

**Document Classification**

**Cisco Highly Confidential**

Cisco Systems, Inc.

Corporate Headquarters

170 West Tasman Drive

San Jose, CA 95134-1706 USA

Phone: +1 408-526-4000

Toll Free: +1 800-553-NETS (6387)

Fax: +1 408-526-4100

ASML

ACI Health Check - TPE

29 Apr 2024

Version 1.0

Contents

[Contents 2](#_Toc165363167)

[Document Summary 5](#_Toc165363168)

[History 5](#_Toc165363169)

[Review 5](#_Toc165363170)

[About this Document 6](#_Toc165363171)

[Purpose 6](#_Toc165363172)

[Audience 6](#_Toc165363173)

[Scope 6](#_Toc165363174)

[Assumptions 6](#_Toc165363175)

[Related Documents 6](#_Toc165363176)

[Summary of Recommendations 7](#_Toc165363177)

[1 Fabric Overview 8](#_Toc165363178)

[1.1 Hardware 8](#_Toc165363179)

[1.2 Firmware 8](#_Toc165363180)

[1.3 Tenant overview 8](#_Toc165363181)

[Bridge Domain overview 8](#_Toc165363182)

[Fault Summary 9](#_Toc165363183)

[2 Findings 10](#_Toc165363184)

[2.1 Operational Health 10](#_Toc165363185)

[FN72464 - Nexus 9300 Switches Can Experience Memory Failures 10](#_Toc165363186)

[2.2 Best Practices 11](#_Toc165363187)

[Backup to Remote Location 11](#_Toc165363188)

[Digital Optical Monitoring (DOM) 12](#_Toc165363189)

[Domain Validation 13](#_Toc165363190)

[Mis-Cabling Protocol (MCP) – Interface Configuration 14](#_Toc165363191)

[Disable Remote EP Learning 16](#_Toc165363192)

[Upgrade Groups 17](#_Toc165363193)

[2.3 Lifecycle 18](#_Toc165363194)

[End of Life 18](#_Toc165363195)

[3 Fault Review 19](#_Toc165363196)

[4 Scale 23](#_Toc165363197)

[4.1 Overview 23](#_Toc165363198)

[4.2 Current Scale 23](#_Toc165363199)

[Fabric-Wide Scale 23](#_Toc165363200)

[Per-Device Managed Object Scale 23](#_Toc165363201)

[Per-Device Endpoint Scale 24](#_Toc165363202)

[5 Other Checks Performed - Info Only 25](#_Toc165363203)

[5.1 Operational Health 25](#_Toc165363204)

[Algosec App 25](#_Toc165363205)

[Switch Node Bootflash Space 26](#_Toc165363206)

[APIC Cluster Health 27](#_Toc165363207)

[High-Risk Faults 28](#_Toc165363208)

[Configuration accepted with IP address mismatch for a given VLAN on the same node 29](#_Toc165363209)

[Multiple Firmware Versions 30](#_Toc165363210)

[FN72145 - SSD Failure After 3.2 Years 31](#_Toc165363211)

[Infra VLAN Consistency 32](#_Toc165363212)

[Multi-pod ISIS Metric 33](#_Toc165363213)

[Overlapping VLANs 34](#_Toc165363214)

[APIC Disk Utilization 35](#_Toc165363215)

[SSD Faults 36](#_Toc165363216)

[5.2 Potential Misconfiguration 37](#_Toc165363217)

[Bridge Domain configurations 37](#_Toc165363218)

[5.3 Best Practices 38](#_Toc165363219)

[BFD on Fabric-Facing Interfaces 38](#_Toc165363220)

[Common Tenant Duplicate Names 39](#_Toc165363221)

[COOP Strict Mode 40](#_Toc165363222)

[Duplicate VLAN usage on EPGs 41](#_Toc165363223)

[Encrypted Backups 42](#_Toc165363224)

[Enforce Subnet Check 43](#_Toc165363225)

[EP Loop Protection 44](#_Toc165363226)

[Fabric ID Check 45](#_Toc165363227)

[Ingress Policy Enforcement 46](#_Toc165363228)

[IP Aging 47](#_Toc165363229)

[L3out Redundancy 48](#_Toc165363230)

[L3out Overlapping Subnets 49](#_Toc165363231)

[MisCabling Protocol (MCP) - Global Configuration 50](#_Toc165363232)

[NTP Redundancy 51](#_Toc165363233)

[Leaf and Spine Out-of-band Management 52](#_Toc165363234)

[Port Tracking 53](#_Toc165363235)

[Rogue EP Control 54](#_Toc165363236)

[Route Reflector Redundancy 55](#_Toc165363237)

[Fabric Topology 56](#_Toc165363238)

[vzAny 57](#_Toc165363239)

[5.4 Configuration Cleanup 58](#_Toc165363240)

[AEP associated with domain with invalid pool 58](#_Toc165363241)

[Bridge Domain VRF Associations 59](#_Toc165363242)

[Missing VLAN Pool on L2/L3/Physical Domain 60](#_Toc165363243)

[6 References 61](#_Toc165363244)

[6.1 Online references 61](#_Toc165363245)

[6.2 Loop Detection 61](#_Toc165363246)

[BD Level Tracking (EP dampening, move frequency) 61](#_Toc165363247)

[EP Loop Protection 61](#_Toc165363248)

[Rogue EP Control 62](#_Toc165363249)

[Trademarks and Disclaimers 63](#_Toc165363250)

[Document Acceptance 64](#_Toc165363251)

Document Summary

|  |  |
| --- | --- |
| Rendered by | Michael Van Kleij (mvanklei@cisco.com) |
| Change Authority | Cisco Systems |
| DCP Reference | 1812370 |
| Project ID | B87866 |

History

|  |  |  |
| --- | --- | --- |
| Version | Issue Date | Reason for Change |
| 1.0 | 29 Apr 2024 | First Release |

Review

|  |  |  |  |
| --- | --- | --- | --- |
| Version | Date | Reviewer's Name | Reviewer's Organization |
|  | 29 Apr 2024 |  |  |

About this Document

Purpose

This document outlines the findings from an architectural and health assessment of the ASML TPE fabric. Details and recommendations are described throughout the document, and recommendations are summarized at the end.

Audience

This report is intended for the ASML engineering, architecture, and leadership responsible for ACI design and operational health.

Scope

This review covers only the TPE fabric. The review includes configuration and design review, scalability, and overall network health. Additional recommendations may be added to future health checks, as well as adjustments to the included recommendations based on Cisco Services customer insights. At the same time, the assessment and recommendations in this document have been written in a way where the guidance should continue to be relevant for the foreseeable changes.

Assumptions

Related Documents

Summary of Recommendations

The following table summarizes and prioritizes the recommendations in this document. Priorities are determined through a combination of risk and ease/risk of implementing the recommended change.

**Note** that these priorities are based on automated analysis of the TPE fabric and Cisco knowledge of the environment; however, they may vary significantly between companies, and between fabrics in the same company. These are guidelines to be discussed and prioritized by ASML to determine what’s best for this environment.

|  |
| --- |
| Recommendation |
| **High** |
| Enable MCP on all interfaces (if scale allows). |
| **Medium** |
| Enable Domain Validation. |
| Ensure all member switches of vPC pairs are in different upgrade groups. |
| Replace the expired certificate. |
| **Low** |
| Enable Digital Optical Monitoring (DOM). |
| Re-enable remote endpoint learning. |
| Start Life Cycle Management process for 93180YC-FX, 9364C and APIC-SERVER-L3. |
| There are some invalid configurations present. Remove these to improve stability. |

# Fabric Overview

## Hardware

The following hardware was found in TPE:

|  |  |  |
| --- | --- | --- |
| Model | Role | Count |
| APIC-SERVER-L3 | Controller | 3 |
| N9K-C9364C | Spine | 2 |
| N9K-C93180YC-FX | Leaf | 8 |
| N9K-C9336C-FX2 | Leaf | 8 |

## Firmware

The lowest version of firmware found in TPE is 5.2(8e). Note that the fabric was reviewed for multiple versions of firmware, malformed version strings, etc.

|  |
| --- |
| Lowest Version |
| 5.2(8e) |

## Tenant overview

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tenant | VRFs | BDs | EPGs | L3outs | Contracts |
| common | 2 | 1 | 0 | 1 | 1 |
| infra | 2 | 2 | 2 | 0 | 0 |
| mgmt | 2 | 1 | 0 | 0 | 0 |
| tpe-dc-tenant1 | 56 | 172 | 196 | 72 | 12 |

Bridge Domain overview

|  |  |
| --- | --- |
| BD Configuration | Count |
| Total | 172 |
| **Layer2** | |
| L2 Total | 20 |
| L2 Hardware Proxy | 1 |
| L2 Flooding | 19 |
| **Layer3** | |
| L3 Total | 152 |
| L3 Hardware Proxy | 145 |
| L3 Flooding | 7 |
| **Specific Configurations** | |
| L3 No Subnet | 0 |
| L3 Limit IP Learning to Subnet | 0 |

Fault Summary

|  |  |
| --- | --- |
| Severity | Count |
| Critical | 60 |
| Major | 55 |
| Warning | 8 |
| Minor | 60 |

# Findings

This section outlines findings where action or further assessment is recommended. Details about risk, change process, and further references are included.

## Operational Health

FN72464 - Nexus 9300 Switches Can Experience Memory Failures

Overview

A limited number of Dual In-line Memory Modules (DIMMs) shipped from Cisco are impacted by a known deviation in the memory supplier's manufacturing process. This deviation can result in a higher rate of failure.

Most DIMMs with this manufacturing deviation will exhibit persistent correctable memory errors. If left untreated, the DIMMs can eventually encounter an uncorrectable memory event. If encountered during runtime, uncorrectable errors will cause an unexpected switch reset.

Various DIMM Reliability, Availability, and Serviceability (RAS) features or even operating system features can mask the extent of these correctable errors. It is recommended to check your DIMMs for exposure

More details can be found in the [field notice](https://www.cisco.com/c/en/us/support/docs/field-notices/724/fn72464.html) on cisco.com.

Finding

The following devices are *potentially* exposed to this issue. **Note** that to identify exact risk additional action will need to be taken, i.e. validate the serial number in the online [Serial Number Validation Tool](https://snvui.cisco.com/snv/FN72464).

|  |  |
| --- | --- |
| Model | Potentially Impacted Device |
| APIC-SERVER-L3 | topology/pod-1/node-1 |
| topology/pod-1/node-2 |
| topology/pod-1/node-3 |
| N9K-C93180YC-FX | topology/pod-1/node-101 |
| topology/pod-1/node-102 |
| topology/pod-1/node-103 |
| topology/pod-1/node-104 |
| topology/pod-1/node-105 |
| topology/pod-1/node-106 |
| topology/pod-1/node-111 |
| topology/pod-1/node-112 |

## Best Practices

Backup to Remote Location

Overview

The ACI Fabric configuration should be backed up to the remote location periodically, in case an engineer needs to restore the fabric. With ACI, all components of the ACI Fabric are treated as one entity (leafs, spines, APIC controllers). The ACI Fabric configuration, while made up of different managed objects, is combined into one tar/gz zip file, which greatly improves the configuration backup and restore process.

* This section checks that the ACI Fabric backup is configured to backup to an external destination.
* Please refers to [Creating a Backup for Your APIC Cluster](https://www.cisco.com/c/dam/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/aci-guide-creating-backup-for-apic-cluster.pdf) for detailed information about the process of creating a backup configuration to a remote location.

Finding

There are export configuration policies **without** a remote location.

* tpe\_aci\_fabric\_daily\_backup

Recommendation

* Follow the instructions in the linked reference article to create backup configuration policy with an external destination.

Digital Optical Monitoring (DOM)

Overview

Digital Optical Monitoring (DOM) is an industry standard that provides additional monitoring for optical connections beyond simple up/down. It monitors optic-specific state, e.g. send and receive power, which can protect against situations like impending failure and degraded connectivity.

More details are available in the [DOM section](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/troubleshooting/Cisco-APIC-Troubleshooting-Guide-401/b_APIC_Troubleshooting_4x_chapter_0110.html#id_37684) of the ACI Troubleshooting Guide.

Finding

Digital Optical Monitoring (DOM) is **not** enabled on all interfaces. The following devices do not have DOM enabled:

108, 113, 115, 116, 103, 104, 109, 1101, 111, 107, 110, 1102, 114, 101, 102, 105, 106, 112

Without DOM the fabric will not have access to L1 optical data that can be useful for troubleshooting. It's therefore recommended to enable DOM on all interfaces. Enabling DOM on optics that do not support DOM will have no impact.

DOM can be configured at **Fabric > Fabric Policies > Policies > Monitoring > Fabric Node Controls > default**

No impact is expected from this change.

Domain Validation

Overview

Domain Validation prevents a misconfiguration where static ports are deployed in an EPG without a domain also configured. Without this feature a VLAN can be configured in an EPG without a domain and the VLAN will be deployed with no validation. If a domain is later associated to this domain, VLANs will be validated against the domain's associated VLAN pool, and invalid VLANs will be removed from hardware.

With Domain Validation enabled, static ports are prevented from being deployed until a domain is configured.

More detail about this feature is available in the [ACI Design Guide](https://www.cisco.com/c/en/us/td/docs/dcn/whitepapers/cisco-application-centric-infrastructure-design-guide.html).

Finding

Domain validation is **not** enabled.

Domain Validation protects against a misconfiguration that can result in unexpected traffic loss. It's recommended to consider enabling Enable Domain Validation

**CAUTION:**

Please be aware of the following considerations:

Domain Validation cannot be turned off once it has been enabled, even through backup restore or downgrade. Previous snapshots or backups will not work.

Additionally, if EPGs currently exist with static path bindings and no domain association, these port bindings will stop working with this feature is enabled. Fault F0468 will be raised for invalid path assignments. These will need to be corrected prior to enabling Domain Validation by associating a domain with the required VLANs.

Domain Validation can be configured at **System > System Settings > Fabric-Wide Settings > Enforce Domain Validation**

Mis-Cabling Protocol (MCP) – Interface Configuration

Overview

The mis-cabling protocol (MCP) was designed to handle misconfigurations not detected by Link Layer Discovery Protocol (LLDP) and Spanning Tree Protocol (STP). MCP sends out layer 2 hello packets. If these packets are received on another interface, the ports that form the loop will be disabled.

More details are [available on CCO](https://www.cisco.com/c/dam/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/aci-guide-using-mcp-mis-cabling-protocol.pdf).

Finding

MCP is **not** configured on all interfaces. The following interface profiles do not have MCP configured:

|  |
| --- |
| Disabled Device |
| uni/infra/accportprof-intprof-swtpele101\_102 |
| uni/infra/accportprof-intprof-swtpele101\_102 |
| uni/infra/accportprof-intprof-swtpele101\_102 |
| uni/infra/accportprof-intprof-swtpele101\_102 |
| uni/infra/accportprof-intprof-swtpele101\_102 |
| uni/infra/accportprof-intprof-swtpele101\_102 |
| uni/infra/accportprof-intprof-swtpele101\_102 |
| uni/infra/accportprof-intprof-swtpele101\_102 |
| uni/infra/accportprof-intprof-swtpele101\_102 |
| uni/infra/accportprof-intprof-swtpele111 |
| uni/infra/accportprof-intprof-swtpele112 |
| uni/infra/accportprof-intprof-swtpele111 |
| uni/infra/accportprof-intprof-swtpele112 |
| uni/infra/accportprof-intprof-swtpele111 |
| uni/infra/accportprof-intprof-swtpele112 |
| uni/infra/accportprof-intprof-swtpele111 |
| uni/infra/accportprof-intprof-swtpele112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |
| uni/infra/accportprof-intprof-swtpele111\_112 |

Without MCP ACI is at risk of propagating bridging loop behavior due to STP failure or L1 problems. It's, therefore, recommended to configure MCP on all interfaces whenever feasible.

Enable MCP within the Interface Policy, at **Fabric > Access Policies > Policies > MCP Interface**.

**CAUTION:**

Please ensure your fabric is not running at scale exposed to [CSCvx37709](https://bst.cloudapps.cisco.com/bugsearch/bug/CSCvx37709) before considering enabling MCP.

MCP hellos are transmitted unencapsulated. An L2 misconfiguration can allow MCP hellos to "jump" from one VLAN to another, appearing as though an L2 loop is occurring, and resulting in err-disable. Because of this, it’s extremely important to ensure legacy environments are configured correctly, and specifically that an 802.1q header is applied to unencapsulated traffic.

Disable Remote EP Learning

Overview

Remote EP learning is when a leaf caches the remote location of an endpoint. This functionality isn't strictly necessary, as unknown remote destinations will be proxied to the spine, which will always maintain EP location in the COOP database.

There are several scenarios where remote endpoint learning on a border leaf can result in stale endpoints, and traffic loss. Details about these scenarios are available in the [Remote EP learning section](https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html#_Toc529820938) of the ACI Endpoint Learning Whitepaper.

Finding

**Disable Remote EP Learning** is on.

Spine proxy can be avoided by turning this feature off. Note that this feature may have been previous recommended to be enabled; however, this is no longer required with newer hardware and code.

As a general best practice, Cisco recommends turning this feature back off to take advantage of remote EP learning from the border leaf.

Note that disabling this feature is a comparatively low priority recommendation. The impact of spine proxy at the border leaf is not expected to be a risk to the fabric.

This feature can be configured at **System > System Settings > Fabric Wide Settings > Disable Remote EP Learning**.

While enabled, unknown EPs will be proxied to the spine; however, no impact is expected from modifying this feature.

Upgrade Groups

Overview

Upgrade groups should be designed to upgrade the fabric in stages defined by redundant infrastructure.

Finding

Note that the assessment performed in this section of the document does not indicate an immediate risk; however, upgrade groups should be configured correctly prior to an upgrade.

**Summary:**

|  |  |
| --- | --- |
| Risk | Impacted Nodes |
| VPC switch pairs are not in redundant upgrade groups | 115  116 |

VPCs are not in Redundant Upgrade Groups

VPCs are **not** in Redundant Upgrade Group, i.e., some VPC interfaces are in the same upgrade group.

Loss of both VPC switches during an upgrade will result in loss of connectivity for all hosts attached to this VPC pair. It's recommended to distribute VPC switches across different upgrade groups.

Upgrade Groups can be configured under **Admin > Firmware > Infrastructure > Nodes**

No impact is expected from upgrade group modification.

## Lifecycle

End of Life

As EoL devices pass critical milestones, they will no longer receive patches. After Last Day of Support (LDoS), EoL devices will no longer be supported by TAC.

Finding

The following EoL notices were identified. Please review these in conjunction with your lifecycle management strategy to ensure devices are replaced prior to critical milestones.

|  |  |  |  |
| --- | --- | --- | --- |
| Model | Announced | URL | Impacted Devices |
| APIC-SERVER-L3 | May 1, 2023 | [CCO Announcment](https://www.cisco.com/c/en/us/products/collateral/cloud-systems-management/application-policy-infrastructure-controller-apic/apic-m3-l3-se-node-g2-eol.html) | apictpe01  apictpe02  apictpe03 |
| N9K-C93180YC-FX | August 1, 2023 | [CCO Announcment](https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/nexus-c93180yc-fx-c93108tc-fx-fixed-switches-eol.html) | swtpele101  swtpele102  swtpele103  swtpele104  swtpele105  swtpele106  swtpele111  swtpele112 |
| N9K-C9364C | August 1, 2023 | [CCO Announcment](https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/nexus-9332c-9364c-fixed-spine-switch-eol.html) | swtpesp1101  swtpesp1102 |

# Fault Review

The following faults types were found in the TPE fabric.

|  |  |  |  |
| --- | --- | --- | --- |
| Code | Count | Explanation | Recommendation |
| **Critical Faults** | | | |
| F0532 | 59 | This fault occurs when a port is down and is in use for epg | 1. Check the port connectivity 2. Remove the configuration or administratively shut the port if the port is not in use 3. To remove the static port configuration from EPG go to Tenant->ApplicationProfile->ApplicationEPG->Static Ports and remove the affected port if not being used. 4. For mcp-loop-err-disable, this could be due to a loop in the network. Check the config to resolve any loops 5. For lacp suspended ports, check for following issues - 6. Check whether peer device supports lacp or not. 7. Enable LACP feature in peer device, if peer device requires explicit global configuration. 8. Check vlan range configured part of member interface and port-channel configurations matches. 9. Check whether peer physical interface is added to port-channel as member interface. 10. Check whether lacp is configured on member interface of peer device. 11. If peer switch is of Cisco's, issue "show lacp counters" and verify interfaces of peer device are sending LACP PDU. 12. Using the "show etherchannel summary"/"show port-channel summary" command, ensure that the port-channel shows the S (Layer-2) and U (in use) flags, and that both the interfaces appear in the ports column with the P (bundled) flag. 13. If peer switch or router is not manufactured by Cisco, Please contact customer support of peer device's manufacturer for not sending lacp or recommended configuration guidelines. 14. For reason of suspend(connected) port, check whether VPC domain is configured in GUI Fabric-> External Access/Access Policies ->Virtual Port Channel defaul to tie two switches part of VPC domain if intent is to make two switches part of VPC. 15. If the above actions did not resolve the issue, create a tech-support file and contact Cisco TAC. |
| F3082 | 1 | This fault occurs when the SAML X.509 Certificate has expired. | 1. Update SAML X.509 Certificate by going to Admin->AAA->Authentication->SAML. |
| **Major Faults** | | | |
| F0600 | 4 | This fault occurs when port has been suspended. |  |
| F1296 | 28 | This fault occurs when vpc interface goes down while peer interface is also down. |  |
| F1545 | 13 | This fault occurs when a significant number of packet drops are detected by a configured and enabled Atomic Counter |  |
| F1547 | 10 | This fault occurs when a significant number of excess packets are detected by a configured and enabled Atomic Counter |  |
| **Warning Faults** | | | |
| F112128 | 8 | This fault is caused by "ingress drop packets rate" statistical property crossing threshold level. |  |
| **Minor Faults** | | | |
| F0756 | 49 | This fault occurs when a configured target of a named relationship cannot be resolved. | 1. [Verify that the configuration for the named target is correct and complete, and that it exists Verify the configuration for the specific relationship is correct and complete. If the above actions did not resolve the issue, create a tech-support file and contact Cisco TAC.] |
| F0758 | 1 | This fault occurs when a service graph could not be instantiated. | 1. [Look at the graph configuration for issues For missing-lif, check for fault on LDev or check LDev specified in LDevCtx For function-type-mismatch-failure, check for valid MFunc in the MDev for selected LDev For missing-ldev-ctx, check if LDevCtx is defined under tenant, or check if LDev associated with LDevCtx has proper MDev specified. For missing-lif, check that matching LIF for func connector exists For invalid-filter-referred-by-rsconntoflt, check AbsFuncConn to filter relation is formed For missing-l3ctx, check that RsConnToCtxTerm points to valid tDn For param-cardinality-error, check that configuration parameter has valid cardinality, Graph param/rel can only have at most one instance For invalid-abstract-graph-config-param, check that all abstract graph configurations are valid For param-validation-failed, check that all specified params are valid For invalid-param-scope, check that RsScopeToTerm is configured and is pointing to valid terminal or terminal node. Or check that there is no missing default RsScopeToTerm in AbsConfig For invalid-conn-for-configparam, check that RsCfgToConn is configured in AbsConn and is pointing to a valid connector For invalid-rsmconnatt, check for valid RsMConn in AbsConn For missing-mandatory-param, check that all mandatory graph params and graph relations exist For invalid-relation-target, fix the incorrect service folder/param name] |
| F0849 | 1 | This fault occurs when a infra selector (port selector, card selector, node selector etc.) is incorrectly configured. | 1. [Look at the configuration for issues For invalid-port, check the infra port selector, it should have only leaf host ports or fex host ports. Fabric ports are not allowed to be configured using infra port selector. For port-configured-for-apic, check the infra port selector and the associated policies. Port connected to the controller can be configured for limited policies. Its allowed to be associated with only infra::AccPortGrp. For invalid-card, check the infra card selector, its referring to a card which does not exist. Fault will clear once the card is added to the node. For further details, refer to the documentation for fabric:SelectorIssues Verify the infra selector configuration is correct and complete If the above action did not resolve the issue, create a tech-support file and contact Cisco TAC.] |
| F1546 | 3 | This fault occurs when a small number of packet drops are detected by a configured and enabled Atomic Counter |  |
| F1548 | 5 | This fault occurs when a small number of excess packets are detected by a configured and enabled Atomic Counter |  |
| F1690 | 1 | This fault occurs when there is incomplete or incorrect configuration, intermittent system errors. |  |

# Scale

## Overview

ACI scalability is tested as a single unit, i.e. the entire fabric is tested against multiple scale limits at the same time. This differs from some of the previous, non-ACI scale testing methodologies that would test a scale metric in isolation, e.g. add 100k OSPF routes to a switch and test against the switch in that state. This isolated testing methodology is called unit testing.

The risk associated with primarily relying on unit testing is that it's not very indicative of a real-world environment, and it's increasingly likely that issues related to multiple stressors might not be discovered.

Some ACI scale limits are hard limits, i.e., a firm limit in software or hardware that cannot be exceeded, and some are soft limits that are limits determined during testability. Whether a given limit is a hard or soft limit is not documented in the scalability guide. This is because a soft limit hasn't been tested over scale, and the behavior is unknown, and could lead to impact. For this reason, all of the scale limits should be treated as hard limits. If there's technical justification to scale over the documented limits, please discuss this with your Cisco team to determine the best approach.

The ACI Scalability Guides are available on CCO on the [APIC homepage](https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html), under the "Verified Scalability" section. **Note** that a new guide is not released for every code release--choose the latest guide for the closest code version *before* your release of code.

## Current Scale

Fabric-Wide Scale

|  |  |  |
| --- | --- | --- |
| Metric | Count | Limit |
| EPGs | 198 | 15000 |
| BDs | 176 | 15000 |
| VRFs | 62 | 3000 |
| Tenants | 4 | 3000 |
| Contracts | 13 | 10000 |
| Filters | 11 | 10000 |

Per-Device Managed Object Scale

Some scale limits in ACI are tested limits, and some are hard limits, e.g. TCAM slots. In either case, a fabric in excess of supported scale limits may be denied support until scale is reduced, and may result in unexpected behavior and/or failure.

The following table lists a subset of scale metrics for this fabric. **Note** that this is not a comprehensive review of *all* scale metrics. The [Scalability Guides](https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html) should serve as the final authority on ACI scalability metrics.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Node | EPGs | BDs | VRFs | TCAM | VLANs |
| swtpele101 | 60 of 3960 | 142 of 3500 | 59 of 800 | 370 of 65536 | 203 of 3960 |
| swtpele102 | 60 of 3960 | 142 of 3500 | 59 of 800 | 370 of 65536 | 203 of 3960 |
| swtpele103 | 4 of 3960 | 5 of 3500 | 5 of 800 | 43 of 65536 | 10 of 3960 |
| swtpele104 | 4 of 3960 | 5 of 3500 | 5 of 800 | 43 of 65536 | 10 of 3960 |
| swtpele105 | 114 of 3960 | 89 of 3500 | 48 of 800 | 450 of 65536 | 204 of 3960 |
| swtpele106 | 114 of 3960 | 89 of 3500 | 48 of 800 | 450 of 65536 | 204 of 3960 |
| swtpele107 | 71 of 3960 | 47 of 3500 | 44 of 800 | 240 of 65536 | 119 of 3960 |
| swtpele108 | 71 of 3960 | 47 of 3500 | 44 of 800 | 240 of 65536 | 119 of 3960 |
| swtpele109 | 47 of 3960 | 47 of 3500 | 45 of 800 | 254 of 65536 | 95 of 3960 |
| swtpele110 | 47 of 3960 | 47 of 3500 | 45 of 800 | 254 of 65536 | 95 of 3960 |
| swtpele111 | 2 of 3960 | 2 of 3500 | 5 of 800 | 41 of 65536 | 5 of 3960 |
| swtpele112 | 1 of 3960 | 1 of 3500 | 4 of 800 | 36 of 65536 | 3 of 3960 |
| swtpele113 | 5 of 3960 | 5 of 3500 | 6 of 800 | 46 of 65536 | 11 of 3960 |
| swtpele114 | 5 of 3960 | 5 of 3500 | 6 of 800 | 46 of 65536 | 11 of 3960 |
| swtpele115 | 1 of 3960 | 1 of 3500 | 4 of 800 | 36 of 65536 | 3 of 3960 |
| swtpele116 | 1 of 3960 | 1 of 3500 | 4 of 800 | 36 of 65536 | 3 of 3960 |

Per-Device Endpoint Scale

|  |  |  |
| --- | --- | --- |
| Node | L2 Total | L3 Total |
| swtpele101 | 437 of 24576 | 1684 of 24576 |
| swtpele102 | 441 of 24576 | 1684 of 24576 |
| swtpele103 | 1243 of 24576 | 71 of 24576 |
| swtpele104 | 1247 of 24576 | 71 of 24576 |
| swtpele105 | 2467 of 24576 | 1171 of 24576 |
| swtpele106 | 2467 of 24576 | 1171 of 24576 |
| swtpele107 | 2151 of 24576 | 820 of 24576 |
| swtpele108 | 2151 of 24576 | 820 of 24576 |
| swtpele109 | 425 of 24576 | 806 of 24576 |
| swtpele110 | 427 of 24576 | 805 of 24576 |
| swtpele111 | 35 of 24576 | 62 of 24576 |
| swtpele112 | 34 of 24576 | 61 of 24576 |
| swtpele113 | 198 of 24576 | 194 of 24576 |
| swtpele114 | 198 of 24576 | 194 of 24576 |
| swtpele115 | 10 of 24576 | 3 of 24576 |
| swtpele116 | 10 of 24576 | 3 of 24576 |

Finding

No per-node scalability risks were identified. Note that only a subset of metrics were validated. Please review the scalability guide for the full list of scale limits.

# Other Checks Performed - Info Only

This section provides an overview of best practices and health checks performed where no risks were identified.

## Operational Health

Algosec App

Overview

The AlgoSec app provides policy-based security integration services using AlgoSec's security analytics.

The AlgoSec app may expose the fabric to [CSCvv12524](https://bst.cloudapps.cisco.com/bugsearch/bug/CSCvv12524/?rfs=iqvred). Before attempting any software upgrade of the ACI fabric the following applications from the ACI App Center **must** be deactivated/removed:

Finding

No AlgoSec apps are installed.

Switch Node Bootflash Space

Overview

ACI switch image files are transfered to switch nodes during an upgrade. It is important to ensure that enough space is available in /bootflash on each switch for image files so that upgrades can complete successfully. If /bootflash is more than 50% utilized, it is recommended to remove unnecessary files from that directory to ensure adequate space is available

Additional information regarding /bootflash utilization can be found in the Pre-Upgrade Checklists chapter of the [Cisco APIC Installation and ACI Upgrade and Downgrade Guide](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/all/apic-installation-aci-upgrade-downgrade/Cisco-APIC-Installation-ACI-Upgrade-Downgrade-Guide/m-pre-upgrade-checklists.html).

Finding

All switch nodes have less than or equal to 50% space utilization on /bootflash.

APIC Cluster Health

Overview

A "fully fit" state indicates the database and messaging between APICs is fully synchronizing.

Any state other than "fully fit" indicates data is **not** synchronizing between APICs. This could result in loss of configuration, issues with VMM integration, and various other issues. This is a high risk that should be dealt with immediately.

Finding

The APIC cluster is **fully fit**.

High-Risk Faults

"High risk" faults are a select list of faults that have been observed to result in higher than average impact across a wide range of customer environments. **Note** that this is not an exhaustive list of impactful faults.

Finding

No high-risk faults were identified in this review. **Note** that this does not indicate that there are no potentially impacting issues--only that these specific issues were not found.

Configuration accepted with IP address mismatch for a given VLAN on the same node

Overview

[CSCvh02653](https://bst.cloudapps.cisco.com/bugsearch/bug/CSCvh02653) allows an invalid configuration with an IP address mismatch to be accepted by the APIC. L3Outs exposed to this issue are configured with two or more IP addresses for a given VLAN on the same node. This results in the last IP address programmed to be the active one while the other IP address is no longer present on the node.

Findings

Issue not identified.

Multiple Firmware Versions

Overview

Multiple versions of firmware are tested in ACI upgrade testing; however, this state is expected to run for a limited time, i.e. only during upgrades. Additionally, Cisco recommends no configuration changes in mixed-firmware state.

Finding

No firmware issues were identified.

FN72145 - SSD Failure After 3.2 Years

Overview

After approximately 3.2 years (28,224 accumulated Power On Hours (POH)), a memory buffer overrun condition occurs which triggers the firmware event in the SSD. This causes the drive to become unresponsive until the drive is power-cycled. No data loss will occur when the memory buffer overrun firmware event occurs. A power-cycle restores normal operation of the drive. The drive continues to operate normally for approximately six weeks (1008 additional accumulated power on hours), at which time the drive will become unresponsive again. Power-cycling the drive again will re-initiate the 1008 hour window.

More details can be found in the [field notice](https://www.cisco.com/c/en/us/support/docs/field-notices/721/fn72145.html) on cisco.com.

Finding

No devices exposed to FN72145.

Infra VLAN Consistency

The infra VLAN is used for internal fabric control traffic.

A mismatch in the infra VLAN may result in communication failure between APICs or between a subset of APICs and leaves or spines. It's recommended to correct the APIC with the wrong VLAN ID.

More information about Infrastructure VLAN can be found in the [Cisco ACI Getting Started Guide](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/getting-started/Cisco-APIC-Getting-Started-Guide-421/b-Cisco-APIC-Getting-Started-Guide-421_chapter_010.html) available on cisco.com.

Finding

The infra VLAN is consistent across controllers.

Multi-pod ISIS Metric

Overview

The default ISIS redistribution metric is **63**, which is the maximum configurable metric. When an inter-pod router (IPN) spine is removed from the fabric and then reintroduced a hold down timer is applied to prevent forwarding on the spine until ISIS has fully converged. During this time, the metric is set to the maximum available (63) to prevent forwarding. Unfortunately, this is the same value as the default metric, so traffic begins forwarding immediately, and traffic loss may occur. [CSCvd75131](https://bst.cloudapps.cisco.com/bugsearch/bug/CSCvd75131/?rfs=iqvred) was filed to address this issue.

CSCvd75131 is first addressed in 2.2(4f); however, even with CSCvd75131 in place, the fix for this issue still requires a manual configuration change. The fix in CSCvd75131 is to *enable* the ability to change the metric; however, the metric must be manually updated to a value that addresses the original issue.

Finding

The ISIS metric is set to 32 as per best-practice.

Overlapping VLANs

Overview

In ACI, VNIDs are assigned based on VLAN + VLAN pool ID. The same VLAN ID in different VLANs is provided a different VNID. Ports using the same VLAN from different pools will experience traffic loss in certain L2 dynamics, e.g. the two legs of a vPC, or BPDU forwarding for spanning tree.

Additionally, if an EPG has multiple domains with the same VLAN provided by different pools, the VLAN used is nondeterministic. Because of this a working environment can start experiencing problems on a reload.

Note, that this analysis *only* checks for overlapping VLANs associated with the same AEP.

Finding

No overlapping VLANs were identified.

APIC Disk Utilization

Overview

It is recommended to ensure that the Disk Utilization of the mount points are below 75% capacity.

Additional information for disk utilization fault codes can be found at:

* [Fault Code: F1529](https://pubhub.devnetcloud.com/media/apic-mim-ref-501/docs/FAULT-F1529.html)
* [Fault Code: F1528](https://pubhub.devnetcloud.com/media/apic-mim-ref-501/docs/FAULT-F1528.html)
* [Fault Code: F1527](https://pubhub.devnetcloud.com/media/apic-mim-ref-501/docs/FAULT-F1527.html)

Finding

Disk utilization in all the APIC nodes is below 75% threshold.

SSD Faults

Overview

This section reviews faults related to known SSD issues. Impact of SSD failure varies from a best case that configuration cannot be saved, to a worst case of device failure.

Finding

No SSD-specific faults were identified.

## Potential Misconfiguration

Bridge Domain configurations

Overview

Bridge Domains in ACI have a number of configuration options to allow the administrator to tune the operation in various ways.

This section reviews Bridge Domain configurations for information purposes and to check for risks and best practices.

Please reference the [ACI Endpoint Learning Whitepaper](https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html) for detailed information about the available configuration options and how to optimally apply them to your network requirements.

Finding

No configuration risks were identified.

## Best Practices

BFD on Fabric-Facing Interfaces

Overview

Internal fabric routing in ACI is performed by ISIS and BGP. Bi-directional Forwarding Detection (BFD) is a protocol that improves convergence of routing protocol in certain failure scenarios.

BFD does this by sending frequent (50ms by default) hellos between nodes, and triggering immediate failover on the loss of three hellos.

BFD is not configured on fabric facing interfaces. This is the recommended configuration.

Common Tenant Duplicate Names

Overview

Duplicate names can contribute to misconfiguration where an incorrect VRF, BD, contract, or filter are used. In some cases, e.g. filters, object relationships are defined by name-only, and which object is used is ambiguous.

This can result in misconfiguration, ambiguous configuration, or unexpected behavior due to the wrong object association.

More information about object naming can be found in the [Cisco ACI Object Naming and Numbering: Best Practices Guide](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/kb/b-Cisco-ACI-Naming-and-Numbering.html) on cisco.com.

Finding

No duplicate object names found.

COOP Strict Mode

Overview

COOP strict mode protects against certain exploits that take advantage of the spine accepting unauthenticated updates. Setting the COOP Group setting to Strict hashes updates using MD5. The MD5 keys are changed out every hour and redistributed to switches.

For this reason, strict mode is recommended as a general best practice.

Finding

COOP strict mode is configured, as per best practice.

Duplicate VLAN usage on EPGs

Overview

It is recommended to use a unique VLAN per EPG where possible as unintended flooding between EPGs can otherwise occur. When a VLAN is selected from a VLAN pool by an EPG, it is allocated a VXLAN identifier known as a fabric\_encap. The fabric\_encap is used to forward spanning-tree BPDUs within the ACI fabric for the given VLAN. Reusing the same VLAN from the same VLAN pool on multiple EPGs may cause unintended forwarding of this traffic due to reuse of the fabric\_encap VXLAN identifier.

In situations where the same VLAN ID must be used by multiple EPGs and this flooding behavior is not intended, ensure that the EPGs use separate VLAN pools to ensure unique fabric\_encap allocation. In such cases however, be careful to avoid an overlapping VLAN pool situation where an EPG is associated to two or more access policy domains (e.g., physical domains) with overlapping VLAN pools. For more information on potential overlapping VLAN pool issues, review the [Cisco Community page on overlapping VLAN pools](https://community.cisco.com/t5/data-center-and-cloud-documents/aci-common-migration-issue-overlapping-vlan-pools/ta-p/3362376).

Finding

No duplicate VLAN usage was found.

Encrypted Backups

Overview

ACI backups are **unencrypted** by default. In an unencrypted backup, only non-sensitive configuration data is backed up. In an encrypted backup, passwords are encrypted, and backed up in addition to the standard, unencrypted configuration.

[Encrypted Backups](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/aci-fundamentals/Cisco-ACI-Fundamentals-401/Cisco-ACI-Fundamentals-401_chapter_01011.html#concept_15E2D7F6CCF24A98A40CBCB9A8302B81)

Finding

Encrypted backups are configured as per best-practice.

Enforce Subnet Check

Overview

[The Enforce Subnet Check section](https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html#_Toc529820939) of the Endpoint Learning whitepaper explains the global Enforce Subnet Check options in detail.

There are a few key differences between this and the BD-level "Limit IP Learning to Subnet" config:

1. This feature limits local learns in hardware. For local learns, it functions like "Limit IP Learning to Subnet" where addresses outside of the BD subnets will not be learned.
2. For remote learns, this feature will restrict IP learning at the VRF-level. This validates that a remotely learned IP belongs to a subnet within the source VRF.
3. Enforce Subnet Check is a single, global configuration option.

Finding

**Enforce Subnet Check** is enabled globally as per best practice.

EP Loop Protection

Overview

EP Loop Protection counts a move when a MAC address moves to another port and then back to its original port. It's therefore designed specifically to detect looping behavior. EP Loop Protection can be configured to err-disable the port or to raise a fault.

The default, BD-level loop protection will stop learning in the entire BD, potentially causing wide-spread impact.

Finding

EP Loop Protect is configured as per best practice for this fabric.

Fabric ID Check

Overview

It is best practice to set the ACI Fabric ID to the default value of "1" during the initial APIC setup script on all APICs in the fabric in most cases. At a minimum, the fabric ID should be a consistent value across all APICs if "1" is not used. Inconsistent fabric IDs may be propagated to fabric switches during node discovery and may result in forwarding inconsistencies. Additional information on fabric ID usage can be found in the [Cisco ACI Multi-Site Architecture White Paper](https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739609.html#Day0MultiSiteinfrastructureconfiguration).

Generally speaking, the fabric ID should only be changed from "1" if ACI GOLF is implemented and the GOLF routers are shared between different ACI fabrics with the same BGP ASN using auto-RT. The fabric ID in such cases serves as a unique identifier in VRF route targets to prevent unexpected cross-VRF route exhanges between fabrics.

Finding

All fabric nodes have a fabric ID of "1".

Ingress Policy Enforcement

Overview

Policy can either be enforced at the ingress or egress of the fabric. Software release 1.2 introduced a new policy enforcement model whereby security rules for all flows are enforced on the leaf node to which internal hosts are connected, rather than at the border leaf.

When the direction is set to **Egress**, the contract rules for an L3Out are deployed on both the border-leaf and non–border-leaf switches. In this situation, when there are many EPGs that need to talk to the L3Out, the TCAM resources for contracts on border leaf switches could be a bottle neck. This is because a border leaf deploys all contracts, while contracts on non–border leaf switches are typically distributed to multiple leaf switches. However, when set to **Ingress**, the contract rules are deployed only on non-border leaf switches; hence, this resolves the concern about TCAM resources for contracts on border leaf switches.

More information about Policy Control Enforcement Direction can be found in the [Cisco ACI L3Out Guide](https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/guide-c07-743150.html) and the [ACI Fabric Endpoint Learning White Paper](https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html), both available on cisco.com.

Finding

Ingress policy enforcement is configured on all VRFs.

IP Aging

Overview

By default, ACI only ages MAC address endpoints. For endpoints with multiple IP addresses assigned to the same MAC address, the IP addresses will not age separately. IP aging will apply the endpoint timer to IP addresses as well.

IP Aging is described in detail in the [IP aging section](https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html#_Toc529820940) of the ACI Endpoint Learning Whitepaper.

Finding

**IP Aging** is enabled as per best practice.

L3out Redundancy

Overview

It's generally recommended to configure L3outs across at least two border leaves. This allows for upgrades and other maintenance activities to occur without impacting the L3 topology.

[ACI L3out Configuration Guide](https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/guide-c07-743150.html)

Finding

All L3outs are redundant.

L3out Overlapping Subnets

Overview

ACI allows the same subnet to be used on an L3out and on a BD or EPG. This can lead to ambiguous forwarding behavior.

Finding

No overlapping subnets were identified.

MisCabling Protocol (MCP) - Global Configuration

Overview

The mis-cabling protocol (MCP) was designed to handle misconfigurations not detected by Link Layer Discovery Protocol (LLDP) and Spanning Tree Protocol (STP). MCP sends out layer 2 hello packets. If these packets are received on another interface, the ports that form the loop will be disabled.

More details are [available on CCO](https://www.cisco.com/c/dam/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/aci-guide-using-mcp-mis-cabling-protocol.pdf).

Finding

MCP is enabled globally as per best practice.

Please note that this health check is not currently checking for exposure to [CSCvx37709](https://bst.cloudapps.cisco.com/bugsearch/bug/CSCvx37709). Although, MCP is *generally* recommended, a scale configuration may put this fabric at risk of hitting this issue.

NTP Redundancy

Overview

Network Time Protocol (NTP) provides consistent, reliable time on the fabric. This is critical for a number of critical features, e.g. logging, authentication, encryption, atomic counters, etc.

More details of NTP configuration can be found in [this NTP configuration guide](https://www.cisco.com/c/en/us/support/docs/cloud-systems-management/application-policy-infrastructure-controller-apic/200128-Configuring-NTP-in-ACI-Fabric-Solution.html).

Finding

NTP is configured as per best practice.

Leaf and Spine Out-of-band Management

Overview

It is recommended to enable out-of-band addresses on all leaf and spine switches in the ACI fabric. Having out-of-band access is useful in the event direct access to the switches is needed and the normal management path via the APIC is disrupted. This check passes if either a valid IPv4 or IPv6 address is found.

Steps to configure out-of-band management and additional information can be found in the Management chapter of the Cisco APIC Basic Configuration Guide found on the [APIC homepage](https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html).

Finding

All nodes were found to have out-of-band IP addresses configured.

Port Tracking

Overview

The Port Tracking feature addresses a scenario where a leaf node may lose connectivity to the spine node and where hosts connected to the affected leaf node in an active / standby manner may not be aware of the failure for a period of time.

Finding

Port Tracking is enabled as per best practice.

Rogue EP Control

Overview

Rogue EP Control is a loop and misbehaving endpoint mechanism. It protects the fabric from issues like frequent flaps, loops, etc. As compared to the default BD-level loop protection, Rogue EP Control counts MAC and IP moves separately and only impacts the specific, misbehaving endpoints.

Detection criteria can be configured by using the following values:

* Rogue EP Detection interval: to specify the time in seconds to detect rogue endpoints. The default is 60 seconds. The supported range is 30 to 3600 seconds.
* Rogue EP Detection Multiplication Factor: The endpoint is declared rogue if the endpoint moves more than this number within the Rogue EP Detection interval. The default is 4. The supported range is 2 to 10.
* Hold Interval: the amount of time the endpoint is being handled as rogue and kept as the static endpoint. After this interval, the endpoint is deleted. The default is 1800 seconds (30 minutes). The supported range depends on the release. With ACI releases prior to ACI 5.2(3) the configurable range is 1800 to 3600. Starting with ACI 5.2(3) you can configure a minimum hold interval of 300 seconds (5 minutes).

More information about Rouge Endpoint Control can be found in the [Rogue Endpoint Control](https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html#RogueEPControl) guide available on cisco.com.

**Note** that enabling Rogue EP Control automatically disables the BD-level loop prevention mechanism.

Finding

Rogue EP Control is currently on. This is configured as recommended for this fabric.

Route Reflector Redundancy

Overview

Route reflectors are used internally within ACI for MP-BGP route redistribution. This provides the critical service of ensuring L3out-learned routes are distributed between leaves.

Finding

All route reflectors are redundant, as per best practice.

Fabric Topology

It's expected that each leaf is connected to more than one spine. This allows for spine-level redundancy, e.g. upgrades and other maintenance can be performed on a spine without impact.

Finding

No design issues were identified with the fabric topology.

vzAny

Overview

vzAny is a managed object within ACI that represents all EPGs within a VRF. This object can be used to provide or consume contracts, reducing TCAM utilization from every EPG to a single vzAny relationship.

Finding

No opportunities were identified to adopt vzAny.

## Configuration Cleanup

AEP associated with domain with invalid pool

Overview

Domains, AEP, and VLANs are mandatory to deploy an EPG on a specific port. The domain profile contains both the VLAN instance profile (VLAN pool) and the attachable Access Entity Profile (AEP), which are associated directly with application EPGs. The AEP deploys the associated application EPGs to all the ports to which it is attached, and automates the task of assigning VLANs.

* This section checks for any domains associated with an AEP with an invalid pool.
* Please reference the section "Creating Domains, Attach Entity Profiles, and VLANs to Deploy an EPG on a Specific Port" in [Cisco APIC Basic Configuration Guide, Release 5.2(x)](https://www.cisco.com/c/en/us/td/docs/dcn/aci/apic/5x/basic-configuration/cisco-apic-basic-configuration-guide-52x/m_tenants.html) for detailed information about the requirements and procedure to deploy EPG related to AEPs, Domains, and VLANs.

Finding

All domains associated with an AEP have valid VLAN pools, as per best-practice.

Bridge Domain VRF Associations

Overview

BDs require an explicit VRF configuration. Without this, the BD configuration is not valid and will not forward traffic.

Finding

All BDs are associated to VRFs.

Missing VLAN Pool on L2/L3/Physical Domain

Overview

In most cases, it is recommended that External Bridged (L2), External Routed (L3), and Physical Domains should be associated to a VLAN pool. Without a VLAN pool, VLAN allocation may fail and likely result in fault F0467 triggering on any associated L2Outs, L3Outs, or EPGs.

In the case of L3 Domains, a VLAN pool does not need to be associated if the L3Out(s) using the L3 Domain are only using routed or subinterfaces. Any L3 Domains identified in this check which meet this criteria can be disregarded. Additional information on L3Out encap usage can be found in the [ACI Fabric L3Out Guide](https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/guide-c07-743150.html).

Finding

All access domains are associated to a VLAN pool.

# References

## Online references

* [ACI Best Practices Quick Summary](https://www.cisco.com/c/en/us/td/docs/dcn/whitepapers/cisco-aci-best-practices-quick-summary.html)
* [ACI Configuration Guides](https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html)
* [Fabric Endpoint Learning Whitepaper](https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html)
* [Scalability guides](https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html)
* [Fault reference guide](https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/all/syslog/guide/b_ACI_System_Messages_Guide.html)

## Loop Detection

Frequent endpoint moves can increase CPU utilization and fill up logs, making troubleshooting more difficult. Additionally, rapid endpoint moves can be a symptom of a bridging loop, which can have catastrophic impact. ACI provides several features to protect the network from bridging loops. The following sections describe the behavior of these mechanisms to help illustrate the difference in functionality.

BD Level Tracking (EP dampening, move frequency)

**What types of moves are detected and counted?**

* **MAC move:** "move count" will be 1 + # of IP Addresses linked to this MAC address in the bridge domain, e.g. if the EP has a MAC address and three IP addresses, move count will be 4 on the first MAC move.
* **IP only move:** move not counted
* Only local moves are counted. MAC moves across leaf switches are not counted.

**Timer and Threshold**

* **Detection Time:** 1 sec (fixed)
* **Move count threshold:** 256 by default
* **BD hold interval:** 300s by default

**What happens when move count exceeds threshold within Detection Time?**

* BD learning is disabled for that BD
* EPs in that BD are **not** flushed
* BD learning will be enabled again after BD hold interval.

EP Loop Protection

**What move is detected?**

* **MAC move:** move count (loop count) will be 1 (see details below)
* **IP only move:** move not counted
* Move is counted only when MAC address moves back to its previous port
* Both local moves and moves across leaf switches are counted

**Timer and Threshold**

* **Detection time:** 60s by default
* **Move count threshold:** 4 by default
* Disabled by default

**What happens when move count exceeds threshold within the detection time?**

* BD Learning is disabled for that BD

**-- and/or --**

* Last learned port is err-disabled (epm-learn-err-disable)
* BD Learning will be enabled again after BD hold interval from BD level tracking
* Port err-disable will be recovered if error disabled recovery policy is configured (not configured by default)
* Port err-disable will be recovered by manual shut/no shut.
* EP will be deleted soon from leaf since learned port is disabled
* If EP flap is so rapid that previous port can learn EP again before EP is deleted from err-disabled port, both ports could be err-disabled.
* Above both err-disable situation should be avoided if BD learning disable is enabled as well as port disable.

Rogue EP Control

**What move is detected?**

* **MAC move:** move count will be 1
* **IP only move:** move count will be 1
* MAC moves and IP only moves are counted separately
* Both local moves and moves across leaf switches are counted

**Timer and Threshold**

* **Detection Time:** 60s by default
* **Move count threshold:** 4 by default
* **Rogue EP hold timer:** 1800s by default
* Disabled by default

**What happens when move count exceeds threshold within Detection Time?**

* EP is marked as Rogue.
* Move notification for Rogue EP is ignored
* Rogue EP will be deleted after hold interval

Trademarks and Disclaimers

IF THIS DOCUMENT IS PROVIDED AS A DELIVERABLE IN ACCORDANCE WITH THE CISCO TERMS AND CONDITIONS ASSOCIATED WITH A PURCHASED CISCO SERVICE (“TERMS”) THEN THIS DOCUMENT IS PRESENTED SUBJECT TO THOSE TERMS. IN ALL OTHER EVENTS, THIS DOCUMENT IS PROVIDED “AS-IS” WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED, INCLUDING WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: www.cisco.com/go/trademarks. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)

© 2022 Cisco and/or its affiliates. All rights reserved.

Document Acceptance

|  |  |  |  |
| --- | --- | --- | --- |
| Name |  | Name |  |
| Title |  | Title |  |
| Company |  | Company |  |
| Signature |  | Signature |  |
| Date |  | Date |  |

|  |  |  |  |
| --- | --- | --- | --- |
| Name |  | Name |  |
| Title |  | Title |  |
| Company |  | Company |  |
| Signature |  | Signature |  |
| Date |  | Date |  |

|  |  |  |  |
| --- | --- | --- | --- |
| Name |  | Name |  |
| Title |  | Title |  |
| Company |  | Company |  |
| Signature |  | Signature |  |
| Date |  | Date |  |