIC Topology Fleet Collector and AI settings for more information.
- Nexus Dashboard pulls fresh jobs data:
  - Manually, whenever you click **Resync**
  - Automatically, every 5 minutes

    Jobs started and completed within this window will not get correlated with GPU details.

- When you perform a backup and restore operation, after you have completed the restore process, complete jobs data is read from Slurm and is shown in job monitoring. Historical jobs will not get correlated with GPU details.
- The AI job **Topology** view does not retain the topology from the time the job ran. Instead, it displays the job topology using the currently available topology data and resource associations. If an AI cluster, fabric association, scalable unit, switch, server, NIC, or GPU changes after the job runs, the topology for that job might be incomplete.

Guidelines and limitations: GPU Anomalies

- All GPU anomalies are enabled by default on startup.
  - GPU anomaly levels: By default, only Critical (set at 90%) is enabled.
  - GPU memory/utilization/power: The critical value is set at 90%.
  - GPU temperature: The critical value is set at 90 degrees Celsius.
- You cannot set or modify the GPU power anomaly thresholds from the global alert thresholds page. Only these default settings are in effect:
  - Warning: 0
  - Major: 0
  - Critical: 90
- Internal thresholds are also set for the GPU temperature, which is based on the limits published by the DCGM Exporter. If those thresholds are breached, then an anomaly will be raised irrespective of the user's setting on the global alerts page.
- Warnings are always suppressed. There is no effect if you change the settings for the Warning levels.
- In the **GPU memory anomaly** area, there is a known issue where the **What's wrong** message provides mismatched unit information.

Add Slurm as an integration

In order to view job details for AI fabrics, you must first integrate Slurm (Simple Linux Utility for Resource Management), which is an open-source job scheduler used to manage resources on high-performance computing clusters. For more information, see Working with Integrations in Your Nexus Dashboard.

DCGM Exporter

To collect GPU statistics, DCGM (Nvidia Data Center GPU Manager) Exporter should be running on the host.

DCGM Exporter is a tool based on the Go APIs to NVIDIA DCGM that allows users to gather GPU metrics and understand workload behavior or monitor GPUs in clusters. DCGM Exporter exposes GPU metrics at an HTTP endpoint (/metrics) for monitoring solutions such as Prometheus.

To use the DCGM Exporter, see DCGM Exporter.

- You may want to run the DGNM exporter in secure mode. For that, you must upload a CA certificate and a client certificate/key in the **Certificate**

Management page and attach those certificates to the **AIComputeVisibility** feature. See Collect data in secured mode for more information.

- You should validate that the DCGM Exporter is properly configured. You can make a REST API call to the DCGM Exporter and validate that the hostname is correct and that the counters in the Working with the dcp-metrics-included.csv file section are present in the response.

Working with the dcp-metrics-included.csv file

The dcp-metrics-included.csv file has a list of counters to collect. You will need this file at the start of DCGM Exporter process.

The dcp-metrics-included.csv file must have these entries.

```
# Temperature

DCGM_FI_DEV_GPU_TEMP,       gauge, GPU temperature (in C).

# Utilization (the sample period varies depending on the product)

DCGM_FI_DEV_GPU_UTIL,       gauge, GPU utilization (in %).

# Power

DCGM_FI_DEV_POWER_USAGE,      gauge, Power draw (in W).

#Slowdown temperature for the device.

DCGM_FI_DEV_SLOWDOWN_TEMP,     gauge, Slowdown temperature for the device.

#Shutdown temperature for the device.

DCGM_FI_DEV_SHUTDOWN_TEMP,     gauge, Shutdown temperature for the device.

#Current Power limit for the device.

DCGM_FI_DEV_POWER_MGMT_LIMIT,    gauge, Current power limit for the device.

#Used Frame Buffer in MB.

DCGM_FI_DEV_FB_USED,       gauge, Framebuffer memory used (in MiB).

#Free Frame Buffer in MB.

DCGM_FI_DEV_FB_FREE,       gauge, Framebuffer memory free (in MiB).
```

```
#Total Frame Buffer of the GPU in MB.

DCGM_FI_DEV_FB_TOTAL,        gauge, Framebuffer memory total (in MiB).

#Percentage used of Frame Buffer: 'Used/(Total - Reserved)'.

DCGM_FI_DEV_FB_USED_PERCENT,    gauge, Frame buffer used (in percent).

DCGM_FI_DEV_COUNT,          counter, Number of Devices on the node.

DCGM_FI_DEV_NAME,           label, Name of device

DCGM_FI_DEV_BRAND,        label, Device Brand

DCGM_FI_DEV_SERIAL,       label, Device Serial Number

DCGM_FI_DEV_UUID,        label, UUID

#Effective power limit that the driver enforces after taking into account all limiters.

DCGM_FI_DEV_ENFORCED_POWER_LIMIT,   gauge, Enforced power limit for the device.
```

Configure DOCA Telemetry Service and Fluent Bit

Use the NVIDIA Data Center Infrastructure on a Chip Architecture (DOCA) Telemetry Service (DTS) and Fluent Bit on each GPU server to forward server network interface card (NIC) metrics to the Nexus Dashboard OpenTelemetry collector. DTS exposes server NIC metrics on a local Prometheus endpoint. Fluent Bit scrapes the DTS endpoint every 30 seconds and forwards the metrics to the CollectorPersistent endpoint in Nexus Dashboard.

Prerequisites

- Ensure that Docker is installed on the DGX host.
- Ensure that the DGX host has network connectivity to the Nexus Dashboard cluster.
- Obtain the CollectorPersistent pod IP address and OTLP NIC receiver port from the Nexus Dashboard cluster.
- For secure communication with the OpenTelemetry (OTel) collector using mutual TLS (mTLS), ensure that the CA certificate, server certificate, and server key are available.
- Ensure that Docker is installed on each GPU server that sends server NIC metrics.
- Ensure that the GPU server has network reachability to the Nexus Dashboard data network.
- Retrieve the CollectorPersistent external data IP address from Admin > System Settings > External Data IPs in Nexus Dashboard.
- Use OpenTelemetry Protocol (OTLP) port 4319 for HTTP, or OTLP port 4320 for HTTPS with mutual TLS (mTLS).

- For HTTPS with mTLS, copy the CA certificate, Fluent Bit client certificate, and key to the GPU server.
- Ensure that the hostname reported by Data Center GPU Manager (DCGM) Exporter, DTS, Slurm, and Link Layer Discovery Protocol (LLDP) advertisements matches the server hostname.

Initial configuration

1. Pull and initialize the DOCA Telemetry Service (DTS) container on the DGX host:

   ```
   export DTS_IMAGE=nvcr.io/nvidia/doca/doca_telemetry:1.21.5-doca3.0.0-host
   docker run -v " /opt/mellanox/doca/services/telemetry/config:/config" \
     --rm --name doca-telemetry-init -it $DTS_IMAGE \
     /bin/bash -c " DTS_CONFIG_DIR=host /usr/bin/telemetry-init.sh"
   ```

   The system creates the default DTS configuration files in /opt/mellanox/doca/services/telemetry/config/.

2. Edit the /opt/mellanox/doca/services/telemetry/config/dts_config.ini file on the DGX host. Locate the PROVIDERS section and update it to enable the **sysfs** and **ethtool** providers while disabling unnecessary sub-components. Replace the existing provider lines with the following::

   ```
   # >>> BEGIN — changed section >>>

   enable-provider=sysfs
   # the sysfs.rate component calculates rate of sysfs counters and is usually unnecessary
   disable-provider=sysfs.rate
   # the sysfs.ib_mr_cache component is usually unnecessary
   disable-provider=sysfs.ib_mr_cache
   #disable-provider=sysfs.eth
   disable-provider=sysfs.hwmon

   enable-provider=ethtool

   # <<< END — changed section <<<
   ```

   Ensure the Prometheus endpoint is enabled. Verify that the prometheus line is not commented out. c. Save the file.

   ```
   prometheus=http://0.0.0.0:9100
   ```

   Start or restart DOCA Telemetry Service (DTS) container to apply the updated configuration.

   For the complete dts_config.ini file with these changes applied in Appendix A.

3. Before you configure Fluent Bit, retrieve the following information from the Nexus Dashboard (ND)

cluster:

Parameter	How to obtain	Example
collectorpersistent1 pod IP	This is the IP under admin→system settings→external data ips, (check the one assigned to **collectorpersistent1 pod**) on the ND UI.	10.10.10.5
OTLP NIC receiver port	HTTP: **4319**, HTTPs: **4320** (configured on the OTel collector)	Use **4319** for HTTP, **4320** for HTTPs
mTLS certificates	Generate a fresh bundle using Appendix B. Before generating, update the collector VIP in the script so the OTEL server certificate SAN matches the CollectorPersistent external IP.	Copy to /root/mtls_certs/ on the DGX host using the directory layout shown below

Certificate requirements

Option	Description
HTTP	No certificates are required on the DGX host.
HTTPS with mTLS	Copy the generated Fluent-side CA and client certificate files to the DGX host using the following layout:

```
mkdir -p /root/mtls_certs/ca
mkdir -p /root/mtls_certs/fluentbit_receiver

# Copy these files from the generated bundle:
# /aiinfra-cert/mtls_certs/ca/aiinfra_ca.crt
# /aiinfra-cert/mtls_certs/dcgm_exporter_server/dcgm_exporter_server.crt
# /aiinfra-cert/mtls_certs/dcgm_exporter_server/dcgm_exporter_server.key

cp /aiinfra-cert/mtls_certs/ca/aiinfra_ca.crt /root/mtls_certs/ca/aiinfra_ca.crt
cp /aiinfra-cert/mtls_certs/dcgm_exporter_server/dcgm_exporter_server.crt
/root/mtls_certs/fluentbit_receiver/dcgm_exporter_server.crt
cp /aiinfra-cert/mtls_certs/dcgm_exporter_server/dcgm_exporter_server.key
/root/mtls_certs/fluentbit_receiver/dcgm_exporter_server.key
```

For HTTPS with mTLS on port 4320, enable TLS verification in the Fluent Bit output and specify the CA certificate, client certificate, and client key files.

4. Create the Fluent Bit configuration. Follow these steps to create the Fluent Bit configuration file (/root/fluent-bit-nic.yaml) on the DGX host based on your transport requirements.
   a. Non-TLS (Port 4319): If the collector exposes the NIC receiver on port 4319 without Transport Layer Security (TLS), use the following configuration:

```
pipeline:
  inputs:
    - name: prometheus_scrape
      tag: nic_rdma_metrics
      host: 127.0.0.1
      port: 9100
      scrape_interval: 30s
      metrics_path: /metrics
      buffer_max_size: 32M
      processors:
        metrics:
          - name: metrics_selector
            metric_name:
/^(rx_crc_errors_phy|tx_discards_phy|rx_discards_phy|rx_bytes_phy|rx_packets_ph
y|rx_bytes|tx_bytes_phy|tx_packets_phy|tx_bytes|hw_rnr_nak_retry_err|hw_implied
_nak_seq_err|hw_req_transport_retries_exceeded|hw_local_ack_timeout_err|hw_ou
t_of_buffer|hw_out_of_sequence|port_rcv_constraint_errors|port_rcv_errors|local_li
nk_integrity_errors|port_rcv_remote_physical_errors|symbol_error|port_rcv_switch_r
elay_errors|hw_rp_cnp_handled|hw_np_cnp_sent|hw_port_state|current_link_speed
|max_link_speed|current_link_width|max_link_width|rx_global_pause_duration|tx_gl
obal_pause_duration|rx_global_pause_transition|tx_global_pause_transition|rx_prio1
_pause_duration|tx_prio1_pause_duration|rx_prio2_pause_duration|tx_prio2_pause
_duration|rx_prio3_pause_duration|tx_prio3_pause_duration|rx_prio4_pause_duratio
n|tx_prio4_pause_duration|rx_prio5_pause_duration|tx_prio5_pause_duration|rx_pri
o6_pause_duration|tx_prio6_pause_duration|rx_prio7_pause_duration|tx_prio7_pau
se_duration)$/
            action: include
  outputs:
    - name: opentelemetry
      match: nic_rdma_metrics
      host: <COLLECTOR_PERSISTENT_POD_IP>
      port: 4319
      metrics_uri: /v1/metrics

    - name: stdout
      match: '*'
      format: json_lines
```

   b. TLS with client authentication (Port 4320): If the collector exposes the NIC receiver on port 4320 with TLS client authentication enabled, use the following configuration:

```
pipeline:
  inputs:
    - name: prometheus_scrape
      tag: nic_rdma_metrics
      host: 127.0.0.1
      port: 9100
      scrape_interval: 30s
      metrics_path: /metrics
      buffer_max_size: 32M
      processors:
        metrics:
          - name: metrics_selector
            metric_name:
/^(rx_crc_errors_phy|tx_discards_phy|rx_discards_phy|rx_bytes_phy|rx_packets_ph
y|rx_bytes|tx_bytes_phy|tx_packets_phy|tx_bytes|hw_rnr_nak_retry_err|hw_implied
_nak_seq_err|hw_req_transport_retries_exceeded|hw_local_ack_timeout_err|hw_ou
t_of_buffer|hw_out_of_sequence|port_rcv_constraint_errors|port_rcv_errors|local_li
nk_integrity_errors|port_rcv_remote_physical_errors|symbol_error|port_rcv_switch_r
elay_errors|hw_rp_cnp_handled|hw_np_cnp_sent|hw_port_state|current_link_speed
|max_link_speed|current_link_width|max_link_width|rx_global_pause_duration|tx_gl
obal_pause_duration|rx_global_pause_transition|tx_global_pause_transition|rx_prio1
_pause_duration|tx_prio1_pause_duration|rx_prio2_pause_duration|tx_prio2_pause
_duration|rx_prio3_pause_duration|tx_prio3_pause_duration|rx_prio4_pause_duratio
n|tx_prio4_pause_duration|rx_prio5_pause_duration|tx_prio5_pause_duration|rx_pri
o6_pause_duration|tx_prio6_pause_duration|rx_prio7_pause_duration|tx_prio7_pau
se_duration)$/
            action: include
  outputs:
    - name: opentelemetry
      match: nic_rdma_metrics
      host: <COLLECTOR_PERSISTENT_POD_IP>
      port: 4320
      metrics_uri: /v1/metrics
      tls: on
      tls.verify: on
      tls.ca_file: /root/mtls_certs/ca/aiinfra_ca.crt
      tls.crt_file: /root/mtls_certs/fluentbit_receiver/dcgm_exporter_server.crt
      tls.key_file: /root/mtls_certs/fluentbit_receiver/dcgm_exporter_server.key

    - name: stdout
      match: '*'
      format: json_lines
```

c. Create the /root/fluent-bit-nic.yaml file on the DGX host and paste the content corresponding

to your selected transport method and save the file.

d. Replace the placeholders in the template and select the configuration that matches your transport method:

Placeholder	Value
<COLLECTOR_PERSISTENT_POD_IP>	IP obtained in Step 3

5. Start the Fluent Bit container on the DGX host using the command that matches your chosen transport.

   a. Update the <path> placeholder in the following commands to match the location of your certs folder and fluent-bit-nic.yaml file.

   b. Start the Fluent Bit container using the command that matches your transport method:

      i. Run Fluent-bit for HTTP.

```
docker run -d --name fluent-bit-nic \
  --network=host \
  -v <path>/fluent-bit-nic.yaml:/fluent-bit/etc/fluent-bit.yaml:ro \
  fluent/fluent-bit:4.2.3 \
  /fluent-bit/bin/fluent-bit -c /fluent-bit/etc/fluent-bit.yaml
```

      ii. Run Fluent-bit for HTTPS with mTLS.

```
docker run -d --name fluent-bit-nic \
  --network=host \
  -v <path>/fluent-bit-nic.yaml:/fluent-bit/etc/fluent-bit.yaml:ro \
  -v /root/mtls_certs:/root/mtls_certs:ro \
  fluent/fluent-bit:4.2.3 \
  /fluent-bit/bin/fluent-bit -c /fluent-bit/etc/fluent-bit.yaml
```

For HTTPS with mTLS, also mount the certificate directory.

      iii. --network=host: Allows Fluent Bit to scrape localhost:9100 (the DOCA Telemetry Service (DTS) Prometheus endpoint) and reach the collector pod IP address directly.

      iv. HTTPS with mTLS: You must mount both the Fluent Bit configuration file and the /root/mtls_certs directory as read-only.

Verification

Check the Fluent Bit logs to confirm that metrics are being scraped and forwarded correctly.

1. View the container logs: docker logs -f fluent-bit-nic
2. Verify the following indicators of a successful setup:
   a. Successful startup: The logs show no errors regarding the configuration or TLS.

b. Metrics processing: Standard output (stdout) displays JSON-formatted NIC metrics.

c. Connection stability: There are no repeated connection-refused or TLS handshake errors on the OpenTelemetry output.

3. Verify the Verify the DTS Prometheus endpoint metrics: curl -s http://localhost:9100/metrics | head -50

The DTS endpoint must return server NIC metrics from the sysfs and eth tool providers, and the Fluent Bit logs must not show repeated connection or TLS errors.

If Nexus Dashboard does not receive server NIC metrics through Fluent Bit for more than 15 minutes, the system raises an anomaly. If this anomaly occurs, verify that the DOCA Telemetry Service (DTS) and Fluent Bit are running, confirm that the DTS metrics endpoint is returning data, and check the Fluent Bit logs for connection or TLS errors.

Appendix A: Full dts_config.ini

```
[clx]
################################ GENERAL CONFIGURATION
################################
sync-time-limit=10000
counter-buffer-size=65536
event-buffer-size=65536
verbose=6
update=1000
standalone=true
runtime-configuration-folder=/config/clx_last_runtime_conf
idle-time-limit=0

######################################## PROVIDERS
########################################
enable-explicitly=true

# --- CHANGED SECTION START ---
enable-provider=sysfs
disable-provider=sysfs.rate
disable-provider=sysfs.ib_mr_cache
#disable-provider=sysfs.eth
disable-provider=sysfs.hwmon

enable-provider=ethtool
# --- CHANGED SECTION END ---

#enable-provider=vnic
#enable-provider=ifconfig
#enable-provider=diagnostic_data_low_freq
```

```
#diagnostic-data-yml-file=/config/diagnostic_data_configs/all-single-port.yml
#enable-provider=amber
#amber_devices=DEV1,DEV2,DEV3
#amber_update_interval_sec=30
#enable-provider=ppcc_eth
#ppcc_eth_devices=mt4129_pciconf0,mt4129_pciconf0.1
#ppcc_algo_slot=0
#ppcc_algo_param_index=0
#ppcc_local_port=1
#ppcc_pnat=0
#ppcc_lp_msb=0
#enable-provider=nvidia_smi
#nvidia_smi_max_processes=20
#nvsmi_with_numeric_fields=1
#enable-provider=dcgm
dcgm_events_enable_common_fields=1
dcgm_events_enable_prof_fields=0
dcgm_events_enable_ecc_fields=0
dcgm_events_enable_nvlink_fields=0
dcgm_events_enable_nvswitch_fields=0
dcgm_events_enable_vgpu_fields=0
dcgm_events_profile_period=1000
dcgm_events_monitor_period=5000
dcgm_events_static_period=120000

##################################### DATA OUTPUTS
#####################################
#output=/data
schema-path=/data/schema
data-template={{year}}/{{month}}{{day}}/{{source}}/{{tag}}{{id}}.bin

################################### Prometheus
#######################################
prometheus=http://0.0.0.0:9100
prometheus-ignore-tags=FI_metrics

################################### Fluent Bit
#######################################
fluentbit-config-dir=/config/fluent_bit_configs

################################# IPC transport
#####################################
ipc-sockets-dir=/tmp/ipc_sockets

##################################### NetFlow
```

```
########################################
#netflow-collector-addr=0
#netflow-collector-port=0

################################## Open Telemetry
########################################
#open-telemetry-receiver=http://64.100.43.47:4318/v1/metrics

################################## Remote Write
########################################
#remote-write-receiver=http://0.0.0.0:9090/api/v1/write

################################## Real Time Analysis Exporter
########################
#lua-script-dir=/config/rta/lua_scripts

################################### HTTP APIs
########################################
#enable-http-api=true

################################### Labels
############################################
#level-labels-file=/config/level_labels.ini

################################### Ad Hoc File
########################################
ad-hoc-runner-file=/config/clx_ad_hoc_runner_config.ini
```

Appendix B: Generate new mTLS certificates

Use this script to generate a new CA, a Fluent Bit client certificate, and an OpenTelemetry (OTEL) server certificate for the HTTPS with mTLS flow.

Before running the script, update OTEL_COLLECTOR_VIP_IP to match the current CollectorPersistent external IP address. The SAN of the generated OTEL certificate must match the IP address to which Fluent Bit connects.

The script generates the certificate bundle locally in /aiinfra-cert/mtls_certs. Copy the generated CA and Fluent Bit client files to the DPU host paths specified in Step 3. You must also install the matching OTEL certificate and CA on the collector side.

```
#!/bin/bash
set -euo pipefail

# Base names for the certificate components
CA_BASE_NAME="aiinfra_ca"
```

```
DCGM_EXPORTER_BASE_NAME=" dcgm_exporter_server"
OTEL_COLLECTOR_BASE_NAME=" otelcollector_client"

# Validity in days
CA_VALIDITY_DAYS=3650
SERVER_VALIDITY_DAYS=365
CLIENT_VALIDITY_DAYS=365

# Common Name (CN) for the Root CA
CA_CN=" AIInfra Root CA"

# DCGM Exporter Server Certificate Configuration
DCGM_EXPORTER_SERVICE_NAME=" dcgm-exporter-service"
K8S_NAMESPACE=" default"
K8S_CLUSTER_DOMAIN=" svc.cluster.local"
DCGM_EXPORTER_CN=" ${DCGM_EXPORTER_SERVICE_NAME}.${K8S_NAMESPACE}.${K8
S_CLUSTER_DOMAIN}"
DCGM_EXPORTER_SAN_HOSTS=" DNS:${DCGM_EXPORTER_CN},DNS:${DCGM_EXPORTE
R_SERVICE_NAME},DNS:*.${DCGM_EXPORTER_SERVICE_NAME}.${K8S_NAMESPACE}.${K
8S_CLUSTER_DOMAIN},DNS:*.cisco.com,IP:127.0.0.1,IP:64.100.43.47,IP:64.100.43.48,IP:6
4.100.43.49"

# OTel Collector Dual-Use Certificate Configuration
OTEL_COLLECTOR_CN=" otelcollector.cisco-nir.svc"
OTEL_COLLECTOR_VIP_IP=" 10.127.121.103"
OTEL_COLLECTOR_SAN_HOSTS=" IP:${OTEL_COLLECTOR_VIP_IP},DNS:${OTEL_COLLECT
OR_CN}"

# Encryption algorithm for encrypted private keys
KEY_ENCRYPTION_ALGO=" aes256"

# Passphrases
CA_PASSPHRASE=" aiinfra"
OTEL_PASSPHRASE=" aiinfra"

# Output Directory
OUTPUT_DIR=" /aiinfra-cert"
rm -rf " $OUTPUT_DIR"
mkdir -p " $OUTPUT_DIR/mtls_certs/ca"
" $OUTPUT_DIR/mtls_certs/dcgm_exporter_server"
" $OUTPUT_DIR/mtls_certs/otelcollector_client"
CDIR=" $OUTPUT_DIR/mtls_certs"

echo " Certificates will be generated in: $CDIR"
```

```
openssl genpkey -algorithm RSA -out "$CDIR/ca/${CA_BASE_NAME}.key.tmp"
openssl pkcs8 -topk8 -v2 "${KEY_ENCRYPTION_ALGO}" -in
"$CDIR/ca/${CA_BASE_NAME}.key.tmp" -out "$CDIR/ca/${CA_BASE_NAME}.key"
-passout pass:"$CA_PASSPHRASE"
rm "$CDIR/ca/${CA_BASE_NAME}.key.tmp"
openssl req -x509 -new -nodes -key "$CDIR/ca/${CA_BASE_NAME}.key" -sha256 -days
"$CA_VALIDITY_DAYS" -out "$CDIR/ca/${CA_BASE_NAME}.crt" -subj "/CN=${CA_CN}"
-passin pass:"$CA_PASSPHRASE"

openssl genpkey -algorithm RSA -out
"$CDIR/dcgm_exporter_server/${DCGM_EXPORTER_BASE_NAME}.key"
cat > "$CDIR/dcgm_exporter_server/server_ext.cnf" <<EOF2
[req]
distinguished_name = req_distinguished_name
req_extensions = v3_req
prompt = no

[req_distinguished_name]
C = US
ST = California
L = San Jose
O = MyOrg
OU = DCGM Exporter
CN = ${DCGM_EXPORTER_CN}

[v3_req]
subjectAltName = ${DCGM_EXPORTER_SAN_HOSTS}
extendedKeyUsage = clientAuth,serverAuth
EOF2
openssl req -new -key
"$CDIR/dcgm_exporter_server/${DCGM_EXPORTER_BASE_NAME}.key" -out
"$CDIR/dcgm_exporter_server/${DCGM_EXPORTER_BASE_NAME}.csr" -config
"$CDIR/dcgm_exporter_server/server_ext.cnf"
cat > "$CDIR/dcgm_exporter_server/v3_server.cnf" <<EOF2
authorityKeyIdentifier=keyid,issuer
basicConstraints=CA:FALSE
keyUsage=digitalSignature,keyEncipherment
extendedKeyUsage=clientAuth,serverAuth
subjectAltName=${DCGM_EXPORTER_SAN_HOSTS}
EOF2
openssl x509 -req -in
"$CDIR/dcgm_exporter_server/${DCGM_EXPORTER_BASE_NAME}.csr" -CA
"$CDIR/ca/${CA_BASE_NAME}.crt" -CAkey "$CDIR/ca/${CA_BASE_NAME}.key"
-CAcreateserial -out
"$CDIR/dcgm_exporter_server/${DCGM_EXPORTER_BASE_NAME}.crt" -days
```

```
" $SERVER_VALIDITY_DAYS"  -sha256 -extfile
" $CDIR/dcgm_exporter_server/v3_server.cnf"  -passin pass:" $CA_PASSPHRASE"

openssl genpkey -algorithm RSA -out
" $CDIR/otelcollector_client/${OTEL_COLLECTOR_BASE_NAME}.key.tmp"
openssl pkcs8 -topk8 -v2 " ${KEY_ENCRYPTION_ALGO}"  -in
" $CDIR/otelcollector_client/${OTEL_COLLECTOR_BASE_NAME}.key.tmp"  -out
" $CDIR/otelcollector_client/${OTEL_COLLECTOR_BASE_NAME}.key"  -passout
pass:" $OTEL_PASSPHRASE"
rm " $CDIR/otelcollector_client/${OTEL_COLLECTOR_BASE_NAME}.key.tmp"
cat > " $CDIR/otelcollector_client/client_ext.cnf"  <<EOF2
[req]
distinguished_name = req_distinguished_name
req_extensions = v3_req
prompt = no

[req_distinguished_name]
C = US
ST = California
L = San Jose
O = AllInfra
OU = OTel Collector
CN = ${OTEL_COLLECTOR_CN}

[v3_req]
subjectAltName = ${OTEL_COLLECTOR_SAN_HOSTS}
extendedKeyUsage = clientAuth,serverAuth
EOF2
openssl req -new -key
" $CDIR/otelcollector_client/${OTEL_COLLECTOR_BASE_NAME}.key"  -out
" $CDIR/otelcollector_client/${OTEL_COLLECTOR_BASE_NAME}.csr"  -config
" $CDIR/otelcollector_client/client_ext.cnf"  -passin pass:" $OTEL_PASSPHRASE"
cat > " $CDIR/otelcollector_client/v3_client.cnf"  <<EOF2
authorityKeyIdentifier=keyid,issuer
basicConstraints=CA:FALSE
keyUsage=digitalSignature,keyEncipherment
extendedKeyUsage=clientAuth,serverAuth
subjectAltName=${OTEL_COLLECTOR_SAN_HOSTS}
EOF2
openssl x509 -req -in " $CDIR/otelcollector_client/${OTEL_COLLECTOR_BASE_NAME}.csr"
-CA " $CDIR/ca/${CA_BASE_NAME}.crt"  -CAkey " $CDIR/ca/${CA_BASE_NAME}.key"
-CAcreateserial -out " $CDIR/otelcollector_client/${OTEL_COLLECTOR_BASE_NAME}.crt"
-days " $CLIENT_VALIDITY_DAYS"  -sha256 -extfile
" $CDIR/otelcollector_client/v3_client.cnf"  -passin pass:" $CA_PASSPHRASE"
```

```
rm "$CDIR/dcgm_exporter_server/server_ext.cnf"
"$CDIR/dcgm_exporter_server/v3_server.cnf"  "$CDIR/otelcollector_client/client_ext.cnf"
"$CDIR/otelcollector_client/v3_client.cnf"

printf '\n--- VERIFY CHAIN ---\n'
openssl verify -CAfile "$CDIR/ca/${CA_BASE_NAME}.crt"
"$CDIR/dcgm_exporter_server/${DCGM_EXPORTER_BASE_NAME}.crt"
openssl verify -CAfile "$CDIR/ca/${CA_BASE_NAME}.crt"
"$CDIR/otelcollector_client/${OTEL_COLLECTOR_BASE_NAME}.crt"
printf '\n--- OTEL CERT DETAILS ---\n'
openssl x509 -in "$CDIR/otelcollector_client/${OTEL_COLLECTOR_BASE_NAME}.crt"  -text
-noout | sed -n '/X509v3 Subject Alternative Name/,+3p;/X509v3 Extended Key
Usage/,+2p'
printf '\n--- KEY ENCRYPTION CHECK ---\n'
if grep -q 'BEGIN ENCRYPTED PRIVATE KEY'
"$CDIR/otelcollector_client/${OTEL_COLLECTOR_BASE_NAME}.key" ; then echo
'otel_key_encrypted=yes'; else echo 'otel_key_encrypted=no'; fi
if grep -q 'BEGIN ENCRYPTED PRIVATE KEY'
"$CDIR/dcgm_exporter_server/${DCGM_EXPORTER_BASE_NAME}.key" ; then echo
'dcgm_key_encrypted=yes'; else echo 'dcgm_key_encrypted=no'; fi
printf '\n--- FILES ---\n'
find "$CDIR"  -maxdepth 2 -type f | sort
```

GPU/NIC Topology Fleet Collector

The GPU/NIC Topology Fleet Collector is a toolkit to collect one time information from AI servers for AI infrastructure visibility. On Nexus Dashboard, this information is used to derive physical connections between GPUs and NICs on a given server, which helps in correlating GPU and network issues.

The GPU/NIC Topology Fleet Collector collects:

- GPU information (model, UUID, PCI addresses)
- Network Interface Cards (NICs) with specifications
- GPU↔NIC topology distances.
- Network interface details (IPs, MACs, drivers)
- DCGM monitoring deployment status

Prerequisites

- Before you run the GPU/NIC Topology Fleet Collector script, verify that the DCGM Exporter, NVIDIA DOCA Telemetry Service (DTS), and Fluent Bit script are running on the target GPU servers. If a required service is not running, the script blocks the collection. For more information about GPU metrics, refer to DCGM Exporter.
- On the machine where you are running the script, you will need:

- Bash 4.0 or higher
  - Linux: Usually pre-installed
  - macOS: Install via Homebrew (see installation steps below)
- SSH client
- Standard Unix tools: tar, awk, sed, sha256sum/shasum
- Internet connection for the initial download

- On the target GPU servers:
  - NVIDIA drivers installed (nvidia-smi available)
  - lspci (pciutils package)
  - ip command (iproute2 package)
  - rdma (iproute2 package)
  - ps (procps package)
  - SSH access (key-based or password authentication)

Download the GPU/NIC Topology Fleet Collector script

To download the GPU/NIC Topology Fleet Collector script:

1. Navigate to the **AI settings** page.

   Edit <fabric> Settings > AI settings

2. Download the GPU/NIC Topology Fleet Collector script from the **AI settings** page.

Run the GPU/NIC Topology Fleet Collector script (single-server collection)

Test the script on a single-server collection before using on a fleet-wide deployment.

1. Verify that the DCGM Exporter is running on the GPU servers before running the GPU/NIC Topology Fleet Collector script.

   See DCGM Exporter for more information.

2. Locate the GPU/NIC Topology Fleet Collector script that Cisco provided.
3. Run the GPU/NIC Topology Fleet Collector script on a single-server collection.
   - On a single-server collection, run this command:

     ./gpu_nic_fleet_collector.sh > output.json

   - For debugging, run this command:

     DEBUG=1 ./gpu_nic_fleet_collector.sh > output.json

   - For remote single host, use fleet mode:

     ./gpu_nic_fleet_collector.sh -f <(echo "gpu-server-01") -u root

Run the GPU/NIC Topology Fleet Collector script (fleet collection)

Once you have successfully tested the script on a single-server collection, use it on a fleet-wide deployment. Ensure that the DCGM Exporter, DTS, and Fluent Bit services are running before you execute the script. If a required service is not running, the script blocks the collection. To resolve this, start the required service and run the collection again.

1. Create a text file and add the GPU servers to this text file.

   These GPU servers must be reachable through SSH.

   For example:

   a. On a Windows system, right-click on the screen and choose **New > Text Document**.

   b. Add the GPU servers to this text document, with each GPU server on its own line. For example:

   ```
   gpu-compute1
   gpu-compute2
   gpu-compute3
   gpu-compute4
   ```

   c. Save and close the text file.

2. Verify your configuration without running the script.
   - Linux:

     bash ./gpu_nic_fleet_collector.sh --dry-run -f hosts.txt -u root

   - macOS:

     /opt/homebrew/bin/bash ./gpu_nic_fleet_collector.sh --dry-run -f hosts.txt -u root

3. Run the collection.
   - Basic collection (sequential):
     - Linux:

       bash ./gpu_nic_fleet_collector.sh -f hosts.txt -u root

     - macOS:

       /opt/homebrew/bin/bash ./gpu_nic_fleet_collector.sh -f hosts.txt -u root

   - Recommended: Parallel collection (much faster):
     - Linux:

       bash ./gpu_nic_fleet_collector.sh -f hosts.txt -u root -j 10

     - macOS:

```
/opt/homebrew/bin/bash ./gpu_nic_fleet_collector.sh -f hosts.txt -u root -j 10
```

You should see output similar to this:

```
=== GPU/NIC Topology Fleet Collection ===
Timestamp: 20260128_143022
Output: /path/to/out_20260128_143022
Hosts: 5 (local: no)
Parallelism: 10

[1/5] gpu-server-01 ... ✓
[2/5] gpu-server-02 ... ✓
[3/5] gpu-server-03 ... ✓
[4/5] 192.168.1.100 ... ✓
[5/5] 192.168.1.101 ... ✓

Collection complete!
Bundle: out_20260128_143022/fleet_bundle_20260128_143022.tar.gz
```

4. Review the output.

After the collection completes, you'll have this structure:

```
out_20260128_143022/

├── fleet_bundle_20260128_143022.tar.gz       # Compressed bundle (deliverable)

├── fleet_bundle_20260128_143022.tar.gz.sha256 # Checksum for verification

├── mappings/

│   ├── gpu-server-01_server_mapping.json     # Per-host inventory

│   ├── gpu-server-02_server_mapping.json

│   └── ...

├── logs/

│   ├── gpu-server-01.log                     # Execution logs (for troubleshooting)

│   └── ...

└── metadata/
```

```
├── MANIFEST.txt                # Bundle metadata
├── SHA256SUMS.txt              # File checksums
└── gpu_nic_fleet_collector.sh  # Script copy
```

5. Upload the zip file to **Job monitoring telemetry**.

   This is the zip file that you will use in the **Job monitoring telemetry** area in the **AI settings** when you edit fabric settings for the AI fabric. See AI settings for more information.

   You might also want to keep the zip file locally for troubleshooting.

Collect data in secured mode

Data collection on Nexus Dashboard from AI servers is supported in http and https modes. When you choose the https mode, mutual TLS (mTLS) is used and both the server and Nexus Dashboard to verify each other. This section provides the necessary steps to set up Nexus Dashboard to collect data in the https mode.

In order to accomplish this, you will need these certificates:

- **Server side**: Use the server certificate that you used to start the DCGM exporter. See DCGM Exporter for more information.
- **Client side**: On the Nexus Dashboard side:
  - CA certificate: Use the server certificate that you used to start the DCGM exporter. See DCGM Exporter for more information.
  - Client certificate and key

When using these procedures, we assume that the DCGM Exporter is also started in secure mode with the server certificate and key. See DCGM Exporter for more information.

Configure Nexus Dashboard AI settings to upload mapping file and start collection in secure mode

1. Upload the certificates to Nexus Dashboard.

   See Managing Certificates for more information.

   a. Navigate to **Admin > Authentication and more > Certificate Management**.

   b. Upload the CA certificate.

      i. Click **CA certificates**.

      ii. Click **Actions > Add CA certificate**.

   c. Upload the system certificate, if necessary.

      This will be the system certificate that you will use to connect to the DCGM Exporter.

i. Click **System certificates**.

ii. Click **Actions > Add system certificate**.

2. Attach the system certificate to the feature.

   a. In the **System certificates** area, choose the system certificate that you want to use to connect to the DCGM Exporter, then click **Actions > Manage feature attachment**.

   b. In the **Manage feature attachments** page, click on the drop-down list and choose **AIComputeVisibility** from the list.

   c. Click **Save**.

You can attach only one certificate to the **AIComputeVisibility** feature, which can be used on all the servers.

3. Upload the mapping file.

   a. Navigate to **Manage > Fabrics**.

   b. Click on the AI fabric.

   c. Click **Actions > Edit fabric settings**.

   d. Click **AI settings**, then scroll down to the **Job monitoring telemetry** area.

   e. Upload the mapping file collected from the servers using the fleet collector script.

   See AI settings and GPU/NIC Topology Fleet Collector for more information.

4. Add the system certificate in the **AI settings** page.

   a. In the **Telemetry data pull mode** field, click **Secured SSL**.

   b. In the **Client certificate name** field, enter the name of the system certificate that you attached to the feature in step 2.

   c. Click **Save**.

After you click **Save**, this configuration will be accepted only if the certificate in the **Client certification name** field has been uploaded and attached to the **AIComputeVisibility** feature.

Regardless of whether you chose the **Secured SSL** or the **Unsecured** option in the **Telemetry data pull mode** field, after you click **Save**, Nexus Dashboard starts collecting GPU data from the DCGM Exporter, such as memory, power consumption, and temperature.

Associate a fabric with an AI cluster

An AI cluster is essentially a fabric group. You can associate a fabric with an AI cluster if you see the **Associate with AI cluster** option in that fabric, as described in the steps below.

You must associate at least one fabric with an AI cluster in order to view the information on AI resources, as described in View information on AI resources. If you do not have an AI cluster already created, you will create one as part of the process of associating a fabric with the AI cluster, as described below.

1. Click **Manage > Fabrics**.

2. In the **Fabrics** page, click the fabric in the **Name** column that you want to associate with an AI cluster.

   The **Overview** page for that fabric appears.

3. Click **Actions > Configuration > Associate with AI cluster**.

   The **Associate fabric with AI cluster** page appears.

4. Click the dropdown list in the **Associate fabric with AI cluster** page.
   - If you have an AI cluster already created and you want to associate the fabric with that AI cluster, choose it from the dropdown list.
   - If you want to create a new AI cluster, either because you do not have an AI cluster already created or you want to create a new AI cluster for this fabric:
     a. Click + **Create new AI cluster**.

        The **Create new AI cluster** page appears.

     b. Enter a name for the new AI cluster, then click **Save**.

        The new AI cluster appears in the **Associate fabric with AI cluster** page.

5. Click **Save** to associate the fabric with this AI cluster.

If you want to disassociate a fabric from an AI cluster:

1. Click **Manage > Fabrics**.
2. In the **Fabrics** page, click the fabric in the **Name** column that you want to disassociate from an AI cluster.

   The **Overview** page for that fabric appears.

3. Click **Actions > Configuration > Remove from AI cluster**.

   The AI cluster is also removed when you remove the last remaining fabric that is associated with it.

Associate a switch with a scalable unit

A scalable unit refers to a foundational, repeatable building block of GPU servers typically connected to the same set of leaf switches. These scalable units are connected using spine switches to deliver specific cluster sizes of AI compute capacity.

Key aspects of a scalable unit include:

- It consists of a set of network leaf switches and the GPU compute nodes connected to them.
- Different scalable unit types exist, defined by the Nexus leaf switch model and the number of UCS nodes, which directly impact the GPU cluster size.
- Scalable units come in various sizes. For example:
  - 8x Cisco N9332D-GX2B switches with up to 16 Cisco UCS C885A nodes supporting up to 128 GPUs

    - 8x Cisco N9364D-GX2A switches with up to 32 Cisco UCS C885A nodes supporting up to 256 GPUs
    - 8x Cisco N9364E-SG2 switches with up to 32 Cisco UCS C880A nodes supporting up to 256 GPUs
- Larger AI clusters are built by scaling out these scalable units in a leaf-spine design, ensuring predictable and consistent performance.
- The networks of scalable units can be managed through Nexus Dashboard, providing operational consistency across fabrics.

To associate a switch with a scalable unit:

1. Click **Manage > Fabrics**.
2. In the **Fabrics** page, click the fabric in the **Name** column that you want to associate with an AI cluster.

   The **Overview** page for that fabric appears.
3. Click **Inventory**.

   The **Inventory** page appears, with the **Switches** tab selected and configured switches displayed by default.
4. Click the check boxes next to two leaf switches that you want to associate with the scalable unit.
5. Click **Actions > Configuration > Associate with scalable unit**.

   The **Associate with scalable unit** page appears.
6. Click the dropdown list in the **Associate with scalable unit** page.
    - If you have scalable unit already created and you want to associate the leaf switches with that scalable unit, choose it from the dropdown list.
    - If you want to create a new scalable unit, either because you do not have a scalable unit already created or you want to create a new scalable unit:
        a. Click + **Create new scalable unit**.

           The **Create new scalable unit** page appears.

        b. Enter a name for the new scalable unit, then click **Create**.

           The new scalable unit appears in the **Associate with scalable unit** page.
7. Click **Save** to associate the leaf switches with this scalable unit.

If you want to disassociate leaf switches from a scalable unit:

1. Click **Manage > Fabrics**.
2. In the **Fabrics** page, click the fabric in the **Name** column where you want to disassociate switches from a scalable unit.

   The **Overview** page for that fabric appears.

3. Click **Inventory**.

   The **Inventory** page appears, with the **Switches** tab selected and configured switches displayed by default.

4. Click the check boxes next to the leaf switches that you want to disassociate from the scalable unit.
5. Click **Actions > Configuration > Disassociate with scalable unit**.

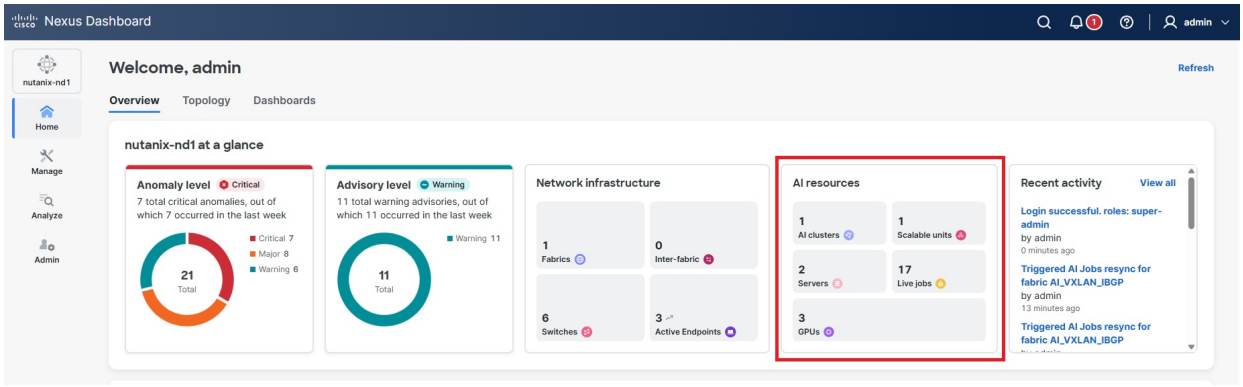
   The scalable unit is also removed when you remove the last remaining leaf switch that is associated with it.

View information on AI resources

You must associate at least one fabric with an AI cluster in order to view the information on AI resources as described below. See Associate a fabric with an AI cluster for more information.

Information on AI resources is provided in two locations:

- A brief summary of the AI resources is provided in the **Overview** page (navigate to **Home > Overview**, then locate the **AI resources** tile), which displays this summary information:
  - AI clusters: Displays the number of AI clusters managed by this system.
  - Scalable units: Displays the number of scalable units.
  - Servers: Displays the number of servers in the AI fabrics.
  - GPUs: Displays the number of GPUs.
  - Jobs: Displays the number of active AI jobs running in the clusters this system manages.



  Click on any area in the **AI resources** tile to bring up a window with more detailed information on that area.

- More detailed information on AI resources is provided in the **Inventory** page (navigate to **Manage > Inventory > AI resources**).

These sections describe the more detailed information provided under **AI resources** in the **Inventory** page.

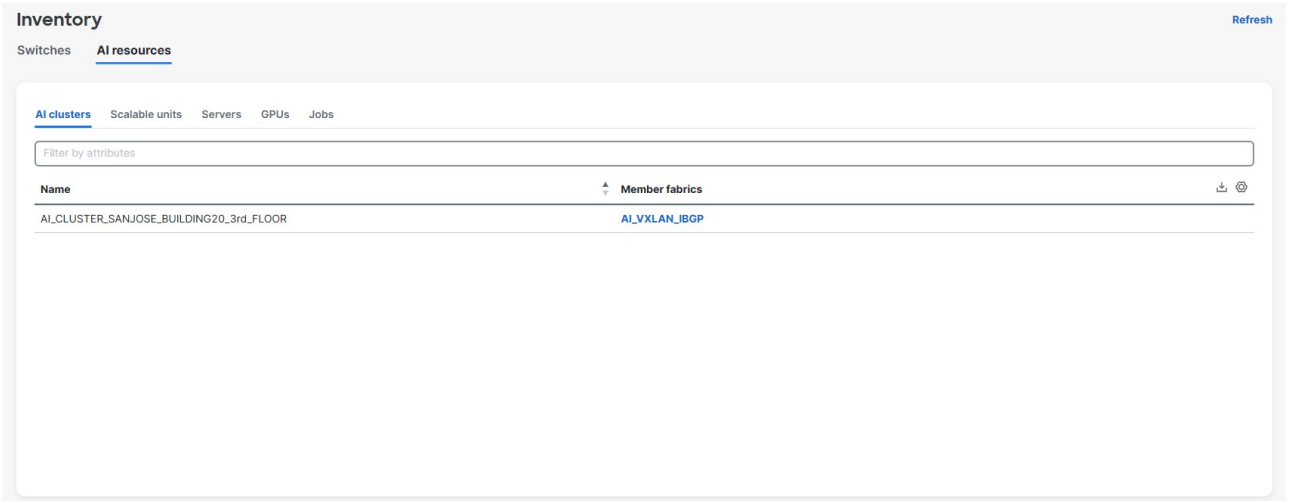
- AI clusters

- Scalable units
- Servers
- GPUs
- Jobs

AI clusters

You must associate at least one fabric with an AI cluster in order to view the information on AI clusters as described below. See Associate a fabric with an AI cluster for more information.

1. Navigate to **Manage > Inventory > AI resources**.
2. Click the **AI clusters** tab to view AI resource information at the cluster level.



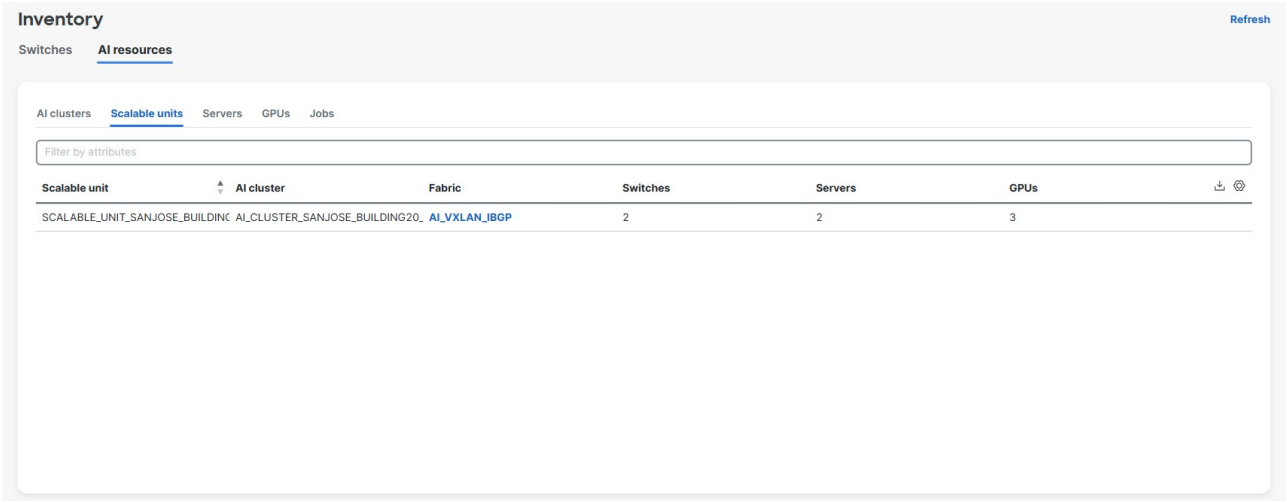
The **AI clusters** tab displays this information:

Field	Description
Name	Displays the name of the AI cluster.
Member fabrics	Displays the fabrics that are members of each AI cluster. Click on a fabric name to navigate to the overview page for that fabric.

Scalable units

You must associate at least two leaf switches with a scalable unit in order to view the information on scalable units as described below. See Associate a switch with a scalable unit for information.

1. Navigate to **Manage > Inventory > AI resources**.
2. Click the **Scalable units** tab to view AI resource information at the scalable units level.



The **Scalable units** tab displays this information:

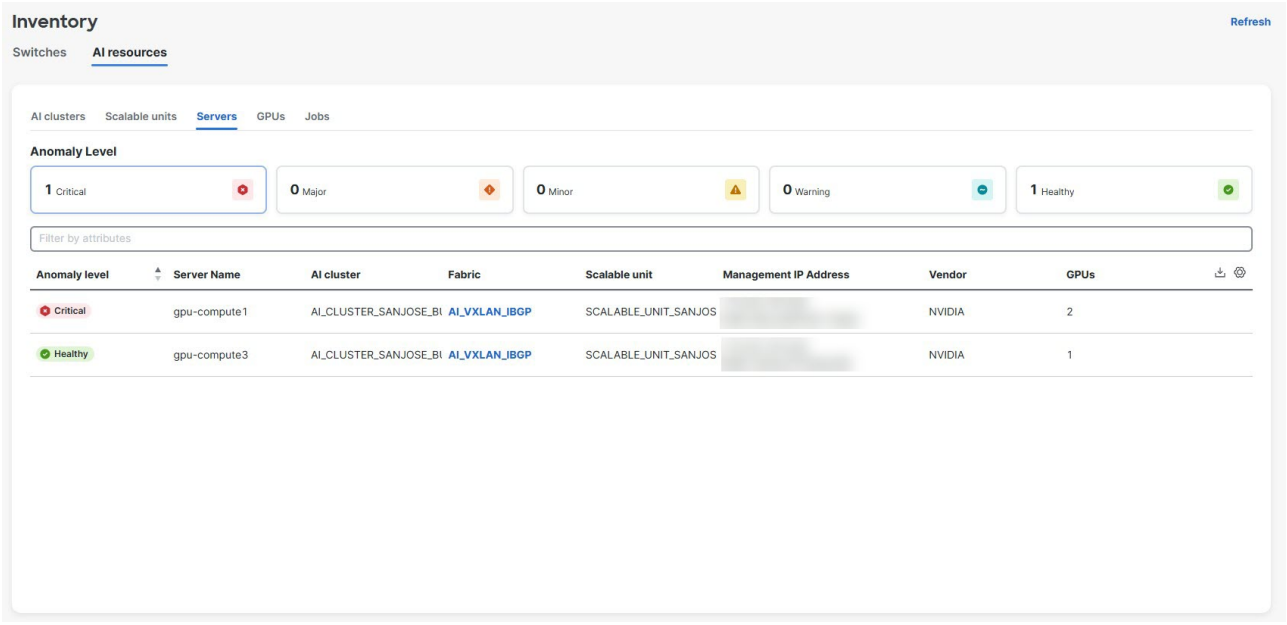
Field	Description
Scalable units	Displays the name of the scalable unit.
Cluster	Displays the cluster that contains the scalable unit.
Fabric	Displays the fabric that contains the scalable unit.
Switches	Displays the number of switches that contains the scalable unit.
Servers	Displays the number of servers that contains the scalable unit.
GPUs	Displays the number of GPUs within the scalable unit.

Servers

When you create an AI fabric, Nexus Dashboard automatically enables the LLDP feature on the network switches and discovers the GPU servers. However, you must also enable LLDP on the GPU servers:

```
root@gpu-compute1:~# apt install lldpd
Reading package lists... Done
Building dependency tree... Done
Reading state information... Done
lldpd is already the newest version (1.0.18-1build3).
0 upgraded, 0 newly installed, 0 to remove and 8 not upgraded.
root@gpu-compute1:~#
```

1. Navigate to **Manage > Inventory > AI resources**.
2. Click the **Servers** tab to view AI resource information at the server level.



The **Servers** tab displays this information:

Field	Description
Anomaly Level	Displays the number of anomalies and their levels for the servers: • Critical • Major • Minor • Warning • Healthy

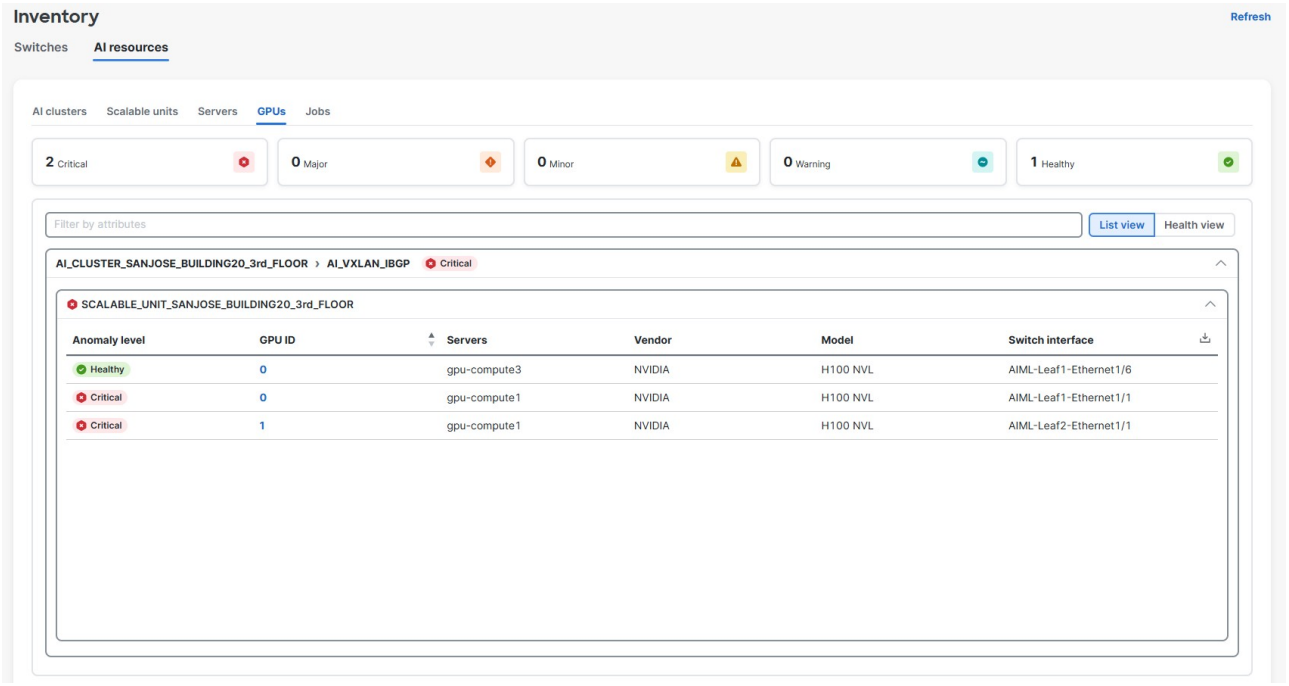
You can also view this information on each of the servers:

Field	Description
Anomaly level	Displays the anomaly level of a specific server with one of these values: • Critical • Major • Minor • Warning • Healthy
Server name	Displays the server name.
Cluster	Displays the cluster that contains the server.
Fabric	Displays the fabric that contains the server.
Scalable unit	Displays the scalable unit that is assigned to the server.
Management IP address	Displays the management IP address of the server.
Vendor	Displays the server's vendor.

Field	Description
GPUs	Displays the number of GPUs in this server.

GPUs

1. Navigate to **Manage > Inventory > AI resources**.
2. Click the **GPUs** tab to view AI resource information at the GPU level.



The **GPUs** tab displays this information:

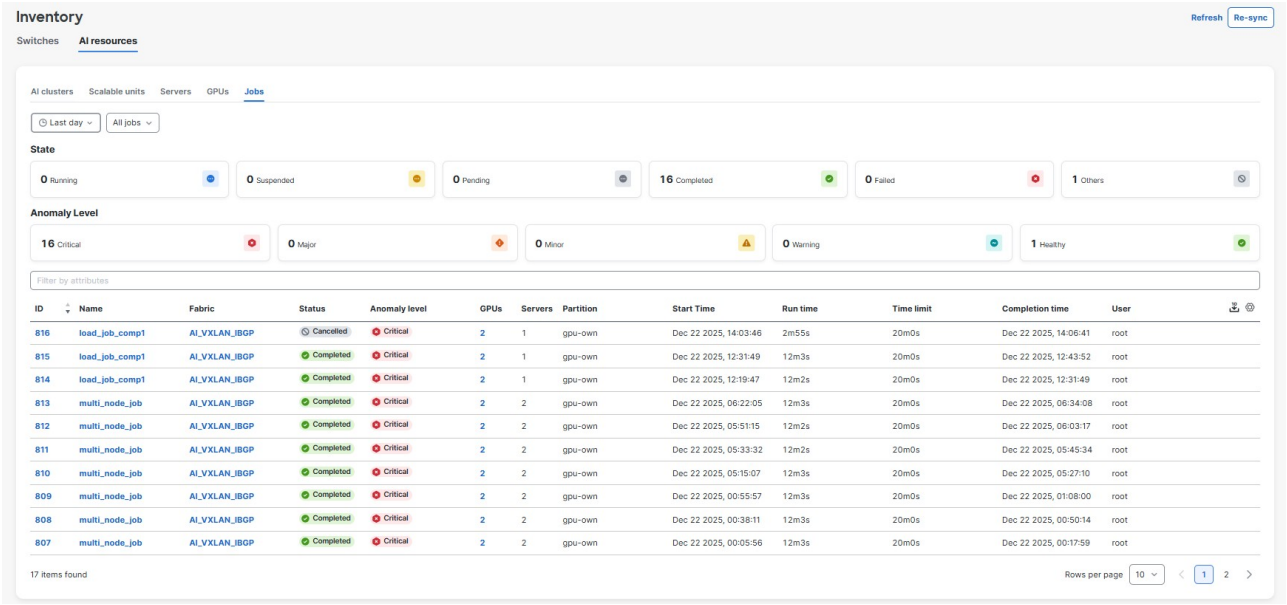
Field	Description
GPU anomaly levels	Displays the number of anomalies and their levels for the GPUs: • Critical • Major • Minor • Warning • Healthy
GPU	Displays critical issues for GPUs in a cluster. Expand the scalable unit to view the connected GPUs.

Field	Description
List view/Health view	Determine how you want the GPU information to be presented. • To view GPU information in a list format, click **List view**. • To view GPU information in a grid format, click **Health view**. In **Health view**, each box represents a server, and each red square is a GPU within the server. Hover over a red square to view general information about a specific GPU. Click the box with the arrow to view more detailed information about a specific GPU. In **Health view**, Nexus Dashboard displays GPUs in good condition with reduced brightness and displays GPUs with health issues at full brightness. A status legend is available from the GPU health title and includes these states: **Empty GPU slot**, **Healthy**, **Warning**, **Minor**, **Major**, and **Critical**. If a server has empty GPU slots, the view shows a separate icon for the empty slots. When you filter by a GPU status, the view shows servers that have at least one GPU in that state and keeps the other GPUs in those servers visible for context.
The following fields become available if you click **List View** above.	

AI scalable unit	Displays if there is a critical issue with one or more AI scalable units. Click the down arrow to expand the view to see the AI scalable units in the GPU, where this information is displayed:
	• Anomaly level: Displays the anomaly level for each GPU:
	  ◦ Critical
	  ◦ Major
	  ◦ Minor
	  ◦ Warning
	  ◦ Healthy
	• GPU ID: The ID of the GPU. Click the number to bring up a window with additional information on a specific GPU.
	• Servers: Displays the server name associated with the GPU.
	• Vendor: Displays the vendor for the GPU.
	• Model: Displays the model number for the GPU.
	• Switch interface: Displays the switch interface associated with the GPU.

Jobs

1. Navigate to **Manage > Inventory > AI resources**.
2. Click the **Jobs** tab to view AI resource information at the jobs level.



The **Jobs** tab displays this information:

Field	Description
State	Displays the state of the AI jobs, with the number of AI jobs that are in these states: • Running • Suspended • Pending • Completed • Failed • Others: Displayed for all others states apart from the above.
Anomaly level	Displays the anomaly level of the AI jobs, with the number of AI jobs that are at these anomaly levels: • Critical • Major • Minor • Warning • Healthy

You can also view this information on each of the AI jobs:

Field	Description
ID	Displays the ID and name of an AI job. Click the ID or name to bring up a window with this additional information on a specific AI job: • Overview: Job level • Topology: Job level • Optics: Job level • Interface analytics: Job level • Anomalies: Job level • Activity logs: Job level
Name	
Fabric	Displays the fabric where the AI job is running.
Status	Displays the status of a specific AI job with one of these values: • Running • Suspended • Pending • Completed • Failed: An AI job might fail if the GPU resources are not sufficient to run that AI job.
Anomaly level	Displays the anomaly level of a specific AI job with one of these values: • Critical • Major • Minor • Warning • Healthy
GPUs	Displays the number of GPUs that are being used by an AI job. Click the number to bring up a window with additional information on the GPUs that are being used by this AI job.
Servers	Displays the number of servers associated with an AI job.
Partition	Displays the partition associated with an AI job.
Start time	Displays the time that the AI job started.
Run time	Displays the amount of time that the AI job has been running.
Time limit	Displays the time limit for the AI job.

Field	Description
Completion time	Displays the time that the AI job was completed. If an AI job has a status of Running, then this field shows the estimated completion time for this AI job based on the time limit that was provided for this AI job.
User	Displays the user who initiated the AI job.

View AI job information

You can use the AI job dashboard to view details of jobs, anomalies, optics, congestion, drops, CRC errors, GPU details, and other job details such as completion time, start and finish time, and so on.

In Nexus Dashboard 4.3.1, the AI job dashboard also shows server NIC statistics for AI jobs when server NIC telemetry is available. You can use the dashboard to correlate a job with GPU health, switch-interface health, and server-interface health.

1. Determine whether you want to see AI job information for all of the fabrics in the system or for a specific fabric.
   - To see AI job information for all of the fabrics in the system, navigate to **Manage > Inventory > AI resources**, then click the **Jobs** tab. See Jobs for more information.
   - To see AI job information for a specific fabric, navigate to **Manage > Fabrics**, then click on the AI fabric and click the **AI jobs** tab.
2. Review the information provided in the AI jobs page.

   You can view AI job information at these levels:

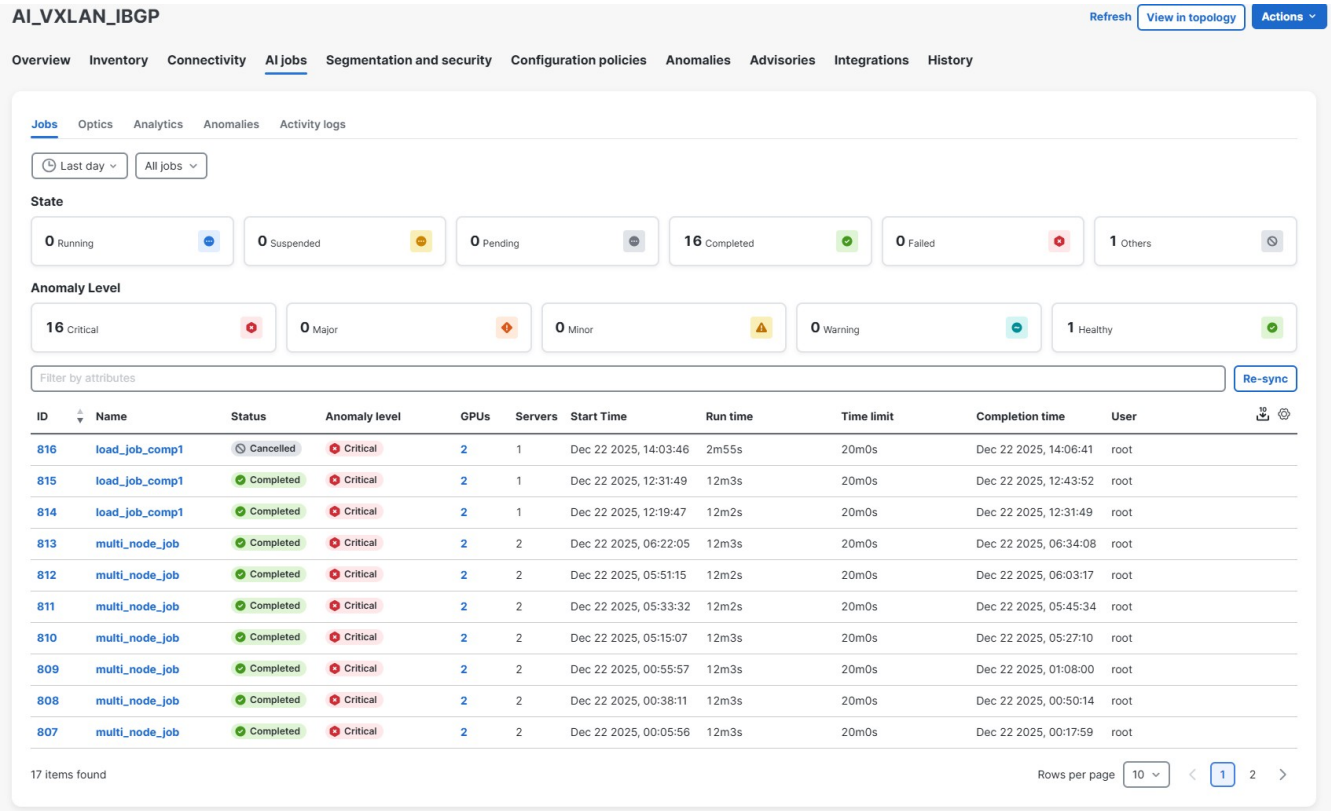
   - View AI job information at the fabric level
   - View information on a specific AI job

View AI job information at the fabric level

- Jobs: Fabric level
- Optics: Fabric level
- Analytics: Fabric level
- Anomalies: Fabric level
- Activity logs: Fabric level

Jobs: Fabric level

The **Jobs** tab displays information on all of the AI jobs associated with this fabric.



To change the frequency and type of information displayed in the **Jobs** tab:

- Click the dropdown on **Last day** to change the frequency that the job information is updated.
- Click the dropdown on **All jobs** to change the type of job information displayed.

The **Jobs** tab displays this information:

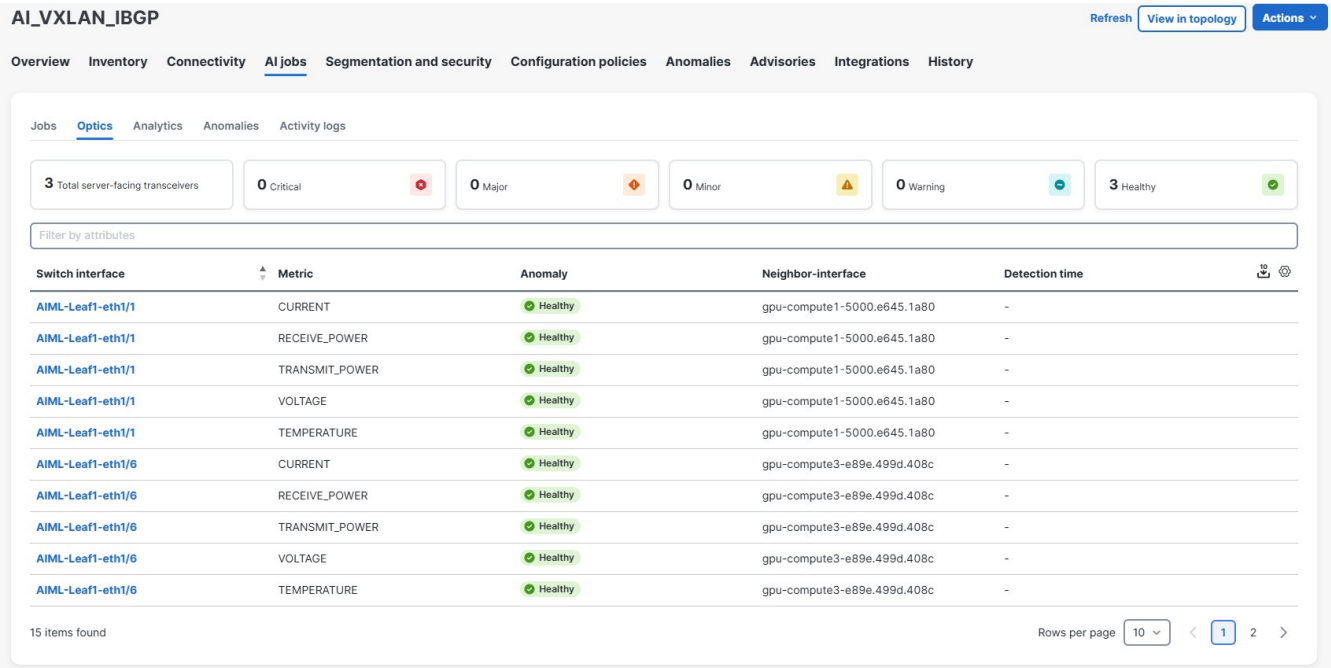
Field	Description
State	Displays the state of the AI jobs, with the number of AI jobs that are in these states: - Running - Suspended - Pending - Completed - Failed: An AI job might fail if the GPU resources are not sufficient to run that AI job.
Anomaly level	Displays the anomaly level of the AI jobs, with the number of AI jobs that are at these anomaly levels: - Critical - Major - Minor - Warning - Healthy

You can also view this information on each of the AI jobs:

Field	Description
ID	Displays the ID and name of an AI job. Click the ID or name to bring up a page with additional information on a specific AI job. See View information on a specific AI job for more information.
Name	
Status	Displays the status of a specific AI job with one of these values: - Running - Suspended - Pending - Completed - Failed
Anomaly level	Displays the anomaly level of a specific AI job with one of these values: - Critical - Major - Minor - Warning - Healthy
GPUs	Displays the number of GPUs that are being used by an AI job. Click the number to bring up a page with additional information on the GPUs that are being used by this AI job.
Servers	Displays the number of servers associated with an AI job.
Start time	Displays the time that the AI job started.
Run time	Displays the amount of time that the AI job has been running.
Time limit	Displays the time limit for the AI job.
Completion time	Displays the time that the AI job was completed. If an AI job has a status of Running, then this field shows the estimated completion time for this AI job based on the time limit that was provided for this AI job.
User	Displays the user who initiated the AI job.

Optics: Fabric level

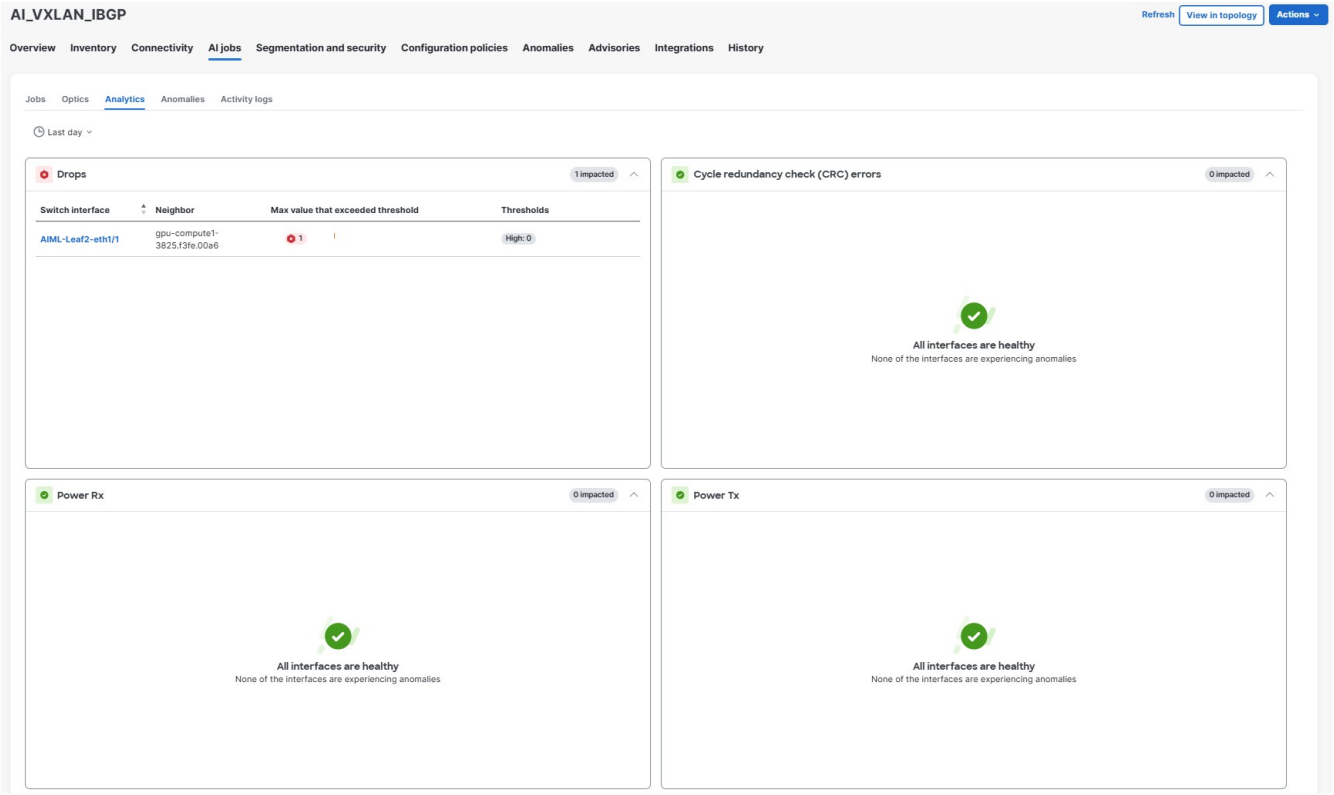
The **Optics** tab displays optics (switch interface) information for AI jobs associated with this fabric.



Field	Description
Switch interface	Displays the names of the switch interfaces. Click the name to bring up a page with this additional information on a specific switch interface.
Metric	Displays the metric for each switch interface.
Anomaly	Displays the anomaly level of a specific switch interface. If multiple anomalies exist for the same metric, the largest severity value appears here, with one of these values: • Critical • Major • Minor • Warning • Healthy
Neighbor-interface	Displays the server-side interface MAC address of the GPU NIC that is connected to the leaf switch, along with the neighbor device hostname and its MAC-based identifier.
Detection time	Displays the detection time, which is the time when an event has been detected for an optic port.

Analytics: Fabric level

Analytics displays analytics information for AI jobs associated with this fabric; relevant sections show only problematic interfaces and remain empty if all interfaces are healthy.

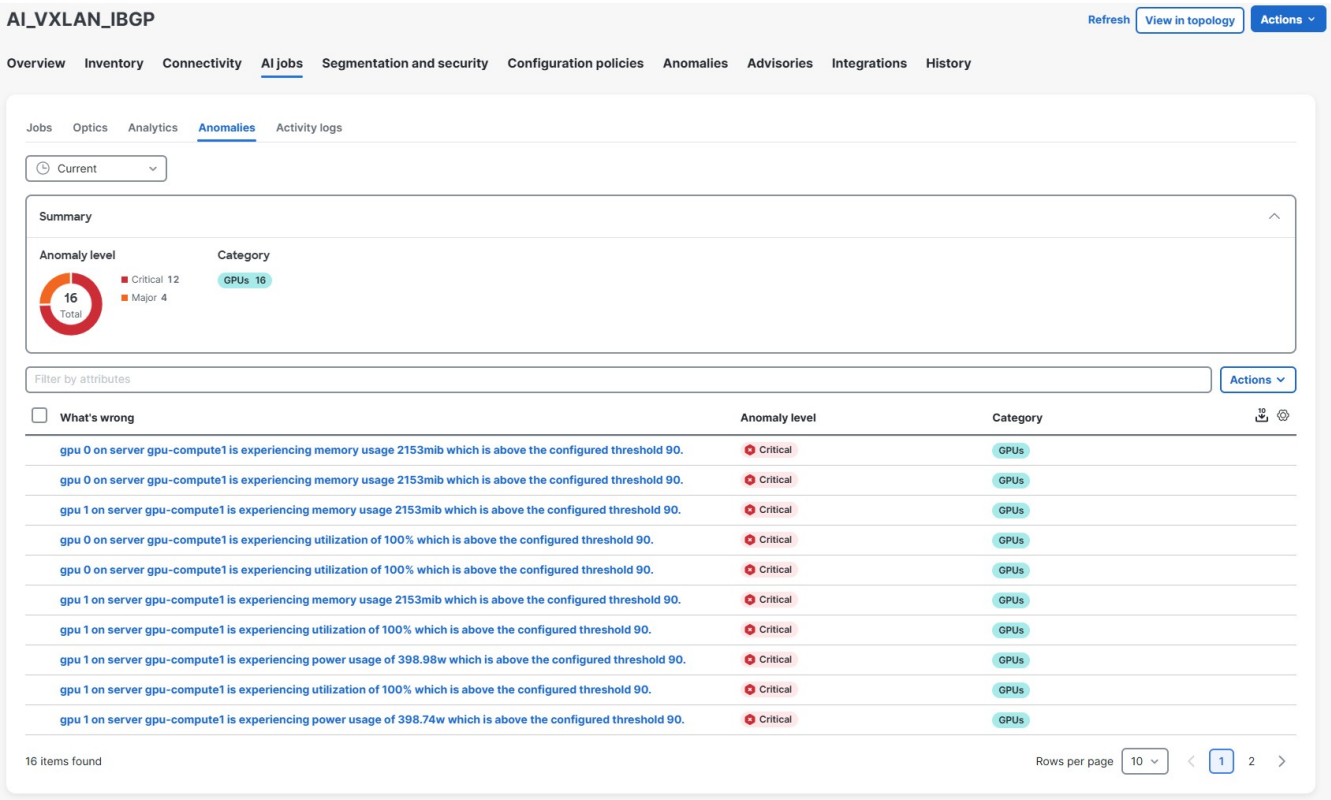


Field	Description
Drops	Shows the number of drops that occurred for this job.
Cycle redundancy error check	Shows the number of cycle redundancy error checks that occurred for this job.
Power Rx	Shows the received (Rx) or transmitted (Tx) power for AI jobs in the fabric. The **Max value that exceeded threshold (dbm)** field shows the power value. The **Thresholds** area shows two sets of low and high thresholds: * Alarm thresholds: Define the high and low values for a critical status. * Warning thresholds: Define the high and low values for a warning status.
Power Tx	
Voltage	Shows the voltage information for for AI jobs associated with this fabric. The **Max value that exceeded threshold** field shows the voltage value, and the **Thresholds** area shows the low and high thresholds for the voltage. A value above the minimum threshold is considered a warning, and a value above the high threshold is considered critical.
Current	Shows the current information for for AI jobs associated with this fabric. The **Max value that exceeded threshold** field shows the current value, and the **Thresholds** area shows the low and high thresholds for the current. A value above the minimum threshold is considered a warning, and a value above the high threshold is considered critical.

Field	Description
Temperature	Shows the temperature information for for AI jobs associated with this fabric. The **Max value that exceeded threshold ©** field shows the temperature value, and the **Thresholds** area shows the low and high thresholds for the temperature. A value above the minimum threshold is considered a warning, and a value above the high threshold is considered critical.

Anomalies: Fabric level

The **Anomalies** tab displays anomaly information for AI jobs associated with this fabric.



To change the frequency for the information that is displayed in the **Anomalies** tab, click the dropdown on **Last day** to change the frequency that the anomaly information is updated:

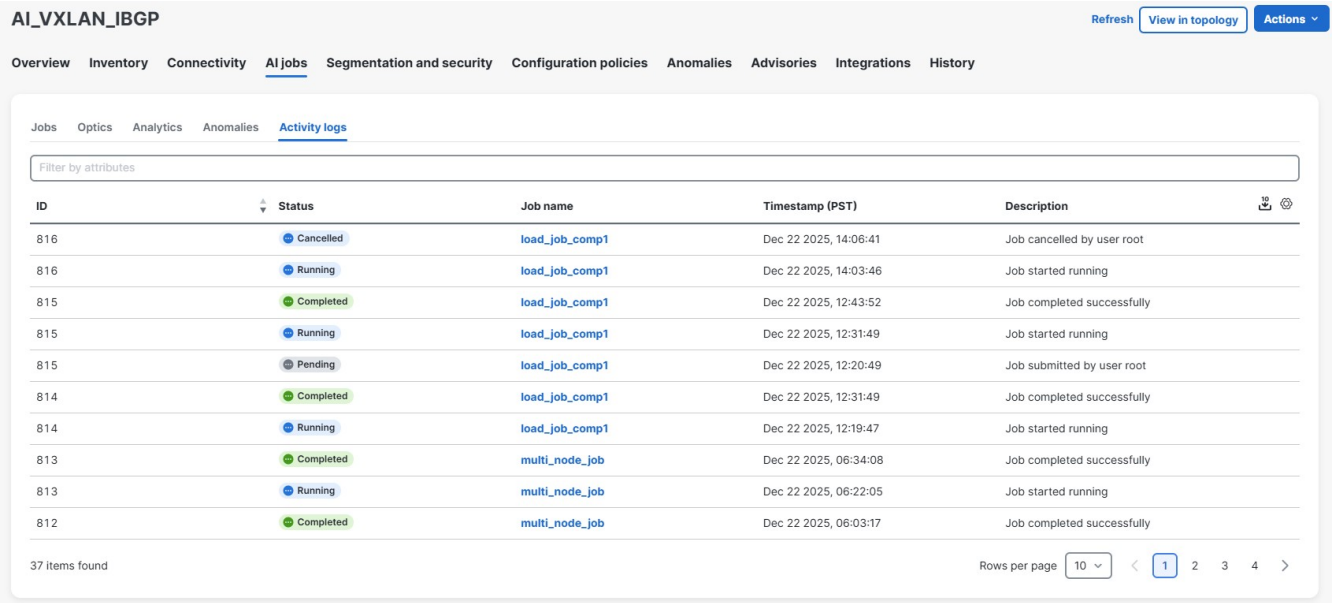
- Current
- Last 15 minutes
- Last hour (default)
- Last 2 hours
- Last 24 hours
- Last week
- Last month
- Custom range

The **Anomalies** tab displays this information:

Field	Description
Summary	Provides a summary of the number of anomalies, and their levels and categories, found with the AI jobs in this fabric.
What's wrong	Provides a description of the anomaly that was found.
Anomaly level	Displays the anomaly level of the AI job at these anomaly levels: • Critical • Major • Minor • Warning • Healthy
Category	Displays the of anomalies grouped by various categories, such as Hardware, Capacity, Compliance, Connectivity, Configuration, Integrations, Active bugs, Telemetry configuration, GPUs, and System.

Activity logs: Fabric level

The **Activity logs** tab displays activity log information for AI jobs associated with this fabric.



Use the time-range selector and the filter by attributes field to narrow the activity log entries that are displayed for the fabric.

The **Activity logs** tab displays this information:

Field	Description
ID	The ID of the activity log.
Status	The status of the activity log: • Completed • Running

Field	Description
Job name	The job that is associated with the activity log. Click the entry in the job name field to bring up information on that AI job. See View information on a specific AI job for more information.
Timestamp	The timestamp on the activity log.
Description	The description of the activity log, such as Job started running or Job completed successfully.
User	The user of the activity log.

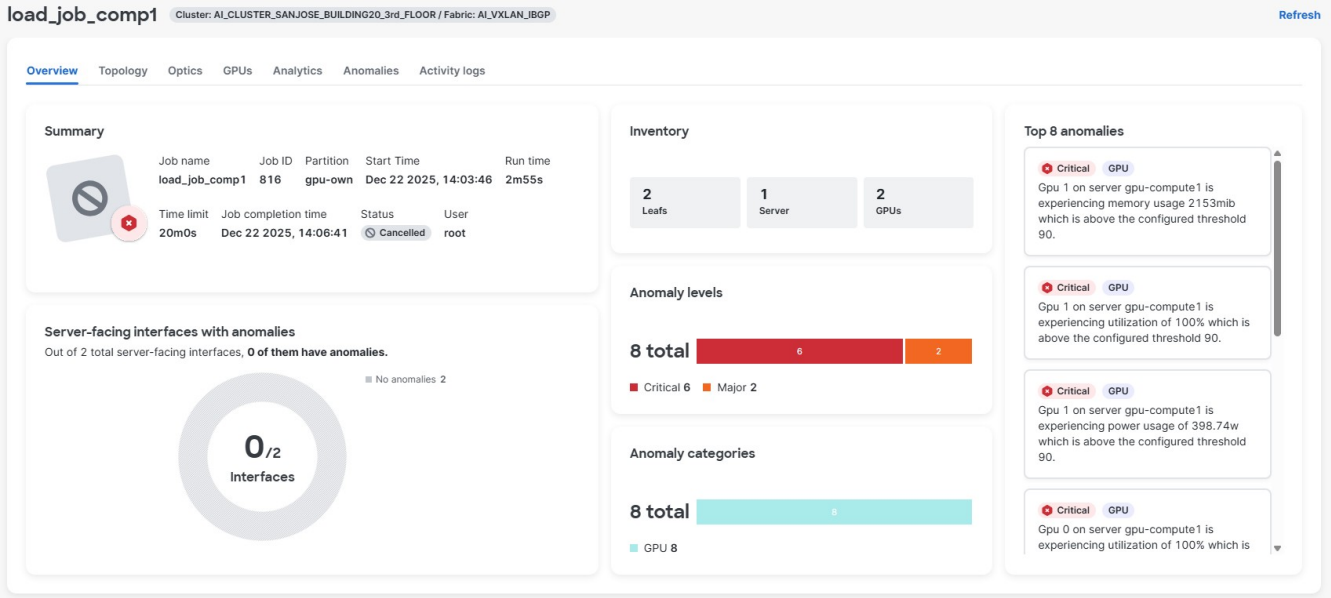
View information on a specific AI job

To view information on a specific AI job:

1. Navigate to the page that displays all the AI jobs at either the system level or at a fabric level.
    - To see AI job information for all of the fabrics in the system, navigate to **Manage > Inventory > AI resources**, then click the **Jobs** tab.
    - To see AI job information for a specific fabric, navigate to **Manage > Fabrics**, then click on the AI fabric and click the **AI jobs** tab.
2. Click on the **Jobs** tab, then click the ID or name of a specific job to bring up a page with this additional information on a specific AI job:
    - Overview: Job level
    - Topology: Job level
    - Optics: Job level
    - GPUs: Job level
    - Interface analytics: Job level
    - Anomalies: Job level
    - Activity logs: Job level

Overview: Job level

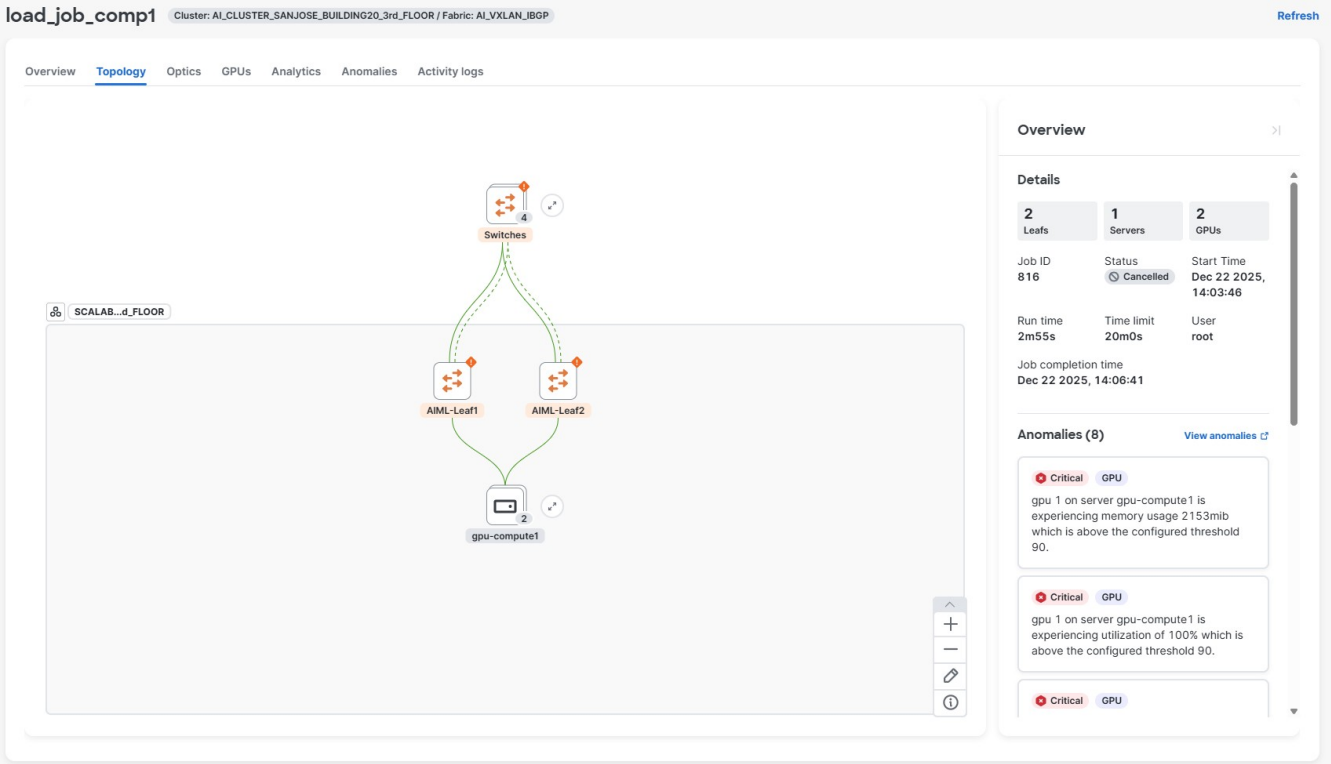
The **Overview** tab displays information as it relates to a specific job.



Field	Description
Summary	Provides a summary of the information available for this AI job, such as the job name and ID, the start and completion times for the job, the status of the job, and so on.
Inventory	Displays this information on the AI job: • **Leaf**: The number of leaf switches associated with this AI job. • **Servers**: The number of servers associated with this AI job. • **GPUs**: The number of GPUs that are being used by this AI job.
Server-facing interfaces with anomalies	Provides information on the number of server-facing interfaces associated with this AI job and how many of the interfaces have anomalies.
Anomaly levels	Displays the number of anomalies, split by level, for the AI job.
Anomaly categories	Displays the number of anomalies by category for the AI job.
Anomalies	Displays the anomalies associated with the AI job.

Topology: Job level

The **Topology** tab displays topology information, similar to the standard **Topology** feature as described in View LAN fabric components in Topology. However, this **Topology** tab provides centralized visualization and management of components as it relates to a specific AI job. For more information, see AI endpoint and topology discovery.

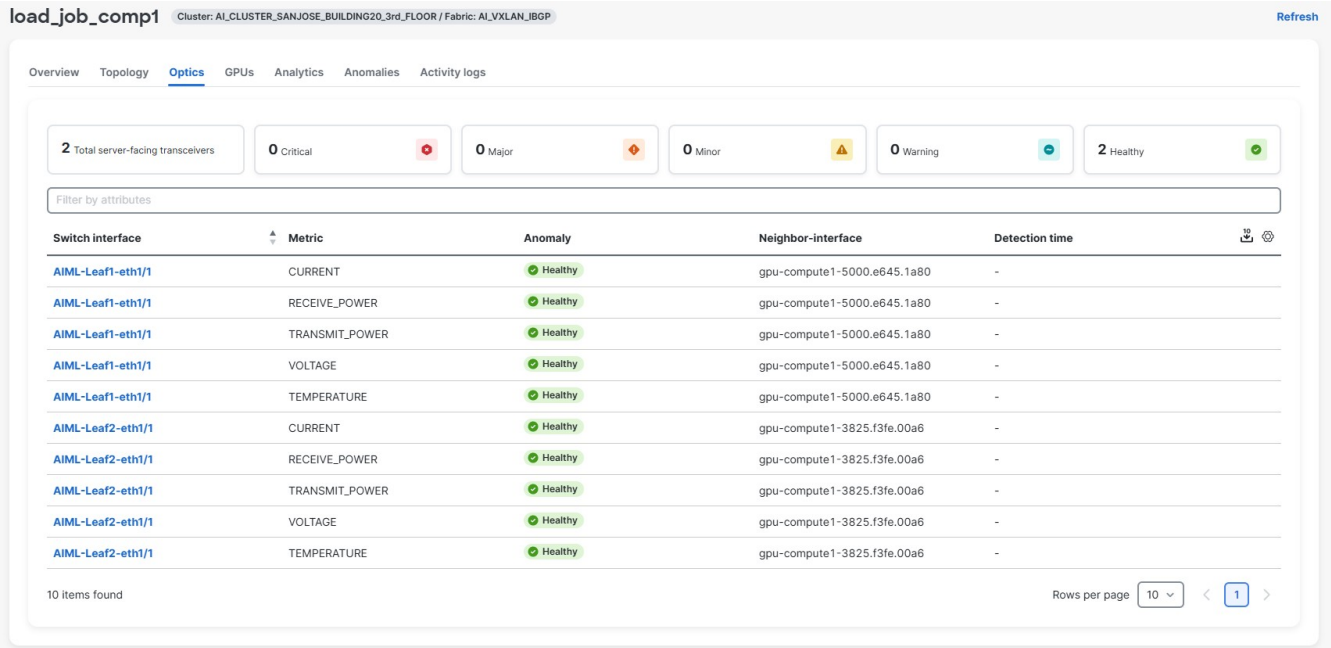


The **Topology** tab includes an overview pane with job details, inventory counts, and a Critical anomalies summary for the selected job. Click **View anomalies** in the overview pane to view the anomalies that are associated with the job.

The AI job **Topology** view does not retain the topology from the time the job ran. Instead, it displays the job topology using the currently available topology data and resource associations. If an AI cluster, fabric association, scalable unit, switch, server, NIC, or GPU changes after the job runs, the topology for that job might be incomplete.

Optics: Job level

The **Optics** tab displays information as it relates to a specific job.



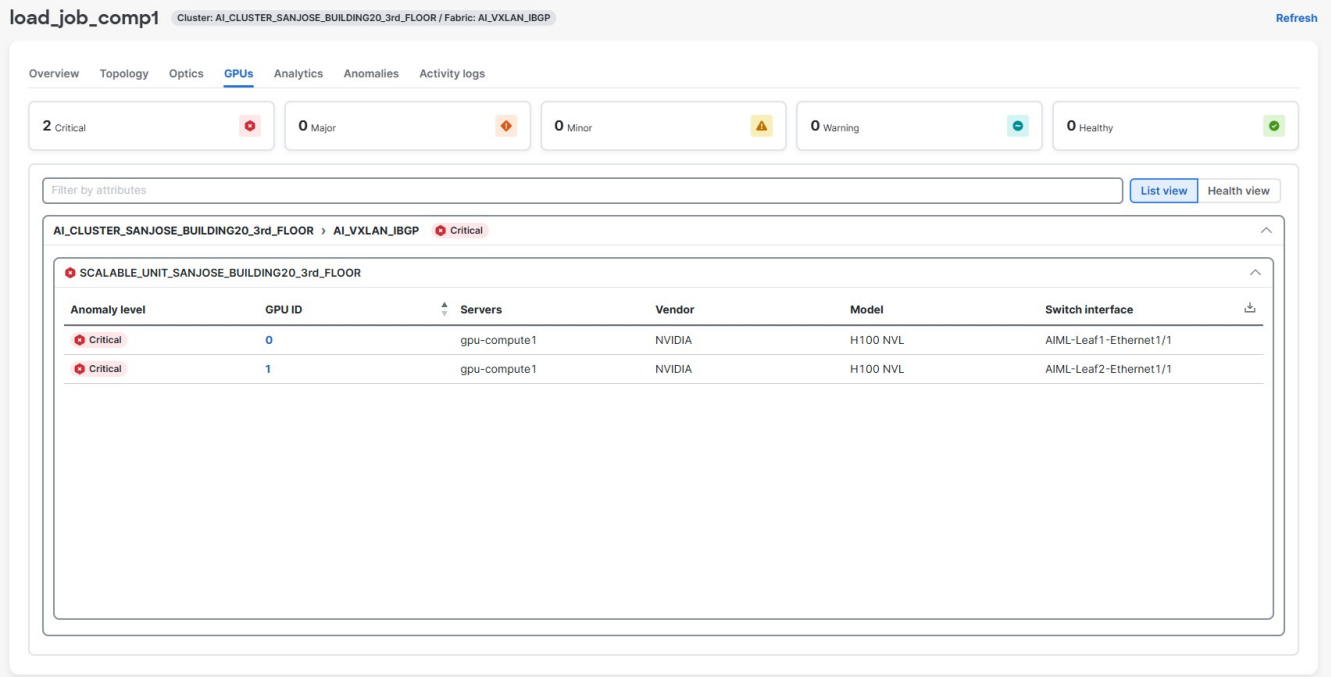
Field	Description
Total server-facing interfaces with anomalies	Provides information on the number of server-facing interfaces associated with this AI job.
Interface anomaly levels	Displays the number of anomalies and their levels for the interfaces: • Critical • Major • Minor • Warning • Healthy

You can also view this information on each of the switch interfaces:

Field	Description
Switch interface	Displays the name of the switch interface. Click the name to bring up a page with this additional information on a specific switch interface.
Metric	Displays the metric for the switch interface.
Anomaly	Displays the anomaly level of a specific switch interface with one of these values: • Critical • Major • Minor • Warning • Healthy
Neighbor-interface	Displays the neighbor interface for the switch interface along with the neighbor device hostname and its MAC-based identifier.
Detection time	Displays the detection time.

GPUs: Job level

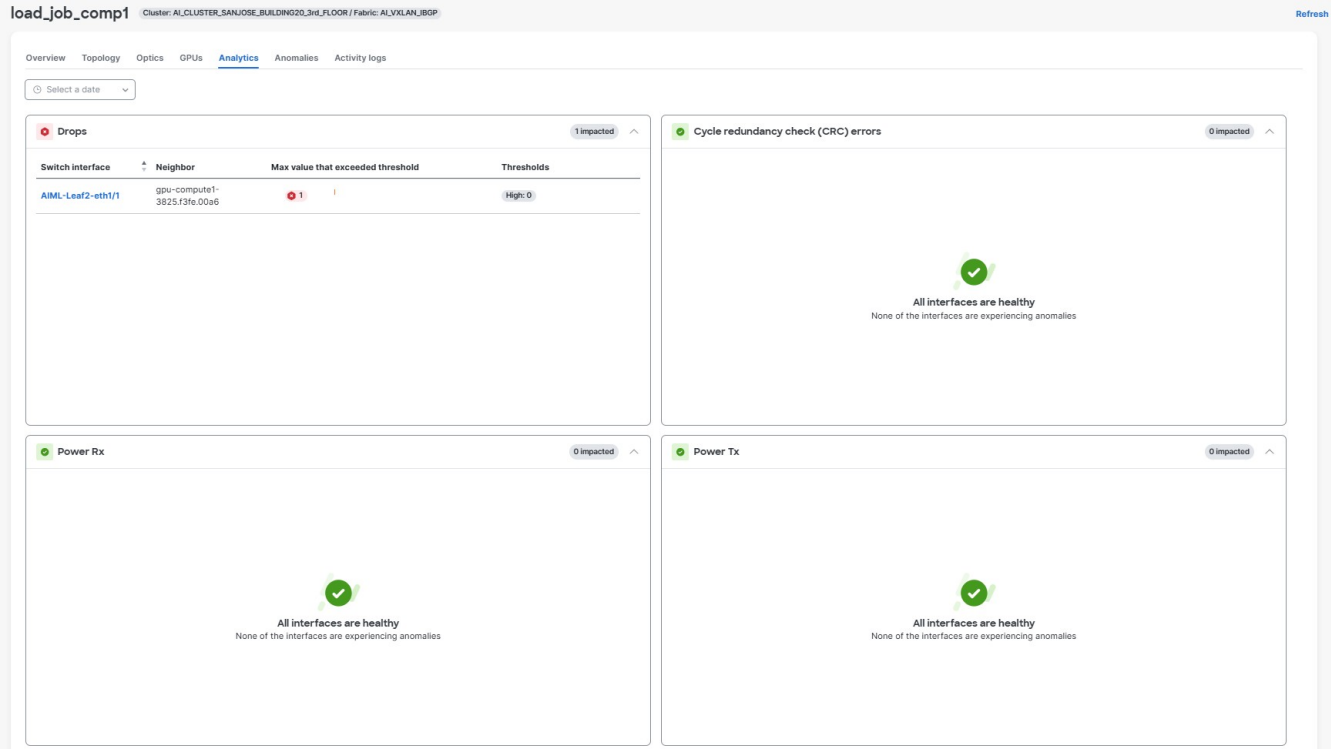
The **GPUs** tab displays information as it relates to a specific job.



Field	Description
Anomaly level	Displays the number of anomalies and their levels for the GPUs: • Critical • Major • Minor • Warning • Healthy
GPU ID	Displays the ID of the GPU in the scalable unit. Click the link to open a new page, which provides details on this GPU, such as which job is running on it, time graph for memory usage, power usage, temperature and utilization.
Servers	Displays the server name associated with the GPU.
Vendor	Displays the vendor for the GPU.
Model	Displays the model number for the GPU.
Switch interface	Displays the switch interface associated with the GPU. The **GPUs** tab includes a **GPU health** summary, a filter, and **List view** and **Health view** options. In Health view, each box represents a server, and each square is a GPU within the server. Hover over a square to view general information about a specific GPU. Click the box with the arrow to view more detailed information about a specific GPU.

Interface analytics: Job level

The information that is provided in the **Analytics** tab is refreshed every 30 minutes or so. Click **Refresh** to refresh the information immediately.



The **Interface analytics** tab also displays server-interface metrics that are correlated with the selected AI job. Server-interface metrics appear in sections such as **Errors** and **Congestion**. Server-interface widgets can include **RDMA Nack - Rx**, **RDMA Nack - Tx**, **CNP Tx**, and **CNP Rx**. Switch-interface widgets continue to show switch-side interface, optics, utilization, and congestion information when switch telemetry is available.

By default, each analytics widget shows the top 10 interfaces for the selected time range. Click **Load all** interfaces to retrieve all available interfaces. Click a server-interface row to open the server interface details page. Click a switch-interface row to open the switch interface details page.

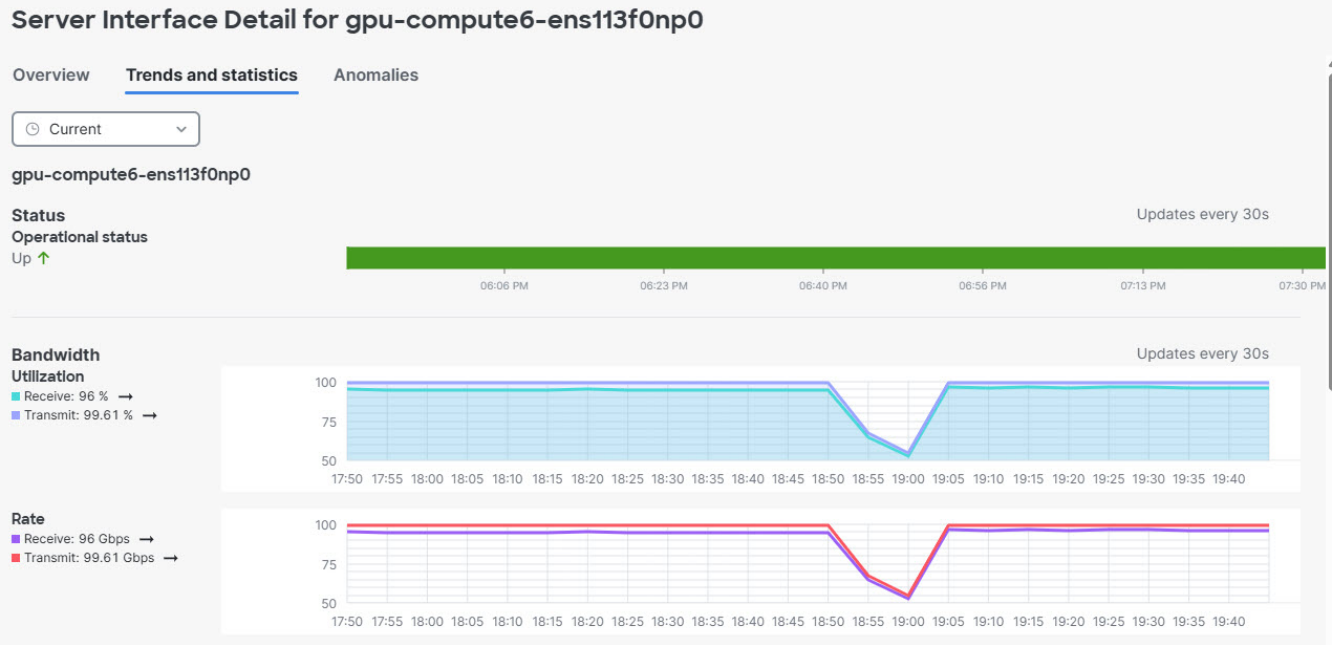
The **Analytics** tab displays this information.

Field	Description
Drops	Shows the number of drops that occurred for this job.
Cycle redundancy check (CRC) errors	Shows the number of cycle redundancy error checks that occurred for this job.
Power Rx	Shows the power Rx or TX information for this job. The **Max value that exceeded threshold (dbm)** field shows the power Rx or TX value, and the **Thresholds** area shows the low and high thresholds for the power Rx or TX. A value above the minimum threshold is considered a warning, and a value above the high threshold is considered critical.
Power Tx	
Voltage	Shows the voltage information for this job. The **Max value that exceeded threshold** field shows the voltage value, and the **Thresholds** area shows the low and high thresholds for the voltage. A value above the minimum threshold is considered a warning, and a value above the high threshold is considered critical.

Field	Description
Current	Shows the current information for this job. The **Max value that exceeded threshold** field shows the current value, and the **Thresholds** area shows the low and high thresholds for the current. A value above the minimum threshold is considered a warning, and a value above the high threshold is considered critical.
Module temperature	Shows the temperature information for this job. The **Max value that exceeded threshold ©** field shows the temperature value, and the **Thresholds** area shows the low and high thresholds for the temperature. A value above the minimum threshold is considered a warning, and a value above the high threshold is considered critical.

Server interface details page

Click a **Server interface** to open the *Server Interface Details* page. This page shows details for the selected server NIC interface.



The server interface page includes these tabs:

- **Overview**: Displays the Anomaly level, General, and artificial intelligence (AI) jobs cards. The Anomaly level card shows the current anomaly status and count. The General card shows interface, NIC, operations speed, IP addresses, operational status, neighbor name, neighbor interface, and neighbor type. The AI jobs card shows jobs that use the server interface, including job name, start time, run time, time limit, job ID, job completion time, and user.
- **Trends and statistics**: Displays time-series data for server NIC status, bandwidth, errors, and congestion. Server NIC statistics in this tab include cyclic redundancy check (CRC) errors, transmit (Tx) discards, received (Rx) discards, bandwidth Rx, bandwidth Tx, negative acknowledgment (NACK) Rx, NACK Tx, congestion notification packet (CNP) Rx, CNP Tx, operational status, and link speed. When the Current toggle is on, the charts display the latest data for the selected interface.

- **Anomalies**: Displays anomalies that are associated with the selected server interface.

The **Overview** tab displays this information.

Field	Description
Anomaly level	Displays the highest anomaly severity and the total number of anomalies for the selected interface.
Interface	Displays the name of the server NIC interface.
NIC	Displays the type of NIC associated with the interface.
Operational speed	Displays the current speed of the interface.
IP addresses	Displays the IP addresses assigned to the interface.
Operational status	Displays whether the interface is operationally up or down.
Neighbor name	Displays the name of the connected neighbor device.
Neighbor interface	Displays the interface name of the connected neighbor device.
Neighbor type	Displays the type of connected neighbor device (for example, switch).
AI jobs	Displays details for AI jobs using this interface, including job name, start time, run time, time limit, job ID, job completion time, and user.

The **Trends and statistics** tab displays this information.

Field	Description
Status	Displays the current operational status of the interface.
Bandwidth Utilization	Displays the receive (Rx) and transmit (Tx) bandwidth utilization as a percentage.
Rate	Displays the receive (Rx) and transmit (Tx) data rates in gigabits per second (Gbps).
Cycle redundancy check (CRC) errors	Displays the number of CRC errors on the interface.
Transmit discard	Displays the number of packets discarded during transmission.
Receive discard	Displays the number of packets discarded during reception.
RDMA NACK – Rx/Tx	Displays the number of RDMA Negative Acknowledgment (NACK) events received (Rx) or transmitted (Tx).
Congestion notification packets (Tx/Rx)	Displays the number of congestion notification packets transmitted (Tx) or received (Rx).
Pause frames (Rx/Tx)	Displays the number of pause frames received (Rx) or transmitted (Tx).
Congestion indicator	Displays the congestion score for the interface.

The **Anomalies** tab displays this information.

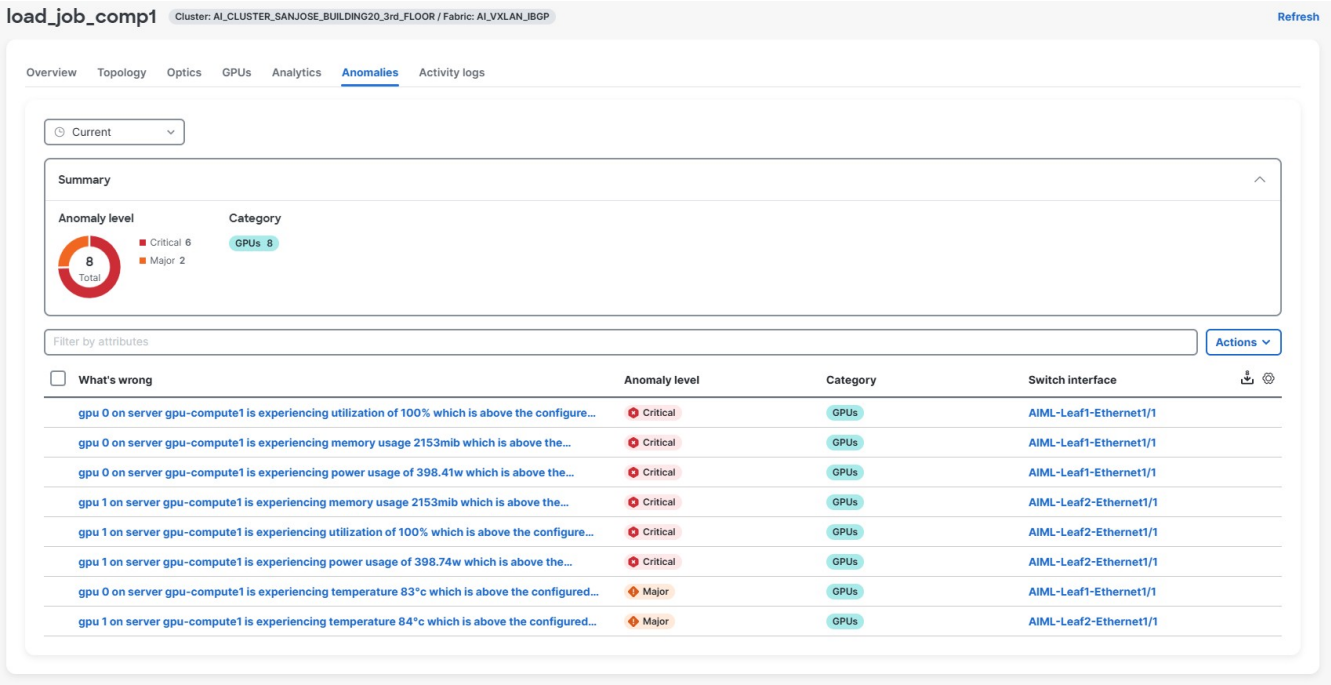
Field	Description
Anomaly type	Displays the anomaly name, such as Interface Ingress Error or Interface Utilization High Threshold.

Field	Description
Anomaly level	Displays the severity of the anomaly.
Category	Displays the anomaly category, such as Connectivity.
Total count	Displays the number of anomalies in the anomaly group.

Anomalies: Job level

To select the time window for which Nexus Dashboard displays **Anomalies**, click the dropdown on **Last day**.

- Current
- Last 15 minutes
- Last hour (default)
- Last 2 hours
- Last 24 hours
- Last week
- Last month
- Custom range



The **Anomalies** tab displays this information.

Field	Description
What's wrong	Provides a description of the anomaly that was found.

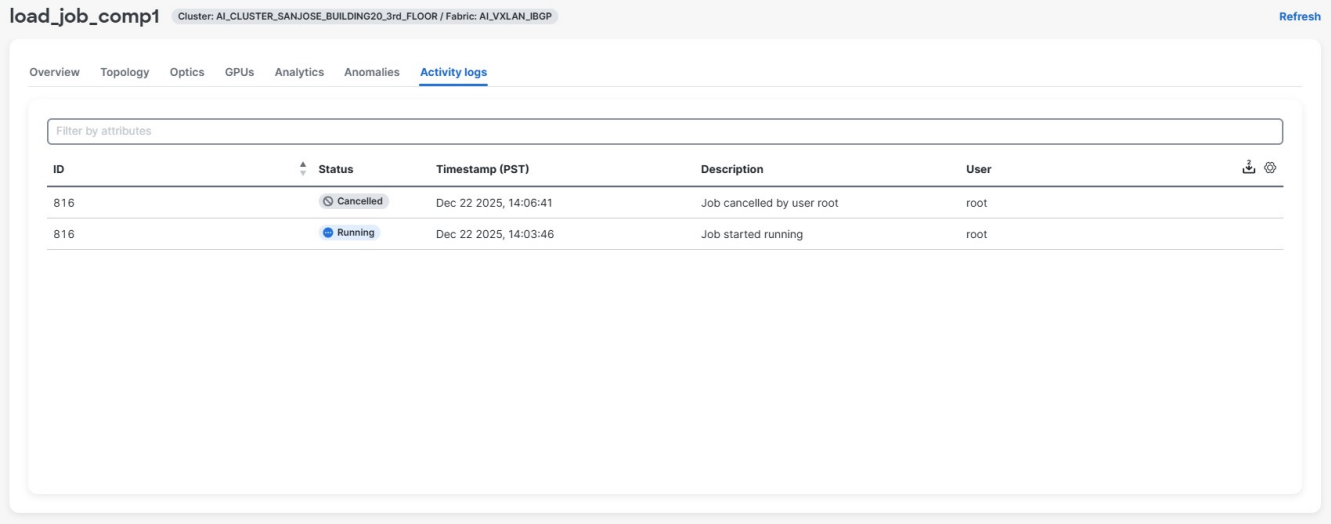
Field	Description
Anomaly level	Displays the anomaly level of the AI job at these anomaly levels: • Critical • Major • Minor • Warning • Healthy
Category	Displays a subset of anomalies that are relevant in the context of AI jobs: • Interface Dom Transmit Power • Interface Dom Receive Power • Interface Dom Current • Interface Dom Temperature • Interface Dom Voltage • Interface FCS Error • Interface Stomped CRC Error • Interface Down • Interface Utilization High Threshold • Network Congestion Indication • Environmental Outbound Discard High
Switch interface	Displays the name of the switch interface. Click the name to bring up a page with this additional information on a specific switch interface.

Server NIC anomalies can also appear in the **Anomalies** tab for metrics such as CRC errors, transmit discards, received discards, NACK Rx, and NACK Tx.

For server NIC anomalies, the **What's wrong** column identifies the affected server interface and describes the issue. Click the description to open the anomaly details page. When available, the **Switch interface** column shows the neighboring switch interface that is associated with the affected server interface.

Activity logs: Job level

The **Activity logs** page provides information on the timeline of the different states a jobs goes through, such as Pending, Running Completed, and so on.



Use the **Filter by attributes** field to narrow the activity log entries that are displayed for the selected job.

The **Activity logs** tab displays this information:

Field	Description
ID	Displays the ID of AI job.
Status	Displays the status of the activity log: • Completed • Running
Timestamp (PST)	Displays the timestamp on the activity log.
Description	Displays the description of the activity log, such as Job started running or Job completed successfully.
User	Displays the user name of the activity log.

Copyright

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706

USA
https://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883