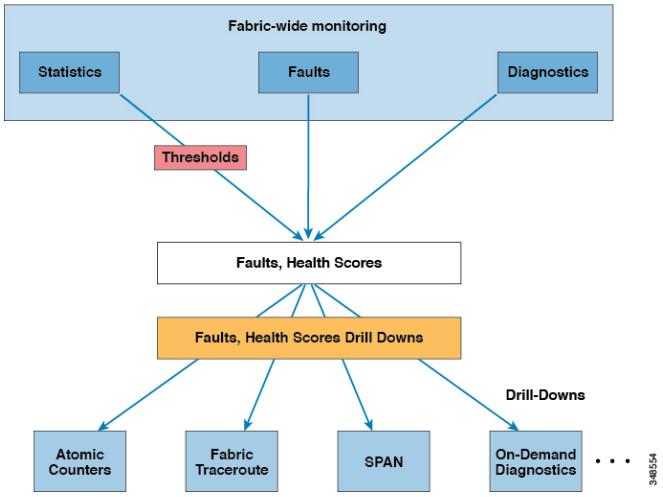
Troubleshooting

This chapter contains these sections:

Troubleshooting

The ACI fabric provides extensive troubleshooting and monitoring tools as shown in the following figure.

Figure 1: Troubleshooting



About ACL Contract Permit and Deny Logs

To log and/or monitor the traffic flow for a contract rule, you can enable and view the logging of packets or flows that were allowed to be sent because of contract permit rules or the logging of packets or flows that were dropped because of:

- Taboo contract deny rules
- Deny actions in contract subjects
- Contract or subject exceptions
- ACL contract permit in the ACI fabric is only supported on Nexus 9000 Series switches with names that end in EX or FX, and all later models. For example, N9K-C93180LC-EX or N9K-C9336C-FX.
- Deny logging in the ACI fabric is supported on all platforms.
- Using log directive on filters in management contracts is not supported. Setting the log directive will cause zoning-rule deployment failure.

For information on standard and taboo contracts and subjects, see *Cisco Application Centric Infrastructure Fundamentals* and *Cisco APIC Basic Configuration Guide*.

EPG Data Included in ACL Permit and Deny Log Output

Up to Cisco APIC, Release 3.2(1), the ACL permit and deny logs did not identify the EPGs associated with the contracts being logged. In release 3.2(1) the source EPG and destination EPG are added to the output of ACI permit and deny logs. ACL permit and deny logs include the relevant EPGs with the following limitations:

- Depending on the position of the EPG in the network, EPG data may not be available for the logs.
- When configuration changes occur, log data may be out of date. In steady state, log data is accurate.

The most accurate EPG data in the permit and deny logs results when the logs are focussed on:

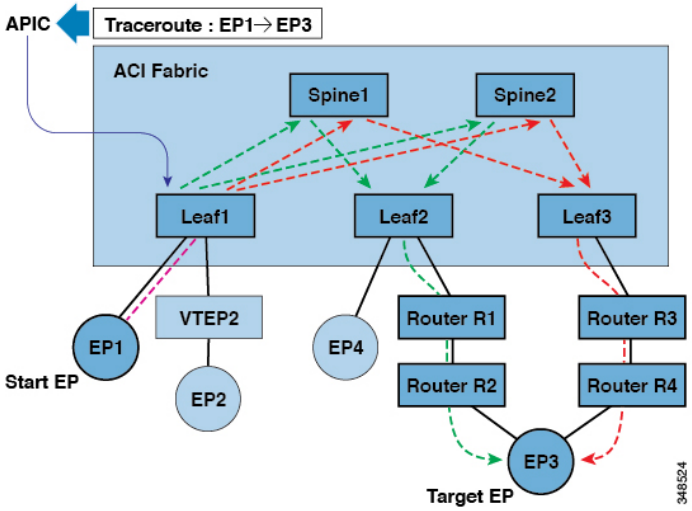
- Flows from EPG to EPG, where the ingress policy is installed at the ingress TOR and the egress policy is installed at the egress TOR.
- Flows from EPG to L3Out, where one policy is applied on the border leaf TOR and the other policy is applied on a non-BL TOR.

EPGs in the log output are not supported for uSeg EPGs or for EPGs used in shared services (including shared L3Outs).

ARPs, ICMP Pings, and Traceroute

ARPs for the default gateway IP address are trapped at the ingress leaf switch. The ingress leaf switch unicasts the ARP request to the destination and the destination sends the ARP response.

Figure 2: APIC Endpoint to Endpoint Traceroute



A traceroute that is initiated from the tenant endpoints shows the default gateway as an intermediate hop appears at the ingress leaf switch.

Traceroute modes include from endpoint to endpoint, and from leaf to leaf (TEP to TEP). Traceroute discovers all paths across the fabric, points of exit for external endpoints, and helps to detect if any path is blocked.

Traceroute works with IPv6 source and destination addresses but configuring source and destination addresses across IPv4 and IPv6 addresses is not allowed. Source (`RsTrEpIpSrc`) and destination (`RsTrEpIpDst`) relations support source and destination of type `fvIp`. At times, multiple IP addresses are learned from the same endpoint. The administrator chooses the desired source and destination addresses.

Atomic Counters

Beginning with Cisco APIC release 6.1(2), Atomic Counters are no longer supported.

Atomic counters detect drops and misrouting in the fabric. The resulting statistics enable quick debugging and isolation of application connectivity issues. Atomic counters require an active fabric Network Time Protocol (NTP) policy. Atomic counters work for either IPv6 or IPv4 source and destination addresses but not across different address families.

For example, an administrator can enable atomic counters on all leaf switches to trace packets from endpoint 1 to endpoint 2. If any leaf switches have nonzero counters, other than the source and destination leaf switches, an administrator can drill down to those leaf switches.

In conventional settings, it is nearly impossible to monitor the amount of traffic from a baremetal NIC to a specific IP address (an endpoint) or to any IP address. Atomic counters allow an administrator to count the number of packets that are received from a baremetal endpoint without any interference to its data path. In addition, atomic counters can monitor per-protocol traffic that is sent to and from an endpoint or an application group.

Leaf-to-leaf (TEP to TEP) atomic counters can provide the following:

- Counts of drops, admits, and excess packets.

- Short-term data collection such as the last 30 seconds, and long-term data collection such as 5 minutes, 15 minutes, or more.
- A breakdown of per-spine traffic is available only when the number of TEPs, leaf or VPC, is less than 64.
- Ongoing monitoring.

Note Leaf-to-leaf (TEP to TEP) atomic counters are cumulative and cannot be cleared. However, because 30-second atomic counters reset at 30-second intervals, they can be used to isolate intermittent or recurring problems.

Tenant atomic counters can provide the following:

- Application-specific counters for traffic across the fabric, including drops, admits, and excess packets
- Modes include the following:
  - Endpoint-to-endpoint MAC address, or endpoint-to-endpoint IP address. Note that a single target endpoint could have multiple IP addresses associated with it.
  - EPG to EPG
  - EPG to endpoint
  - EPG to * (any)
  - Endpoint to external IP address

  Beginning with the 5.2(3) release, endpoint security groups (ESGs) can be used as an alternative for EPGs in these modes.

Note Use of atomic counters is not supported when the endpoints are in different tenants or in different Virtual Routing and Forwarding (VRF) instances (also known as contexts or private networks) within the same tenant. Atomic counters work for IPv6 source and destinations but configuring source and destination IP addresses across IPv4 and IPv6 addresses is not allowed.

Endpoint-to-endpoint atomic counter statistics are not reported for Layer 2 bridged traffic with IPv6 headers when the endpoints belong to the same EPG.

For atomic counters to work for traffic flowing from an EPG or ESG to an L3Out EPG, configure the L3Out EPG with 0/1 and 128/1 to match all prefixes instead of 0/0.

About Digital Optical Monitoring

Real-time digital optical monitoring (DOM) data is collected from SFPs, SFP+, and XFPs periodically and compared with warning and alarm threshold table values. The DOM data collected are transceiver transmit bias current, transceiver transmit power, transceiver receive power, and transceiver power supply voltage.

Health Scores

The ACI fabric uses a policy model to combine data into a health score. Health scores can be aggregated for a variety of areas such as for the system, infrastructure, tenants, applications, or services.

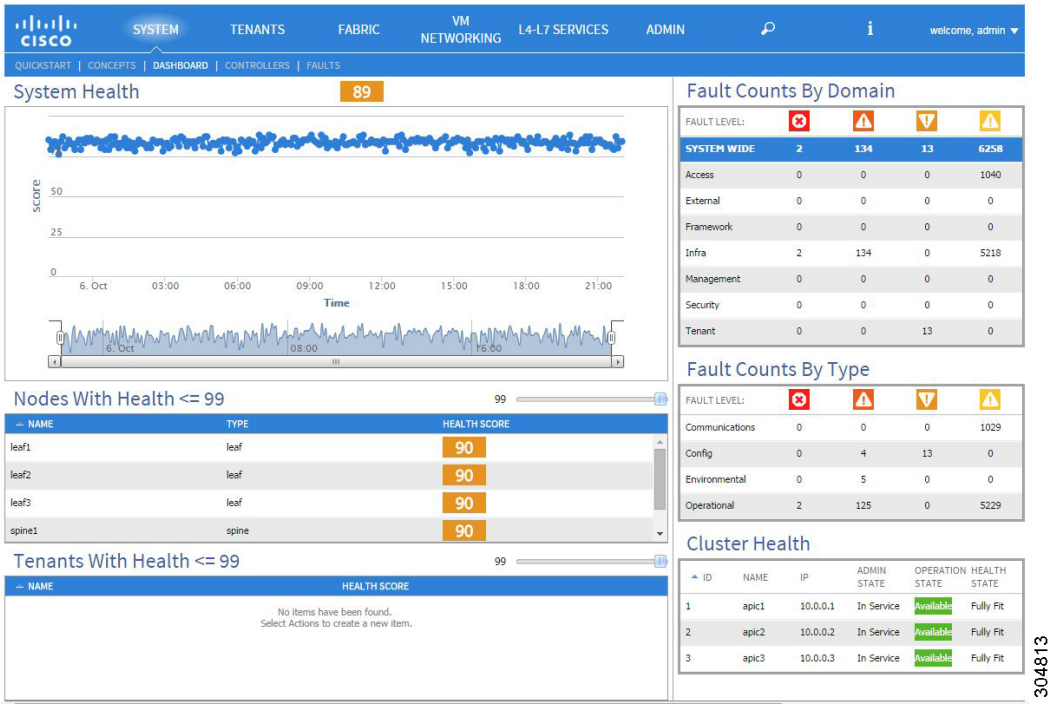
ACI fabric health information is available for the following views of the system:

- System — aggregation of system-wide health, including pod health scores, tenant health scores, system fault counts by domain and type, and the APIC cluster health state.
- Pod — aggregation of health scores for a pod (a group of spine and leaf switches), and pod-wide fault counts by domain and type.
- Tenant — aggregation of health scores for a tenant, including performance data for objects such as applications and EPGs that are specific to a tenant, and tenant-wide fault counts by domain and type.
- Managed Object — health score policies for managed objects (MOs), which includes their dependent and related MOs. These policies can be customized by an administrator.

System and Pod Health Scores

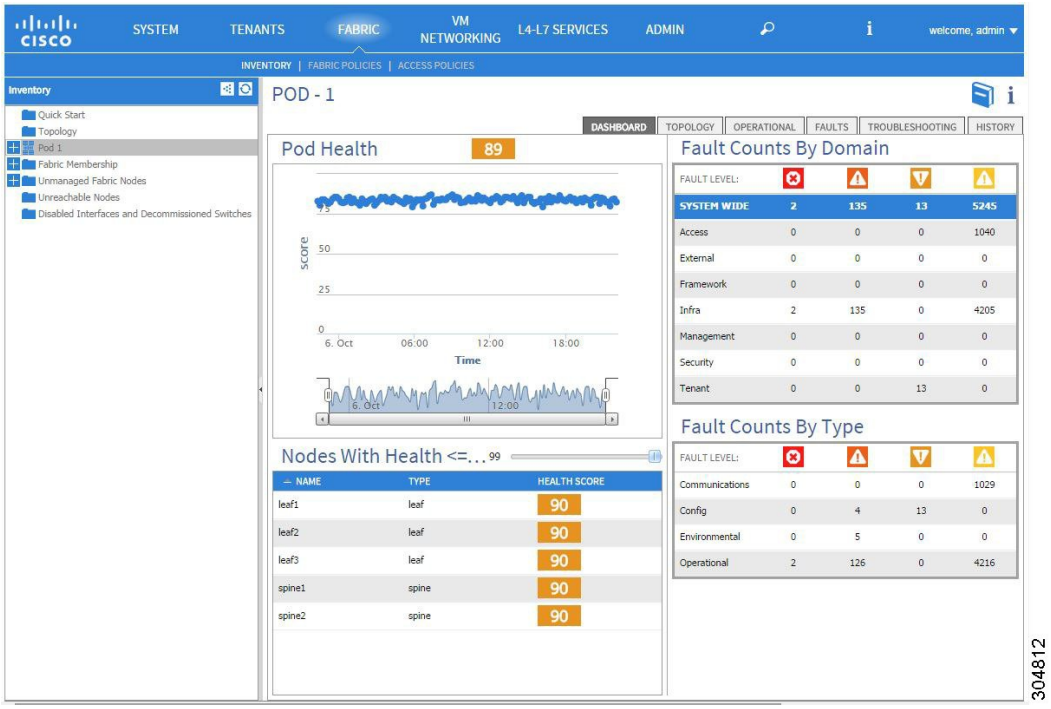
The system and pod health scores are based on the leaf and spine switches health scores as well as the number of end-points learned on the leaf switches. The GUI System Dashboard also displays system-wide fault counts by domain type, along with the APIC cluster per-node admin state, operational state, and health state.

Figure 3: System Health Scores



The pod health scores are based on the leaf and spine switches health scores as well as the number of end-points learnt on the leaf switches. The GUI fabric pod dashboard screen also displays pod-wide fault counts by domain and type.

Figure 4: Pod Health Scores



The system and pod health scores are calculated the same way. The calculation is based on the weighted average of the leaf health scores divided by the total number of learned end points of the leaf switches times the spine coefficient which is derived from the number of spines and their health scores.

The following equation shows how this calculation is done.

Figure 5: System and Pod Health Score Calculation

$$Health_{Fabric} = \frac{\sum_{i=1}^{N_{Leaf}} Health_{Leaf_i} \times Weight_{Leaf_i}}{\sum_{i=1}^{N_{Leaf}} Weight_{Leaf_i}} \times \left(1 - \left(1 - \frac{\sum_{i=1}^{N_{Spine}} Health_{Spine_i}}{N_{Spine} \times 100}\right)^{N_{Spine}}\right)$$

304814

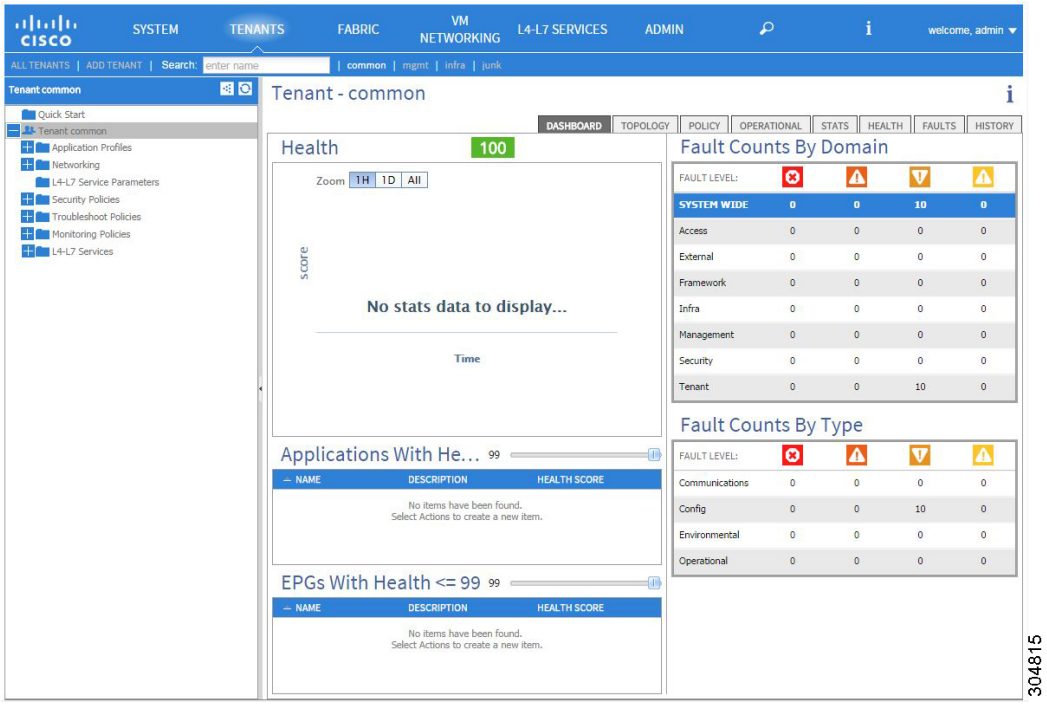
The following legend defines the equation components.

- $Health_{Leaf i}$ is the health score of the leaf switch.
- $Weight_{Leaf i}$ is the number of end-points on the leaf switch.
- N_{Leaf} is the number of leaf switches in the fabric.
- $Health_{Spine i}$ is the health score of the spine switch.
- N_{Spine} is the number of spine switches in the fabric.

Tenant Health Scores

Tenant health scores aggregate the tenant-wide logical objects health scores across the infrastructure they happen to use. The GUI tenant dashboard screen also displays tenant-wide-fault counts by domain and type.

Figure 6: Tenant Health Scores

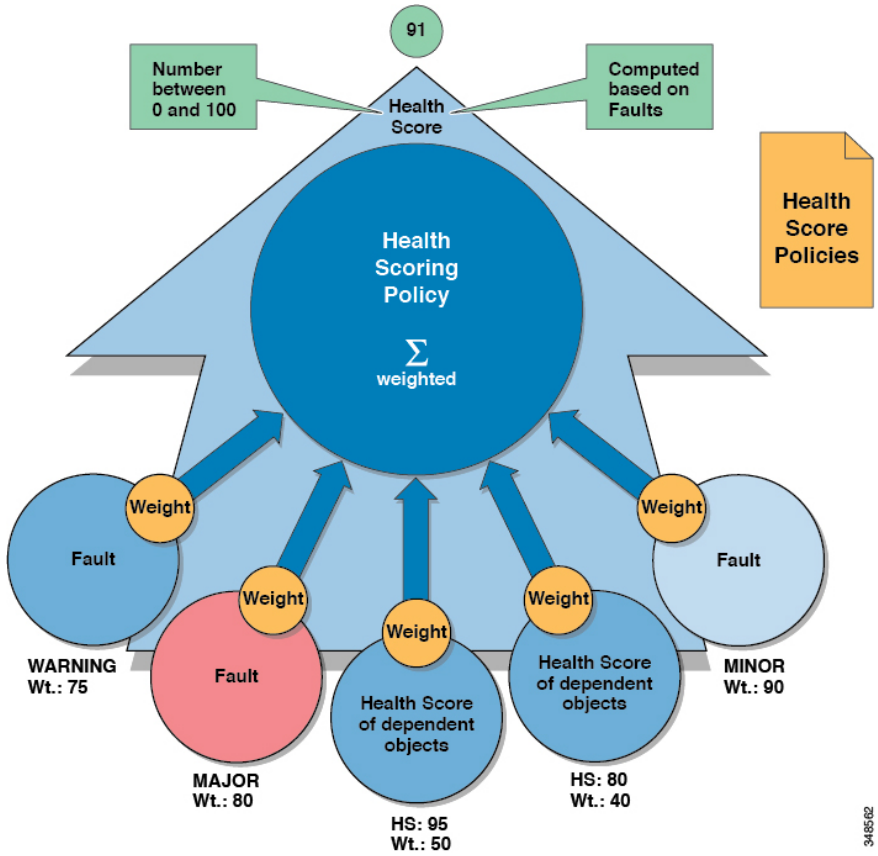


For example, an EPG could be using ports of two leaf switches. Each leaf switch would contain a deployed EPG component. The number of learned endpoints is a weighting factor. Each port could have a different number of learned endpoints. So the EPG health score would be derived by summing the health score of each EPG component times its number of learned endpoints on that leaf, divided by the total number of learned endpoints across the leaf switches the EPG uses.

MO Health Scores

Each managed object (MO) belongs to a health score category. By default, the health score category of an MO is the same as its MO class name.

Figure 7: MO Health Score



Each health score category is assigned an impact level. The five health score impact levels are Maximum, High, Medium, Low, and None. For example, the default impact level of fabric ports is Maximum and the default impact level of leaf ports is High. Certain categories of children MOs can be excluded from health score calculations of its parent MO by assigning a health score impact of None. These impact levels between objects are user configurable. However, if the default impact level is None, the administrator cannot override it.

The following factors are the various impact levels:

Maximum: 100% High: 80% Medium: 50% Low: 20% None: 0%

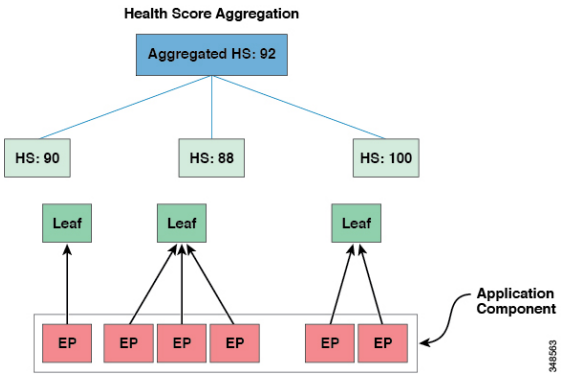
The category health score is calculated using an Lp -Norm formula. The health score penalty equals 100 minus the health score. The health score penalty represents the overall health score penalties of a set of MOs that belong to a given category and are children or direct relatives of the MO for which a health score is being calculated.

The health score category of an MO class can be changed by using a policy. For example, the default health score category of a leaf port is `eqpt:LeafP` and the default health score category of fabric ports is `eqpt:FabP`. However, a policy that includes both leaf ports and fabric ports can be made to be part of the same category called ports.

Health Score Aggregation and Impact

The health score of an application component can be distributed across multiple leaf switches as shown in the following figure.

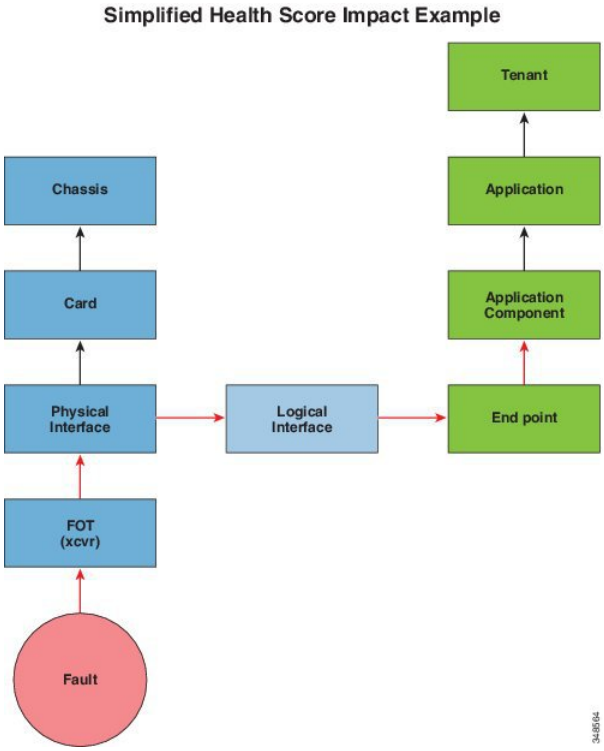
Figure 8: Health Score Aggregation



The aggregated health score is computed at the APIC.

In the following figure, a hardware fault impacts the health score of an application component.

Figure 9: Simplified Health Score Impact Example



About SPAN

You can use the Switched Port Analyzer (SPAN) utility to perform detailed troubleshooting or to take a sample of traffic from a particular application host for proactive monitoring and analysis.

SPAN copies traffic from one or more ports, VLANs, or endpoint groups (EPGs) and sends the copied traffic to one or more destinations for analysis by a network analyzer. The process is nondisruptive to any connected devices and is facilitated in the hardware, which prevents any unnecessary CPU load.

You can configure SPAN sessions to monitor traffic received by the source (ingress traffic), traffic transmitted from the source (egress traffic), or both. By default, SPAN monitors all traffic, but you can configure filters to monitor only selected traffic.

You can configure SPAN on a tenant or on a switch. When configured on a switch, you can configure SPAN as a fabric policy or an access policy.

APIC supports the encapsulated remote extension of SPAN (ERSPAN).

Multinode SPAN

APIC traffic monitoring policies can SPAN policies at the appropriate places to track members of each application group and where they are connected. If any member moves, APIC automatically pushes the policy to the new leaf switch. For example, when an endpoint VMotions to a new leaf switch, the SPAN configuration automatically adjusts.

Additional Information

Refer to the *Cisco APIC Troubleshooting Guide* for detailed information about the configuration, use, and limitations of SPAN.

About SNMP

The Cisco Application Centric Infrastructure (ACI) provides extensive SNMPv1, v2, and v3 support, including Management Information Bases (MIBs) and notifications (traps). The SNMP standard allows any third-party applications that support the different MIBs to manage and monitor the Cisco ACI fabric.

SNMPv3 provides extended security. Each SNMPv3 device can be selectively enabled or disabled for SNMP service. In addition, each device can be configured with a method of handling SNMPv1 and v2 requests.

Beginning in the 5.1(1) release, SNMPv3 supports the Secure Hash Algorithm-2 (SHA-2) authentication type.

For more information about using SNMP, see the *Cisco ACI MIB Quick Reference*.

About Syslog

During operation, a fault or event in the Cisco Application Centric Infrastructure (ACI) system can trigger the sending of a system log (syslog) message to the console, to a local file, and to a logging server on another system. A system log message typically contains a subset of information about the fault or event. A system log message can also contain audit log and session log entries.

Note For a list of syslog messages that the APIC and the fabric nodes can generate, see *Cisco ACI System Messages Reference Guide*.

Many system log messages are specific to the action that a user is performing or the object that a user is configuring or administering. These messages can be the following:

- Informational messages, providing assistance and tips about the action being performed
- Warning messages, providing information about system errors related to an object, such as a user account or service profile, that the user is configuring or administering

In order to receive and monitor system log messages, you must specify a syslog destination, which can be the console, a local file, or one or more remote hosts running a syslog server. In addition, you can specify the minimum severity level of messages to be displayed on the console or captured by the file or host. The local file for receiving syslog messages is `/var/log/external/messages`.

A syslog source can be any object for which an object monitoring policy can be applied. You can specify the minimum severity level of messages to be sent, the items to be included in the syslog messages, and the syslog destination.

You can change the display format for the Syslogs to NX-OS style format.

Additional details about the faults or events that generate these system messages are described in the *Cisco APIC Faults, Events, and System Messages Management Guide*, and system log messages are listed in the *Cisco ACI System Messages Reference Guide*.

Note Not all system log messages indicate problems with your system. Some messages are purely informational, while others may help diagnose problems with communications lines, internal hardware, or the system software.

About the Troubleshooting Wizard

The Troubleshooting Wizard allows you to understand and visualize how your network is behaving, which can ease your networking concerns should issues arise. For example, you might have two endpoints that are having intermittent packet loss, but you do not understand why. Using the Troubleshooting Wizard, you can evaluate the issue so that you can effectively resolve the issue rather than logging onto each machine that you suspect to be causing this faulty behavior.

This wizard allows you (the administrative user) to troubleshoot issues that occur during specific time frames for the chosen source and destination. You can define a time window in which you want to perform the debug, and you can generate a troubleshooting report that you can send to TAC.

Related Topics

Getting Started with the Troubleshooting Wizard
Topology in the Troubleshooting Wizard

About Cisco Nexus 9000 Switch Secure Erase

Cisco Nexus 9000 switches utilize persistent storage to maintain system software images, switch configuration, software logs, and operational history. Each of these areas can contain user-specific information such as details on network architecture and design, and potential target vectors for would-be attackers. The secure erase feature enables you comprehensively to erase this information, which you can do when you return a switch with return merchandise authorization (RMA), upgrade or replace a switch, or decommission a system that has reached its end-of-life.

Secure erase is supported from Cisco APIC release 6.0(x). All the leaf and spine switches in the fabric must be APIC release 6.0(x) or later.

This feature erases user data in the following storage devices:

- SSD
- EMMC
- MTD
- CMOS
- NVRAM

Note Not every switch model has all these storage devices.