Basic Operations

Troubleshooting APIC Crash Scenarios

Cluster Troubleshooting Scenarios

The following table summarizes common cluster troubleshooting scenarios for the Cisco APIC.

Problem	Solution
An APIC node fails within the cluster. For example, node 2 of a cluster of 5 APICs fails.	There are two available solutions: • Leave the target size and replace the APIC. • Reduce the cluster size to 4, decommission controller 5, and recommission it as APIC 2.The target size remains 4, and the operational size is 4 when the reconfigured APIC becomes active. **Note** You can add a replacement APIC to the cluster and expand the target and operational size. For instructions on how to add a new APIC, refer to the *Cisco APIC Management, Installation, Upgrade, and Downgrade Guide*..

Problem	Solution
A new APIC connects to the fabric and loses connection to a leaf switch.	Use the following commands to check for an infra (infrastructure) VLAN mismatch: • cat /mit/sys/lldp/inst/if-\[eth1--1\]/ctrlradj/summary—Displays the VLAN configured on the leaf switch. • cat /mit/sys/lldp/inst/if-\[eth1--1\]/ctrlradj/summary—Displays the infra (infrastructure) VLANs advertised by connected APICs. If the output of these commands shows different VLANs, the new APIC is not configured with the correct infra (infrastructure) VLAN. To correct this issue, follow these steps: • Log in to the APIC using rescue-user. **Note** Admin credentials do not work because the APIC is not part of the fabric. • Erase the configuration and reboot the APIC using the **acidiag touch setup** command. • Reconfigure the APIC. Verify that the fabric name, TEP addresses, and infra (infrastructure) VLAN match the APICs in the cluster. • Reload the leaf node.
Two APICs cannot communicate after a reboot.	The issue can occur after the following sequence of events: • APIC1 and APIC2 discover each other. • APIC1 reboots and becomes active with a new ChassisID (APIC1a) • The two APICs no longer communicate. In this scenario, APIC1a discovers APIC2, but APIC2 is unavailable because it is in a cluster with APIC1, which appears to be offline. As a result, APIC1a does not accept messages from APIC2. To resolve the issue, decommission APIC1 on APIC2, and commission APIC1 again.
A decommissioned APIC joins a cluster.	The issue can occur after the following sequence of events: • A member of the cluster becomes unavailable or the cluster splits. • An APIC is decommissioned. • After the cluster recovers, the decommissioned APIC is automatically commissioned. To resolve the issue, decommission the APIC after the cluster recovers.

Problem	Solution
Mismatched ChassisID following reboot.	The issue occurs when an APIC boots with a ChassisID different from the ChassisID registered in the cluster. As a result, messages from this APIC are discarded. To resolve the issue, ensure that you decommission the APIC before rebooting.
The APIC displays faults during changes to cluster size.	A variety of conditions can prevent a cluster from extending the OperationalClusterSize to meet the AdminstrativeClusterSize. For more information, inspect the fault and review the "Cluster Faults" section in the *Cisco APIC Basic Configuration Guide*.
An APIC is unable to join a cluster.	The issue occurs when two APICs are configured with the same ClusterID when a cluster expands. As a result, one of the two APICs cannot join the cluster and displays an expansion-contender-chassis-id-mismatch fault. To resolve the issue, configure the APIC outside the cluster with a new cluster ID.
APIC unreachable in cluster.	Check the following settings to diagnose the issue: • Verify that fabric discovery is complete. • Identify the switch that is missing from the fabric. • Check whether the switch has requested and received an IP address from an APIC. • Verify that the switch has loaded a software image. • Verify how long the switch has been active. • Verify that all processes are running on the switch. For more information, see the "acidiag Command" section in the *Cisco APIC Basic Configuration Guide*. • Confirm that the missing switch has the correct date and time. • Confirm that the switch can communicate with other APICs.

Problem	Solution
Cluster does not expand.	The issue occurs under the following circumstances: • The OperationalClusterSize is smaller than the number of APICs. • No expansion contender (for example, the admin size is 5 and there is not an APIC with a clusterID of 4. • There is no connectivity between the cluster and a new APIC • Heartbeat messages are rejected by the new APIC • System is not healthy. • An unavailable appliance is carrying a data subset that is related to relocation. • Service is down on an appliance with a data subset that is related to relocation. • Unhealthy data subset related to relocation.
An APIC is down.	Check the following: • Connectivity issue—Verify connectivity using ping. • Interface type mismatch—Confirm that all APICs are set to in-band communication. • Fabric connectivity—Confirm that fabric connectivity is normal and that fabric discovery is complete. • Heartbeat rejected—Check the fltInfraIICIMsgSrcOutsider fault. Common errors include operational cluster size, mismatched ChassisID, source ID outside of the operational cluster size, source not commissioned, and fabric domain mismatch.

Cluster Faults

The APIC supports a variety of faults to help diagnose cluster problems. The following sections describe the two major cluster fault types.

Discard Faults

The APIC discards cluster messages that are not from a current cluster peer or cluster expansion candidate. If the APIC discards a message, it raises a fault that contains the originating APIC's serial number, cluster ID, and a timestamp. The following table summarizes the faults for discarded messages:

Fault	Meaning
expansion-contender-chassis-id-mismatch	The ChassisID of the transmitting APIC does not match the ChassisID learned by the cluster for expansion.
expansion-contender-fabric-domain-mismatch	The FabricID of the transmitting APIC does not match the FabricID learned by the cluster for expansion.

Fault	Meaning
expansion-contender-id-is-not-next-to-oper-cluster-size	The transmitting APIC has an inappropriate cluster ID for expansion. The value should be one greater than the current OperationalClusterSize.
expansion-contender-message-is-not-heartbeat	The transmitting APIC does not transmit continuous heartbeat messages.
fabric-domain-mismatch	The FabricID of the transmitting APIC does not match the FabricID of the cluster.
operational-cluster-size-distance-cannot-be-bridged	The transmitting APIC has an OperationalClusterSize that is different from that of the receiving APIC by more than 1. The receiving APIC rejects the request.
source-chassis-id-mismatch	The ChassisID of the transmitting APIC does not match the ChassisID registered with the cluster.
source-cluster-id-illegal	The transmitting APIC has a clusterID value that is not permitted.
source-has-mismatched-target-chassis-id	The target ChassisID of the transmitting APIC does not match the Chassis ID of the receiving APIC.
source-id-is-outside-operational-cluster-size	The transmitting APIC has a cluster ID that is outside of the OperationalClusterSize for the cluster.
source-is-not-commissioned	The transmitting APIC has a cluster ID that is currently decommissioned in the cluster.

Cluster Change Faults

The following faults apply when there is an error during a change to the APIC cluster size.

Fault	Meaning
cluster-is-stuck-at-size-2	This fault is issued if the OperationalClusterSize remains at 2 for an extended period. To resolve the issue, restore the cluster target size.
most-right-appliance-remains-commissioned	The last APIC within a cluster is still in service, which prevents the cluster from shrinking.
no-expansion-contender	The cluster cannot detect an APIC with a higher cluster ID, preventing the cluster from expanding.
service-down-on-appliance-carrying-replica-related-to-relocation	The data subset to be relocated has a copy on a service that is experiencing a failure. Indicates that there are multiple such failures on the APIC.
unavailable-appliance-carrying-replica-related-to-relocation	The data subset to be relocated has a copy on an unavailable APIC. To resolve the fault, restore the unavailable APIC.
unhealthy-replica-related-to-relocation	The data subset to be relocated has a copy on an APIC that is not healthy. To resolve the fault, determine the root cause of the failure.

APIC Unavailable

The following cluster faults can apply when an APIC is unavailable:

Fault	Meaning
fltInfraReplicaReplicaState	The cluster is unable to bring up a data subset.
fltInfraReplicaDatabaseState	Indicates a corruption in the data store service.
fltInfraServiceHealth	Indicates that a data subset is not fully functional.
fltInfraWiNodeHealth	Indicates that an APIC is not fully functional.

Troubleshooting Fabric Node and Process Crash

The ACI switch node has numerous processes which control various functional aspects on the system. If the system has a software failure in a particular process, a core file will be generated and the process will be reloaded.

If the process is a Data Management Engine (DME) process, the DME process will restart automatically. If the process is a non-DME process, it will not restart automatically and the switch will reboot to recover.

This section presents an overview of the various processes, how to detect that a process has cored, and what actions should be taken when this occurs

DME Processes

The essential processes running on an APIC can be found through the CLI. Unlike the APIC, the processes that can be seen via the GUI in **FABRIC** > **INVENTORY** > **Pod 1** > ***node*** shows all processes running on the leaf.

Through the **ps-ef | grep svc_ifc**:

```
rtp_leaf1# ps -ef |grep svc_ifc
root 3990 3087 1 Oct13 ? 00:43:36 /isan/bin/svc_ifc_policyelem --x
root 4039 3087 1 Oct13 ? 00:42:00 /isan/bin/svc_ifc_eventmgr --x
root 4261 3087 1 Oct13 ? 00:40:05 /isan/bin/svc_ifc_opflexelem --x -v
dptcp:8000
root 4271 3087 1 Oct13 ? 00:44:21 /isan/bin/svc_ifc_observerelem --x
root 4277 3087 1 Oct13 ? 00:40:42 /isan/bin/svc_ifc_dbgrelem --x
root 4279 3087 1 Oct13 ? 00:41:02 /isan/bin/svc_ifc_confelem --x
rtp_leaf1#
```

Each of the processes running on the switch writes activity to a log file on the system. These log files are bundled as part of the techsupport file but can be found via CLI access in /tmp/logs/ directory. For example, the Policy Element process log output is written into /tmp/logs/svc_ifc_policyelem.log.

The following is a brief description of the DME processes running on the system. This can help in understanding which log files to reference when troubleshooting a particular process or understand the impact to the system if a process crashed:

Process	Function
policyelem	Policy Element: Process logical MO from APIC and push concrete model to the switch
eventmgr	Event Manager: Processes local faults, events, health score
opflexelem	Opflex Element: Opflex server on switch

Process	Function
observerelem	Observer Element: Process local stats sent to APIC
dbgrelem	Debugger Element: Core handler
nginx	Web server handling traffic between the switch and APIC

Identify When a Process Crashes

When a process crashes and a core file is generated, a fault as well as an event is generated. The fault for the particular process is shown as a "process-crash" as shown in this syslog output from the APIC:

```
Oct 16 03:54:35 apic3 %LOG_LOCAL7-3-SYSTEM_MSG [E4208395][process-crash][major]
[subj-[dbgs/cores/node-102-card-1-svc-policyelem-ts-2014-10-16T03:54:55.000+00:00]/
rec-12884905092]Process policyelem cored
```

When the process on the switch crashes, the core file is compressed and copied to the APIC. The syslog message notification comes from the APIC.

The fault that is generated when the process crashes is cleared when the process is Troubleshooting Cisco Application Centric Infrastructure 275 restarted. The fault can be viewed via the GUI in the fabric history tab at **FABRIC** > **INVENTORY** > **Pod 1**.

Collecting the Core Files

The APIC GUI provides a central location to collect the core files for the fabric nodes.

An export policy can be created from **ADMIN** > **IMPORT/EXPORT** > **Export Policies** > **Core**. However, there is a default core policy where files can be downloaded directly.

The core files can be accessed via SSH/SCP through the APIC at /data/techsupport on the APIC where the core file is located. Note that the core file will be available at /data/ techsupport on one APIC in the cluster, the exact APIC that the core file resides can be found by the Export Location path as shown in the GUI. For example, if the Export Location begins with “files/3/”, the file is located on node 3 (APIC3).

APIC Process Crash Verification and Restart

Symptom 1

Process on switch fabric crashes. Either the process restarts automatically or the switch reloads to recover.

- **Verification**:

  As indicated in the overview section, if a DME process crashes, it should restart automatically without the switch restarting. If a non-DME process crashes, the process will not automatically restart and the switch will reboot to recover.

  Depending on which process crashes, the impact of the process core will vary.

  When a non-DME process crashes, this will typical lead to a HAP reset as seen on the console:

  ```
  [ 1130.593388] nvram_klm wrote rr=16 rr_str=ntp hap reset to nvram
  [ 1130.599990] obfl_klm writing reset reason 16, ntp hap reset
  [ 1130.612558] Collected 8 ext4 filesystems
  ```

- **Check Process Log**:

  The process which crashes should have at some level of log output prior to the crash. The output of the logs on the switch are written into the /tmp/logs directory. The process name will be part of the file name. For example, for the Policy Element process, the file is svc_ifc_policyelem.log

```
rtp_leaf2# ls -l |grep policyelem
-rw-r--r-- 2 root root 13767569 Oct 16 00:37 svc_ifc_policyelem.log
-rw-r--r-- 1 root root 1413246 Oct 14 22:10 svc_ifc_policyelem.log.1.gz
-rw-r--r-- 1 root root 1276434 Oct 14 22:15 svc_ifc_policyelem.log.2.gz
-rw-r--r-- 1 root root 1588816 Oct 14 23:12 svc_ifc_policyelem.log.3.gz
-rw-r--r-- 1 root root 2124876 Oct 15 14:34 svc_ifc_policyelem.log.4.gz
-rw-r--r-- 1 root root 1354160 Oct 15 22:30 svc_ifc_policyelem.log.5.gz
-rw-r--r-- 2 root root 13767569 Oct 16 00:37 svc_ifc_policyelem.log.6
-rw-rw-rw- 1 root root 2 Oct 14 22:06 svc_ifc_policyelem.log.PRESERVED
-rw-rw-rw- 1 root root 209 Oct 14 22:06 svc_ifc_policyelem.log.stderr
rtp_leaf2#
```

  There will be several files for each process located at /tmp/logs. As the log file increases in size, it will be compressed and older log files will be rotated off. Check the core file creation time (as shown in the GUI and the core file name) to understand where to look in the file. Also, when the process first attempts to come up, there be an entry in the log file that indicates “Process is restarting after a crash” that can be used to search backwards as to what might have happened prior to the crash.

- **Check Activity**:

  A process which has been running has had some change which then caused it to crash. In many cases the changes may have been some configuration activity on the system. What activity occurred on the system can be found in the audit log history of the system.

- **Contact TAC**:

  A process crashing should not normally occur. In order to understand better why beyond the above steps it will be necessary to decode the core file. At this point, the file will need to be collected and provided to the TAC for further processing.

  Collect the core file (as indicated above how to do this) and open up a case with the TAC.

Symptom 2

Fabric switch continuously reloads or is stuck at the BIOS loader prompt.

- **Verification**:

  If a DME process crashes, it should restart automatically without the switch restarting. If a non-DME process crashes, the process will not automatically restart and the switch will reboot to recover. However in either case if the process continuously crashes, the switch may get into a continuous reload loop or end up in the BIOS loader prompt.

```
[ 1130.593388] nvram_klm wrote rr=16 rr_str=policyelem hap reset to nvram
[ 1130.599990] obfl_klm writing reset reason 16, policyelem hap reset
[ 1130.612558] Collected 8 ext4 filesystems
```

- **Break the HAP Reset Loop**:

  First step is to attempt to get the switch back into a state where further information can be collected.

  If the switch is continuously rebooting, when the switch is booting up, break into the BIOS loader prompt through the console by typing CTRL C when the switch is first part of the boot cycle.

  Once the switch is at the loader prompt, enter in the following commands:

- cmdline no_hap_reset
- boot

The cmdline command will prevent the switch from reloading with a hap reset is called. The second command will boot the system. Note that the boot command is needed instead of a reload at the loader as a reload will remove the cmdline option entered.

Though the system should now remain up to allow better access to collect data, whatever process is crashing will impact the functionality of the switch.

As in the previous table, check the process log, activity, and contact TAC steps.

Troubleshooting an APIC Process Crash

The APIC has a series of Data Management Engine (DME) processes which control various functional aspects on the system. When the system has a software failure in a particular process, a core file will be generated and the process will be reloaded.

The following sections cover potential issues involving system processes crashes or software failures, beginning with an overview of the various system processes, how to detect that a process has cored, and what actions should be taken when this occurs. The displays taken on a working healthy system can then be used to identify processes that may have terminated abruptly.

DME Processes

The essential processes running on an APIC can be found either through the GUI or the CLI. Using the GUI, the processes and the process ID running is found in **System** > **Controllers** > **Processes**.

Using the CLI, the processes and the process ID are found in the summary file at /aci/system/controllers/1/processes (for APIC1):

```
admin@RTP_Apic1:processes> cat summary
processes:
process-id process-name max-memory-allocated state
---------- ---------------- ------------------- -------------------
0 KERNEL 0 interruptible-sleep
331 dhcpd 108920832 interruptible-sleep
336 vmmmgr 334442496 interruptible-sleep
554 neo 398274560 interruptible-sleep
1034 ae 153690112 interruptible-sleep
1214 eventmgr 514793472 interruptible-sleep
2541 bootmgr 292020224 interruptible-sleep
4390 snoopy 28499968 interruptible-sleep
5832 scripthandler 254308352 interruptible-sleep
19204 dbgr 648941568 interruptible-sleep
21863 nginx 4312199168 interruptible-sleep
32192 appliancedirector 136732672 interruptible-sleep
32197 sshd 1228800 interruptible-sleep
32202 perfwatch 19345408 interruptible-sleep
32203 observer 724484096 interruptible-sleep
32205 lldpad 1200128 interruptible-sleep
32209 topomgr 280576000 interruptible-sleep
32210 xinetd 99258368 interruptible-sleep
32213 policymgr 673251328 interruptible-sleep
32215 reader 258940928 interruptible-sleep
32216 logwatch 266596352 interruptible-sleep
32218 idmgr 246824960 interruptible-sleep
```

```
32416 keyhole 15233024 interruptible-sleep
admin@apic1:processes>
```

Each of the processes running on the APIC writes to a log file on the system. These log files can be bundled as part of the APIC techsupport file but can also be observed through SSH shell access in /var/log/dme/log. For example, the Policy Manager process log output is written into /var/log/dme/log/svc_ifc_policymgr.bin.log.

The following is a brief description of the processes running on the system. This can help in understanding which log files to reference when troubleshooting a particular process or understand the impact to the system if a process crashed:

Process	Function
KERNEL	Linux kernel
dhcpd	DHCP process running for APIC to assign infra addresses
vmmmgr	Handles process between APIC and Hypervisors
neo	Shell CLI Interpreter
ae	Handles the state and inventory of local APIC appliance
eventmgr	Handles all events and faults on the system
bootmgr	Controls boot and firmware updates on fabric nodes
snoopy	Shell CLI help, tab command completion
scripthandler	Handles the L4-L7 device scripts and communication
dbgr	Generates core files when process crashes
nginx	Web service handling GUI and REST API access
appliancedirector	Handles formation and control of APIC cluster
sshd	Enabled SSH access into the APIC
perfwatch	Monitors Linux cgroup resource usage
observer	Monitors the fabric system and data handling of state, stats, health
lldpad	LLDP Agent
topomgr	Maintains fabric topology and inventory

Cisco APIC Troubleshooting Operations

Shutting Down the Cisco APIC System

This procedure shuts down the Cisco Application Policy Infrastructure Controller (APIC) system. After you shut down the system, you will relocate the entire fabric and power it up, then update the time zone and/or NTP servers accordingly.

Before you begin

Ensure cluster health is fully fit.

Procedure

Step 1 On the menu bar, choose **System** > **Controllers**.
Step 2 In the Navigation pane, choose **Controllers** > ***apic_name***.
Step 3 Right-click the Cisco APIC and choose **Shutdown**.
Step 4 Relocate the Cisco APIC, then power it up.
Step 5 Confirm that the cluster has fully converged.
Step 6 Repeat this procedure for the next Cisco APIC.

Shutting Down a Cisco APIC Using the GUI

This procedure shuts down a Cisco Application Policy Infrastructure Controller (APIC). This procedure shuts down only one Cisco APIC, not the entire Cisco APIC system itself. Following this procedure causes the controller to shut down immediately. Use caution in performing a shutdown because the only way to bring the controller back up is to do so from the actual machine. If you need to access the machine, see Controlling the LED Locator Using the GUI, on page 12.

Note If possible, move Cisco APICs one at a time. As long as there are at least two Cisco APICs in the cluster online, there is read/write access. If you need to relocate more than one Cisco APIC at a time, this results in one or no remaining controllers online, and the fabric will go into a read-only mode when they are shut down. During this time, there can be no policy changes including endpoint moves (which includes virtual machine movement).

Procedure

Step 1 On the menu bar, choose **System** > **Controllers**.
Step 2 In the Navigation pane, choose **Controllers** > ***apic_name***.

Step 3 Right-click the Cisco APIC and choose **Shutdown**.

Step 4 Relocate the Cisco APIC, then power it up.

Step 5 Confirm that the cluster has fully converged.

Using the APIC Reload Option Using the GUI

This procedure reloads the Cisco Application Policy Infrastructure Controller (APIC), not the entire Cisco APIC system, using the GUI.

Procedure

Step 1 On the menu bar, choose **System** > **Controllers**.

Step 2 In the Navigation pane, choose **Controllers** > ***apic_name***.

Step 3 Right-click the Cisco APIC and choose **Reload**.

Controlling the LED Locator Using the GUI

This procedure turns on or off the LED locator for the Cisco Application Policy Infrastructure Controller (APIC) using the GUI.

Procedure

Step 1 On the menu bar, choose **System** > **Controllers**.

Step 2 In the Navigation pane, choose **Controllers** > ***apic_name***.

Step 3 Right-click the Cisco APIC and choose **Turn On Locator LED** or **Turn On Locator LED** as appropriate.

Powering Down the Fabric Using the GUI

This procedure powers down the fabric for power maintenance using the Cisco Application Policy Infrastructure Controller (APIC) GUI and Cisco Integrated Management Controller (IMC) GUI.

Procedure

Step 1 Shut down all Cisco APICs except the last one using the Cisco APIC GUI.

a) Log in to a Cisco APIC.

b) On the menu bar, choose **System** > **Controllers**.

c) Choose one of the Cisco APICs. In the Navigation pane, choose **Controllers** > ***apic_name***.

d) Right-click the Cisco APIC and choose **Shutdown**.
e) Repeat steps 1.c, on page 12 and 1.d, on page 13 for all other Cisco APICs, except the last one.

Step 2 Shut down the last Cisco APIC using the Cisco IMC GUI.

a) Log in to the Cisco IMC GUI of the last Cisco APIC.
b) In the Navigation pane, click the **Chassis** menu.
c) In the **Chassis** menu, choose **Summary**.
d) In the toolbar above the work pane, choose **Host Power** > **Shut Down**.

You must shut down the last Cisco APIC using the Cisco IMC GUI because this server will be in the read-only mode and will not process a shutdown request through Cisco APIC GUI.

Step 3 After you shut down all Cisco APICs, power down the switches by turning off their power supply.

Powering Up the Fabric Using the GUI

This procedure powers up the fabric using the Cisco Integrated Management Controller (IMC) GUI.

Procedure

Step 1 Power on Cisco APICs using the Cisco IMC GUI.

a) Log in to the Cisco IMC GUI of a Cisco APIC.
b) In the Navigation pane, click the **Chassis** menu.
c) In the **Chassis** menu, choose **Summary**.
d) In the toolbar above the work pane, choose **Host Power** > **Power On**.
e) Repeat these substeps for all Cisco APICs.

Step 2 Power on the leaf switches that are connected directly to the Cisco APIC.

Step 3 Power on the spine switches approximately a minute after you powered on the leaf switches.

Step 4 Power on the remaining leaf switches in the fabric.

The Cisco APICs discover the leaf switches directly connected to them through LLDP, followed by discovering the spine switches and remaining leaf switches. The discovery is automatic because the Cisco APICs retain the configurations and fabric memberships across reloads and shutdowns. The cluster comes up in the fully fit state after the Cisco APICs discover all leaf switches connected to them and discover the spine switches.

Switch Operations

Manually Removing Disabled Interfaces and Decommissioned Switches from the GUI

In a scenario where a fabric port is shut down then brought back up, it is possible that the port entry will remain disabled in the GUI. If this occurs, no operations can be performed on the port. To resolve this, you must manually remove the port from the GUI.

Procedure

Step 1 From the **Fabric** tab, click **Inventory**.

Step 2 In the **Navigation** pane, click **Disabled Interfaces and Decommissioned Switches**.
The list of disabled interfaces and decommissioned switches appears in a summary table in the **Work** pane.

Step 3 From the **Work** pane, right-click on the interface or switch that you want to remove and choose **Delete**.

Decommissioning and Recommissioning Switches

To decommission and recommission all the nodes in a pod, perform this procedure. One use case for this is to change the node IDs to a more logical, scalable numbering convention.

Procedure

Step 1 Decommission the nodes in the pod by following these steps for each one:

a) Navigate to **Fabric** > **Inventory** and expand the **Pod**.

b) Select the switch, right-click on it, and choose **Remove from Controller**.

c) Confirm the action and click **OK**.

The process takes about 10 minutes. The node is automatically wiped and reloaded. In addition, the node configuration is removed from the controller.

d) If a decommissioned node had the port profile feature deployed on it, some port configurations are not removed with the rest of the configuration. It is necessary to manually delete the configurations after the decommission for the ports to return to the default state. To do this, log on to the switch, run the **setup-clean-config.sh** script, and wait for it to run. Then, enter the **reload** command.

Step 2 When all the switches have been decommissioned from the pod, verify they are all physically connected and booted in the desired configuration.

Step 3 Perform the following actions to recommission each node.

Note

Before recommissioning a node with a port profile configuration as a new node, for the 6.0(7) and earlier releases, you must run the **setup-clean-config.sh script**, or for the 6.0(8) and later releases you must run the **setup-clean-config.sh -d** script to restore the port configuration to the default settings.

a) Navigate to **Fabric** > **Inventory**, expand **Quick Start**, and click **Node or Pod Setup**.
b) Click **Setup Node**.
c) In the **Pod ID** field, choose the pod ID.
d) Click the + to open the **Nodes** table.
e) Enter the node ID, serial number, Switch name, TEP Pool ID, and Role (**leaf** or **spine**) for the switch.
f) Click **Update**.

Step 4 Verify the nodes are all set up by navigating to **Fabric** > **Inventory** > **Fabric Membership**.

What to do next

If the pod is one of the pods in a multipod topology, reconfigure multipod for this pod and the nodes. For more information, see *Multipod* in the *Cisco APIC Layer 3 Networking Configuration Guide*.

Clean Reloading a Cisco ACI-Mode Switch

This procedure performs a clean reload of Cisco ACI-mode switches. A clean reload erases the configuration on the switch. After the switch boots up, the switch gets its configuration from the Cisco Application Policy Infrastructure Controller (APIC).

Procedure

Step 1 Log into a switch that you want to clean reload.

Step 2 Run the **setup-clean-config.sh** script with the **-k** argument.

Example:

```
switch1# setup-clean-config.sh -k
```

Step 3 Reload the switch.

Example:

```
switch1# reload
```

Recovering a Disconnected Leaf

If all fabric interfaces on a leaf are disabled (interfaces connecting a leaf to the spine) due to a configuration pushed to the leaf, connectivity to the leaf is lost forever and the leaf becomes inactive in the fabric. Trying to push a configuration to the leaf does not work because connectivity has been lost. This chapter describes how to recover a disconnected leaf.

Recovering a Disconnected Leaf Using the NX-OS-Style CLI

This this procedure enables fabric interfaces using the Cisco Application Policy Infrastructure Controller (APIC) NX-OS-style CLI. Use this procedure if you do not have any external tools from which you can make REST API calls.

Note This procedure assumes that 1/31 is one of the leaf switch ports connecting to the spine switch.

Procedure

Step 1 Using Cisco APIC NX-OS-style CLI, remove the block list policy.

Example:

```
apic1# podId='1'
apic1# nodeId='103'
apic1# interface='eth1/31'
apic1# icurl -sX POST 'http://127.0.0.1:7777/api/mo/.json' -d '{"fabricRsOosPath":{"attributes":

{"dn":"uni/fabric/outofsvc/rsoosPath-[topology/pod-'$podId'/paths-'$nodeId'/pathep-['$interface']]","status":"deleted"}}}'
```

Step 2 Using the CLI of a leaf or spine switch, set the port in service to bring up the port on the leaf switch.

Example:

```
switch1# podId='1'
switch1# nodeId='103'
switch1# interface='eth1/31'
switch1# icurl -X POST
'http://127.0.0.1:7777/api/node/mo/topology/pod-'$podId'/node-'$nodeId'/sys/action.json'
  -d
'{"actionLSubj":{"attributes":{"oDn":"sys/phys-['$interface']"},"children":[{"l1EthIfSetInServiceLTask":

  {"attributes":{"adminSt":"start"}}}]}}'
```

Recovering a Disconnected Leaf Using the REST API

To recover a disconnected leaf switch, you must enable at least one of the fabric interfaces using this procedure. You can enable the remaining interfaces using the GUI, REST API, or CLI.

To enable the first interface, post a policy using the REST API to delete the policy posted and bring the fabric ports Out-of-Service. You can post a policy to the leaf switch to bring the port that is Out-of-Service to In-Service as follows:

Note This procedure assumes that 1/49 is one of the leaf switch ports connecting to the spine switch.

Procedure

Step 1 Clear the block list policy from the Cisco APIC using the REST API.

Example:

```
$APIC_Address/api/policymgr/mo/.xml
<polUni>
    <fabricInst>
        <fabricOOServicePol>
            <fabricRsOosPath tDn="topology/pod-1/paths-$LEAF_Id/pathep-[eth1/49]" lc="blacklist"
status ="deleted"/>
        </fabricOOServicePol>
    </fabricInst>
</polUni>
```

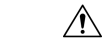
Step 2 Post a local task to the node itself to bring up the interfaces you want using **l1EthIfSetInServiceLTask**.

Example:

```
$LEAF_Address/api/node/mo/topology/pod-1/node-$LEAF_Id/sys/action.xml
<actionLSubj oDn="sys/phys-[eth1/49]">
    <l1EthIfSetInServiceLTask adminSt='start'/>
</actionLSubj>
```

Performing a Rebuild of the Fabric

Rebuilding the Fabric



Caution This procedure is extremely disruptive. It eliminates the existing fabric and recreates a new one.

This procedure allows you to rebuild (reinitialize) your fabric, which you may need to do for any of the following reasons:

- To change the TEP IPs
- To change the Infra VLAN
- To change the fabric name
- To perform TAC troubleshooting tasks

Deleting the APICs erases the configuration on them and brings them up in the startup script. Performing this on the APICs can be done in any order, but ensure that you perform the procedure on all of them (every leaf and spine in the fabric).

Before you begin

Ensure that the following is in place:

- Regularly scheduled backups of the configuration
- Console access to the leaves and spines
- A configured and reachable CIMC, which is necessary for KVM console access
- No Java issues

Procedure

Step 1 If you would like to retain your current configuration, you can perform a configuration export. For more information, see the *Cisco ACI Configuration Files: Import and Export* document: https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html

Step 2 Erase the configuration on the APICs by connecting to the KVM console and entering the following commands:

a) **>acidiag touch clean**
b) **>acidiag touch setup**
c) **>acidiag reboot**

Ensure that each node boots up in fabric discovery mode and is not part of the previously configured fabric.

Note
The **acidiag touch** command alone is not useful for this procedure, because it does not bring the APIC up in the startup script.

Caution
It is extremely important that you ensure that all previous fabric configurations have been removed. If any previous fabric configuration exists on even a single node, the fabric cannot be rebuilt.

Step 3 When all previous configurations have been removed, run the startup script for all APICs. At this point, you can change any of the above values, TEP, TEP Vlan, and/or Fabric Name. Ensure that these are consistent across all APICs. For more information, refer to the *Cisco APIC Getting Started Guide*: https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html.

Step 4 To clean reboot the fabric nodes, log in to each fabric node and execute the following:

a) **>setup-clean-config.sh**
b) **>reload**

Step 5 Log in to apic1 and perform a configuration import. For more information, see the *Cisco ACI Configuration Files: Import and Export* document: https://www.cisco.com/c/en/us/support/cloud-systems-management/application-policy-infrastructure-controller-apic/tsd-products-support-series-home.html.

Step 6 Wait for a few minutes as the fabric now uses the previous fabric registration policies to rebuild the fabric over the nodes. (Depending on the size of the fabric, this step may take awhile.)

Troubleshooting a Loopback Failure

Identifying a Failed Line Card

This section expains how to identify a failed line card when getting a loopback failure.

Before you begin

You should have created a On-Demand TechSupport policy for the fabric node. If you have not already created an On-Demand TechSupport policy, see the "Sending an On-Demand Tech Support File Using the GUI" section in the *Cisco APIC Basic Configuration Guide*.

Procedure

Step 1 Collect the Logs Location file of the On-Demand TechSupport policy for the fabric node. To initiate the collection:

a) In the menu bar, click **Admin**.
b) In the submenu bar, click **Import/Export**.
c) In the **Navigation** pane, expand **Export Policies** and right-click the On-Demand TechSupport policy for the fabric node.
A list of options appears.
d) Choose **Collect Tech Supports**.
The **Collect Tech Supports** dialog box appears.
e) In the **Collect Tech Supports** dialog box, click **Yes** to begin collecting tech support information.

Step 2 Download the the Logs Location file of the On-Demand TechSupport policy for the fabric node. To download the Logs Location file:

a) From the On-Demand TechSupport policy window in the **Work** pane, click the **Operational** tab.
A summary table appears in the On-Demand TechSupport policy window with several columns, including the **Logs Location** column.
b) Click the URL in the **Logs Location** column.

Step 3 Inside the Logs Location file, go to the /var/sysmgr/tmp_logs/ directory and unzip the svc_ifc_techsup_nxos.tar file.

```
-bash-4.1$ tar xopf svc_ifc_techsup_nxos.tar
```

The show_tech_info directory is created.

Step 4 Run **zgrep "fclc-conn failed" show-tech-sup-output.gz | less**.

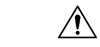
```
-bash-4.1$ zgrep "fclc-conn failed" show-tech-sup-output.gz | less
[103] diag_port_lb_fail_module: Bringing down the module 25 for Loopback test failed. Packets possibly
 lost on the switch SPINE or LC fabric (fclc-conn failed)
[103] diag_port_lb_fail_module: Bringing down the module 24 for Loopback test failed. Packets possibly
 lost on the switch SPINE or LC fabric (fclc-conn failed)
```

Note
The **fclc-conn failed** message indicates a failed line card.

Step 5 Power cycle the currently failed fabric cards and ensure the fabric cards come online.

Step 6 If the fabric cards fail to come online, or after the fabric cards go offline again, immediately collect the diag_port_lb.log file and send the file to the TAC team. The diag_port_lb.log file is located in the /var/sysmgr/tmp_logs/ directory of the Logs Location file.

Removing Unwanted _ui_ Objects



Caution Changes made through the APIC Basic GUI can be seen, but cannot be modified in the Advanced GUI, and changes made in the Advanced GUI cannot be rendered in the Basic GUI. The Basic GUI is kept synchronized with the NX-OS style CLI, so that if you make a change from the NX-OS style CLI, these changes are rendered in the Basic GUI, and changes made in the Basic GUI are rendered in the NX-OS style CLI, but the same synchronization does not occur between the Advanced GUI and the NX-OS style CLI. See the following examples:

- Do not mix Basic and Advanced GUI modes. If you apply an interface policy to two ports using Advanced mode and then change the settings of one port using Basic mode, your changes might be applied to both ports.
- Do not mix the Advanced GUI and the CLI, when doing per-interface configuration on APIC. Configurations performed in the GUI, may only partially work in the NX-OS CLI.

  For example, if you configure a switch port in the GUI at **Tenants** > ***tenant-name*** > **Application Profiles** > ***application-profile-name*** > **Application EPGs** > ***EPG-name*** > **Static Ports** > **Deploy Static EPG on PC, VPC, or Interface**

  Then you use the show running-config command in the NX-OS style CLI, you receive output such as:

  ```
  leaf 102
  interface ethernet 1/15
  switchport trunk allowed vlan 201 tenant t1 application ap1 epg ep1
   exit
   exit
  ```

  If you use these commands to configure a static port in the NX-OS style CLI, the following error occurs:

  ```
  apic1(config)# leaf 102
  apic1(config-leaf)# interface ethernet 1/15
  apic1(config-leaf-if)# switchport trunk allowed vlan 201 tenant t1 application ap1 epg
   ep1
  No vlan-domain associated to node 102 interface ethernet1/15 encap vlan-201
  ```

  This occurs because the CLI has validations that are not performed by the APIC GUI. For the commands from the show running-config command to function in the NX-OS CLI, a vlan-domain must have been previously configured. The order of configuration is not enforced in the GUI.
- Do not make changes with the Basic GUI or the NX-OS CLI before using the Advanced GUI. This may also inadvertantly cause objects to be created (with names prepended with `_ui_`) which cannot be changed or deleted in the Advanced GUI.

If you make changes with the Basic GUI or the NX-OS CLI before using the Advanced GUI, this may inadvertently cause objects to be created (with names prepended with `_ui_`) which cannot be changed or deleted in the Advanced GUI.

For the steps to remove such objects, see Removing Unwanted _ui_ Objects Using the REST API, on page 21.

Removing Unwanted _ui_ Objects Using the REST API

If you make changes with the Cisco NX-OS-Style CLI before using the Cisco APIC GUI, and objects appear in the Cisco APIC GUI (with names prepended with `_ui_`), these objects can be removed by performing a REST API request to the API, containing the following:

- The Class name, for example **infraAccPortGrp**
- The Dn attribute, for example **dn="uni/infra/funcprof/accportgrp-__ui_l101_eth1--31"**
- The Status attribute set to **status="deleted"**

Perform the POST to the API with the following steps:

Procedure

Step 1 Log on to a user account with write access to the object to be removed.

Step 2 Send a POST to the API such as the following example:

```
POST https://192.168.20.123/api/mo/uni.xml
Payload:<infraAccPortGrp dn="uni/infra/funcprof/accportgrp-__ui_l101_eth1--31" status="deleted"/>
```

Cisco APIC SSD Replacement

Use this procedure to replace the Solid-State Drive (SSD) in Cisco APIC.

Note This procedure should only be performed when there is at least one APIC with a healthy SSD in the cluster, that is fully fit. If all the APIC controllers in the cluster have SSDs that have failed, open a case with the Cisco Technical Assistance Center (TAC).

Replacing the Solid-State Drive in Cisco APIC

Before you begin

- If your Cisco IMC release is earlier than 2.0(9c), you must upgrade the Cisco IMC software before replacing the solid-state drive (SSD). Refer to the release notes of the target Cisco IMC release to determine the recommended upgrade path from your currrent release to the target release. Follow the instructions in the current version of the *Cisco Host Upgrade Utility (HUU) User Guide* at this link to perform the upgrade.

- In the Cisco IMC BIOS, verify that the Trusted Platform Module (TPM) state is set to "Enabled." Using the KVM console to access the BIOS settings, you can view and configure the TPM state under **Advanced > Trusted Computing > TPM State**.

Note APIC will fail to boot if the TPM state is "Disabled."

- Obtain an APIC `.iso` image from the Cisco Software Download site.

Note The release version of the APIC `.iso` image must be the same version as the other APIC controllers in the cluster.

Procedure

Step 1 From another APIC in the cluster, decommission the APIC whose SSD is to be replaced.

a) On the menu bar, choose **System** > **Controllers**.
b) In the **Navigation** pane, expand **Controllers** > **apic_controller_name** > **Cluster as Seen by Node**. For the **apic_controller_name**, specify an APIC controller that is not being decommissioned.
c) In the **Work** pane, verify that the **Health State** in the **Active Controllers** summary table indicates the cluster is **Fully Fit** before continuing.
d) In the same **Work** pane, select the controller to be decommissioned and click **Actions** > **Decommission**.
e) Click **Yes**.
The decommissioned controller displays **Unregistered** in the **Operational State** column. The controller is then taken out of service and is no longer visible in the **Work** pane.

Step 2 Physically remove the old SSD, if any, and add the new SSD.

Step 3 In the Cisco IMC, create a RAID volume using the newly installed SSD.

Refer to the *Cisco UCS C-Series Integrated Management Controller GUI Configuration Guide* for your Cisco IMC release. In the "Managing Storage Adapters" chapter, follow the instructions in the procedure "Creating Virtual Drive from Unused Physical Drives" to create and initialize a RAID 0 virtual drive.

Step 4 In the Cisco IMC, install the APIC image using the virtual media. In this step, the SSD is partitioned and the APIC software is installed on the HDD.

Note
For a fresh install of Cisco APIC Release 4.x or later, see the *Cisco APIC Installation, Upgrade, and Downgrade Guide*.

a) Mount the APIC `.iso` image using the Cisco IMC vMedia functionality.
b) Boot or power cycle the APIC controller.
c) During the boot process press **F6** to select the **Cisco vKVM-Mapped vDVD** as the one-time boot device. You may be required to enter the BIOS password. The default password is 'password'.
d) During the initial bringup, a configuration script runs. Follow the onscreen instructions to configure the initial settings of the APIC software.
e) After the installation is completed, un-map the virtual media mount.

Step 5 From an APIC in the cluster, commission the decommissioned APIC.

a) Select any other APIC that is part of the cluster. From the menu bar, choose **System** > **Controllers**.
b) In the **Navigation** pane, expand **Controllers** > **apic_controller_name** > **Cluster as Seen by Node**. For the **apic_controller_name**, specify any active controller that is part of the cluster.
c) From the **Work** pane, click the decommissioned controller that displays **Unregistered** in the **Operational State** column.
d) From the **Work** pane, click **Actions** > **Commission**.
e) In the **Confirmation** dialog box, click **Yes**.

The commissioned controller displays the Health state as **Fully-fit** and the operational state as **Available**. The controller should now be visible in the **Work** pane.

Viewing CRC Error Counters

Viewing CRC and Stomped CRC Error Counters

Beginning in Cisco APIC Release 4.2(3), CRC errors are split into two categories: CRC errors and stomped CRC errors. CRC errors are corrupted frames that were dropped locally and stomped CRC errors are corrupted frames that were cut-through switched. This differentiation can make it easier to identify the actual interface impacted by CRC errors and troubleshoot physical layer issues within the fabric.

This section demonstrates how to view the CRC and stomped CRC errors.

Viewing CRC Errors Using the GUI

This section demonstrates how to view CRC and stomped CRC error counters using the GUI.

SUMMARY STEPS

1. On the menu bar, choose **Fabric > Inventory**.
2. In the **Navigation** pane, click to expand a pod.
3. Click to expand **Interfaces**.
4. Click to choose an interface.
5. In the **Work** pane, click the **Error Counters** tab.

DETAILED STEPS

Procedure

Step 1 On the menu bar, choose **Fabric > Inventory**.

Step 2 In the **Navigation** pane, click to expand a pod.

Step 3 Click to expand **Interfaces**.
A list of interfaces appear in the **Navigation** pane.

Step 4 Click to choose an interface.

A list of tabs appear in the **Work** pane.

Step 5 In the **Work** pane, click the **Error Counters** tab.
A list of error categories appears including **CRC Errors (FCS Errors)** and **Stomped CRC Errors (packets)**.

Viewing CRC Errors Using the CLI

This section demonstrates how to view CRC and stomped CRC error counters using the CLI

Procedure

To view CRC and stomped CRC errors:

Example:

```
Switch# show interface ethernet 1/1
Ethernet1/1 is up
admin state is up, Dedicated Interface
  Belongs to po4
  Hardware: 100/1000/10000/25000/auto Ethernet, address: 00a6.cab6.bda5 (bia 00a6.cab6.bda5)
  MTU 9000 bytes, BW 10000000 Kbit, DLY 1 usec
  reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA, medium is broadcast
  Port mode is trunk
  full-duplex, 10 Gb/s, media type is 10G
  FEC (forward-error-correction) : disable-fec
^[[B  Beacon is turned off
  Auto-Negotiation is turned on
  Input flow-control is off, output flow-control is off
  Auto-mdix is turned off
  Rate mode is dedicated
  Switchport monitor is off
  EtherType is 0x8100
  EEE (efficient-ethernet) : n/a
  Last link flapped 3d02h
  Last clearing of "show interface" counters never
  1 interface resets
  30 seconds input rate 0 bits/sec, 0 packets/sec
  30 seconds output rate 4992 bits/sec, 8 packets/sec
  Load-Interval #2: 5 minute (300 seconds)
    input rate 0 bps, 0 pps; output rate 4536 bps, 8 pps
  RX
    0 unicast packets  200563 multicast packets  0 broadcast packets
    200563 input packets  27949761 bytes
    0 jumbo packets  0 storm suppression bytes
    0 runts  0 giants  0 CRC  0 Stomped CRC  0 no buffer
    0 input error  0 short frame  0 overrun   0 underrun  0 ignored
    0 watchdog  0 bad etype drop  0 bad proto drop  0 if down drop
    0 input with dribble  0 input discard
    0 input buffer drop  0 input total drop
    0 Rx pause
  TX
    0 unicast packets  2156812 multicast packets  0 broadcast packets
    2156812 output packets  151413837 bytes
    0 jumbo packets
    0 output error  0 collision  0 deferred  0 late collision
    0 lost carrier  0 no carrier  0 babble  0 output discard
```

```
0 output buffer drops  0 output total drops
0 Tx pause
```