



Cisco Advanced Services

ADP

ACI Health Check & Best Practices Report

January 21, 2021

Version 1.0

Cisco Systems, Inc.
Corporate Headquarters
170 West Tasman Drive
San Jose, CA 95134-1706 USA
<http://www.cisco.com>
Tel: 408 526-4000 Toll Free: 800 553-NETS (6387)
Fax: 408 526-4100

Contents

CONTENTS	2
LIST OF FIGURES AND TABLES	4
ABOUT THIS ACI HEALTH CHECK & BEST PRACTICES REPORT	6
HISTORY	6
REVIEW	6
1 INTRODUCTION.....	7
1.1 PURPOSE.....	7
1.2 ASSESSMENTS SCOPE.....	7
1.3 AUDIENCE	7
1.4 ASSUMPTIONS.....	7
1.5 REFERENCES.....	7
1.6 PRIORITY AND SEVERITY	8
2 ACI INFRASTRUCTURE AND INVENTORY.....	9
2.1 HIGH LEVEL DESIGN OVERVIEW	9
2.2 INVENTORY DETAILS OF ACI FABRIC	9
2.3 END OF LIFE MILESTONES.....	10
2.4 AUTOMATION AND ORCHESTRATION	11
3 DETAILED FINDINGS	12
4 FABRIC POLICIES	13
4.1 POD POLICIES.....	13
4.2 POD POLICIES IS-IS	14
4.3 POD POLICIES BGP ROUTE-REFLECTOR.....	15
4.4 POD POLICIES DATE/TIME (NTP)	16
4.5 POD POLICIES COMMUNITY/MANAGEMENT POLICY	17
4.6 POD POLICIES SNMP	19
5 ACCESS POLICIES	20
5.1 VLAN POOLS.....	20
5.2 DOMAIN-VLAN POOL -AAEP CHECK.....	26
5.3 INTERFACE PROFILES/POLICIES/POLICY GROUPS	30
6 TENANT POLICIES – NETWORKING	32
6.1 TENANT	32
6.2 VRF INGRESS POLICY ENFORCEMENT.....	32
6.3 BRIDGE DOMAINS.....	34
6.4 BRIDGE DOMAINS CONFIGURATION OVERVIEW	38
6.5 END POINT GROUPS (EPGs).....	40
6.6 L3OUTS	43
7 TENANT POLICIES – SECURITY POLICIES	46
7.1 CONTRACTS & VZANY	46
8 ADMIN	51
8.1 AAA-FALLBACK DOMAIN	51
8.2 FIRMWARE	51
8.3 ENCRYPTED BACKUPS	51

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

8.4 FABRIC EXPORT POLICIES	52
9 DESIGN BEST PRACTICES	54
9.1 POD POLICIES COOP POLICY	54
9.2 BFD FOR FABRIC FACING INTERFACES	55
9.3 DOMAIN VALIDATION	55
9.4 GENERAL STABILITY	56
9.4.1 <i>Mis-Cabling Protocol (MCP)</i>	56
9.4.2 <i>Digital Optical Monitoring (DOM)</i>	57
9.4.3 <i>Port Tracking</i>	59
10 END POINT LEARNING	61
10.1 ENFORCE SUBNET CHECK	62
10.1.1 <i>Remote EP Learning</i>	63
10.2 IP AGING POLICY	64
10.3 LIMIT IP LEARNING TO SUBNET	65
10.4 LOOP DETECTION	66
11 FABRIC HEALTH DC4	69
11.1 HEALTH SCORE	69
11.2 FABRIC HEALTH AND PERSISTENT FAULTS	70
11.3 FAULT MANAGEMENT	70
11.4 CISCO RECOMMENDATION	71
11.5 FAULTS	71
11.5.1 <i>Critical Faults</i>	72
11.5.2 <i>Major Faults</i>	72
11.5.3 <i>Minor Faults</i>	73
11.5.4 <i>Warning Faults</i>	73
12 FABRIC HEALTH DC5	76
12.1 HEALTH SCORE	76
12.2 CISCO RECOMMENDATION	77
12.2.1 <i>Critical Faults</i>	77
12.2.2 <i>Major Faults</i>	77
12.2.3 <i>Minor Faults</i>	79
12.2.4 <i>Warning Faults</i>	80
13 CURRENT SCALE PER SWITCH	82
13.1 ACI SCALABILITY MATRIX AND COMPLIANCE	82
13.2 DC4 PER-DEVICE SCALE LIMITS	84
13.3 DC5 PER-DEVICE SCALE LIMITS	84
TRADEMARKS AND DISCLAIMERS	85
DOCUMENT ACCEPTANCE	86

List of Figures and Tables

Figure 1: DC4 and DC5 Fabric Network Infrastructure Overview	9
Figure 2: Fabric Policies structure.....	13
Figure 3: The Hierarchical access policy relationships.....	20
Figure 4: ARP Gleaning Mechanism in ACI.....	36
Figure 5: L2 Connection to Fabric with External Gateways	38
Figure 6: TCAM entries per EPG pair	46
Figure 7: Multiple EPGs consuming a single contract.....	47
Figure 8: vzAny consuming a contract.....	47
Figure 9: Bi-Directional Contracts - Regular Configuration.....	48
Figure 10: Unidirectional Contracts.....	48
Figure 11: Use of vzAny with "Established" Contract	49
Figure 12: Leaf-Spine Connectivity Loss - Impact on Hosts	59
Figure 13: Leaf-Spine Connectivity Loss – Port Tracking Feature	59

Table 1: ADP Hardware DC4	9
Table 2: ADP Hardware DC5	9
Table 3: ADP Software	9
Table 4: End of Life Milestones.....	10
Table 5: Overall Findings and Recommendations	12
Table 6: ADP Pod Policy DC4 and DC5	13
Table 7: ADP IS-IS policy DC4 and DC5.....	14
Table 8: IS-IS pod policy descriptions	14
Table 9: ADP BGP RR current configuration DC4.....	15
Table 10: ADP BGP RR current configuration DC5	16
Table 11: Current NTP configurations DC4.....	16
Table 12: Current NTP configurations DC5.....	16
Table 13: Current DC4 policy configurations	17
Table 14: Current DC5 policy configurations	17
Table 15: ADP SNMP policy configurations DC4.....	19
Table 16: ADP SNMP policy configurations DC5	19
Table 17: VLAN Pool Example	21
Table 18: DC4 VLAN pool allocation	21
Table 19: DC5 VLAN pool allocation	23
Table 20: VLAN pools with overlapping ranges DC4.....	24
Table 21 - VLAN pools with overlapping ranges DC5.....	25
Table 22: DC4 Domain-Pool allocation	27
Table 23 - VLAN pools with overlapping ranges in the same AAEP DC4	28
Table 24 - VLAN pools with overlapping ranges in the same AAEP DC5	29
Table 25: DC4 – CTX.....	33
Table 26: DC5 – CTX.....	34
Table 27 - L3 BDs with no ARP glean	39
Table 28 - L3 BDs with no ARP glean	40
Table 29: DC4 Fabric Export policy	53
Table 30: DC5 Fabric Export policy	53
Table 31: DC4 and DC5 COOP policy configurations	54
Table 32: Global Mis-cabling Status and Configuration.....	57
Table 33: Limit IP Learning to Subnet configuration status at DC4 and DC5.....	65
Table 34: Per-device health score	69
Table 35: Per-device health score	76
Table 36: ACI General Scalability limit	82

Table 37: Overview of DC4 Configuration	83
Table 38: Overview of DC5 Configuration	83

About This ACI Health Check & Best Practices Report

Author Mateusz Jarosz
Change Authority Cisco Systems
DCP Reference
Project ID

History

Version	Issue Date	Reason for Change
1.0	January 21, 2021	First Release

Review

Version	Date	Reviewer's Name	Reviewer's Organization
1.0	January 21, 2021	Kaja Rog	Cisco Proactive Services

Document Conventions



Note: Alerts readers to take note. Notes contain helpful suggestions or references to material not covered in the document.



Caution: Alerts readers to be careful. In this situation, you might do something that could result in equipment damage or loss of data.



Warning: Alerts readers of a situation that could cause bodily injury. They need to be aware of the hazards involved with electrical circuitry and familiarize themselves with standard practices for preventing accidents.



Timesaver: Alerts the reader that they can save time by performing the action described in the paragraph affixed to this icon.



Tip: Alerts the reader that the information affixed to this icon will help them solve a problem. The information might not be troubleshooting or even an action, but it could be useful information similar to a Timesaver.

1 Introduction

1.1 Purpose

The purpose of this document is to provide the analysis of the overall health of the ACI Fabric and Best Practice Conformance for the ADP fabric.

The Assessment summarizes the configuration analysis and compares against ACI configuration best practices

1.2 Assessments Scope

The scope of this configuration Assessment is the ADP Application Centric Infrastructure already running deployment

1.3 Audience

The intended audience of this document is the ADP network teams and Cisco CX Team.

1.4 Assumptions

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.

1.5 References

Cisco Application Centric Infrastructure Best Practices Guide

https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/ACI_Best_Practices/b_ACI_Best_Practices.html

Cisco Application Centric Infrastructure Fundamentals

https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/1-x/aci-fundamentals/b_ACI-Fundamentals.html





Cisco Application Centric Infrastructure Design Guide White Paper

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-737909.html>

The document uses a color-scheme for the criticality and compliance of findings, shown in the tables below.

1.6 Priority and Severity

Table 1: Color scheme of Criticality

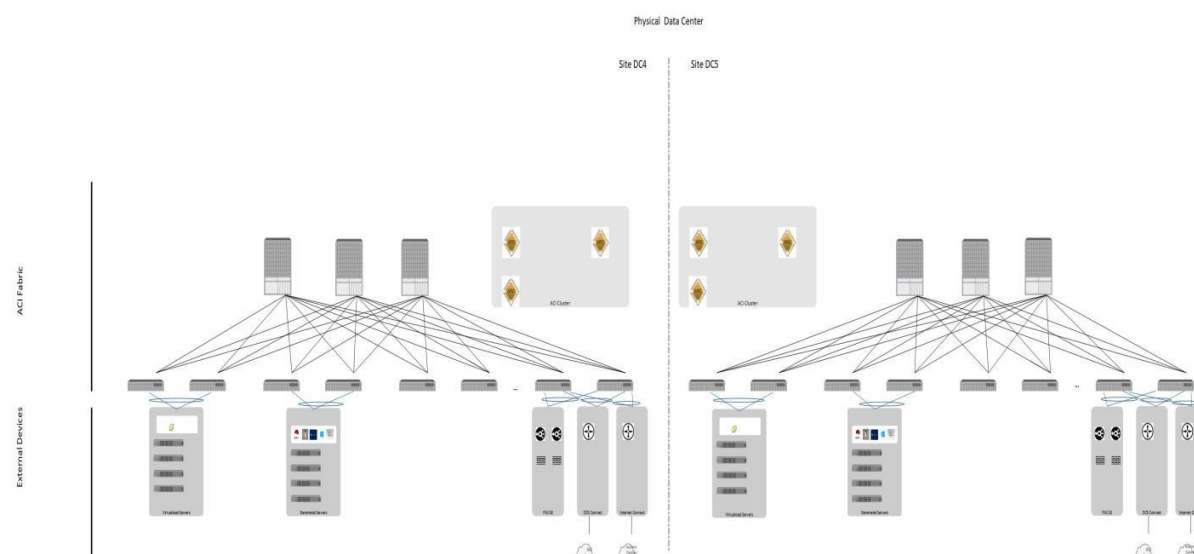
	GREEN is assigned for a rule with low criticality.
	ORANGE is assigned for a rule with medium criticality
	RED is assigned for a rule with high criticality
	BLUE is assigned when the findings or observations are for informational reference

2 ACI Infrastructure and Inventory

2.1 High level design overview

ADP currently has two ACI data centers in Europe. DC4 located in France and DC5 located in Spain are compute data centers, which host servers and applications. ADP has deployed DC4 as an active and DC5 as disaster recovery data centers, which provides backup services in the event of a failure in DC4 data center.

Figure 1: DC4 and DC5 Fabric Network Infrastructure Overview



2.2 Inventory Details of ACI Fabric

Table 1: ADP Hardware DC4

Device Role	PID	Quantity
Fabric Spine	N9K-C9504	3
Fabric Leaf	N9K-C93180YC-FX	8
Fabric Leaf	N9K-C93180YC-EX	52
APIC Controllers	APIC-SERVER-L2	3

Table 2: ADP Hardware DC5

Device Role	PID	Quantity
Fabric Spine	N9K-C9504	3
Fabric Leaf	N9K-C93180YC-EX	40
APIC Controllers	APIC-SERVER-L2	3

Table 3: ADP Software

Platform	Software
APIC	3.2(9h)

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

Fabric Spine	13.2(9h)
Fabric Leaf	13.2(9h)

2.3 End of Life Milestones

Table 4: End of Life Milestones

Product ID	Product Bulletin URL	End-of-Life Announcement Date	End of SW Maintenance Releases	End of New Service Attachment	Last Date of Support
APIC-SERVER-L2	https://www.cisco.com/c/en/us/products/collateral/cloud-systems-management/application-policy-infrastructure-controller-apic/eos-eol-notice-c51-742143.html	03/21/2019	06/19/2021	06/19/2020	06/30/2024
N9K-SUP-A	https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/eos-eol-notice-c51-743848.html	05/29/2020	11/27/2021	11/27/2021	11/30/2025



End-of-Life Announcement Date: The date the document that announces the end-of-sale and end-of-life of a product is distributed to the general public.



End of SW Maintenance Releases: The last date that Cisco Engineering may release any software maintenance releases or bug fixes to the software product. After this date, Cisco Engineering will no longer develop, repair, maintain, or test the product software.



End of New Service Attachment Date: For equipment and software that is not covered by a service-and-support contract, this is the last date to order a new service-and-support contract or add the equipment and/or software to an existing service-and-support contract.



Last Date of Support: The last date to receive service and support for the product. After this date, all support services for the product are unavailable, and the product becomes obsolete. Note: For pricing bundles, the last date of support is specified in the EoL Plan for the individual products, not the bundle.



Best Practice

All devices should be running on same software version.
It is recommended to configure OOB mgmt for all the devices

Observations:

- All Devices are defined with OOB MGMT address
- The current APICs were announced end of life on March 21, 2019. The announcement with the EoL milestones can be found here:
<https://www.cisco.com/c/en/us/products/collateral/cloud-systems-management/application-policy-infrastructure-controller-apic/eos-eol-notice-c51-742143.html>
- Also Cisco announces the end-of-sale and end-of-life dates for the Cisco Nexus 9500 4-Core/4-Thread Supervisor on May 29 2020:
<https://www.cisco.com/c/en/us/products/collateral/switches/nexus-9000-series-switches/eos-eol-notice-c51-743848.html>

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

- Due to the current positioning in the EoL timeline, it's recommended that these be included in upcoming life cycle management activity.
- No End of life milestones were announced on the Switch Models and line cards used by ADP.

2.4 Automation and Orchestration

One of the key advantages of ACI is a single API providing a programable interface to the network infrastructure. ACI presents a hierarchy of configurable managed objects through a REST API, allowing unlimited programmability through a variety of commercial, open source, and custom tooling.

Cisco can assist with ACI automation and orchestration in several ways:

1. Ongoing programmability support through an ACI Business Critical Services contract.
2. Advanced SOW support can be provided for specific, scoped efforts, e.g. building Ansible playbooks, various 3rd party integration, etc.
3. Cisco provides a number of industry leading automation solutions, including [NSO](#) (formerly ConfD), [CloudCenter](#) (formerly CliQr), etc.

3 Detailed Findings

This section describes the detailed findings of the Assessment of ADP Data Center network.

Table 5: Overall Findings and Recommendations

Findings	Criticality	Compliance
Fabric Health and Faults	High	Partially Conform
Pod Policies	Info Only	Conform
Pod Policies ISIS	Low	Conform
Pod Policies BGP RR	Medium	Partially conforms
Pod Policies DataTime	Medium	Partially Conform
Pod Policies CommPolicy	Medium	Partially conforms
Pod Policies SNMP	Medium	Not Conform
VLAN Pools	High	Partially Conform
Domain-VLAN Pool Allocation-AAEP Check	High	Partially Conform
Interface Profile/Policies/Policy Groups	Medium	Conform
VRF Ingress Policy Enforcement	Medium	Conform
Bridge Domains	High	Not Conform
End Point Groups(EPGs)	High	Partially Conform
L3Outs	High	Conform
Contracts & VzAny	Low	Partially Conform
AAA-Fallback Domain	Medium	Conform
Encrypted Backups	Medium	Conform
Fabric Export Policies	High	Conform
Pod Policies COOP Policy	Info Only	Partially Conform
BFD for Fabric Facing Interfaces	Low	Not Conform
Domain Validation	Low	Not Conform
Mis-Cabling Protocol (MCP)	High	Not Conform
Digital Optical Monitoring	Medium	Not Conform
Port Tracking	Medium	Not Conform
Enforce Subnet Check	Medium	Not Conform
Disable Remote End Point Learn	Medium	Conform
IP Aging Policy	Medium	Not Conform
Limit IP Learning to Subnet	Low	Partially Conform
Loop detection	Medium	Not Conform

4 Fabric Policies

Fabric policies apply to the fabric as a whole. Fabric policies are configured once and then modified infrequently after that. Fabric policies are found on the main Fabric tab in the APIC GUI and are concerned with configuration of the fabric itself (for example, IS-IS, management access, MP-BGP, and fabric MTU). Many of the fabric policies should not be modified under normal circumstances, and the recommendation is to not change them.

Fabric policies are grouped into the following categories and configure interfaces that connect spine and leaf switches:

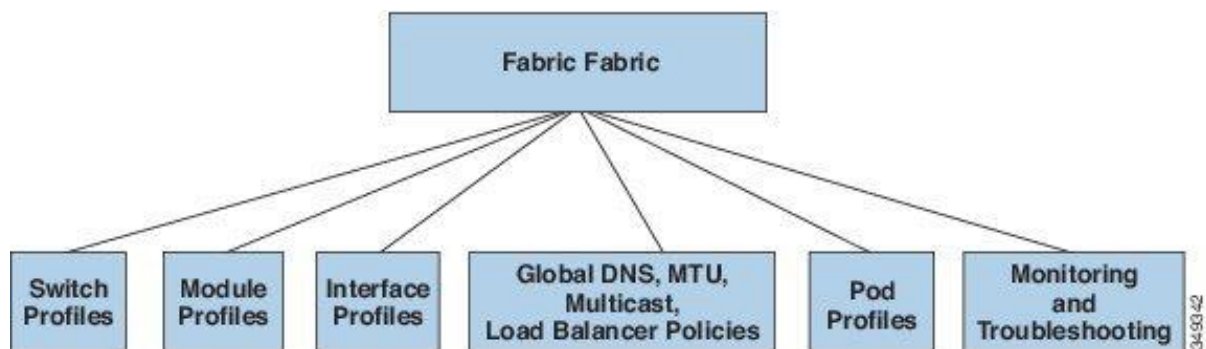


Figure 2: Fabric Policies structure

Pod profiles specify data and time, SNMP, council of oracle protocol (COOP), IS-IS and BGP route reflector policies.

4.1 POD Policies

Criticality/Compliance

Info Only	Conform
-----------	---------

These policies are currently found in the ADP fabric. All pod policies outlined below are in the default state.

Table 6: ADP Pod Policy DC4 and DC5

Pod Policy	Isis	BGP	Date/Time	Communications	COOP	SNMP
DC4-Services-Policy-Grp	default	default	NTP-SYNC-POLICY	default	default	DC4-SNMP-POLICY
DC5-Services-Policy-Grp	default	default	NTP-SYNC-POLICY	default	default	DC5-SNMP-POLICY

4.2 POD Policies IS-IS

Criticality/Compliance

Low	Conform
-----	---------

IS-IS builds a routing table of VTEP ip addresses using the information discovered via LLDP.



Best Practice

For IS-IS POD Policies the Best Practice is to leave configuration as defined by default values and settings.

Table 7: ADP IS-IS policy DC4 and DC5

Name	MTU	Metric for redist. routes	LSP Fast Flood	LSP Gen Init	LSP Max Interval	LSP Gen 2nd Wait	SPF Init Compute Wait	SPF Compute Freq Max Wait	SPF Compute Freq 2nd Wait
default	1492	63	enabled	50	8000	50	50	8000	50

Table 8: IS-IS pod policy descriptions

Property	Description
Name	The IS-IS Domain policy name.
ISIS MTU	The IS-IS Domain policy LSP MTU. The MTU is from 128 to 4352. The default is 1492.
ISIS metric for redistributed routes	The IS-IS metric that is used for all imported routes into IS-IS. The values available are from 1-63. The default value is 63.
LSP Fast Flood Mode	<p>The IS-IS Fast-Flooding of LSPs Using the fast-flood Command feature improves IS-IS convergence time when new link-state packets (LSPs) are generated in the network and shortest path first (SPF) is triggered by the new LSPs.</p> <ul style="list-style-type: none">• Enabled• Disabled <p>The default is Enabled.</p>
LSP Generate Initial wait interval	The LSP generation initial wait interval. This is used in the LSP generation interval for the LSP MTU. The default is 50.
LSP generation maximum wait interval	The LSP generation maximum wait interval. This is used in the LSP generation interval for the LSP MTU. The default is 8000.
LSP generation second wait interval	The LSP generation second wait interval. This is used in the LSP generation interval for the LSP MTU. The default is 50.

SPF computation frequency initial wait interval	The SPF computation frequency initial wait interval. This is used in the SPF computations for the LSP MTU. The default is 50.
SPF computation frequency maximum wait interval	The SPF computation frequency maximum wait interval. This is used in the SPF computations for the LSP MTU. The default is 8000.
SPF computation frequency second wait interval	The SPF computation frequency second wait interval. This is used in the SPF computations for the LSP MTU. The default is 50.

Observations:

ADP fabric IS-IS policy is configured per the Best Practices, that is to refer to default settings.

Recommendation:

Reduce the ISIS metric to 32. This change bounces the multi-pod ISIS routes, so may impact multi-pod traffic. The metric can be configured at **Fabric > Fabric Policies > Policies > Pod > ISIS Policy default > ISIS metric for redistributed routes**.

Note that this is recommended even though ADP currently has no multi-pod configured in DC4 and DC5. This will ensure that if multi-pod is considered in the future, this issue will already be addressed.

4.3 Pod Policies BGP Route-Reflector

Criticality/Compliance

Medium	Partially conforms
--------	--------------------

The ACI fabric route reflectors use multiprotocol BGP (MP-BGP) to distribute external routes within the fabric. To enable route reflectors in the ACI fabric, the fabric administrator must select the spine switches that will be the route reflectors, and provide the autonomous system (AS) number. For redundancy purposes, more than one spine is configured as a router reflector node (one primary and one secondary reflector).



Best Practice

For BGP Route-Reflector POD Policies, the Best Practice is to configure at least two Spine Switches as RR for redundancy purposes.

Table 9: ADP BGP RR current configuration DC4

BGP	ASN	RRs
default	65000	101, 103

Table 10: ADP BGP RR current configuration DC5

BGP	ASN	RRs
default	65001	101, 103

Observations:

No Spine with Even ID is defined as Route Reflector. ADP is advised to add node 102 as a route reflector in both DC4 and DC5.



If you have maintenance groups separated by odd and even, you will lose external prefixes when Odd Maintenance window spines reboot

4.4 Pod Policies Date/Time (NTP)**Criticality/Compliance**

Medium	Partially conforms
--------	--------------------

Within ACI fabric, time synchronization is a crucial capability upon which many of the monitoring, operational, and troubleshooting tasks depends. Clock synchronization is important for proper analysis of traffic flows as well as for correlating debug and fault timestamps across multiple fabric nodes.

An offset present on one or more devices can hamper the ability to properly diagnose and resolve many common operational issues. In addition, clock synchronization allows for the full utilization of the atomic counter capability that is built in to the ACI, upon which the application health scores depend.

**Best Practice**

For NTP Pod Policies, Best Practice is to have multiple NTP sources configured for redundancy, fabric should be synchronized, and authentication should be enabled.

Table 11: Current NTP configurations DC4

Date/Time	State	auth	Pol Desc	Server	Min Poll	Max Poll	Key ID	Preferred	EPG
NTP-SYNC-POLICY	enabled	disabled		7.128.16.2	4	6	0	no	oob-default
NTP-SYNC-POLICY	enabled	disabled		7.128.16.1	4	6	0	yes	oob-default

Table 12: Current NTP configurations DC5

Date/Time	State	auth	Pol Desc	Server	Min Poll	Max Poll	Key ID	Preferred	EPG
NTP-SYNC-POLICY	enabled	disabled		7.130.16.1	4	6	0	yes	oob-default
NTP-SYNC-POLICY	enabled	disabled		7.130.16.2	4	6	0	no	oob-default

Observations:

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

- Cisco CX has found that ADP has configured the Date and Time policy and servers are configured for redundancy which is a good practice.
- Authentication is seen disabled in the policy; it is recommended to enable them for security purpose.

4.5 Pod Policies CommPolicy/Management Policy

Criticality/Compliance

Medium

Partially conforms



Best Practice:

For enhanced security TELNET and HTTP access should be disabled, only SSH and HTTPS should be used on the fabric.

Table 13: Current DC4 policy configurations

Comm Policy	Telnet/Port	SSH/Port	Web SSH/Port	HTTP/Port	HTTP Redirect	HTTPS/Port	SSL Protocols	Ciphers	SSL Key Ring
default	disabled /23	Enabled /22	Disabled /4200	Disabled /80	disabled	Enabled /443	TLSv1.1,TLSv1.2	EECDH,EECDH+aRSA+SHA256,EECDH+aRSA+SHA384,EECDH+aRSA,EECDH+aRSA+AE SGCM	default

Table 14: Current DC5 policy configurations

Comm Policy	Telnet/Port	SSH/Port	Web SSH/Port	HTTP/Port	HTTP Redirect	HTTPS/Port	SSL Protocols	Ciphers	SSL Key Ring
Telnet Policy	enabled /23	enabled/22	disabled/4200	disabled/80	disabled	enabled/443	TLSv1.1,TLSv1.2	EDH+aRSA,EECDH+aRSA+SHA384,EECDH+aRSA+SHA256,EECDH,EECDH+aRSA+AES GCM	default
default	disabled /23	enabled/22	disabled/4200	disabled/80	disabled	enabled/443	TLSv1.1,TLSv1.2	EECDH,EECDH+aRSA+SHA384,EECDH+aRSA+SHA256,EECDH+aRSA+AES GCM	default

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

Observations:

ADP partially conform to Cisco's ACI BPs for CommPolicy/Management Policy. It is not advised to use telnet for remote connectivity and management.

4.6 Pod Policies SNMP

Criticality/Compliance

Medium

Not Conform

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 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.

SNMP support in ACI is as follows:

- SNMP read queries (Get, Next, Bulk, Walk) are supported by leaf and spine switches and by APIC
- SNMP write commands (Set) are not supported by leaf and spine switches or by APIC
- SNMP traps (v1, v2c, and v3) are supported by leaf and spine switches and by APIC
- SNMPv3 is supported by leaf and spine switches and by APIC



Best Practice

SNMP Policy should restrict SNMP Clients, when possible SNMPv3 should be used.

Table 15: ADP SNMP policy configurations DC4

Pod SNMP	State	ver
SNMP_WEP113-POLICY	disabled	v2c
DC4-SNMP-POLICY	enabled	v2c

Table 16: ADP SNMP policy configurations DC5

Pod SNMP	State	ver
DC5-SNMP-POLICY	enabled	v2c

Observations:

Within the Pod Policies, ADP DC pods has been configured with SNMP v2c. It is advisable to use SNMP v3 instead of SNMPv2c, since SNMPv3 has enhanced security features.

5 Access Policies

Access policies are responsible for the physical configuration of the ports to which devices are attached. Access policies are configured in a hierarchical manner and built to be modular so isolated details can be changed without having to modify or re-create the entire policy.

Once access policies are defined, the configuration stays dormant until a tenant policy is triggered. When the tenant policy is applied, the port is configured with the physical characteristics defined in the access policy.

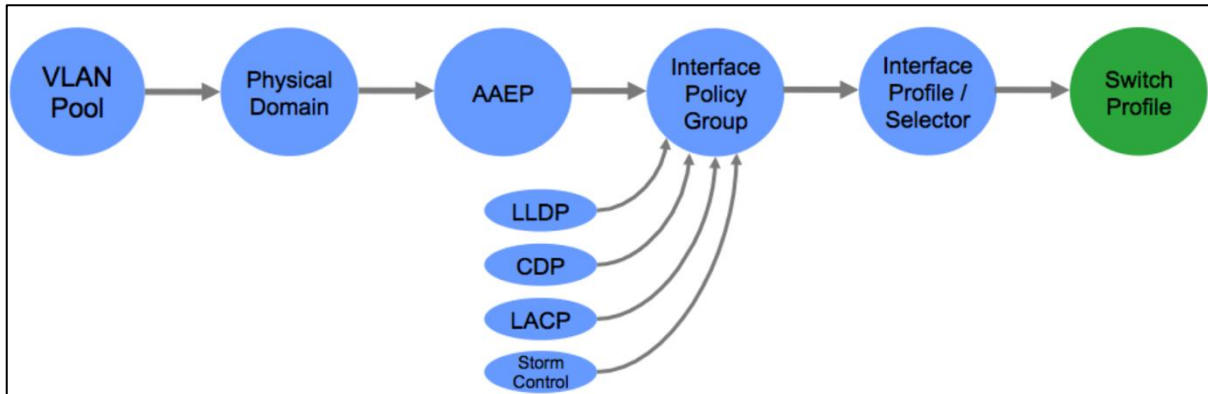


Figure 3: The Hierarchical access policy relationships

5.1 VLAN Pools

Criticality/Compliance

High	Partially Conform
------	-------------------

A VLAN pool is used to define a range of VLAN numbers that will ultimately be applied on specific ports on one or more leaf nodes.

A pool is a shared resource and can be consumed by multiple domains such as VMM and Layer 4-7 services.

A VLAN pool is a container that is comprised of encapsulation blocks, which contain the actual VLAN definitions.

Guidelines for VLAN pools are as follows:

- VLAN pools with an allocation mode of Dynamic are typically used for VMM integration deployments. VMM integration generally does not require explicit VLAN assignment, so a dynamic pool allows the system to pull free resources as needed.
- VLAN pools with an allocation mode Static are typical for the majority of other deployment scenarios including static paths, L2Out and L3Out out definitions.
- A dynamic VLAN pool can have a static encapsulation block defined within it. This is generally only done for the specific case of utilizing the "pre-provision" resolution immediacy.

- A static VLAN pool **cannot** have a dynamic encapsulation block. This will be rejected by the Application Policy Infrastructure Controller (APIC), as there are no features that utilize this configuration.
- It is a common practice to divide VLAN pools into functional groups, as shown in the following table.

Table 17: VLAN Pool Example

VLAN Range	Type	Use
1000 - 1100	Static	Bare-metal hosts
1101 - 1200	Static	Firewalls
1201 - 1300	Static	External WAN routers
1301 - 1400	Dynamic	Virtual machines



VLAN pools containing overlapping encapsulation block definitions should *not* be associated to the same AAEP (and subsequently the same leaf nodes). This can cause issues with BPDU forwarding through the fabric if the domains associated to an EPG have overlapped VLAN block definitions. When possible divide VLAN pools into functional groups.

Table 18: DC4 VLAN pool allocation

Name	From vlan	To vlan	Allocation mode
BCKP-VLAN-POOL	vlan-1750	vlan-1751	static
	vlan-1752	vlan-1753	inherit
	vlan-1897	vlan-1897	static
	vlan-2490	vlan-2490	inherit
DC4T8-VLAN-POOL	vlan-1000	vlan-1001	static
	vlan-1100	vlan-1100	static
	vlan-1500	vlan-1500	static
	vlan-1600	vlan-1600	static
	vlan-1602	vlan-1603	static
	vlan-1625	vlan-1625	static
	vlan-1650	vlan-1650	static
	vlan-1652	vlan-1652	static
	vlan-1751	vlan-1751	static
	vlan-2000	vlan-2001	static
	vlan-2004	vlan-2004	static
	vlan-2250	vlan-2251	static
HBE1-VLAN-POOL	vlan-1000	vlan-1099	static
HBE2-VLAN-POOL	vlan-1100	vlan-1199	static
HBE3-VLAN-POOL	vlan-1200	vlan-1300	static
HWT1-VLAN-POOL	vlan-2000	vlan-2139	static
	vlan-2150	vlan-2150	static

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

	vlan-2151	vlan-2246	static
	vlan-2247	vlan-2249	static
IBE1-VLAN-POLL	vlan-1600	vlan-1649	static
IBE2-VLAN-POOL	vlan-1556	vlan-1556	static
	vlan-1650	vlan-1700	inherit
	vlan-2449	vlan-2449	static
IWT1-VLAN-POOL	vlan-2250	vlan-2389	static
	vlan-2390	vlan-2393	static
	vlan-2449	vlan-2449	inherit
	vlan-2465	vlan-2465	static
	vlan-2466	vlan-2466	static
	vlan-2467	vlan-2469	static
	vlan-2470	vlan-2470	static
	vlan-2471	vlan-2473	static
	vlan-2474	vlan-2474	static
	vlan-2475	vlan-2477	static
	vlan-2479	vlan-2481	static
	vlan-2482	vlan-2482	static
	vlan-2483	vlan-2485	static
	vlan-2487	vlan-2489	static
	vlan-2491	vlan-2493	static
	vlan-2494	vlan-2496	static
	vlan-2497	vlan-2499	static
L3OUT-FW-ACI_POOL	vlan-1896	vlan-1896	inherit
	vlan-1897	vlan-1897	static
	vlan-1898	vlan-1898	static
	vlan-1899	vlan-1899	static
L3OUT-VRF-STORAGE_POOL	vlan-1582	vlan-1585	inherit
SHAR-VLAN-POOL	vlan-1499	vlan-1549	static
	vlan-1558	vlan-1558	static
VCE0411_VM_POOL	vlan-1000	vlan-1099	static
	vlan-1100	vlan-1199	static
	vlan-1200	vlan-1300	static
	vlan-1499	vlan-1505	static
	vlan-1538	vlan-1549	static
	vlan-1600	vlan-1649	static
	vlan-1650	vlan-1699	static
	vlan-1700	vlan-1719	inherit
	vlan-1750	vlan-1751	static
	vlan-1897	vlan-1897	static
	vlan-1898	vlan-1898	static
	vlan-1899	vlan-1899	static
	vlan-2000	vlan-2139	static

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

	vlan-2247	vlan-2249	static
	vlan-2250	vlan-2255	static
	vlan-2387	vlan-2389	static
	vlan-2497	vlan-2499	static

Table 19: DC5 VLAN pool allocation

Name	From vlan	To vlan	Allocation mode
IBE1-VLAN-POOL	vlan-3100	vlan-3149	inherit
POC_VCE0511_VM_POOL	vlan-3191	vlan-3191	static
	vlan-3194	vlan-3194	static
	vlan-3154	vlan-3154	static
	vlan-3188	vlan-3188	static
	vlan-3178	vlan-3178	static
	vlan-3152	vlan-3152	static
	vlan-3999	vlan-3999	static
	vlan-2700	vlan-2800	static
	vlan-3399	vlan-3399	static
	vlan-2500	vlan-2599	static
	vlan-3250	vlan-3250	static
	vlan-3179	vlan-3179	static
	vlan-7	vlan-7	static
	vlan-3190	vlan-3190	static
	vlan-3175	vlan-3175	static
	vlan-3398	vlan-3398	static
	vlan-3161	vlan-3161	static
	vlan-3177	vlan-3177	static
	vlan-3500	vlan-3639	static
	vlan-3150	vlan-3150	static
	vlan-3200	vlan-3219	inherit
	vlan-2999	vlan-3005	static
	vlan-3748	vlan-3755	static
	vlan-3195	vlan-3195	static
	vlan-3193	vlan-3193	static
	vlan-3038	vlan-3049	static
	vlan-3155	vlan-3155	static
	vlan-3192	vlan-3192	static
	vlan-3189	vlan-3189	static
	vlan-3158	vlan-3158	static
	vlan-3176	vlan-3176	static
	vlan-2600	vlan-2699	static
	vlan-3197	vlan-3197	static
	vlan-3251	vlan-3251	static
	vlan-3747	vlan-3747	static

	vlan-3397	vlan-3397	static
	vlan-3885	vlan-3890	static
	vlan-3151	vlan-3151	static
	vlan-3153	vlan-3153	static
	vlan-3100	vlan-3149	static
	vlan-3199	vlan-3199	static
IWT1-VLAN-POOL	vlan-3652	vlan-3655	inherit
	vlan-3748	vlan-3999	inherit
BCKP-VLAN-POOL	vlan-3397	vlan-3397	static
	vlan-3250	vlan-3251	inherit
HBE2-VLAN-POOL	vlan-2600	vlan-2699	static
HBE3-VLAN-POOL	vlan-2700	vlan-2800	static
HBE1-VLAN-POOL	vlan-2500	vlan-2599	static
L3OUT-FW-ACI_POOL	vlan-3396	vlan-3396	inherit
HWT1-VLAN-POO	vlan-3640	vlan-3746	static
	vlan-3747	vlan-3749	static
	vlan-3500	vlan-3639	static
L3OUT-VRF-STORAGE_POOL	vlan-3082	vlan-3085	inherit
IBE2-VLAN-POOL	vlan-3200	vlan-3200	inherit
	vlan-3150	vlan-3199	inherit
	vlan-3399	vlan-3399	inherit
	vlan-3397	vlan-3398	inherit
SHAR-VLAN-POOL	vlan-2999	vlan-3049	inherit
Test_clone_01-POOL	vlan-4001	vlan-4002	inherit
POC-IBE2-VM-POOL	vlan-7	vlan-9	static
	vlan-3152	vlan-3152	static
	vlan-3199	vlan-3199	static
	vlan-3151	vlan-3151	static
	vlan-3150	vlan-3150	static
	vlan-3350	vlan-3399	inherit
	vlan-3197	vlan-3197	static
	vlan-3198	vlan-3198	static

DC4 Observations:

- Number of static pool: 12
- Number of dynamic pool: 1

Table 20: VLAN pools with overlapping ranges DC4

Vlan pool 1	Vlan pool 2	Vlans
BCKP-VLAN-POOL	DC4T8-VLAN-POOL	1751
BCKP-VLAN-POOL	L3OUT-FW-ACI_POOL, VCE0411-DOM	1897
DC4T8-VLAN-POOL	HBE1-VLAN-POOL	1000-1001
DC4T8-VLAN-POOL	HBE2-VLAN-POOL	1100
DC4T8-VLAN-POOL	HWT1-VLAN-POOL	2000,2001,2004

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

DC4T8-VLAN-POOL	IBE1-VLAN-POLL	1600,1602-1603,1625
DC4T8-VLAN-POOL	IBE2-VLAN-POOL	1650,1652
DC4T8-VLAN-POOL	IWT1-VLAN-POOL	2250, 2251
DC4T8-VLAN-POOL	SHAR-VLAN-POOL	1500
IBE2-VLAN-POOL	IWT1-VLAN-POOL	2449
VCE0411_VM_POOL	DC4T8-VLAN-POOL	1000-1001,1500, 1600,1602-1603,1625,1650,1652,1751,2000-2001,2004,2250-2251
VCE0411_VM_POOL	L3OUT-FW-ACI_POOL,	1897-1899,
VCE0411_VM_POOL	HBE1-VLAN-POOL	1000-1099
VCE0411_VM_POOL	HBE2-VLAN-POOL	1100-1199
VCE0411_VM_POOL	HBE3-VLAN-POOL	1200-1300
VCE0411_VM_POOL	SHAR-VLAN-POOL	1499-1505, 1538-1549
VCE0411_VM_POOL	IBE1-VLAN-POLL	1600-1649
VCE0411_VM_POOL	IBE2-VLAN-POLL	1650-1700
VCE0411_VM_POOL	BCKP-VLAN-POOL	1750-1751, 1897
VCE0411_VM_POOL	HWT1-VLAN-POOL	2000-2139, 2247-2249
VCE0411_VM_POOL	IWT1-VLAN-POOL	2250-2255,2387-2389,2497-2499

DC5 Observations:

- Number of static pool: 12
- Number of dynamic pool: 2
- static vlan pool: Test_clone_01-POOL not used
- dynamic vlan pool: POC-IBE2-VM-POOL not used – assigned to POC-IBE2-VDS-6-5 domain
- dynamic vlan pool: POC_VCE0511_VM_POOL not used, assigned to domains POC-VCE0511-Credentials, VCE0531-DOM

Table 21 - VLAN pools with overlapping ranges DC5

Vlan pool 1	Vlan pool 2	Vlans
BCKP-VLAN-POOL	IBE2-VLAN-POOL	3397
HWT1-VLAN-POOL	IWT1-VLAN-POOL	1652-1655, 3748,3749
POC_VCE0511_VM_POOL	POC-IBE2-VM-POOL	7,3150-3152,3197,3199,3397-3399
POC_VCE0511_VM_POOL	HBE1-VLAN-POOL	2500-2599
POC_VCE0511_VM_POOL	HBE2-VLAN-POOL	2600-2699
POC_VCE0511_VM_POOL	HBE3-VLAN-POOL	2700-2800
POC_VCE0511_VM_POOL	SHAR-VLAN-POOL	2999-3005, 3038-3049
POC_VCE0511_VM_POOL	IBE1-VLAN-POOL	3100-3149
POC_VCE0511_VM_POOL	IBE2-VLAN-POOL	3150-3155,3158,3161,3175-3179, 3188-3195, 3197,3199,3200,3397-3399
POC_VCE0511_VM_POOL	BCKP-VLAN-POOL	3250-3251, 3397
POC_VCE0511_VM_POOL	HWT1-VLAN-POOL	3500-3639, 3747-3749
POC_VCE0511_VM_POOL	IWT1-VLAN-POOL	3748-3755, 3885-3890, 3999
POC-IBE2-VM-POOL	IBE2-VLAN-POOL	3150-3152,3197-3199,3397-3399

Recommendations:

- ADP partially conform to VLAN Pool Best Practices
- As can be seen in the VLAN Pool Check findings, there are multiple VLAN Pools used by multiple domains.
- There are VLAN pools created but not being used i.e there is no mapping of these VLANS pools to any domain. ADP to check if this is really needed.
- As can be seen in the VLAN Pool Check findings, there are multiple VLAN Pools with overlapping encapsulation blocks. Some of these VLAN Pools with overlapping ranges are associated to same AAEP (as shown in below findings). This is recommended to avoid a misconfiguration where multiple domains with overlapping VLANs are associated to the same EPG. If different domains with the same VLAN are attached to the same EPG, BD-wide BPDU flooding will be nondeterministic, which can result in bridging loops, high CPU, and intermittent traffic.

5.2 Domain-VLAN Pool -AAEP Check

Criticality/Compliance

High	Partially Conform
------	-------------------

A **domain** is used to define the scope of VLANs in the Cisco ACI fabric: in other words, where and how a VLAN pool will be used.

There are a number of domain types: physical, virtual (VMM domains), external Layer 2, and external Layer 3.

Physical domain profile defines the physical resources (ports and port-channels) and encapsulation resources (VLAN pools) that should be used for tenant or endpoint groups associated with this domain.

Guidelines for physical and external domain policies are as follows:

- Build one physical domain per tenant for bare metal servers or servers without hypervisor integration requiring similar treatment.
- Build one external routed/bridged domain per tenant for external connectivity.
- For VMM domains, if both DVS and AVS is in use, create a separate VMM domain to support each environment.
- For large deployments where domains (physical/VMM/etc) need to be leveraged across multiple tenants, a single physical domain or VMM domain can be created and associated with all leaf ports where services are connected.

An **Attachable Access Entity Profile (AAEP)** represents a group of external entities with similar infrastructure policy requirements. The infrastructure policies consist of physical interface policies, such as Cisco Discovery Protocol (CDP), Link Layer Discovery Protocol (LLDP), or Link Aggregation Control Protocol (LACP).

An AAEP is required to deploy VLAN pools on leaf switches. Encapsulation pools (and associated VLAN) are reusable across leaf switches. An AAEP implicitly provides the scope of the VLAN pool to the physical infrastructure.



Best Practices:

- It is a recommendation to keep a simple configuration to reduce complexity, when possible have a 1:1 mapping between a VLAN pool and a domain.
- Avoid having AAEPs with overlapping VLANs domains.
- Although only one AAEP can be used for simplicity, it is recommended to have more than one to be able to limit the scope of VLANs across the Fabric Infrastructure.
- Cisco recommends preventing VMM domains from utilizing the same AAEP that is also associated to Layer-3 interfaces.
- Objects not being used or needed should be removed from the config to avoid confusion and complexity.

Table 22: DC4 Domain-Pool allocation

Domain	VLAN Pool
BCKP-L2-OUT-DOM	BCKP-VLAN-POOL
DC4T8-DOM	DC4T8-VLAN-POOL
DD1314-OUT-DOM	BCKP-VLAN-POOL
HBE1-L2-OUT-DOM	HBE1-VLAN-POOL
HBE2-L2-OUT-DOM	HBE2-VLAN-POOL
HBE3-L2-OUT-DOM	HBE3-VLAN-POOL
HWT1-L2-OUT-DOM	HWT1-VLAN-POOL
IBE1-L2-OUT-DOM	IBE1-VLAN-POLL
INS251-DOM	IBE2-VLAN-POOL
IWT1-L2-OUT-DOM	IWT1-VLAN-POOL
L2OUT-IBE2-DOM	IBE2-VLAN-POOL
L3OUT-FW_ACI-DOM	L3OUT-FW-ACI_POOL
L3OUT-VRF-STORAGE_DOM	L3OUT-VRF-STORAGE_POOL
NA077_BCKP-L2-OUT-DOM	BCKP-VLAN-POOL
NA077_IBE1-L2-OUT-DOM	IBE1-VLAN-POLL
NA077_IBE2-L2-OUT-DOM	IBE2-VLAN-POOL
NA077_IWT1-L2-OUT-DOM	IWT1-VLAN-POOL
NA077_SHA-L2-OUT-DOM	SHAR-VLAN-POOL
NA078_BCKP-L2-OUT-DOM	BCKP-VLAN-POOL
NA078_IBE1-L2-OUT-DOM	IBE1-VLAN-POLL

NA078_IBE2-L2-OUT-DOM	IBE2-VLAN-POOL
NA078_IWT1-L2-OUT-DOM	IWT1-VLAN-POOL
NA078_SHA-L2-OUT-DOM	SHAR-VLAN-POOL
SAN-DOM	BCKP-VLAN-POOL
SHAR-L2-OUT-DOM	SHAR-VLAN-POOL
V169HBE3-L2-OUT-DOM	HBE3-VLAN-POOL
VCE0411_DA-DOM	VCE0411_VM_POOL
VCE0411-DOM	VCE0411_VM_POOL
VCE0421-DOM	VCE0411_VM_POOL

DC4 Observations:

- num of Physical Domain: 25
- All L3 Domain: 2
- All L2 Domain: 0
- All VMM domain: 3
- All AEP: 32

Physical Domain: phys has no vlan pool
Physical Domain: phys not used by AEP

AEP: DC4T8-AEP not used by interface policy group
AEP: VCE0411_DA-AEP not used by interface policy group
AEP: INS251-EAP not used by interface policy group
AEP: VCE0421-AEP not used by interface policy group
AEP: NA115_116-AEP not used by interface policy group
AEP: default has no domain
AEP: default not used by interface policy group
AEP: SAN-AE not used by interface policy group
AEP: V169HBE3-AEP not used by interface policy group

Table 23 - VLAN pools with overlapping ranges in the same AAEP DC4

AAEP	Domain 1	Domain 2	note
NA115_116-AEP	IWT1-L2-OUT-DOM	L2OUT-IBE2-DOM	Vlan 2449
ALL_PODS-AEP	IWT1-L2-OUT-DOM	L2OUT-IBE2-DOM	Vlan 2449
L2OUT-IBE2-AEP	INS251-DOM	L2OUT-IBE2-DOM	The same vlan pool

DC5 Observations:

- number of Physical Domain: 20
- All L3 Domain: 2
- All L2 Domain: 0
- All VMM domain: 3

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

- All AEP: 29

AEP: default has no domain

AEP: default not used by interface policy group

AEP: VCE0531-AEP not used by interface policy group

Table 24 - VLAN pools with overlapping ranges in the same AAEP DC5

AAEP	Domain 1	Domain 2	notes
ALL_PODS-AEP	BCKP-L2-OUT-DOM/ BCKP-VLAN-POOL	POC-IBE2-L2-OUT-DOM / IBE2-VLAN-POOL	Vlan 3397
ALL_PODS-AEP	BCKP-L2-OUT-DOM/ BCKP-VLAN-POOL	IWT1-L2-OUT-DOM / IWT1-VLAN-POOL	Vlan 3397
ALL_PODS-AEP	HWT1-L2-OUT-DOM / HWT1-VLAN-POOL	IWT1-L2-OUT-DOM / IWT1-VLAN-POOL	3652-3655,3748-3749
NA615_616-AEP	BCKP-L2-OUT-DOM	POC-IBE2-L2-OUT-DOM / IBE2-VLAN-POOL	Vlan 3397
NA615_616-AEP	BCKP-L2-OUT-DOM	IWT1-L2-OUT-DOM / IWT1-VLAN-POOL	Vlan 3397

Recommendations:

ADP partially comply with Cisco Best Practices because:

- There are AAEPs with overlapping VLAN domains. Review the overlapping VLAN pools identified in the previous section to ensure there's no potential for these to be deployed in the same EPG.
- There is no 1:1 mapping between VLAN pools and a domain.
- Validate if existing non-linked objects in the config are still needed, delete non-needed ones to keep simplicity and avoid confusion.

5.3 Interface Profiles/Policies/Policy Groups

Criticality/Compliance

Medium	Conform
--------	---------

The interface policies are a common set of interface level policies such as LLDP, CDP, LACP, interface speed settings. This is where port behavior is configured.



Best Practices:

- Do not use the default setting for interface policies, if possible.
- Re-use policies whenever possible. For example, create new separate interface policies for LACP active, passive, and mac-pinning; for 1-GE port speed and 10-GE port speed; and for CDP and LLDP policies.
- When naming interface policies, use names that clearly describe the setting. For example, a policy that enables LACP in mode active could be called "LACP-Active". There are many default policies out of the box. However, it can be hard to remember what all the defaults are, which is why policies should be clearly named to avoid making a mistake when adding new devices to the fabric.

DC4 Observations:

Number of interface profile: 56

Num of VPC policy group 822

Num of unused interface policy groups: 3 F3240C-TIN1-04G-SHARED-C1U1-DB-INT-POLICY-GRP, TESTOLIVIER-INT-POLICY-GRP, F3240C-TIN1-04G-SHARED-C1U2-DB-INT-POLICY-GRP

No re-used VPC policy group (one vpc policy per VPC)

Num of PC policy group 0

No unused PC policy group

Num of access policy group 14

No unused Access port policy group

DC5 Observations:

Num of interface profile: 36

Num of VPC policy group 480

Num of unused interface policy groups: F3240C-TIN1-04G-SHARED-C1U1-DB-INT-POLICY-GRP, V2400-BCN1-01S-CORE-C1U1-IBE1-INT-POLICY-GRP, DC5PRHCIHST0113-INT-POLICY-GRP, MDOM-INT-POLICY-GRP-02, V2400-BCN1-01S-CORE-C1U2-IBE1-INT-POLICY-GRP, F3240C-BCN1-01S-SHARED-C1U2-DB-INT-POLICY-GRP, DC5PRHCIHST0114-INT-POLICY-GRP.

No re-used VPC policy group (one vpc policy per VPC)

Num of PC policy group 0

No unused PC policy group

Num of access policy group 18

No unused Access port policy group

Current Interface Policy Groups in ADP



Recommendations

- ADP has used Interface profiles as per the best practices. However, there are some unused interface policy groups, as highlighted above which needs further investigation.
- Validate objects present in the configuration that may no longer be needed. Keep only needed objects and relationships in place for clarity and consistency.

6 Tenant Policies – Networking

6.1 Tenant

DC4 Observations:

num of Tenant: 5
num of VRF: 9
num of BD: 334
num of App Profile: 6
num of EPG: 332
num of contract: 10
num of imported contract: 3
num of filters: 52
No duplicate BD name between common tenant and user tenant
No duplicate Vrf name between common tenant and user tenant
No duplicate filter name between common tenant and user tenant
No duplicate Contract name between common tenant and user tenant

DC5 Observations:

num of Tenant: 6
num of VRF: 11
num of BD: 331
num of App Profile: 8
num of EPG: 328
num of contract: 13
num of imported contract: 3
num of filters: 60
No duplicate BD name between common tenant and user tenant
No duplicate Vrf name between common tenant and user tenant
Filter name defined in both common tenant and user tenant (1): ['icmp', 'LEGACY']
No duplicate Contract name between common tenant and user tenant

6.2 VRF Ingress Policy Enforcement

Criticality/Compliance

Medium

Conform

When a policy is enforced between endpoint groups, it can be enforced on the ingress leaf switch or on the egress leaf switch for internal endpoint groups. On ACI releases prior to 1.2(1), the policy for traffic from an internal endpoint group to an external endpoint group (L3Out endpoint group) is enforced on the egress leaf switch where the L3Out is deployed.

A common network design has a large number leaf switches connecting to the compute environment, but only a pair of border leaf switches. Because internal to external policy enforcement is done on the egress switch (border leaf), this can create a resource (TCAM) bottleneck on the border leaf switch.

The ingress policy enforcement feature is a configurable option to enable ingress policy enforcement for internal to external communications. With ingress policy enforcement, the destination class lookup for the destination prefix can be done on the ingress leaf switch. This distributes the enforcement of the policy across more switches since there are typically more compute leaf switches than border leaf switches, reducing the likelihood of a bottleneck at the border leaf switches.

Use the ingress policy enforcement when there are a large number of prefixes and external endpoint groups configured at the border leaf switches. Ingress policy enforcement is implemented at the VRF level and applies to all L3Outs that are configured within that VRF.

This feature was introduced in release 1.2(1) and is the default setting for VRFs created in the 1.2(1) release and later. Any VRFs created prior to the release 1.2(1) are set to egress policy enforcement by default and must be manually changed to use ingress policy enforcement.



Best Practices:

1. Recommendation is to use ingress policy enforcement whenever possible. For ingress enforce, please disable remote learning on BL.
2. Ingress policy enforcement does not apply to the following cases:
 - a. Transit routing; the rules are already applied at ingress for transit routing
 - b. When a vzAny contract is used
 - c. When a taboo contract is used

DC4 VRF CHECK:

num of vrf: 9, enforced: 9, unenforced: 0

num of ingress enforced: 9, egress enforced: 0

DC5 VRF CHECK:

num of vrf: 11 enforced: 11, unenforced: 0

num of ingress enforced: 11, egress enforced: 0

Observations:

For all the VRFs the Policy Control Enforcement Direction is set as Ingress and is confirm with Cisco Best Practices Recommendation.

- BD "[default]" have empty vrf name. Cisco CX recommends to check this

Table 25: DC4 – CTX

Tenant Name	CTX Name	Mcast	Enforcement Direction	Enforcement
common	default	permit	ingress	Yes
common	copy	permit	ingress	Yes
infra	overlay-1	permit	ingress	Yes
infra	ave-ctrl	permit	ingress	Yes
LEGACY	VRF-STORAGE	permit	ingress	Yes
LEGACY	VRF-LEGACY	permit	ingress	Yes
MDOM	VRF-MDOM	permit	ingress	Yes
mgmt	oob	permit	ingress	Yes
mgmt	inb	permit	ingress	Yes

Table 26: DC5 – CTX

Tenant Name	CTX Name	Mcast	Enforcement Direction	Enforcement
common	default	permit	ingress	Yes
common	POC-IBE2-VRF	permit	ingress	Yes
common	copy	permit	ingress	Yes
infra	ave-ctrl	permit	ingress	Yes
infra	overlay-1	permit	ingress	Yes
LEGACY	VRF-STORAGE	permit	ingress	Yes
LEGACY	VRF-LEGACY	permit	ingress	Yes
MDOM	VRF-MDOM	permit	ingress	Yes
mgmt	oob	permit	ingress	Yes

6.3 Bridge Domains

Criticality/Compliance

High	Not Conform
------	-------------

Within a private network, one or more bridge domains must be defined. A bridge domain is a Layer 2 forwarding construct within the fabric, used to constrain broadcast and multicast traffic.

Bridge Domains in ACI have a number of configuration options to allow the administrator to tune the operation in various ways. The configuration options can be summarized as follows:

- **L2 Unknown Unicast:** This option can be set to either Flood or Hardware Proxy. If this option is set to Flood, layer 2 unknown unicast traffic will be flooded inside the fabric. If the Hardware Proxy option is set, the fabric mapping database will be queried for L2 unknown unicast traffic. Note that this option does not have any impact on what the mapping database actually learns – the mapping database is always populated for L2 entries regardless of this configuration.
- **ARP Flooding:** If ARP flooding is enabled, ARP traffic will be flooded inside the fabric as per regular ARP handling in traditional networks. If this option is disabled, the fabric will attempt to unicast the ARP traffic to the destination. Note that this option only applies if Unicast Routing is enabled on the Bridge Domain. If Unicast Routing is disabled, ARP traffic is always flooded, regardless of the status of the ARP Flooding option.
- **Unicast Routing:** This option enables learning of IP addresses in the fabric mapping database. Note that MAC addresses are always learned by the mapping database. Use of the Unicast Routing option is generally recommended – even when only L2 traffic is present – to assist troubleshooting (e.g. Traceroute tool) and allow advanced functionality such as dynamic end point attachment with L4-7 services. Note also that

enabling unicast routing helps to reduce flooding in a bridge domain as disabling ARP flooding depends upon it. When considering unicast routing, consideration must be given to the desired topology. If an external device (such as a firewall) is acting as the default gateway and routing between two bridge domains, enabling unicast routing might cause traffic to be routed on the fabric and bypass the external device.

- **Enforce Subnet Check for IP Learning:** If this option is checked, the fabric will not learn IP addresses from a subnet other than the one configured on the Bridge Domain. For example, if a Bridge Domain is configured with a subnet address of 10.1.1.0/24, the fabric would not learn the IP address of an end point using an address outside of this range (e.g. 20.1.1.1/24). Note that this feature does not affect the data path; in other words, it will not drop packets coming from the “wrong” subnet. The feature simply prevents the fabric from learning end point information in this scenario.

Given the above options, it may not be immediately obvious how a bridge domain should be configured. The following sections attempt to explain when and why particular options should be selected.

Bridge Domain Scenario 1: IP, Routed Traffic

Scenario 1: My bridge domain will contain IP based, routed traffic. I won't be connecting firewalls / load balancers to this BD and it will not contain clusters or similar which may rely on “floating” IP addresses (i.e. IP addresses that may move to different MACs). I am not expecting silent hosts in my BD.

Given the above requirements, the recommended Bridge Domain settings are as follows:

- **L2 Unknown Unicast:** Hardware Proxy
- **ARP Flooding:** Disabled
- **Unicast Routing:** Enabled
- **Subnet Configured:** Yes, if required
- **Enforce Subnet Check for IP Learning:** Yes

In this scenario, most of the BD settings can be left at their default, optimized values. A subnet (i.e. a gateway address) should be configured as required and it is recommended to enforce the subnet check for IP learning.

Bridge Domain Scenario 2: IP, Routed Traffic, Possible Silent Hosts

Scenario 2: My bridge domain will contain IP based, routed traffic. I won't be connecting firewalls / load balancers to this BD and it will not contain clusters or similar which may rely on “floating” IP addresses (i.e. IP addresses that may move to different MACs). I *might* have some “silent hosts” connected to the BD.

Given the above requirements, the recommended Bridge Domain settings are as follows:

- **L2 Unknown Unicast:** *Hardware Proxy*
- **ARP Flooding:** *Disabled*
- **Unicast Routing:** *Enabled*
- **Subnet Configured:** *Yes*
- **Enforce Subnet Check for IP Learning:** *Yes*

The Bridge Domain settings for this scenario are similar to scenario 1; however, in this case the subnet address should be configured. As silent hosts may exist within the Bridge Domain, a mechanism must exist to ensure those hosts are learnt correctly inside the ACI fabric. ACI implements an ARP gleaning mechanism that allows the spine switches to generate an ARP request for an end point using the subnet IP as the source address. This ARP gleaning mechanism ensures that silent hosts are always learnt, even when using optimized BD features such as hardware proxy.

The following figure shows the ARP gleaning mechanism when end points are not present in the mapping database.

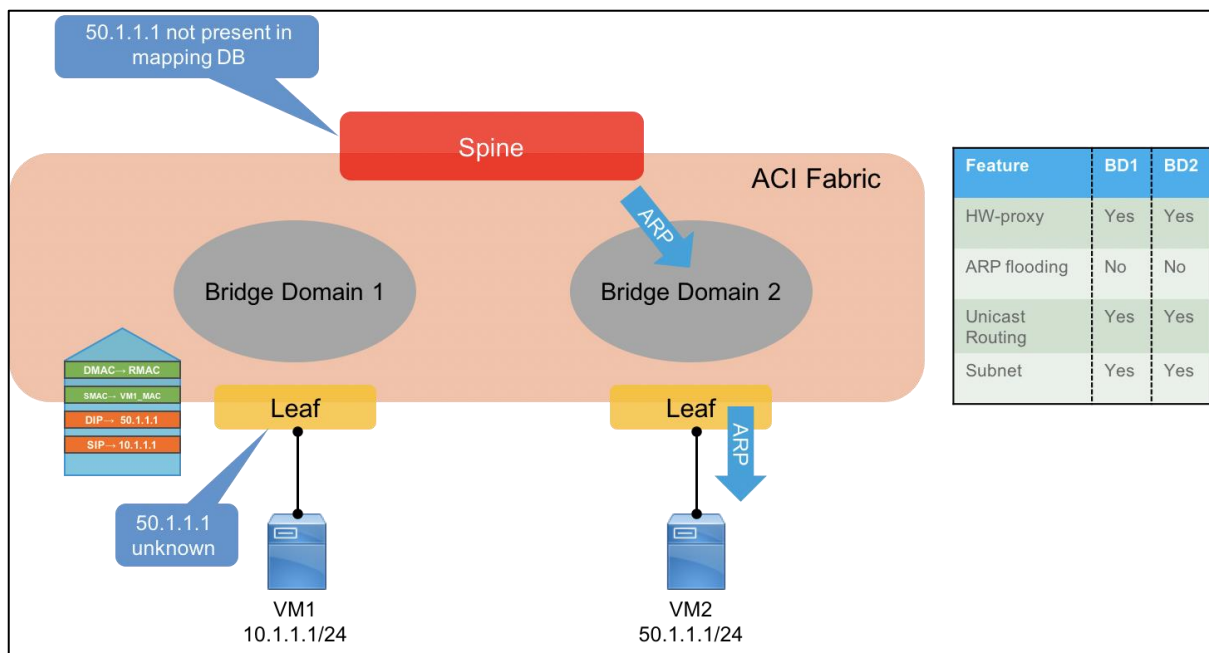


Figure 4: ARP Gleaning Mechanism in ACI

If a subnet IP address cannot be configured for any reason, ARP flooding should be enabled as an alternative to allow learning of silent hosts.

Bridge Domain Scenario 3: Non-IP, Switched Traffic, Possible Silent Hosts

Scenario 3: My bridge domain will contain non-IP based, switched traffic. I *might* have some “silent hosts” connected to the BD.

Given the above requirements, the recommended Bridge Domain settings are as follows:

- **L2 Unknown Unicast:** Flood
- **ARP Flooding:** N/A (enabled automatically due to no unicast routing)
- **Unicast Routing:** Disabled
- **Subnet Configured:** No
- **Enforce Subnet Check for IP Learning:** N/A

In this scenario, all optimizations inside the Bridge Domain are disabled and the BD is operating in a “traditional” manner. Silent hosts are dealt with through normal ARP flooding, which is always enabled when unicast routing is turned off.

It should also be noted that when operating the Bridge Domain in a “traditional” mode, the size of the BD should be kept manageable, i.e. limit the subnet size / number of hosts as you would in a regular VLAN environment.

Bridge Domain Scenario 4: non-IP or IP, Routed or Switched Traffic, Possible “Floating” IP Addresses

Scenario 4: My bridge domain will contain hosts or devices where the IP address may “float” between MAC addresses.

Given the above requirements, the recommended Bridge Domain settings are as follows:

- **L2 Unknown Unicast:** *Hardware Proxy*
- **ARP Flooding:** *Enabled*
- **Unicast Routing:** *Enabled*
- **Subnet Configured:** *Yes*
- **Enforce Subnet Check for IP Learning:** *Yes*

In this scenario, the bridge domain contains devices where the IP address may move from one device to another (in other words, the IP address moves to a new MAC address). This may be the case where routed firewalls are operating in active / standby mode, or where server clustering is used. Where this is a requirement, it is useful for gratuitous ARPs to be flooded inside the bridge domains in order to update the ARP cache of other hosts.

In this example, Unicast Routing and Subnet configuration are enabled for troubleshooting purposes (e.g. Traceroute), or for advanced features that require it (such as dynamic end point attachment).

Bridge Domain Scenario 5: Migrating to ACI, Legacy Network Connected via L2 Extension, Gateways on Legacy Network

Scenario 5: I am in the process of migrating to ACI. I am extending layer 2 from ACI to my legacy network – layer 3 gateways still reside on the legacy network.

Given the above requirements, the recommended Bridge Domain settings are as follows:

- **L2 Unknown Unicast:** *Hardware Proxy*
- **ARP Flooding:** *Enabled*
- **Unicast Routing:** *Enabled*
- **Subnet Configured:** *If Required*
- **Enforce Subnet Check for IP Learning:** *If Required*

In this scenario, the user is migrating hosts and services from the legacy network into the ACI fabric. A layer 2 connection has been set up between the two environments and the layer 3 gateway functionality will continue to exist in the legacy network for some time.

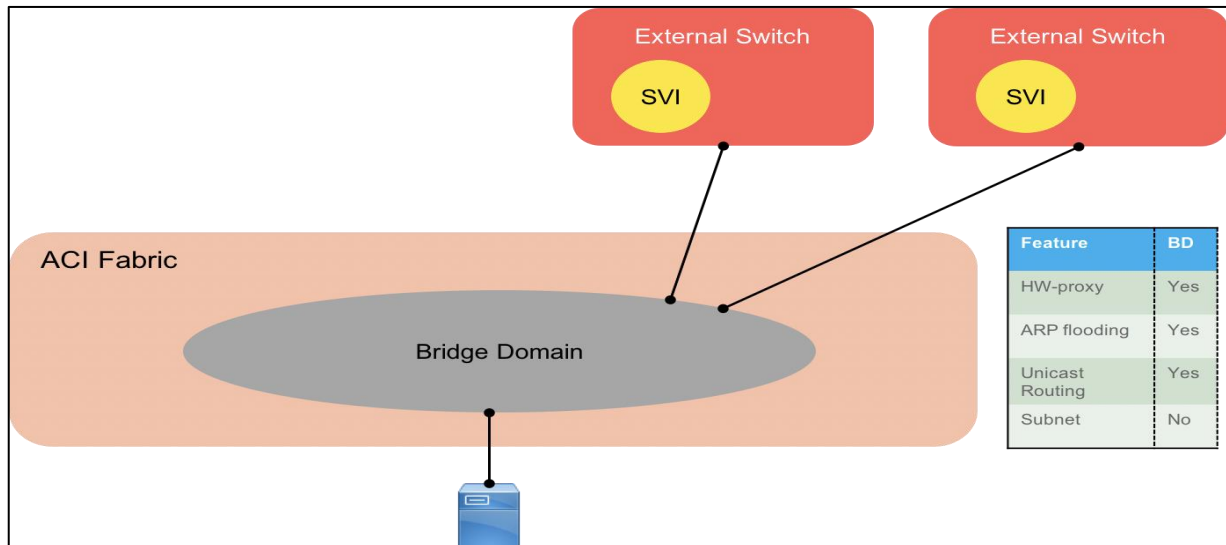


Figure 5: L2 Connection to Fabric with External Gateways

In this situation, ensure ARP flooding is enabled in the bridge domain.

6.4 Bridge Domains configuration overview

DC4:

num of BD: 334
 num of BD in common tenant: 1
 1 unused (no epG) BDs: ['mgmt:inb']

VRF-LEGACY has 165 I3 BDs with subnet and 153 BDs enabling routing but no subnet. Be aware of asymmetric routing



Microsoft Excel
Worksheet

vrf common: has 1 I3outs'default' and 1 BDs 'default' enabling routing but no subnet. Be aware of asymmetric routing

Only if you have silent hosts on a subnet and you don't have an IP address set on the Bridge Domain, will you need to enable ARP flooding.

Total L2 BDs: 0

Total L3 BDs: 334

I3 BD with HW Proxy: 15

I3 BD with Flood: 319

I3 BD no subnet: 156

I3 BD no subnet with HW proxy: 3

I3 BD no subnet with HW proxy, ARP flooding: 0

I3 BD no subnet with HW proxy, no ARP flooding (ARP gleaning not working!!): 3 ['ave-ctrl', 'default', 'inb']

I3 BD no subnet with Flood: 153

I3 BD with subnet: 178

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

Page 38 of 86

I3 bd with subnet but not limit subnet: 1 : ['default']
 I3 bd with subnet limit subnet: 177
 I3 BD with subnet and Flood: 166 :

I3 BD with subnet and HW Proxy: 12 ['BD-MDOM-Vmotion', 'default', 'BD-LEG-VL1649', 'BD-LEG-VL1099', 'BD-MDOM-VSAN', 'BD-MDOM-VS-MGMT', 'BD-LEG-VL1199', 'BD-LEG-VL1299', 'BD1699-VMW_VMO', 'BD-LEG-VL2249', 'BD-LEG-VL2499', 'BD-LEG-VL1549']

I3 BD with subnet with HW Proxy and ARP flood: 10 : ['BD-MDOM-Vmotion', 'BD-LEG-VL1649', 'BD-LEG-VL1099', 'BD-MDOM-VSAN', 'BD-MDOM-VS-MGMT', 'BD-LEG-VL1199', 'BD-LEG-VL1299', 'BD-LEG-VL2249', 'BD-LEG-VL2499', 'BD-LEG-VL1549']

I3 BD with subnet with HW Proxy and No ARP flood: 2 : ['default', 'BD1699-VMW_VMO']

Table 27 - L3 BDs with no ARP glean

Tenant	BD name
common	default
mgmt	inb

DC5:

!!WARN vrf LEGACY:VRF-LEGACY has 161 I3 BDs with subnet and 151 BDs enabling routing but no subnet. Be aware of asymmetric routing:



Microsoft Excel
Worksheet

Polly-VRF has 3 I3 BDs 'BD00016-STORAGE-TEST', 'Polly-Test-BD02', 'Polly-Test-BD01' with subnet and 1 BDs (['BD-test']) enabling routing but no subnet. Be aware of asymmetric routing.

Vrf common: has 1 I3outs 'default' and 1 BDs 'default' enabling routing but no subnet. Be aware of asymmetric routing.

2 unused (no epg) BDs: ['mgmt:inb', 'Polly-Test:BD00016-STORAGE-TEST']

num of BD: 331

num of BD in common tenant: 1

Total L2 BDs: 0

Total L3 BDs: 331

I3 BD with HW Proxy: 20

I3 BD with Flood: 311

I3 BD no subnet: 155

I3 BD no subnet with HW proxy: 4

I3 BD no subnet with HW proxy, ARP flooding: 0

I3 BD no subnet with HW proxy, no ARP flooding (ARP glean not working!!): 4 ['inb', 'default', 'BD-test', 'ave-ctrl']

I3 BD no subnet with Flood: 151

I3 BD with subnet: 176

I3 bd with subnet but not limit subnet: 1 : ['default']

I3 bd with subnet limit subnet: 175

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

Page 39 of 86

Table 28 - L3 BDs with no ARP glean

Tenant	BD name
Polly-Test	BD-test
common	default
mgmt	inb

Recommendations:

- Recommendation is to match ADP business requirements accordingly as per mentioned above scenarios.
- There are L3 BD's with routing enabled but no subnet they might lead to asymmetric routing, ADP is advised to check these BD's
- Validate BDs and other objects/constructs present on Fabric, the ones that are not required should be removed to keep consistency and simplicity.
- Review L3 BDs with no ARP glean.
The following BD configuration will not perform ARP glean, and therefore, will not learn silent hosts:
 - L3 / Hardware Proxy
 - No subnet configured
 - ARP flooding is disabled
- If a subnet cannot be configured on these BDs, ARP flooding should be enabled to ensure EP discovery will function properly (This might be a service impacting change).
- Avoid IP addressing overlapping between different BDs on the same VRF.

6.5 End Point Groups (EPGs)

Criticality/Compliance

High	Partially Conform
------	-------------------

ACI Endpoint Groups (EPGs) provide a new model for mapping applications to the network. Rather than using forwarding constructs such as addressing or VLANs to apply connectivity and policy, EPGs use a grouping of application endpoints. EPGs act as a container for collections of applications, or application components and tiers that can be used to apply forwarding and policy logic. They allow the separation of network policy, security, and forwarding from addressing and instead apply it to logical application boundaries.

Within an EPG separate endpoints can exist in one or more subnets, and subnets could be applied to one or more EPGs based on several other design considerations. Layer 2 forwarding behavior can then be applied independently of the Layer 3 addressing. Figure 4 shows the relationship between EPGs and subnets.

Endpoint groups not only allow for better mapping of applications to the network itself, but also for better mapping of the network to application owners and developers. Rather than application owners and developers being required to maintain mappings to IP addressing and subnets they can group applications or application components to logical EPGs.

EPGs are designed as flexible containers for endpoints that require common policy. Several methods exist for defining endpoints and placing them in EPGs. Once grouped, policy is applied based on the logical grouping rather than addressing and forwarding constructs. The use of EPGs can and will differ across customer environments and even across a single fabric deployment.

DC4 EPG CHECK:

num of EPG: 332

Num of physical-only EPGs :174

Num of virtual-only EPGs :3

Num of hybrid EPGs :153

Num of EPGs with phys domain, but no physical path bindings: 0

Unused EPGs 2 (no domain, no path):, ['infra|ave-ctrl|ave-ctrl', 'infra|access|default']

EPGs with physical path bindings but missing domain (2): MDOM|MDOM-APP-PROFIL|EPG-MDOM-VSAN, MDOM|MDOM-APP-PROFIL|EPG-MDOM-Vmotion

4 MGMT EPGs, VMM domain not using pre-provision:

- MDOM|MDOM-APP-PROFIL|EPG-MDOM-Vmotion, uni/vmmp-VMware/dom-VCE0531-DOM
- MDOM|MDOM-APP-PROFIL|EPG-MDOM-Vmotion, uni/vmmp-VMware/dom-POC-VCE0511-Credentials
- MDOM|MDOM-APP-PROFIL|EPG-MDOM-VS-MGMT, uni/vmmp-VMware/dom-POC-VCE0511-Credentials
- MDOM|MDOM-APP-PROFIL|EPG-MDOM-VS-MGMT, uni/vmmp-VMware/dom-VCE0531-DOM

Cannot find bundle policy group for interface: V172IBE2-INT-POLICY-GROUP does not exist.

Cannot find AEP for topology/pod-1/protopaths-1103-1203/pathep-[V172IBE2-INT-POLICY-GROUP], used by EPG:

LEGACY|INTERNAL-HOSTING|EA1651-BEN_DLA
LEGACY|INTERNAL-HOSTING|EA1652-RND_DLA
LEGACY|INTERNAL-HOSTING|EA1653-STG_DLA
LEGACY|INTERNAL-HOSTING|EA1655-LAB_ITS
LEGACY|INTERNAL-HOSTING|EA1675-STG_DBA
LEGACY|INTERNAL-HOSTING|EA1676-BEN_DBA
LEGACY|INTERNAL-HOSTING|EA1677-RND_DBA
LEGACY|INTERNAL-HOSTING|EA1678-STG_DBA
LEGACY|INTERNAL-HOSTING|EA1688-PRV_UNI
LEGACY|INTERNAL-HOSTING|EA1689-PRV_WIN
LEGACY|INTERNAL-HOSTING|EA1691-IBE2-ORA_RAC
LEGACY|INTERNAL-HOSTING|EA1692-ORA_RAC
LEGACY|INTERNAL-HOSTING|EA1697-VMW_VMM
LEGACY|INTERNAL-HOSTING|EA1698-VMW_STO
MDOM|MDOM-APP-PROFIL|EA1699-VMW_VMO

EPG MDOM|MDOM-APP-PROFIL|EPG-MDOM-VS-MGMT with Domain ['BCKP-L2-OUT-DOM'] binding Vlan {'vlan-1897'} Path topology/pod-1/protopaths-1121-1221/pathep-[252 different policy groups] not allowed by interface AEP VCE0411-DOM-AEP (allowed domain [])

EPG MDOM|MDOM-APP-PROFIL|EPG-MDOM-VSAN with Domain [] binding Vlan {'vlan-1898'} Path topology/pod-1/protopaths-1102-1202/pathep-[31 different policy groups] not allowed by interface AEP VCE0411-DOM-AEP (allowed domain [])

EPG MDOM|MDOM-APP-PROFIL|EPG-MDOM-Vmotion with Domain [] binding Vlan {'vlan-1899'} Path topology/pod-1/protopaths-1115-1215/pathep-[31 different policy groups] not allowed by interface AEP VCE0411-DOM-AEP (allowed domain [])



Microsoft Excel
Worksheet

DC5 EPG CHECK:

num of EPG: 328

Num of physical-only EPGs :165

Num of virtual-only EPGs :3

Num of hybrid EPGs :157

Num of EPGs with phys domain, but no physical path bindings: 0

Unused EPGs 3 (no domain, no path): LEGACY|INTERNAL-HOSTING|EA1234-IBEx-RAC_DB, infra|ave-ctrl|ave-ctrl, infra|access|default

4 MGMT EPGs,VMM domain not using pre-provision:

- MDOM|MDOM-APP-PROFIL|EPG-MDOM-Vmotion, uni/vmmp-VMware/dom-VCE0531-DOM,
- MDOM|MDOM-APP-PROFIL|EPG-MDOM-Vmotion, uni/vmmp-VMware/dom-POC-VCE0511-Credentials,
- MDOM|MDOM-APP-PROFIL|EPG-MDOM-VS-MGMT,uni/vmmp-VMware/dom-POC-VCE0511-Credentials,
- MDOM|MDOM-APP-PROFIL|EPG-MDOM-VS-MGMT, uni/vmmp-VMware/dom-VCE0531-DOM

The following EPGs have wrong static binding. They cannot find an AEP that links to the right interfaces or the PC/VPC policy groups have been deleted. Remediate it or delete the no longer valid EPG static bindings:

EPG LEGACY|INTERNAL-HOSTING|EA3110-IBE1_VOICE with Domain ['IBE1-L2-OUT-DOM'] binding Vlan {'vlan-3110'} Path topology/pod-1/protopaths-1110-1210/pathep-[NA578_IBE1-INT-POLICY-GROUP] not allowed by interface AEP NA578_IBE1-AEP (allowed domain ['NA578_IBE1-L2-OUT-DOM'])

EPG LEGACY|INTERNAL-HOSTING|EA3110-IBE1_VOICE with Domain ['IBE1-L2-OUT-DOM'] binding Vlan {'vlan-3110'} Path topology/pod-1/protopaths-1110-1210/pathep-[NA577_IBE1-INT-POLICY-GROUP] not allowed by interface AEP NA577_IBE1-AEP (allowed domain ['NA577_IBE1-L2-OUT-DOM'])

EPG LEGACY|INTERNAL-HOSTING|EA3144-IBE1-LBtoFW with Domain ['IBE1-L2-OUT-DOM'] binding Vlan {'vlan-3144'} Path topology/pod-1/protopaths-1104-1204/pathep-[F3100D-BCN1-01S-BACKUP-C1U2-IBE2-INT-POLICY-GRP] not allowed by interface AEP POC-IBE2-AEP (allowed domain ['POC-IBE2-L2-OUT-DOM'])

EPG LEGACY|INTERNAL-HOSTING|EA3161-STG_DBA with Domain ['POC-IBE2-L2-OUT-DOM'] binding Vlan {'vlan-3161'} Path topology/pod-1/protopaths-1110-1210/pathep-[NA578_IBE2-INT-POLCY-GROUP] not allowed by interface AEP NA578_IBE2-AEP (allowed domain ['NA578_IBE2-L2-OUT-DOM'])

EPG LEGACY|INTERNAL-HOSTING|EA3161-STG_DBA with Domain ['POC-IBE2-L2-OUT-DOM'] binding Vlan {'vlan-3161'} Path topology/pod-1/protopaths-1110-1210/pathep-[NA577_IBE2-INT-POLICY-GROUP] not allowed by interface AEP NA577_IBE2-AEP (allowed domain ['NA577_IBE2-L2-OUT-DOM'])

EPG LEGACY|INTERNAL-HOSTING|EA3197-VMW_VMM with Domain ['POC-IBE2-L2-OUT-DOM'] binding Vlan {'vlan-3197'} Path topology/pod-1/protopaths-1103-1203/pathep-[ESXi-VMM-DOM-INT-POLICY-GRP] not allowed by interface AEP VMM-6.5-DOM-AEP (allowed domain [])

EPG MDOM|MDOM-APP-PROFIL|EPG-LEG-VL3199 with Domain ['POC-IBE2-L2-OUT-DOM'] binding Vlan {'vlan-3199'} Path topology/pod-1/protopaths-1103-1203/pathep-[ESXi-VMM-DOM-INT-POLICY-GRP] not allowed by interface AEP VMM-6.5-DOM-AEP (allowed domain [])

EPG MDOM|MDOM-APP-PROFIL|EPG-MDOM-VS-MGMT with Domain ['POC-IBE2-L2-OUT-DOM', 'BCKP-L2-OUT-DOM'] binding Vlan {'vlan-3397'} Path topology/pod-1/protopaths-1102-1202/pathep-[89 interface policy groups] not allowed by interface AEP VMM-VCE0511-DOM-AEP (allowed domain [])

EPG MDOM|MDOM-APP-PROFIL|EPG-MDOM-VSAN with Domain ['POC-IBE2-L2-OUT-DOM'] binding Vlan {'vlan-3398'} Path topology/pod-1/protopaths-1102-1202/pathep-[9 INT-POLICY-GRP] not allowed by interface AEP VMM-VCE0511-DOM-AEP (allowed domain [])

EPG MDOM|MDOM-APP-PROFIL|EPG-MDOM-Vmotion with Domain ['POC-IBE2-L2-OUT-DOM'] binding Vlan {'vlan-3399'} Path topology/pod-1/protopaths-1102-1202/pathep-[9 INT-POLICY-GRP] not allowed by interface AEP VMM-VCE0511-DOM-AEP (allowed domain [])



Microsoft Excel
Worksheet

Recommendations:

- From the “EPG Check Findings”, there are a number of EPGs with configurations of “physical domain associated with AEP having virtual domain”, “no domain and no path bindings”
- Some EPGs have domain defined but these domains are not associated to any VLAN pools.
- There are some EPG’s created but not used, ADP is advised to validate EPGs and other objects/constructs present on Fabric, the ones that are not required should be removed to not waste resources, keep consistency and simplicity.
- Make sure all Access Policies Objects and relationships are properly in place for all ports.

6.6 L3Outs

Criticality/Compliance

High	Conform
------	---------

When configuring an L3Out policy in Cisco Application Centric Infrastructure (ACI) for external connectivity, there are a number of managed objects that are created as part of the L3Out configuration.

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

The Logical Node Profile managed object is used to identify the nodes (leaf switches) where the L3Out will be instantiated. The Node managed object is where the node and router ID is configured. In ACI, the router ID that is specified in the node profile is always configured as a manual router ID under the protocol that is configured for the L3Out.

The following recommendation apply when assigning L3Out router IDs:

- **Do not create a loopback interface with a router ID for OSPF, EIGRP, and static L3Outs.**
The node profile also has an option to create a loopback interface with the same value as the router ID. This option is only needed for BGP if you are establishing BGP peering sessions from a loopback interface with the router ID value. For OSPF and EIGRP, you should disable this option.
- **Create a loopback interface for BGP multi-hop peering between loopback addresses.**
For BGP, this option can be enabled if you are peering to the loopback address (BGP multi-hop) and are using the router ID address for the peering. You are not required to peer to the router ID address. You can also establish BGP peers to a loopback address that is not the router ID. For this configuration, disable the Use Router ID as Loopback Address option and specify a loopback address that is different than the router ID.
- **Each node (leaf switch) should use a unique router ID.**
Do not use the same router ID on different nodes in a single routing domain. Duplicate router IDs can cause routing issues. When configuring L3Outs on multiple border leaf switches, each switch (node profile) should have a unique router ID.
- **Use the same router ID value for all L3Outs on the same node within the same VRF.**
When configuring multiple L3Outs on the same node, you must use the same router ID value on all L3Outs. Using different router IDs is not supported. A fault will be raised if different router IDs are configured for L3Outs on the same node.
- **Configure a router ID for static L3Outs.**
The router ID is a mandatory field for the node policy. It must be specified even if no dynamic routing protocol is used for the L3Out. When creating an L3Out for a static route, you must still specify a router ID value. The Use Router ID as Loopback Address check box should be unchecked and the same rules apply regarding router ID value: use the same router ID for all L3Outs on the same node in the same VRF and different router ID for different nodes in the same VRF.



Best Practices:

- For redundancy purposes, at least 2 Border Leaf switches should take part of every L3Out.
- Make sure Subnets in L3Outs are not in conflict with BD subnets.
- Do not create a loopback interface with a router ID for OSPF, EIGRP, and static L3Outs.
- Each node (leaf switch) should use a unique router ID
- Use the same router ID value for all L3Outs on the same node within the same VRF
- Configure a router ID for static L3Outs

Table 29 DC4 External connectivity:

L3out Name:	L3out Name: Tenant/VRF	Protocol	Nodes
L3OUT_CORE_STORAGE	LEGACY:VRF-STORA	OSPF	1201, 1101
L3OUT-FW_ACI	MDOM:VRF-MDOM	Static	1201, 1101

January 21, 2021

L3out subnets not conflicting with BD subnet
One unique router-id per node per VRF. All L3outs are redundant in DC4.

Table 30 DC5 External connectivity:

L3out Name:	L3out Name: Tenant/VRF	Protocol	Nodes
L3OUT_CORE_STORAGE	LEGACY:VRF-STORA	OSPF	1201, 1101
L3OUT-FW_ACI	MDOM:VRF-MDOM	Static	1201, 1101

L3out subnets not conflicting with BD subnet
One unique router-id per node per VRF. All L3outs are redundant in DC5.

Recommendations:

- ADP Conforms with the Best Practices.
- Loopback using Route-ID should not be used in case of static/OSPF L3Outs to avoid wasting unnecessary resources.



It's generally recommended to configure L3outs across at least two nodes. This will ensure redundancy during upgrades and in the event of a border leaf failure. Although this is a general best practice, there are designs where this isn't necessary, e.g. redundancy provided across separate L3outs.

7 Tenant Policies – Security Policies

7.1 Contracts & VzAny

Criticality/Compliance

Low	Partially conform
-----	-------------------

By default, a VRF is in enforced mode, which means that without a contract, different endpoint groups are unable to communicate to each other. Endpoint groups associate to a contract with provider/consumer relationships.

ACLs, rules, and filters are created in the leaf switches to realize the intent of contracts that will be programmed on the ternary content-addressable memory (TCAM).

vzAny Overview

Policy information in ACI is programmed into two TCAM tables:

- *Policy TCAM* contains entries for the allowed EPG to EPG traffic.
- *App TCAM* contains shared destination L4 port ranges.

The size of the policy TCAM depends on the generation of Cisco ASIC in use. For ALE-based systems, policy TCAM provides 4k entries. For LSE-based systems, 64k hardware entries are available. In larger scale environments with larger policy scale it's important to take policy TCAM usage into account and ensure that the limits are not exceeded. As of 3.x, TCAM utilization is reported in the capacity dashboard and available via the API in 5 minute measurement intervals.

TCAM entries are generally specific to each EPG pair, i.e. even if the same contract is reused, new TCAM entries are installed for every pair of EPGs. The following diagram illustrates this behavior.

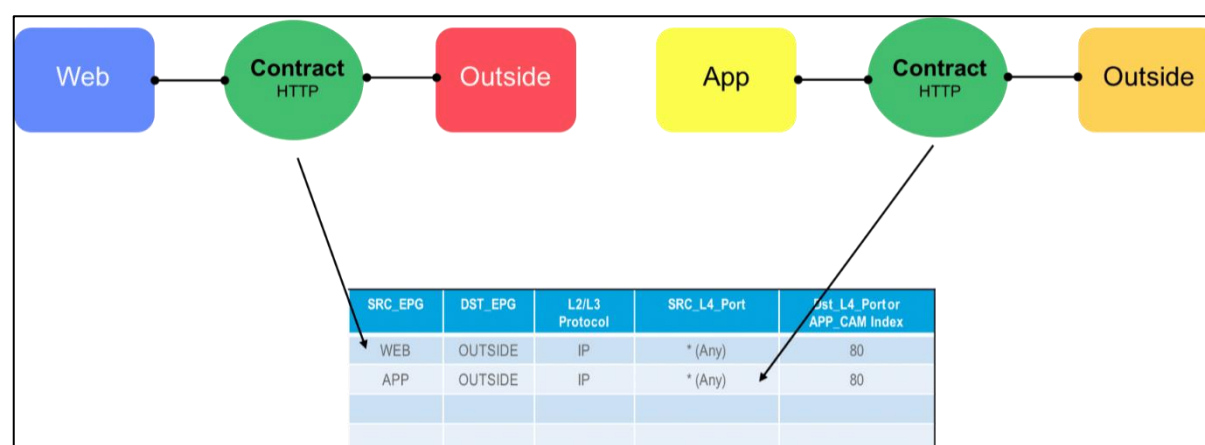


Figure 6: TCAM entries per EPG pair

An approximate calculation for the number of TCAM entries is as follows:

Number of entries in a contract * Number of Consumer EPGs * Number of Provider EPGs * 2

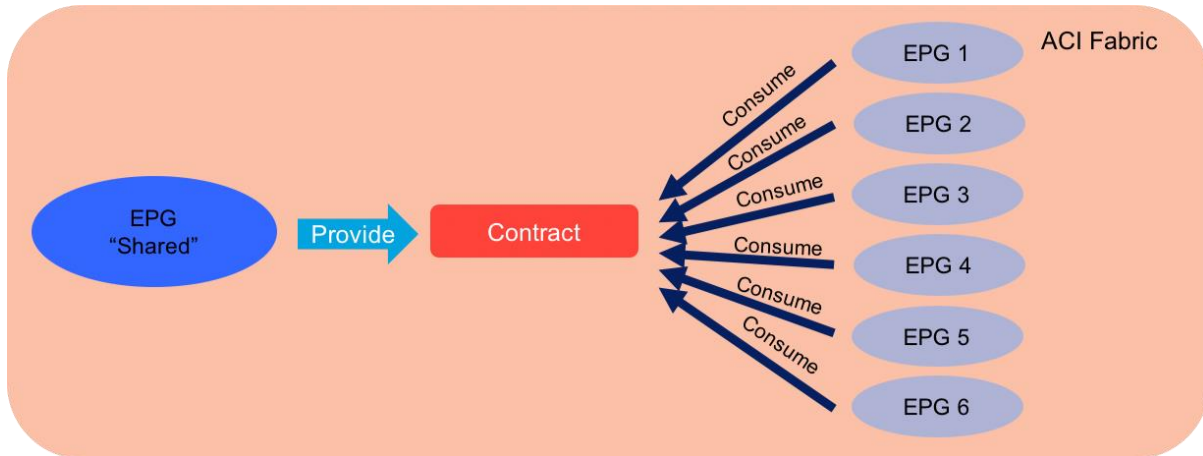


Figure 7: Multiple EPGs consuming a single contract

In this scenario, a single EPG named “Shared” is providing a contract, with multiple EPGs consuming that contract. This works; however, it has some drawbacks. Firstly, the administrative burden increases as each EPG must be configured separately to consume the contract. Secondly, the number of hardware TCAM entries increases each time an EPG associates with a contract. A very high number of EPGs all providing or consuming a contract can, in extreme cases, lead to exhaustion of the hardware resources.

To overcome these issues, the “vzAny” object may be used – vzAny is simply a managed object within ACI that represents all EPGs within a VRF. This object can be used to provide or consume contracts, so in our example above, we can consume the contract from vzAny with the same results, as shown in the following figure.

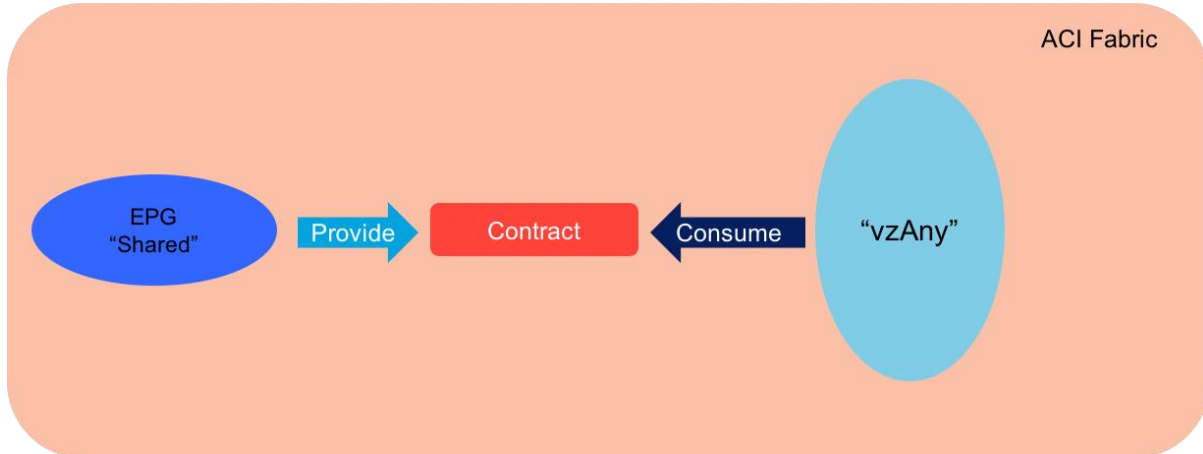


Figure 8: vzAny consuming a contract

This is not only easier to configure (although automation may render that point moot), but also represents the most efficient use of fabric hardware resources. vzAny is recommended to be used in cases where every EPG within a VRF must consume or provide a given contract.

Whenever use of the vzAny object is being considered, the administrator must plan for its use carefully. Once the vzAny object is configured to provide or consume a contract, any new EPGs associated with the VRF will inherit the policy (i.e. a new EPG added to the VRF will provide or consume the same contract(s) configured under vzAny). If it is likely that new EPGs will need to be added later and which may not need to consume the same contract as every other EPG in the VRF, then vzAny may not be the most suitable choice. This should be carefully considered before the use of vzAny is sanctioned.

Using vzAny with the “Established” flag

An additional example of the use of the vzAny policy to reduce resource consumption is to use it in conjunction with the “established” flag. By doing so, it is possible to configure contracts as ‘unidirectional’ in nature, which further reduces hardware resource consumption.

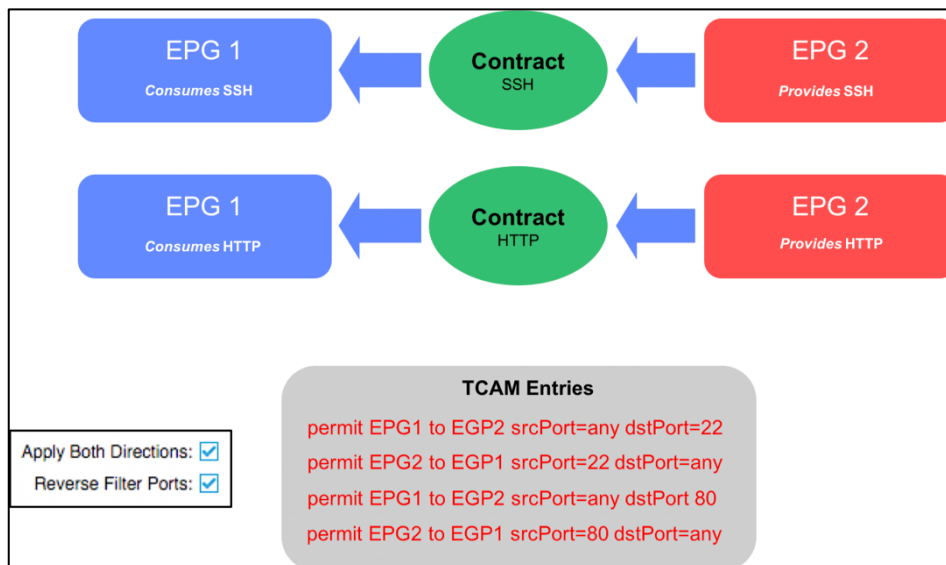


Figure 9: Bi-Directional Contracts - Regular Configuration

In this example, two contracts are configured (for SSH and HTTP) – both contracts are provided by EPG2 and consumed by EPG1. The “Apply Both Directions” and “Reverse Filter Ports” options are checked, resulting in the four TCAM entries shown.

It is possible to reduce the TCAM utilization by half by making the contract unidirectional, as shown in the following figure.

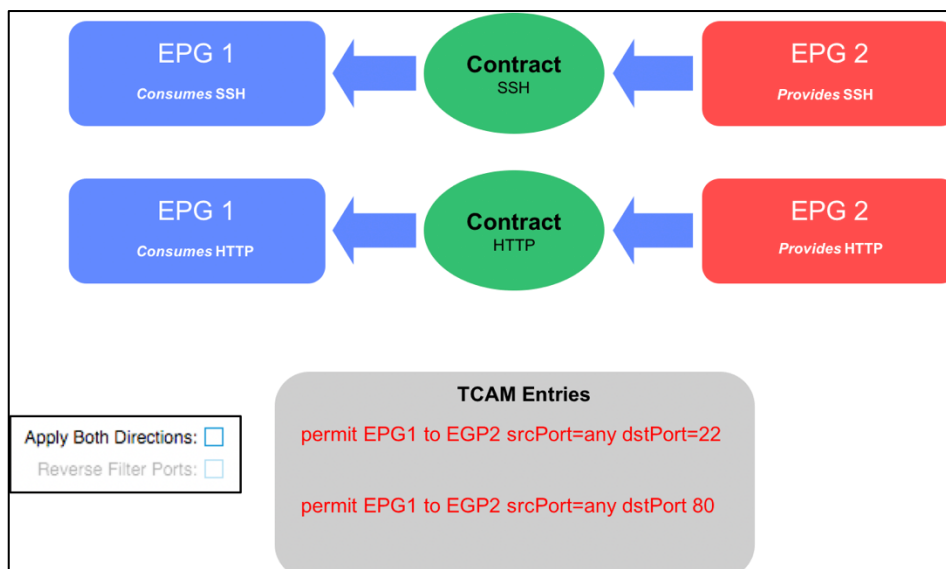


Figure 10: Unidirectional Contracts

However, the above scenario presents a problem – return traffic is not allowed in the contracts, therefore the connections cannot be completed and traffic fails. In order to allow return traffic to pass, we can configure a rule that allows traffic between all ports with the “established” flag. We can take advantage

of vzAny in this case to configure a single contract for the “established” traffic and apply it to the entire VRF.

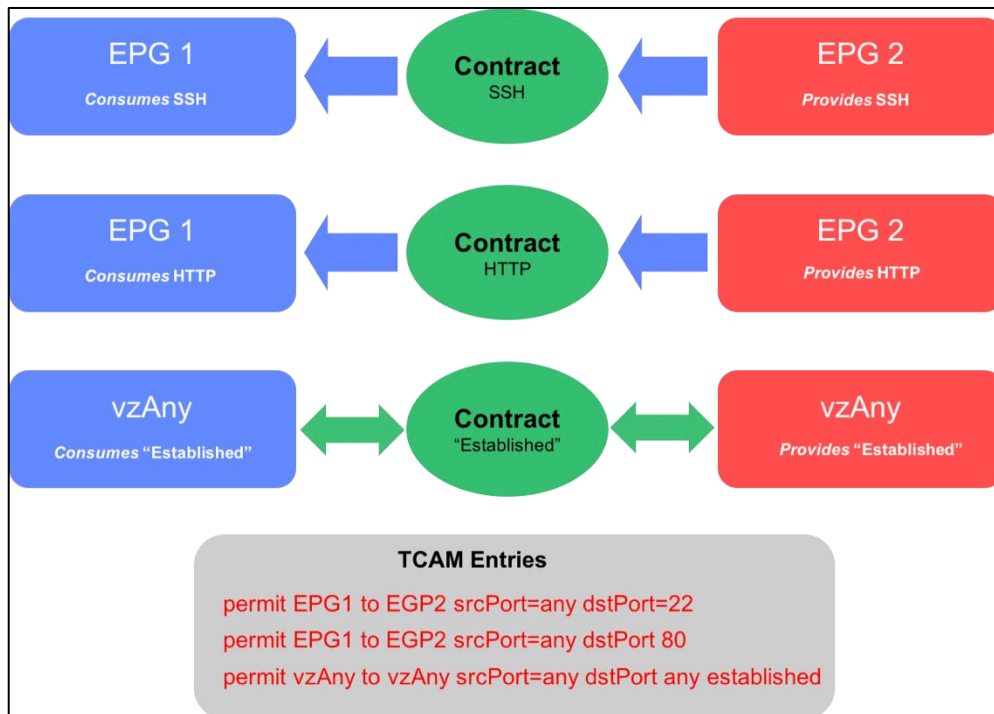


Figure 11: Use of vzAny with "Established" Contract

In an environment with a large number of contracts being consumed and provided, this can reduce the number of TCAM entries significantly. It's recommended to use vzAny in situations where contract enforcement will be uniform throughout a VRF.



Best Practices:

- Large production systems can easily have thousands of contract relationships in a single VRF. In such cases, vzAny could eliminate half or more of the contract relationships.
- This simplification not only makes the configuration much easier to maintain but can also save on switch TCAM buffer consumption.

DC4 Contracts & VzAny Checks:

num of VRF: 9
 num of Enforced VRF: 9
 num of Unenforced VRF: 0
 num of contract: 10
 num of imported contract: 3
 num of filters: 52

Tenant common has contract interface default

No error found in contract interface definition

non-consumed contract interfaces (1): [('common', 'default')]
 Nonused filters (2): [('common', 'arp'), ('common', 'est')]

Num of unused contracts: 2
 No contracts missing provider

No contracts is provided-only

Unused contracts:

Tenant	Contract
MDOM	CT-DEFAULT
common	default

9 out of total 9 VRFs are not using vzAny

DC5 Contracts & VzAny Checks Findings:

num of VRF: 11

num of Enforced VRF: 11

num of Unenforced VRF: 0

num of contract: 13

num of imported contract: 3

num of filters: 60

Tenant common has contract interface default

No error found in contract interface definition

non-consumed contract interfaces (1): [('common', 'default')]

Num of unused filters: 13

Num of unused contracts: 5

No contracts missing provider

No contracts is provided-only

Unused contracts:

Tenant	Contract
LEGACY	CT-STORAGE-PERMITANY
common	OSPF
common	default
common	test
MDOM	CT-DEFAULT

11 out of total 11 VRFs are not using vzAny

Observations:

- As shown in the findings, there are contracts not being used. ADP is advised to check these and take appropriate action.
- Consider using vzAny for ubiquitous contracts

8 Admin

8.1 AAA-Fallback Domain

Criticality/Compliance

Medium

Conform

A login domain defines the authentication domain for a user. Login domains can be set to the Local, LDAP, RADIUS, or TACACS+ authentication mechanisms



Best Practices:

It is recommended to leave the "fallback" domain as local authentication in case an issue arises with the remote authentication server.

8.2 Firmware

The following recommendations apply to ADP when upgrading the fabric:

When creating the maintenance group, verify the following items:

- vPC or active and standby pair of leaf switches are in two different groups so that while one of the switches is upgrading, the other switch can still pass the traffic.
- Spine switches that are configured as MP-BGP router reflectors are in two different groups, otherwise you will lose external connectivity during the upgrade.
- Divide switches into two or more groups - Upgrade one group at a time.

In the Run Mode field, choose the "*Pause only Upon Upgrade Failure*" radio button which is the default mode.



Best Practices:

- Different Firmware Maintenance Groups should be used for Spine Switches
- Different Firmware Maintenance Groups should be used for vPC or active and standby pair Leaf Switches
- Different Firmware Maintenance Groups should be used for Border Leaf Switch

8.3 Encrypted Backups

Criticality/Compliance

Medium

Conform

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.

Without encrypted backups, all passwords and sensitive data would have to be manually configured on restore. This can break logins and integration with external systems, such as AAA, SNMP, and VMM integration, significantly increasing MTTR. Because of this, it's recommended to run encrypted backups at regular scheduled intervals.

Note that encrypted backups do not encrypt the full backup, but only sensitive data—non-sensitive data remains unencrypted.



Best Practice:

Configure encrypted backups. Encrypted backups can be configured by selecting one of the existing backup jobs:

Admin > Import/Export > Export Policies > Configuration > [job name]

Then select the link next to **Modify Global AES Encryption Settings**. This is a global setting and will apply to all backups.

Observations:

Encrypted backups are currently enabled in the DC4 and DC5 fabric. No action is required.

8.4 Fabric Export Policies

Criticality/Compliance

High	Conform
------	---------

Fabric Export Policies are imperative for successful operation of the fabric as well as for analysing fabric configuration.

Export policies enable you to create an archive of configuration information, logs, and diagnostic data. Export policies also enable you to process core files and debug information from the fabric to an external host. You can configure an export policy that allows either scheduled or immediate backups to a remote server, and you can configure policy details such as transfer protocol, compression algorithm, and frequency of transfer.

The Export Policies panel contains a row of tabs. Each tab represents a different type of export policy. When a tab is clicked, you can view, create, and delete policies of the chosen policy type.

Fabric Export Policies in addition to Configuration Snapshots are configured under Admin > Import/Export



Best Practice:

- Fabric Export Policies should be configured to run on a schedule.
- Export Hashed Secure Properties should be Enabled, failure to do this will cause passwords to be left blank causing issues later at the time of a possible rollback.

Table 31: DC4 Fabric Export policy

Export Name	Remote Path	Scheduler	Description
BKACI	SRV_LX3058	BKACI_LX3058	""
Onetime-bkp	SRV_LX3058		""

Table 32: DC5 Fabric Export policy

Export Name	Remote Path	Scheduler	Description
BKACI_DC5	SRV_LX3058B	BKACI_LX3058B	"Production Back ups"
exp0005	wep157b		"automaticDataADP (Passphrase used for AES encryption)"

Observations:

ADP does conform to Cisco's ACI BPs for Fabric Export Policies.

9 Design Best Practices

The following sections cover a number of general ACI best practices and recommendations for the ADP ACI fabric.

9.1 Pod Policies COOP Policy

Criticality/Compliance

Info Only

Partially Conform

Council of Oracle Protocol (COOP) is used to communicate the mapping information (location and identity) to the spine proxy.

A leaf switch forwards endpoint address information to the spine switch 'Oracle' using Zero Message Queue (ZMQ). COOP running on the spine nodes will ensure all spine nodes maintain a consistent copy of endpoint address and location information and additionally maintain the distributed hash table (DHT) repository of endpoint identity to location mapping database.

COOP data path communication provides high priority to transport using secured connections. COOP is enhanced to leverage the MD5 option to protect COOP messages from malicious traffic injection.

The APIC controller and switches support COOP protocol authentication.

COOP protocol is enhanced to support two ZMQ authentication modes: strict and compatible.

- **Strict mode:** COOP allows MD5 authenticated ZMQ connections only.
- **Compatible mode:** COOP accepts both MD5 authenticated and non-authenticated ZMQ connections for message transportation.



Best Practice

Configure COOP strict authentication. For COOP POD Policy the Best Practice is to leave configuration as defined by default values and settings.

Table 33: DC4 and DC5 COOP policy configurations

COOP Policy	Type
default	compatible

Observations:

- Current COOP configuration in DC4 and DC5: **compatible**
- Recommendation is to **Configure COOP strict authentication**
This can be configured at **System > System Settings > COOP Group**
- ADP partially conform to Cisco's ACI BPs for COOP Pod Policy

9.2 BFD for Fabric Facing Interfaces

Criticality/Compliance

Low	Not Conform
-----	-------------

Cisco ACI Software Release 1.2(2g) added support for bidirectional forwarding detection (BFD), which is a software feature used to provide fast failure detection and notification to decrease the convergence times experienced in a failure scenario. BFD is particularly useful in environments where Layer 3 routing protocols are running over shared Layer 2 connections, or where the physical media does not provide reliable failure detection mechanisms. Some of the benefits of using BFD are as follows:

It provides sub second Layer 3 failure detection.

It supports multiple client protocols (for example, OSFP, BGP, EIGRP).

It is less CPU-intensive than routing protocol hello messages (the BFD echo function uses the data plane).

The client protocol is notified when BFD detects a failure. The client protocol does not need to run low hello timers.

Starting from APIC Release 3.1(1), BFD between leaf and spine switches is supported on fabric-interfaces for IS-IS. In addition, BFD feature on spine switch is supported for OSPF and static routes.

BFD is supported on modular spine switches that have -EX and -FX line cards (or newer versions), and BFD is also supported on the Nexus 9364C non-modular spine switch (or newer versions).

Current BFD ISIS configuration in DC4 and DC5: **disabled**



Best Practice recommendation – It is recommended to have BFD enabled for fabric facing interfaces.

This can be configured at **Fabric>Fabric Policies>L3 Interface>default>BFD ISIS Policy Configuration**

9.3 Domain Validation

Criticality/Compliance

Low	Not Conform
-----	-------------

It is recommended as a best practice to have Domain Validation enabled before a static path is added to an EPG to ensure the domain exists before adding the EPG to the static path.

Feature verifies whether the VLAN used in an EPG matches the AEP configured, that there are no overlaps, and so on.

Current value for Domain Validation in DC4 and DC5 : **disabled**



Best Practice recommendation – Enable Domain Validation

This can be configured at **System> System Settings > Fabric Wide Setting** and select “Enforce Domain Validation”



Once turned on, you can't turn it off.



If overlapping VLAN pools already exist and this parameter is checked, the system returns an error. You must assign VLAN pools that are not overlapping to the EPGs before choosing this feature.



If this parameter is checked and an attempt is made to add an overlapping VLAN pool to an EPG, the system returns an error.

9.4 General Stability

9.4.1 Mis-Cabling Protocol (MCP)

Criticality/Compliance

High

Not Conform

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.

Faults and events are generated when a port is disabled by MCP. MCP can be enabled globally and per-interface. By default, MCP is disabled globally and is enabled on each port. For MCP to work, it must be enabled globally, regardless of the per-interface configuration.

More details are available in the [Mis-Cabling Protocol section](#) of the ACI Best Practices guide.



Best Practices:

MCP should be enabled globally. It is recommended to enable MCP on all ports facing external switches or similar devices. Option “Enable MCP PDU per VLAN” should be enabled for MCP to work as intended.

Enable MCP at **Fabric > Access Policies > Global > MCP Instance Policy default**



Best Practice Recommendation– Enable PDU per VLAN

Enable PDU per VLAN at **Fabric > Access Policies > Global > MCP Instance Policy default**

Required to be enabled in order to set MCP interface policy, or else MCP interface policy has no effect.

Observations in DC4 and DC5:

Interface MCP state: **Enabled**

Global MCP state: **Disabled**

PDU Per VLAN: **disabled**

Table 34: Global Mis-cabling Status and Configuration

MCP Global Status	Auth	Transmit (hz)	Init Delay (s)	Loop Detect Factor	Action
Disabled	no	2	180	3	port-disable

Portgroup level Mis-cabling Protocol Status



DC4 MCP Checks:

- MCP is Globally disabled
- 2 out 822 policy group have empty mcp policy, using default. State: enabled
- 820 out 822 policy group using mcp policy default, State: enabled

DC5 MCP Checks:

- MCP is Globally disabled
- 1 out 480 policy group have empty mcp policy, using default. State: enabled
- 479 out 480 policy group using mcp policy default, State: enabled

Recommendation:

Please enable MCP as it was discussed with advanced services, following their advisory.

9.4.2 Digital Optical Monitoring (DOM)

Criticality/Compliance

Medium	Not Conform
--------	-------------

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](#) of the ACI Troubleshooting Guide.



Best Practice recommendation - Configure DOM

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

Observations:

DOM is not currently enabled on all the switches. As a best practice, enable DOM

9.4.3 Port Tracking

Criticality/Compliance

Medium	Not Conform
--------	-------------

The Port Tracking feature (first available in release 1.2(2g)) 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. This scenario is shown in the following figure.

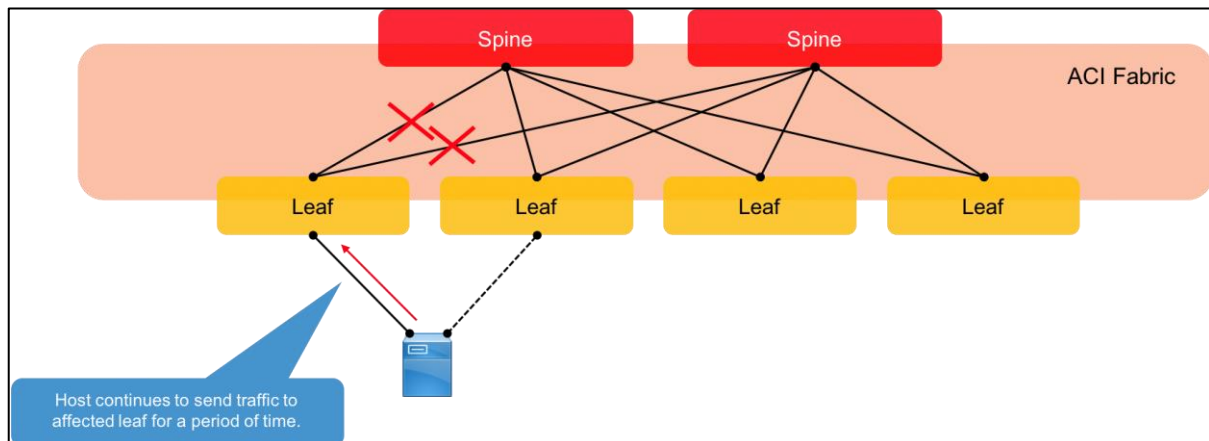


Figure 12: Leaf-Spine Connectivity Loss - Impact on Hosts

The Port Tracking feature detects a loss of fabric connectivity on a leaf node and brings down the host facing ports. This allows the host to fail over to the second link, as shown in the following figure.

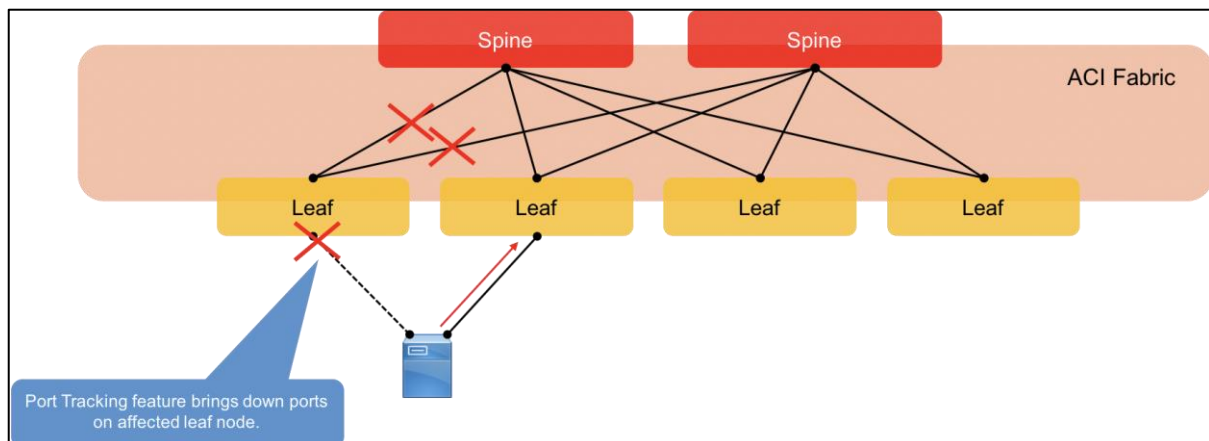


Figure 13: Leaf-Spine Connectivity Loss – Port Tracking Feature

The Port Tracking feature is configured under **System > System Settings > Port Tracking**, and it is a recommended feature to enable where active/standby NIC teaming hosts are connected to the fabric. Additionally, a port tracking setting of 0 uplinks is recommended as a general best practice. This will keep front-facing interfaces in a down state while there are no uplinks from a leaf, preventing some hosts from forwarding on an inactive interface.

Note that the preferred host connectivity to the ACI fabric is vPC wherever possible. Port tracking is useful in situations where hosts are connected using active/standby NIC teaming. The configuration parameters are as follows:

- **Delay Restore Timer:** This timer controls how long the fabric waits before bringing host ports up after the leaf spine links converge.
- **Number of Active Spine Links:** This value specifies how many links must be connected before Port Tracking takes action. In this example, the value is '0', which means Port Tracking will take action (i.e. disable host ports) only once the number of active links to the spine reaches zero.

More details about port tracking are available in the [Port Tracking section](#) of the ACI Best Practices guide.

Observations:

Port tracking state in DC4 and DC5 is not enabled



Best Practice recommendation - Configure port tracking.

The default setting of 120s delay restore timer and 0 uplinks will keep front-facing ports down on an isolated (no uplinks) switch and will prevent traffic loss in multiple scenarios. This is recommended as a general best practice.

10 End Point Learning

Cisco ACI fundamentally handles endpoint learning in a different manner than traditional network devices. This difference gives Cisco ACI the unique advantage of being able to limit flooding of ARP, unknown unicast, and other traffic types.

As Cisco ACI has evolved, the best way to configure ACI has evolved as well. This section presents a list of recommended configurations for endpoint learning that you should use, depending on the hardware that you have installed.

For optimal fabric operations, you should use settings that cause ACI to learn only IP addresses that are configured on a bridge domain subnet.

The options you use to enable the desired behavior depend on the generation of Cisco ACI leaf switches in your fabric.

First-generation leaf switches

For first-generation leaf switches, the following configurations are recommended for optimal endpoint update and forwarding behavior:

- Bridge domain–level configurations
 - Limit IP Learning To Subnet
- Fabric-level configurations
 - IP Aging Policy
 - Disable Remote EP Learn (on border leaf)
 - Prerequisite is to set Tenant > Networking > VRFs > Policy Control Enforcement to Ingress on your VRF instances

Second-generation leaf switches

For second-generation leaf switches, the following configurations are recommended for optimal endpoint update and forwarding behavior:

- Fabric-level configurations
 - IP Aging Policy
 - Disable Remote EP Learn (on border leaf)
 - Prerequisite is to set Tenant > Networking > VRFs > Policy Control Enforcement to Ingress on your VRF instances
 - Enforce Subnet Check

Fabrics with both first- and second-generation leaf switches

For fabrics with a mix of first- and second-generation leaf switches, the following configurations are recommended for optimal endpoint update and forwarding behavior:

- Bridge domain–level configurations
 - Limit IP Learning To Subnet
- Fabric-level configurations
 - IP Aging
 - Disable Remote EP Learn (on border leaf)
 - Prerequisite is to set Tenant > Networking > VRFs > Policy Control Enforcement to Ingress on your VRF instances
 - Enforce Subnet Check

Extensive details about endpoint learning are available in the [ACI Fabric Endpoint Learning Whitepaper](#).

10.1 Enforce Subnet Check

Criticality/Compliance

Medium	Not Conform
--------	-------------

The Enforce Subnet Check option was first introduced in APIC Releases 2.2(2q) and 3.0(2h) with the following enhancement:

CSCvf43074: ACI knob to limit IP EP learning to available BD subnets under the same VRF

In APIC Release 2.2(2q), the option is located at Fabric > Access Policies > Global Policies > Fabric Wide Setting Policy.

In APIC Release 3.0(2h) and later, it is located at System > System Settings > Fabric Wide Setting

This feature is available only on second-generation leaf switches.

This feature enforces subnet checks at the VRF level, when Cisco ACI learns the IP address as an endpoint from the data plane. Although the subnet check scope is the VRF instance, this feature can be enabled and disabled only globally under Fabric Wide Setting Policy. You cannot enable this option only in one VRF instance.

This feature is disabled by default.

Enforce Subnet Check works as follows:

On the ingress leaf (local endpoint learning): The option enforces bridge domain-level subnet checks for local endpoint learning. When this feature is enabled, the Cisco ACI leaf learns an IP address and MAC address as a new local endpoint only when the source IP address of the incoming packet belongs to one of the ingress bridge domain subnets.

This behavior is almost the same as Limit IP Learning To Subnet option under the bridge domain. The difference is that Limit IP Learning To Subnet limits only IP learning if the source IP address of a packet doesn't belong to an ingress bridge domain subnet, whereas this feature limits learning of both the MAC address and IP address if the source IP address doesn't belong to an ingress bridge domain subnet. Thus, Enforce Subnet Check enables slightly stronger checks than Limit IP Learning To Subnet. This check will be enabled on all bridge domains, and you cannot turn the checks on and off per bridge domain. Therefore, Limit IP Learning To Subnet is not required when this feature is enabled.

On the egress leaf (remote endpoint learning): This option enforces VRF-level subnet checks for remote endpoint learning. When this feature is enabled, the Cisco ACI leaf will learn an IP address as a remote endpoint only when the source IP address of the incoming packet belongs to any bridge domain subnet in the same VRF instance on the egress leaf.

This behavior prevents IP spoofing scenarios, in which an endpoint sends a packet with an unexpected source IP address that does not belong to any of the bridge domains on the VRF instance, such as an IP address that exists behind the L3Out connection.

When this feature is enabled, Cisco ACI flushes all local IP endpoints outside bridge domain subnets and all remote IP endpoints.



Best Practices:

Enforce Subnet Check should be enabled.

The following CCO document explains the global Enforce Subnet Check options

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html> - Toc529820939

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 similar to “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.

Note that this feature only works on Gen2 or later hardware. It’s safe to enable this on Gen1 hardware; however, Gen1 devices will continue to learn off-subnet IP addresses regardless of this configuration setting.

DC4 and DC5 Observations:

- ADP does not Conform, Enforce Subnet Check is not enabled.
- Enable Enforce Subnet Check (**System > System Settings > Fabric Wide Settings > Enforce Subnet Check**)



Because it operates at the VRF level, legitimate spoofing, e.g. load balancers, etc, should not be impacted by this feature; however, this should still be enabled with care. The Enhanced Endpoint Tracker provides visibility into off-subnet learns and is a good methodology to validate any potential impact prior to enabling this feature.



As this feature only impacts learning, enabling the feature will have no *immediate* impact. The feature will take effect once endpoints time out and have to be relearned on the fabric

10.1.1 Remote EP Learning

Criticality/Compliance

Medium

Conform

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](#) of the ACI Endpoint Learning Whitepaper.

Note that, most of these scenarios are specific to ALE-based (“gen 1”) hardware; however, there are also some less-common scenarios that can also impact -EX and -FX series switches.

The following configurations provide a mechanism to protect the fabric against IP spoofing. IP spoofing is when a packet is received from a source that doesn’t belong to the subnets associated with that source. This kind of traffic can be the result of misconfiguration, e.g. a server is configured in the incorrect VRF, BD, EPG, etc., or it can be the result of malicious traffic, e.g. an attacker attempting to hide their source address.

There are also legitimate reasons for off-subnet traffic to be forwarded into ACI. Some appliances, e.g. transparent firewalls and load balancers will spoof traffic. An example of this is a transparent load balancer proxying traffic from a VIP. There are also some non-appliance server and application designs where this behavior may be intentional.

ACI provides two different mechanisms for validating the subnet of received packet. Note that both of these features only impact whether the source address of a packet is learned by ACI, not if the packet will be forwarded. If the necessary contracts are in place, a packet could fail the following checks and still be forwarded across the fabric, which may still result in a remote learn on the receiving side.



Best Practices:

New Recommendation is that for Gen2 and above leaves DO NOT check “Disable Remote EP Learn” box.

DC4 and DC5 Observations:

Cisco Advanced services recommends to have this feature unchecked or disabled. ADP conforms to the best practices.

10.2 IP Aging Policy

Criticality/Compliance

Medium

Not Conform

The IP Aging Policy was first introduced in APIC Release 2.1(1h) with the following enhancement:

CSCut23815 ACI: unused local IP endpoint should be aged out separately from its MAC endpoint

This configuration is disabled by default to keep the same behavior with the older release.

For APIC Release 2.0, this option is located at Fabric > Access Policies > Global Policies > IP Aging Policy.

For APIC Release 3.0(1k) and later, it is located at System > System Settings > Endpoint Controls > IP Aging.

The IP aging policy tracks and ages unused IP addresses on an endpoint. Tracking is performed by using the endpoint retention policy, which is configured for the bridge domain to send ARP requests (for IPv4) and neighbor solicitations (for IPv6) at 75 percent of the local endpoint aging interval. When no response is received from an IP address, that IP address is aged out.

IP Aging provides a mechanism to age the IP address of an endpoint separately from a MAC address. Prior to this configuration, if traffic continued to flow from an endpoint (MAC), all IP addresses would be retained indefinitely. With this feature enabled, IP addresses are aged and timed out independently of the parent MAC address. This is an important capability for certain designs, such as load balancers, where a long-running MAC address may retain many hundreds or even thousands of IP addresses.

This feature is described in the following document:

<https://www.cisco.com/c/en/us/solutions/collateral/data-center-virtualization/application-centric-infrastructure/white-paper-c11-739989.html> - [Toc529820940](#)



Best Practices:

IP Aging Policy should be enabled to avoid endpoints (such as an interface on a virtual machine) having unused IP addresses stuck on the same MAC address.

DC4 and DC5 Observations:

- IP Aging Policy is not enabled.
- Configure IP Aging:
IP aging can be configured at **System > System Settings > Endpoint Controls > Ip Aging**. As this feature only impacts *inactive* IP addresses, there is no expected impact from this change.

10.3 Limit IP Learning to Subnet

Criticality/Compliance

Low	Partially Conform
-----	-------------------

Limit IP Learning to Subnet is a BD-level configuration. This feature restricts learning to IP addresses in subnets associated to the BD. This feature is useful to protect against spoofed EP learning, but has some drawbacks:

1. Legitimate spoofed traffic may be impacted. This feature has to be enabled with care to ensure that firewalls, load balancers, and other by-design spoofed traffic isn't impacted.
2. The VNID in the VXLAN header does not include BD information. Even if a packet isn't learned on a local leaf, it may still be learned on a remote leaf.



Best Practices:

Best Practice is to enable Limit IP Learning to Subnet at a BD level.

Table 35: Limit IP Learning to Subnet configuration status at DC4 and DC5

name	tenant	vrf	limit_ip_learn_to_subnet
inb	mgmt	inb	no
default	common		no
default	infra	overlay-1	no

DC4 and DC5 Observations:

- ADP has almost all the Bridge domains configured with this feature. Only three of them are not enabled with it.
- ADP should evaluate if it is as per their design or make appropriate changes

10.4 Loop Detection

Criticality/Compliance

Medium	Not Conform
--------	-------------

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 behaviour of these mechanisms.

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:** 128 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

Findings:

The current configuration in the ADP fabrics is as follows:

BD level tracking: enabled
EP loop protection: disabled
Rogue EP control: disabled

When rogue EP control is configured, this disables the default BD-level endpoint retention settings. A major advantage to this is the separation of IP and MAC tracking. This protects the network from scenarios where a single MAC with multiple IP addresses, e.g. a network appliance, could prematurely trigger the BD-level policy, suspend learning, and cause an outage.

Use BD Learn Disable only if you are using prior to Release 3.2 of ACI. From 3.2 onwards use Rogue End Point Detection Feature Instead. "Rogue End Point Detection" turned on, will render "BD Learn Disable" unapplicable, even though it's configured.

Recommendations:

- Recommendation is to leave EP loop protection disabled (**System > System Settings > Endpoint Controls > Ep Loop Protection**) and use rogue EP Control instead.
- Recommendation - Enable Rogue EP control (**System > System Settings > Endpoint Controls > Rogue EP Control**)

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

Page 67 of 86

- As a starting point, it's recommended to use the default timers and to disable (uncheck) both actions. This will still generate a fault if a loop is detected but will avoid causing unexpected impact in the event of a false positive. If faults are generated, the root cause can be identified, and these actions can potentially be enabled in the future. Rogue EP control and MCP will provide, additional hands-off protection in the event of an actual bridging loop.
- The following settings were discussed with ADP and are recommended:
 - Rogue EP Detection Interval: **30s** (default is 60s)
 - Rogue EP Multiplication Factor: **6** (default is 10)
 - Hold Interval: **1800s** (default)



Rogue EP control is recommended to be **disabled** during upgrades.



Note that the Rogue EP Control was first introduced in 3.2(1). Although the feature is available in the GUI and API in earlier releases of 3.x, it is not fully functional, and not recommended to be used. CUSTOMER will have to upgrade the fabric before enabling this feature.



Note that rogue EP control does **not** need to be **disabled** during upgrades if it was enabled in version 3.2 or later. If rogue EP control was enabled prior to 3.2, then it needs to be disabled during the next upgrade to avoid issues with vPC peer devices reloading.

11 Fabric Health DC4

11.1 Health Score

The fabric health score is calculated based on number and severity of faults. This number is useful to get a quick assessment of the state of the fabric, e.g. pre and post change or if a concern is raised about network behavior. In addition, individual device health scores can help quickly identify particular nodes of concern.

Health score: 97

Table 36: Per-device health score

Device	Health Score
NX9504-TIN1-04G-DC_SPINE-C1U01	100
NX9504-TIN1-04G-DC_SPINE-C1U02	100
NX9504-TIN1-04G-DC_SPINE-C1U03	100
NX93180-TIN1-04G-DC_LEAF-C1U01	78
NX93180-TIN1-04G-DC_LEAF-C1U03	100
NX93180-TIN1-04G-DC_LEAF-C1U05	91
NX93180-TIN1-04G-DC_LEAF-C1U07	99
NX93180-TIN1-04G-DC_LEAF-C1U09	100
NX93180-TIN1-04G-DC_LEAF-C1U11	78
NX93180-TIN1-04G-DC_LEAF-C1U13	99
NX93180-TIN1-04G-DC_LEAF-C1U15	86
NX93180-TIN1-04G-DC_LEAF-C1U17	78
NX93180-TIN1-04G-DC_LEAF-C1U19	99
NX93180-TIN1-04G-DC_LEAF-C1U21	100
NX93180-TIN1-04G-DC_LEAF-C1U23	78
NX93180-TIN1-04G-DC_LEAF-C1U25	93
NX93180-TIN1-04G-DC_LEAF-C1U27	68
NX93180-TIN1-04G-DC_LEAF-C1U29	99
NX93180-TIN1-04G-DC_LEAF-C1U31	100
NX93180-TIN1-04G-DC_LEAF-C1U33	95
NX93180-TIN1-04G-DC_LEAF-C1U35	92
NX93180-TIN1-04G-DC_LEAF-C1U37	100
NX93180-TIN1-04G-DC_LEAF-C1U39	100
NX93180-TIN1-04G-DC_LEAF-C1U41	100
NX93180-TIN1-04G-DC_LEAF-C1U43	100
NX93180-TIN1-04G-DC_LEAF-C1U45	100
NX93180-TIN1-04G-DC_LEAF-C1U47	100
NX93180-TIN1-04G-DC_LEAF-C1U49	100
NX93180-TIN1-04G-DC_LEAF-C1U51	100
NX93180-TIN1-04G-DC_LEAF-C1U53	100
NX93180-TIN1-04G-DC_LEAF-C1U55	100
NX93180-TIN1-04G-DC_LEAF-C1U57	78
NX93180-TIN1-04G-DC_LEAF-C1U59	100
NX93180-TIN1-04G-DC_LEAF-C1U02	85
NX93180-TIN1-04G-DC_LEAF-C1U04	99
NX93180-TIN1-04G-DC_LEAF-C1U06	78
NX93180-TIN1-04G-DC_LEAF-C1U08	99
NX93180-TIN1-04G-DC_LEAF-C1U10	100
NX93180-TIN1-04G-DC_LEAF-C1U12	100
NX93180-TIN1-04G-DC_LEAF-C1U14	99

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

NX93180-TIN1-04G-DC_LEAF-C1U16	78
NX93180-TIN1-04G-DC_LEAF-C1U18	94
NX93180-TIN1-04G-DC_LEAF-C1U20	52
NX93180-TIN1-04G-DC_LEAF-C1U22	78
NX93180-TIN1-04G-DC_LEAF-C1U24	78
NX93180-TIN1-04G-DC_LEAF-C1U26	100
NX93180-TIN1-04G-DC_LEAF-C1U28	68
NX93180-TIN1-04G-DC_LEAF-C1U30	98
NX93180-TIN1-04G-DC_LEAF-C1U32	100
NX93180-TIN1-04G-DC_LEAF-C1U34	95
NX93180-TIN1-04G-DC_LEAF-C1U36	92
NX93180-TIN1-04G-DC_LEAF-C1U38	100
NX93180-TIN1-04G-DC_LEAF-C1U40	100
NX93180-TIN1-04G-DC_LEAF-C1U42	100
NX93180-TIN1-04G-DC_LEAF-C1U44	100
NX93180-TIN1-04G-DC_LEAF-C1U46	100
NX93180-TIN1-04G-DC_LEAF-C1U48	100
NX93180-TIN1-04G-DC_LEAF-C1U50	100
NX93180-TIN1-04G-DC_LEAF-C1U52	100
NX93180-TIN1-04G-DC_LEAF-C1U54	100
NX93180-TIN1-04G-DC_LEAF-C1U56	100
NX93180-TIN1-04G-DC_LEAF-C1U58	78
NX93180-TIN1-04G-DC_LEAF-C1U60	100

11.2 Fabric Health and Persistent Faults

Criticality/Compliance

High	Partially Conform
------	-------------------

The Application Policy Infrastructure Controller (APIC) maintains a comprehensive, up-to-date run-time representation of the administrative and operational state of the Cisco Application Centric Infrastructure (ACI) fabric system in a collection of managed objects (MOs). In this model, a fault is represented as a mutable, stateful, and persistent MO. When a specific condition occurs, such as a component failure or an alarm, the system creates a fault MO as a child object to the MO that is primarily associated with the fault. For a fault object class, the fault conditions are defined by the fault rules of the parent object class.

In most cases, a fault MO is automatically created, escalated, de-escalated, and deleted by the system as specific conditions are detected. If the same condition is detected multiple times while the corresponding fault MO is active, no additional instances of the fault MO are created. A fault MO remains in the system until the fault condition is cleared. The fault MO is deleted according to the settings in the fault collection and fault retention policies. A fault MO is read-only unless it is in the cleared and retained state, when it can be deleted by the user by acknowledging it.

The creation of a fault MO can be triggered by internal processes such as finite state machine (FSM) transitions or detected component failures, or by conditions specified by various fault policies, some of which are user configurable. For example, you can set fault thresholds on statistical measurements such as health scores, data traffic, or temperatures.

11.3 Fault Management

Detailed operational recommendations are outside of the scope of this assessment; however, Cisco recommends integrating ACI fault management into existing operational process. Typically, this would include two key components:

1. Integrating real time alarming with operational tools to enable real time alarming upon a fault condition
2. An operational playbook including most common faults and process for operations to address the issue

This strategy has been shown to help address common, low-risk faults more rapidly and consistently, which helps maintain a higher health score and makes it easier to quickly identify high risk issues.

11.4 Cisco Recommendation

Cisco recommendation is to review the current faults and correct the problems identified. Many of these are provisioning faults e.g. F1296. Pay particular attention to F0135, F3019

Review current faults, taking one of the following actions:

- Correct the issue, clearing the fault.
- Adjust the fault policy to squelch any faults that will *always be irrelevant* for ADP DC, e.g. by design or unique network requirements.
- If a fault is expected and should **not** be squelched, i.e. it may indicate an issue in the future, document the fault and the reason it is being ignored. This documentation should be reviewed periodically, e.g. 3 to 6 months, to ensure the scenarios are still relevant.



Note that in the following sections, the fault description column is a truncated sample taken from one of the active faults. Each specific fault would provide the details to aid in diagnosing exact root cause, e.g. DN, EPG, etc.

Below are some of the observations from Fabric Health Perspective:

- ADP Fabric has some un-acknowledged faults, and these faults can impact ACI Fabric health score. This can impact the day to day operation, leading to operational issues.
- These issues could also be part of normal network behavior, but may indicate a hardware or cable problem, malfunctioning host, or other undiagnosed issue.
- ADP is advised to review the faults and take appropriate action as per the Best Practices.

11.5 Faults

At the time of this assessment, there were different fault types found in the DC4 fabric. The following sections cover each of these faults. The severity of these faults is taken into account with calculating the fabric health score, therefore, major and critical faults should typically be addressed first.

Recommended actions for each of the faults can be found in the [ACI System Messages Guide](#).

Where the resolution can't be easily identified, Cisco Services and/or TAC can assist with diagnosing and clearing the faults.

Detailed information about faults in DC4:



DC4_Fault_Report.xlsx

11.5.1 Critical Faults

Code	Count	Description	Comments
F0532	83	Port is down, reason:noOperMembers(connected), used by:EPG	This fault occurs when a port is down and is in use for epG.

If you see this fault, take the following actions

- Check the port connectivity
- Remove the configuration or administratively shut the port if the port is not in use
- For mcp-loop-err-disable, this could be due to a loop in the network. Check the config to resolve any loops
- For lacp suspended ports, check for following issues -
- Check whether peer device supports lacp or not.
- Enable LACP feature in peer device if peer device requires explicit global configuration.
- Check vlan range configured part of member interface and port-channel configurations matches.
- Check whether peer physical interface is added to port-channel as member interface.
- Check whether lacp is configured on member interface of peer device.
- If peer switch is a cisco, issue "show lacp counters" and verify interfaces of peer device are sending LACP PDU.
- Try the above and then using the "show etherchannel summary"/"show port-channel summary" command ensure the port-channel shows the S (Layer-2) and U (in use) flags, and that both interface appear in the ports column with the P (bundled) flag.
- If peer switch or router is non-cisco, Please contact customer support of peer device's manufacturer for not sending lacp or recommended configuration guidelines.
- For reason of suspend(connected) port, check whether VPC domain is configured in GUI Fabric-> External Access/Access Policies ->Virtual Port Channel default to tie two switches part of VPC domain if intent is to make two switches part of VPC.
- If the above actions did not resolve the issue, create a tech-support file and contact Cisco TAC.

11.5.2 Major Faults

Code	Count	Description	Comments
F3062	1	APIC Controller product is not registered with CSSM and 90 days evaluation period is expired.	This fault is raised when product license evaluation period (90 days) is expired. Create a registration token from customer's license account in CSSM and register APIC Controller product using the registration token.

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

F2705	17	vPC F511800_ECS_U1-INT-POLICY-GRP is impaired.	This fault occurs when vpc interface goes down while peer interface is up Please verify the port connectivity.
F1296	44	vPC SSE3200U2_HWT-INT-POLICY-GRP is down.	This fault occurs when vpc interface goes down while peer interface is also down. Please verify the port connectivity
F0135	154	Unsupported remote operation detected on EPG: uni/tn-LEGACY/ap-EXTERNAL-HOSTING/epg-EA1000-HBE1_NS_DLA detected in Controller: 11.128.242.157 with name VCE0411_TIN in datacenter NX-Pantin in domain VCE0411-DOM , error: [Insufficient permission to create/modify port group]	This fault is raised when remote or external disruptive EPG operations are performed on the Controller. Restore EPG to its original state on the Controller.
F3019	2	Operational issues detected for retrieving Tagging information from VMM controller: 11.128.242.157 with name VCE0411_TIN in datacenter NX-Pantin in domain: VCE0411-DOM due to error: Failed to retrieve all Tag information.	This fault is raised when ACI controller failed to retrieve complete Tagging information. Check vCenter and validate that Tagging information is available. Verify that all Tag and Category entries have non-empty Description

11.5.3 Minor Faults

Code	Count	Description	
F0603	28	Port is operationally individual.	This fault occurs when port becomes operationally individual. Please verify the configuration on both ends of the port-channel.
F0467	66	Configuration failed for uni/tn-MDOM/ap-MDOM-APP-PROFIL/epg-EPG-MDOM-VS-MGMT due to DHCP Provider Not Reachable, debug message:	This fault occurs when an End Point Group / End Point Security Group is incompletely or incorrectly configured.
F1394	3	Port is down, reason:sfpAbsent(connected), used by:Fabric	This fault occurs when a port is down and is in use for fabric Check the port connectivity Administratively shut the port if the port is not in use

11.5.4 Warning Faults

Code	Count	Description	
F119936	8	TCA: egress drop packets rate(qosmEgrPkts5min:dropRate) value 62 raised above threshold 1	This fault is caused by "egress drop packets rate" statistical property crossing threshold level.

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

			Recommended Action: Check the values of "egress drop packets" statistical counter and either correct the conditions that cause the counter values to cross certain threshold levels, or adjust the threshold values in monitoring policies.
F3057	1	APIC Controller product is not registered with CSSM and the product is in 90 days evaluation period.	his fault is raised when APIC Controller product is not registered with Cisco Smart Software Manager (CSSM). Create a registration token from customer's license account in CSSM and register APIC Controller product using the registration token.
F1192	8	Port-channel(po14) membership configuration failure for eth1/13	This fault is caused by a hardware programming failure Recommended Action: The system will periodically retry to program the hardware. If the failure is due to lack of hardware resources, free up resources by removing or simplifying existing configuration. One of the reasons for lack of hardware resources is exhaustion of VLANs in ToR. These hardware resources can be examined by looking at capacity dashboard under OPERATIONS tab. Please check entry against the number of VLANs in output of command show vlan summary issued at node where fault is generated to check whether it exceeds above advertised node capacity.
F3525	14	High SSD usage observed. Please check switch activity and contact Cisco Technical Support about high SSD usage.	This warning occurs when high SSD usage is observed Check switch activity and contact Cisco TAC.
F0546	54	Port is down, reason:sfpAbsent(connected), used by:Discovery	This fault occurs when a port goes down If you see this fault, take the following actions Check the port connectivity Administratively shut the port if the port is not in use
F1186	8	Port configuration failure. Reason: 2 Config: l1:Physlfspeed_failed_flag	This fault is caused by a hardware programming failure
F2740	8	Port speed configured on node-1103 NX93180-TIN1-04G-DC_LEAF-C1U05 interface eth1/13 is invalid or unsupported	This fault occurs when port speed is configured to an invalid/unsupported value

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

			To recover from this fault, correct speed in link-level policy used in TOR or change the speed as "inherit"
--	--	--	---

12 Fabric Health DC5

Detailed information about faults in DC5:



DC5_Fault_Report.xlsx

12.1 Health Score

Health score on DC5 fabric is **92**.

Table 37: Per-device health score

Device	Health Score
NX9504-BCN1-01S-DC_SPINE-C1U01	100
NX9504-BCN1-01S-DC_SPINE-C1U02	100
NX9504-BCN1-01S-DC_SPINE-C1U03	100
NX93180-BCN1-01S-DC_LEAF-C1U01	83
NX93180-BCN1-01S-DC_LEAF-C1U03	94
NX93180-BCN1-01S-DC_LEAF-C1U05	78
NX93180-BCN1-01S-DC_LEAF-C1U07	100
NX93180-BCN1-01S-DC_LEAF-C1U09	100
NX93180-BCN1-01S-DC_LEAF-C1U11	72
NX93180-BCN1-01S-DC_LEAF-C1U13	100
NX93180-BCN1-01S-DC_LEAF-C1U15	80
NX93180-BCN1-01S-DC_LEAF-C1U17	99
NX93180-BCN1-01S-DC_LEAF-C1U19	100
NX93180-BCN1-01S-DC_LEAF-C1U21	90
NX93180-BCN1-01S-DC_LEAF-C1U23	100
NX93180-BCN1-01S-DC_LEAF-C1U25	94
NX93180-BCN1-01S-DC_LEAF-C1U27	100
NX93180-BCN1-01S-DC_LEAF-C1U29	68
NX93180-BCN1-01S-DC_LEAF-C1U31	60
NX93180-BCN1-01S-DC_LEAF-C1U33	65
NX93180-BCN1-01S-DC_LEAF-C1U35	60
NX93180-BCN1-01S-DC_LEAF-C1U37	100
NX93180-BCN1-01S-DC_LEAF-C1U39	100
NX93180-BCN1-01S-DC_LEAF-C1U02	83
NX93180-BCN1-01S-DC_LEAF-C1U04	94
NX93180-BCN1-01S-DC_LEAF-C1U06	78
NX93180-BCN1-01S-DC_LEAF-C1U08	100
NX93180-BCN1-01S-DC_LEAF-C1U10	100
NX93180-BCN1-01S-DC_LEAF-C1U12	72
NX93180-BCN1-01S-DC_LEAF-C1U14	100
NX93180-BCN1-01S-DC_LEAF-C1U16	80
NX93180-BCN1-01S-DC_LEAF-C1U18	93
NX93180-BCN1-01S-DC_LEAF-C1U20	52
NX93180-BCN1-01S-DC_LEAF-C1U22	90
NX93180-BCN1-01S-DC_LEAF-C1U24	100
NX93180-BCN1-01S-DC_LEAF-C1U26	94
NX93180-BCN1-01S-DC_LEAF-C1U28	92
NX93180-BCN1-01S-DC_LEAF-C1U30	68

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

NX93180-BCN1-01S-DC_LEAF-C1U32	62
NX93180-BCN1-01S-DC_LEAF-C1U34	65
NX93180-BCN1-01S-DC_LEAF-C1U36	54
NX93180-BCN1-01S-DC_LEAF-C1U38	100
NX93180-BCN1-01S-DC_LEAF-C1U40	100

12.2 Cisco Recommendation

Cisco recommendation is to review the current faults and correct the problems identified.

12.2.1 Critical Faults

Code	Count	Description	Comments
F0532	15483	Port is down, reason:noOperMembers(connected), used by:EPG	This fault occurs when a port is down and is in use for epg.

12.2.2 Major Faults

Code	Count	Description	Comments
F0133	1	Unsupported remote operation on controller: 11.130.242.157 with name VCE0511_BCN in datacenter NX-Barcelona in domain POC-VCE0511-Credentials detected, error: [LACP update for Portgroup failed.]	This fault is raised when deployment of given configuration fails for a Controller. Try to delete and add the policy again.
F3062	1	APIC Controller product is not registered with CSSM and 90 days evaluation period is expired.	This fault is raised when product license evaluation period (90 days) is expired. Create a registration token from customer's license account in CSSM and register APIC Controller product using the registration token.
F1296	92	vPC SSE3200U1_HWT-INT-POLICY-GRP is down.	This fault occurs when vpc interface goes down while peer interface is also down. Please verify the port connectivity.
F2705	10	vPC LX4839B-INT-POLICY-GRP is impaired.	This fault occurs when vpc interface goes down while peer interface is up. Please verify the port connectivity.
F1547	12	100% of packets were received in excess during the last collection interval	This fault occurs when a significant number of excess packets are detected by a configured and enabled Atomic Counter Recommended Action: If you see this fault, take the following actions: Check the network configuration in the packet path where this Atomic Counter Policy is configured.

			Check interface counters and rates in the packet path where this Atomic Counter Policy is configured to make sure there is no oversubscription.
F1545	11	100% of packets were dropped during the last collection interval	This fault occurs when a significant number of packet drops are detected by a configured and enabled Atomic Counter Check the network configuration in the packet path where this Atomic Counter Policy is configured. Check for any CRC, etc errors in the packet path where this Atomic Counter Policy is configured. Check interface counters and rates in the packet path where this Atomic Counter Policy is configured to make sure there is no oversubscription.
F6076 03	1	[FSM:FAILED]: Task for updating FNV(TASK:ifc:dhcpd:DhcpClientUpdateF nv)	
F0132	2	Operational issues detected for VMM controller: 11.130.240.92 with name POC-IBE2-VCSARD02-Profil in datacenter DC5_LAB in domain: POC-IBE2-VDS-6-5 due to error: Last inventory pull did not complete for a few hosts or VMs or no hosts found. Please verify the Hosts and VMs are in connected state in the VMM controller and manually re-trigger inventory sync on APIC. Please ignore this fault if there are no hosts in the VMM controller.	This fault is raised when Remote or External disruptive operations on VMM Controller are detected. Ensure all vCenter hosts are in connected state and online Ensure at least one VM exists within the vCenter Inventory Ensure there are no VMs in an 'Unreachable' state Manually retrigger a VMM inventory sync by going to the VM networking -> Inventory -> Expand 'VMware' -> Expand Domain -> Expand 'Controllers' -> Right click on the controller and select "Trigger Inventory Sync"
F6063 91	4	[FSM:FAILED]: Get LLDP/CDP adjacency information for the physical adapters on the host: v1058ibe2.ehc.adp.com(TASK:ifc:vmmm gr:CompHvGetHpNicAdj)	
F0135	152	Unsupported remote operation detected on EPG: uni/tn-LEGACY/ap-EXTERNAL-HOSTING/epg-EA2500-HBE1_NS_DLA detected in Controller: 11.130.242.157 with name VCE0511_BCN in datacenter NX-Barcelona in domain POC-VCE0511-Credentials , error: [Insufficient permission to create/modify port group]	This fault is raised when remote or external disruptive EPG operations are performed on the Controller. Restore EPG to its original state on the Controller.
F2840	2	Operational issues detected for Host dc5prhchist0500.ehc.adp.com in VMM controller: 11.130.242.157 with name VCE0511_BCN in datacenter NX-Barcelona in domain: POC-VCE0511-	This fault is raised when ACI controller failed to update the properties of a VMware hypervisor. Check in VCenter whether there is any faults raised for the Hypervisor

		Credentials due to error: ESX Host is disconnected or not responding.	in question and resolve the faults via VCenter
--	--	---	--

12.2.3 Minor Faults

Code	Count	Description	
F0849	2	Configuration failed due to invalid-port for node NX93180-BCN1-01S-DC_LEAF-C1U03	his fault occurs when a infra selector (port selector, card selector, node selector etc.) is incorrectly configured. 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.
F0467	2	Configuration failed for uni/tn-MDOM/ap-MDOM-APP-PROFIL/epg-EPG-MDOM-VS-MGMT due to DHCP Provider Not Reachable, debug message:	This fault occurs when an End Point Group / End Point Security Group is incompletely or incorrectly configured.
F0053	2	Configuration backup/restore job 2018-11-13T10-53-08 failed with error: Upload failed, Timeout was reached	his fault occurs when a configuration export/import job fails. Verify that the Controllers can reach the remote destination without issues and that the remote file is correct Verify that the remote destination is healthy. To clear this fault, you can acknowledge the job. In GUI: Navigate to the Operational tab of the config policy Select the failed job, and choose Delete from the Actions menu
F0756	1	Could not resolve the target hintfpol-25G-link-autoneg-ON to form a named relationship. Using a default target uni/infra/hintfpol-default instead	This fault occurs when a configured target of a named relationship cannot be resolved. 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.
F2563	41	Callhome message failed to send to destination . Please check smtp server config	This fault occurs when sending a callhome message to destination does not succeed. Verify destination information is correct.

			Ensure it is reachable (ping/ssh) from the fabric and DNS entries have been configured correctly. You may remove and reconfigure the callhome destination. Disable RFC compliant option if enabled.
F0603	44	Port is operationally individual.	This fault occurs when port becomes operationally individual. Please verify the configuration on both ends of the port-channel.

12.2.4 Warning Faults

Code	Count	Description	
F3057	1	APIC Controller product is not registered with CSSM and the product is in 90 days evaluation period.	This fault is raised when APIC Controller product is not registered with Cisco Smart Software Manager (CSSM). Create a registration token from customer's license account in CSSM and register APIC Controller product using the registration token.
F0546	20	Port is down, reason:noOperMembers(connected), used by:Discovery	This fault occurs when a port goes down. Check the port connectivity. Administratively shut the port if the port is not in use.
F3525	3	High SSD usage observed. Please check switch activity and contact Cisco Technical Support about high SSD usage.	This warning occurs when high SSD usage is observed. Check switch activity and contact Cisco TAC.
F1011	2	Failed to form relation to MO uni/infra/funcprof/accportgrp-Polly_Test_Grp of class infraAccBaseGrp	The object refers to an object that was not found. Make sure that referenced object exists and the name is spelled correctly in the relation object.
F0956	2	Failed to form relation to MO uni/phys-NA613_614-DOM of class physDomP	The object refers to an object that was not found. Make sure that referenced object exists and the name is spelled correctly in the relation object.

13 Current Scale per Switch

This section compares the maximum verified scalability limits for ACI parameters for the Cisco APIC and Cisco Nexus 9000 Series ACI-Mode Switches with ADP fabric.

13.1 ACI Scalability Matrix and Compliance



Recommended ACI Scalability

These values are based on a profile where each feature was scaled to the numbers specified in the tables. These numbers do not represent the theoretically possible ACI fabric scale.

Table 38: ACI General Scalability limit

Configurable Options	L2 Fabric	L3 Fabric
Number of APIC controllers Note * denotes preferred cluster size. While the higher number of controllers is supported, the preferred size is based on the number of leaf switches in the environment.	3* or 4 node APIC cluster	3* or 4 node APIC cluster
Number of leaf switches	80	80 for 3-node cluster 200 for 4-node cluster
Number of spine switches	Maximum spines per pod: 6. Total spines per fabric: 24.	Maximum spines per pod: 6. Total spines per fabric: 24.
Number of FEXs	20 FEXs per leaf switch 576 ports per leaf switch 650 FEXs per fabric	20 FEXs per leaf switch 576 ports per leaf switch 650 FEXs per fabric
Number of tenants	1000	1000
Number of Layer 3 (L3) contexts (VRFs)	N/A	1000
Number of contracts/filters	N/A	10,000 contracts 10,000 filters
Number of endpoint groups (EPGs)	For a fabric with a single Tenant: 4,000 For a fabric with multiple Tenants: 500 per Tenant, up to 21,000 total across all Tenants	For a fabric with a single Tenant: 4,000 For a fabric with multiple Tenants: 500 per Tenant, up to 15,000 total across all Tenants
Number of Isolation enabled EPGs	400	400
Number of bridge domains (BDs)	21,000	15,000
Number of BGP + number of OSPF sessions + EIGRP (for external connection)	N/A	3,000
Number of Multicast routes	N/A	32,000

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

Number of Multicast routes per VRF	N/A	32,000
Number of vCenters	N/A	200 VDS 50 AVS 50 Cisco ACI Virtual Edge
Number of Service Chains	N/A	1000
Number of L4 - L7 devices	N/A	30 managed or 50 unmanaged physical HA pairs, 1,200 virtual HA pairs (1,200 maximum per fabric)
Number of ESXi hosts - VDS	N/A	3200
Number of ESXi hosts - AVS	N/A	3200 (Only 1 AVS instance per host)
Number of ESXi hosts - AVE	N/A	3200 (Only 1 AVE instance per host)
Number of VMs	N/A	Depends upon server scale
Number of configuration zones per fabric	30	30
Number of BFD sessions	256 per Leaf switch	256 per Leaf switch
Multi-Pod NOTE: * is preferred cluster size	3* or 4 node APIC cluster 6 pods 80 leaf switches overall	3* or 4 node APIC cluster 6 pods 80 for 3-node cluster 200 for 4-node cluster
L3 EVPN Services over Fabric WAN - GOLF (with and without OpFlex)	N/A	1000 VRFs, 60,000 routes in a fabric
Layer 3 Multicast routes	N/A	32,000

For detailed scalability report, please refer to Verified Scalability Guide for Cisco ACI and Cisco Nexus 9000 Series ACI-Mode Switches located at the URL below:

<https://www.cisco.com/c/en/us/td/docs/switches/datacenter/aci/apic/sw/4-x/verified-scalability/Cisco-ACI-Verified-Scalability-Guide-422.html>

Observations:

Below are the major components currently used in the ADP APIC:

Table 39: Overview of DC4 Configuration

Tenant	VRFs	BDs	Contracts	EPGs	L3Outs
common	2	1	1	0	1
infra	2	2	0	2	0
LEGACY	2	319	2	319	1
MDOM	1	11	7	11	1
mgmt	2	1	0	0	0

Table 40: Overview of DC5 Configuration

Tenant	VRFs	BDs	Contracts	EPGs	L3Outs
common	3	1	3	0	1
infra	2	2	0	2	0
LEGACY	2	312	2	312	1

January 21, 2021

Cisco Highly Confidential. All printed copies and duplicate soft copies are considered uncontrolled and the original online version should be referred to for the latest version.

MDOM	1	11	7	11	1
mgmt	2	1	0	0	0
Polly-Test	1	4	1	3	0

Leaf Capacity Compliance:



Note that the following utilization percentages are determined via the APIC API. The scalability guide contains many interdependent limits, e.g. BDs per VRF per leaf, many of which are not measured by this assessment. The scalability guides on CCO should serve as the final authority on whether your fabric is within current certified scaling limits.

13.2 DC4 Per-device scale limits



Microsoft Excel
Worksheet

13.3 DC5 Per-device scale limits



Microsoft Excel
Worksheet

Trademarks and Disclaimers

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR CISCO REPRESENTATIVE FOR A COPY.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THIRD PARTY SUPPLIERS ARE PROVIDED "AS IS" WITH ALL FAULTS. CISCO AND THIRD PARTY SUPPLIERS DISCLAIM ALL WARRANTIES, EXPRESSED 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.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Cisco and the Cisco Logo are trademarks of Cisco Systems, Inc. and/or its affiliates in the U.S. and other countries. A listing of Cisco's trademarks can be found at 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.

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

©2020 Cisco Systems, Inc. All rights reserved.

Document Acceptance

Name _____

Title _____

Company _____

Signature _____

Date _____

Name _____

Title _____

Company _____

Signature _____

Date _____

Name _____

Title _____

Company _____

Signature _____

Date _____

Name _____

Title _____

Company _____

Signature _____

Date _____

Name _____

Title _____

Company _____

Signature _____

Date _____

Name _____

Title _____

Company _____

Signature _____

Date _____