

Troubleshoot ACI Fabric Discovery - Device Replacement

Contents

[Introduction](#)

[Background Information](#)

[Overview](#)

[Procedures and verification](#)

[Hardware replacement](#)

[Leaf](#)

[Spine](#)

[APIC](#)

[IPN device replacement](#)

[Clean reload of APIC/leaf/spine](#)

[Troubleshooting scenarios](#)

[Problem: Arrives in NXOS mode](#)

[Verification](#)

[Solution](#)

[Problem: Leaf/Spine EPLD/FPGA not correct, F1582](#)

[Verification](#)

[EPLD notes](#)

[Solution](#)

Introduction

This document describes steps to understand and troubleshoot device replacement scenarios in ACI.

Background Information

The material from this document was extracted from the [Troubleshooting Cisco Application Centric Infrastructure, Second Edition](#) book, specifically the **Fabric Discovery - Device Replacement** chapter.

Overview

During the evolution of an ACI fabric, it will become necessary to replace various components including: APICs, leaf switches, spine switches, and IPN devices. The most common reasons for replacement include RMAs and hardware upgrades. These procedures are well documented in the Cisco Install/Upgrade guides and the most recent guide should be read prior to replacement. This section will include a deeper look into how the procedures work under the hood; as well as walk through several of the most common troubleshooting scenarios.

Procedures and verification

Note: As of ACI Switch version 5.2(3), NXOS switches plugged into a discovered ACI Fabric switch can use [POAP to convert into an ACI Switch](#).

Hardware replacement

Leaf

A leaf from the RMA depot will arrive running NXOS software. Please reference to the below section called 'Problem: Arrives in NXOS mode' to properly convert the leaf to ACI mode. If using a leaf from a different fabric or with previous configuration, make sure to use the commands 'acidiag touch clean' and 'reload'.

After the above steps are completed and the new leaf switch is ready for registration, remove the leaf to be replaced from the fabric via the 'Remove from Controller' option.

The 'Remove from Controller' option will completely remove the node from the APIC, releasing the node ID, SN association, and TEP address which was assigned by the APIC. These processes are desired when replacing a switch node. The 'Decommission' option is only used when the expectation is that the same node will rejoin the fabric with the same node ID and SN.

When the leaf switch to be replaced is no longer seen on the **Fabric Membership** page, the new leaf can be connected to the fabric via the spine interfaces. Once the leaf is discovered by the APIC, it will show up in the Fabric Inventory and be ready for registration. If the device to be replaced has not yet released its node ID, and a new switch is registered with the same node ID, a fault will be thrown referencing the fact that the ID is already associated to another leaf node. The fault should clear after some time. If the new node does not show up on the Fabric Membership submenu, there could be a cabling issue; this can be verified by viewing the LLDP neighbors via the 'show lldp neighbors detail' command on the spine switches connecting to the newly attached leaf switch. For more detail on the Fabric Discovery process, please reference the "Initial fabric setup" chapter.

If the infra VLAN is modified, all leaf nodes must be clean rebooted at the same time. If all leaf switches are not cleaned at the same time, a clean reloaded switch will come online and receive the old infra VLAN via LLDP from a not-yet-cleaned leaf, and the clean reloaded leaf will fail to register with the APIC. See the "Initial fabric setup" chapter for more details.

Due to platform limitations, VPC pairs cannot be a mix of Gen1 and Gen2 or higher leaf switches. However, at the time of writing, any Gen2 leaf and higher can mix with any other Gen2 leaf or higher.

Spine

Like a leaf, depending on the HW of the spine (such as modular spine) it could arrive in NXOS mode. Use the procedure "Problem: Arrives in NXOS mode" under the scenarios to perform the conversion.

When replacing a spine switch, the user must consider the **BGP Route Reflector** functionality. As a best practice there must be at least two spine switches configured as BGP Route Reflectors for

a Layer 3 Cisco ACI fabric. This configuration is located at 'System > System Settings > BGP Route Reflectors' under Route Reflector Nodes. When replacing or removing a spine switch, ensure the appropriate configuration changes are made to maintain one active Route Reflector, and ensure at least two active Route Reflectors after the changes are completed.

Refer to section "Pod Policies — BGP RR / Date&Time / SNMP" in chapter "Management and core services" for more information on the BGP Route Reflectors.

APIC

The most important consideration when performing an APIC replacement is the health of the existing APIC cluster. Prior to the replacement, all APICs in the cluster should be reported as Fully Fit. In 4.2, an additional tool was introduced to verify the health of the APIC cluster via CLI:

```
apicl# acidiag cluster
Admin password:
Product-name = APIC-SERVER-L2
Serial-number = FCH2206W0RK
Running...

Checking Core Generation: OK
Checking Wiring and UUID: OK
Checking AD Processes: Running
Checking All Apics in Commission State: OK
Checking All Apics in Active State: OK
Checking Fabric Nodes: OK
Checking Apic Fully-Fit: OK
Checking Shard Convergence: OK
Checking Leadership Degration: Optimal leader for all shards
Ping OOB IPs:
APIC-1: 192.168.4.20 - OK
Ping Infra IPs:
APIC-1: 10.0.0.1 - OK
Checking APIC Versions: Same (4.2(1i))
Checking SSL: OK

Done!
```

When replacing an APIC, make sure to note the initial setup variables of the APIC to be replaced, before performing a decommission of the APIC.

```
apicl# cat /data/data_admin/sam_exported.config
Setup for Active and Standby APIC
fabricDomain = POD37
fabricID = 1
systemName =apicl
controllerID = 1
tepPool = 10.0.0.0/16
infraVlan = 3937
GIPO = 225.0.0.0/15
clusterSize = 3
standbyApic = NO
enableIPv4 = Y
enableIPv6 = N
firmwareVersion = 4.2(1i)
ifcIpAddr = 10.0.0.1
apicX = NO
podId = 1
```

oobIpAddr = 10.48.176.57/24

Prepare the new APIC with the correct software version and re-enter the initial setup values referenced earlier. When the initial setup is complete and the APIC is fully booted, recommission it to the fabric from UI of one of the other APICs in the cluster.

IPN device replacement

In a Multi-Pod environment, it might be necessary to replace one of the devices being used for the IPN (Inter-Pod Network). Prior to the replacement, the IPN network must have **PIM Bidirectional Rendezvous Point Redundancy** configured in the form of **Phantom RPs**. Without Phantom RPs in place, if the node replaced was the RP, there would be a PIM convergence and packet loss would be seen for all BUM traffic sent across the IPN.

Please refer to "RP configuration" in "Multi-Pod Discovery" chapter for more information on how to configure Phantom RP.

Clean reload of APIC/leaf/spine

In certain scenarios, the best option for recovering a leaf/spine that won't join the fabric is to perform a clean reload of the device.

It is not recommended to perform a clean reload on a device that is waiting for its turn to upgrade. Clean reload of any device can take an extended period of time.

The 'acidiag touch' command has two options, clean and setup. The **clean** option removes all policy data while retaining the APIC network configuration (such as fabric name, IP address, login). The **setup** option removes both policy data and the APIC network configuration. The setup option is most commonly used when moving devices across Pods, as the Pod ID must be changed, and normally the management network will need to update as well.

APIC

```
fab1-apic1# acidiag touch clean  
This command will wipe out this device, Proceed? [y/N] y
```

```
fab1-apic1# acidiag reboot  
This command will restart this device, Proceed? [y/N] y
```

Leaf/Spine

```
fab1-leaf101# acidiag touch clean  
This command will wipe out this device, Proceed? [y/N] y
```

```
fab1-leaf101# reload  
This command will reload the chassis, Proceed (y/n)? [n]: y
```

The 'acidiag touch clean' command works by putting a hidden file on the leaf in /mnt/pss called .clean. When the leaf is booted, a shell script runs that checks to see if .clean file is present. In the event that .clean file exists under /mnt/pss, policy configuration is wiped and configuration is redownloaded from the APIC. If this command is entered and the node is not reloaded, then the file will still be present and the policy will still be wiped upon the next reload, no matter how much time has elapsed since the touch clean was entered.

Troubleshooting scenarios

Problem: Arrives in NXOS mode

Verification

Sometimes when a switch is shipped via RMA, it can arrive with NXOS software that has not yet been configured via the Power On Auto Provisioning (POAP) process. When the user consoles into this device they will see some form of the following message:

Abort Auto Provisioning and continue with normal setup?(yes/no)

If the device has already gone through POAP, the simplest way to determine if a leaf is running standalone NXOS code is to look for the 'NXOS image file' line in the 'show version' output. If such output is present, the leaf is running standalone code and will need to be converted to ACI mode. The presence of Kickstart and system images can be verified and will only be present on a leaf running an ACI image, by looking at the image itself, which will be n9000 on standalone and aci-n9000 on ACI.

Standalone NXOS

```
nxos-n9k# show version
Cisco Nexus Operating System (NX-OS) Software
.
.
.
Software
  BIOS: version 07.17
  NXOS: version 6.1(2)I3(4)
  BIOS compile time: 09/10/2014
  NXOS image file is: bootflash:///n9000-dk9.6.1.2.I3.4.bin
  NXOS compile time: 3/18/2015 0:00:00 [03/18/2015 07:49:10]
```

ACI

```
aci-leaf101# show version
Cisco Nexus Operating System (NX-OS) Software
.
.
.
Software
  BIOS: version 07.66
  kickstart: version 14.2(1i) [build 14.2(1i)]
  system: version 14.2(1i) [build 14.2(1i)]
  PE: version 4.2(1i)
  BIOS compile time: 06/11/2019
  kickstart image file is: /bootflash/aci-n9000-dk9.14.2.1i.bin
  kickstart compile time: 09/07/2019 10:25:16 [09/07/2019 10:25:16]
  system image file is: /bootflash/auto-s
  system compile time: 09/07/2019 10:25:16 [09/07/2019 10:25:16]
```

Solution

If the switch was shipped running NXOS code, it will need to be converted to ACI mode. The

switch should be shipped with both the NXOS and the ACI image in the bootflash, although this is not always the case. The ACI image will start with 'aci-n9000'. If the ACI image is not present, then it will need to be manually loaded onto the bootflash. This can be performed via the USB connection (local access needed) or via SCP from the APIC directly (assuming both devices are connected via a management network). Here are the instructions to copy the image via SCP:

```
1. nexus-9000(config)# feature scp-server
2. apic1# scp -r /firmware/fwrepos/fwrepo/switch-image-name admin@standalone_switch:switch-image-name
```

The leaf will then need to be configured to not boot the NXOS image, save the configuration, change the boot statements to ACI.

```
1. (config)# no boot nxos
2. (config)# copy run start
3. (config)# boot aci bootflash:<aci-image-name>
4. (config)# reload
```

Problem: Leaf/Spine EPLD/FPGA not correct, F1582

Verification

The following faults will be seen in the Faults for the Nexus 9000 ACI switch.

F1582 FPGA version mismatch detected. Running version:0x(z) Expected version:0x(y)

From the APIC CLI, search for all instances of Fault F1582:

```
apic1# moquery -c faultInst -f 'fault.Inst.code=="F1582"'
```

EPLD notes

The Cisco Nexus 9000 Series ACI-mode switches contain several programmable logical devices (PLDs) that provide hardware functionalities in all modules. Cisco provides electronic programmable logic device (EPLD) image upgrades to enhance hardware functionality or to resolve known issues. PLDs include electronic programmable logic devices (EPLDs), field programmable gate arrays (FPGAs), and complex programmable logic devices (CPLDs), but they do not include ASICs.

The term EPLD is used to cover both FPGA and CPLDs.

The advantage of having EPLDs for some module functions is that when those functions need to be upgraded, just upgrade their software images instead of replacing their hardware.

EPLD image upgrades for an I/O module disrupt the traffic going through the module because the module must power down briefly during the upgrade. In a modular chassis, the system performs EPLD upgrades on one module at a time, so at any one time the upgrade disrupts only the traffic going through one module.

Cisco provides the latest EPLD images with each release. Typically, these images are the same as provided in earlier releases but occasionally some of these images are updated. These EPLD image updates are not mandatory unless otherwise specified. When Cisco makes an EPLD image upgrade available, these release notes announce their availability, and they can be downloaded

from the Cisco web site.

When new EPLD images are available, the upgrades are always recommended if the network environment allows for a maintenance period in which some level of traffic disruption is acceptable. In general, EPLD upgrades will be needed when new hardware functionality is added as a result of a software upgrade.

There may also be various reasons for the need to upgrade EPLD firmware while already in ACI Mode:

1. EPLD versions required an upgrade prior to a Cisco NX-OS to ACI Boot Mode conversion and the FPGA/EPLDs were NOT upgraded.
2. Leaf/Spine was upgraded manually (instead of a policy upgrade from the APIC), which does not include an EPLD upgrade.

Once the leaf or spine is added to the fabric, then the EPLD will be automatically upgraded with any policy upgrade (normal upgrade initiated from the APIC firmware tab) where a new version of EPLD is available.

Solution

In older versions of ACI, it was necessary to downgrade and then upgrade the leaf/spine in question, but as of 11.2(1m), there are two shell scripts available to the admin user which greatly simplify the process.

```
fab1-leaf101# /bin/check-fpga.sh FpGaDoWnGrAdE
fab1-leaf101# /usr/sbin/chassis-power-cycle.sh
```

The '/usr/sbin/chassis-power-cycle.sh' script hard resets power, as compared to a 'reload' which is simply a software restart. When upgrading EPLD, the power needs to be removed entirely to reprogram the firmware on the line cards. If '/usr/sbin/chassis-power-cycle.sh' is not available or does not work, power cables need to be removed for at least 30 seconds and then re-attached to restore power.