APPENDIX **B**

Cluster Management and Recovery Scenarios

This appendix includes information on cluster management and recovery procedures that are used when one or more switches in a Cisco IOA cluster is offline or when you want to change the master switch assignment from one switch to another switch.

Note The procedures in this appendix describe troubleshooting solutions that use the CLI.

Note The Cisco IOA cluster configuration for an offline switch must be done using the CLI. Cisco IOA cluster configuration for an online switch can be done using Fabric Manager or the CLI.

This appendix includes the following sections;

- Cluster Quorum and Master Switch Election, page B-1
- In-Service Software Upgrade (ISSU) in a Two-Node Cluster, page B-4
- Supported Topologies, page B-5
- Cluster Recovery Scenarios, page B-5

Cluster Quorum and Master Switch Election

This section describes the Cisco IOA cluster quorum and the process for electing the master switch in a cluster.

Node ID

Every switch in a cluster has a node ID. Cisco IOA assigns a node ID to every new switch as it is added to the cluster. The switch where the cluster is created is assigned the node ID of 1. This is the master switch. When a new switch is added to the cluster, it is assigned the next available higher node ID. For example, when a second switch is added to the cluster it gets the node ID of 2 and the third switch gets the node ID of 3, and so on.

Cluster View

The cluster view is the set of switches that are part of the operational cluster.

Cluster Quorum

For a cluster to be operational, it must include more than half the number of configured switches in the cluster view. In an N-node cluster, N/2 + 1 nodes form a cluster quorum.

If N is even, the cluster quorum requires N/2 nodes and also, the presence of the switch with the lowest node ID.

The quorum logic ensures that in the event of cluster partitions at least one partition can be operational. All other switches are nonoperational. This guarantees the consistency of the cluster.

Master Switch Election

When a cluster is created, the switch on which the cluster is created becomes the cluster master switch. When the master switch fails or is rebooted, another switch takes over as the master switch. The master election logic uses the node ID and the latest cluster configuration to determine which switch in the cluster will become the master switch. The master election logic is described as follows:

- If the master switch fails in an operational cluster, the switch with the next lowest node ID takes over as the master switch. Note that in an operational cluster, all the switches run the same cluster configuration.

  - When the previous master switch comes back online and joins the cluster, it does not immediately become the master.

- When all the switches of a cluster are coming up, the switch that has the latest cluster configuration becomes the master switch. If there are multiple switches with the same configuration, the switch with the lowest node ID is chosen to be the master switch.

  - Once a master switch is chosen and the cluster is operational (there is a quorum), even if a switch with a lower node ID joins the cluster at a later time, the master switch does not change.

    For example, there are three switches S1, S2, and S3 with node IDs 1, 2, and 3, respectively. If switches S2 and S3 form a quorum then switch S2 becomes the master switch. Even if switch S1 with the node ID of 1 comes up and joins the cluster at a later time, switch S2 continues to be the master. However, if switch S2 goes down for any reason, switch S1 will become the master switch.

Two-Switch Cluster Scenarios

According to the cluster quorum logic, a cluster with two configured switches can be operational if both switches are operational or the switch with the lowest node ID is operational.

In the latter case, the switch with the lowest node ID is the master of the one-switch cluster. The other switch could have failed or simply lost connectivity to the operational switch. In either case, the switch with the higher node ID would become nonoperational. If the node with the lower node ID failed, the other switch cannot form an operational cluster.

The examples that follow describe these scenarios. The first three examples consider single switch failures.

1. Assume that in a two-switch cluster with switches S1 (node ID 1) and S2 (node ID 2), S1 is the master (the master has the lower node ID).

   When the switches lose connectivity between them, the master switch S1 continues to be operational since it has the lower node ID and can form an (N/2) switch cluster. Switch S2 becomes nonoperational.

2. Assume that in a two-switch cluster with switches S1 (node ID 1) and S2 (node ID 2), S2 is the master (note that the master has the higher node ID because it has the latest configuration when both the switches came online).

   When the switches lose connectivity between them, switch S2 becomes nonoperational and S1 takes over as the master to form a 1-switch cluster. This is consistent with the quorum logic in a two-switch cluster (N/2 with lowest node ID).

3. Assume that in a two-switch cluster with switches S1 (node ID 1) and S2 (node ID 2). If S1 fails (regardless of which switch was the master), S2 will also become non-operational as long as S1 is down.

   When S1 comes up, S1 and S2 will form a two-switch cluster.

The next set of examples describe reboots of both switches (S1 with node ID 1 and S2 with node ID 2):

> ⚠
> **Caution**   If you perform any configuration change on a cluster, you must save the running configuration to the startup configuration by entering the **copy running-config startup-config** CLI command on all switches before rebooting them. Otherwise, the cluster may not form correctly after the reboot (see example 3.).

1. After a reboot, if both switches S1 and S2 come up about the same time, a two-switch cluster will be formed.

   a. If the cluster configurations are the same, S1 (with the lower node ID) will become the master.

   b. If the cluster configurations are different, the switch with the latest cluster configuration will become the master.

2. After a reboot, if switch S2 comes up first, it will not be able to form a cluster until S1 also comes up. After that, the algorithm explained in the previous case will be used.

3. After a reboot, if switch S1 comes up first, it will form a one-switch cluster (N/2 with lowest node ID). When S2 comes up, it will join the cluster to form a two-switch cluster.

   When S2 comes up and if it happens to have the latest cluster configuration in the startup configuration (this can happen if you did not save the running configuration to the startup configuration on S1 but did so on S2), it will not be able to join the cluster formed by S1.

> ⚠
> **Caution**   It is critical that you save the running configuration on all switches before a reboot.

Three-Switch Cluster Scenarios

In a three-switch cluster, the quorum requires two switches to be in the cluster view (N/2 + 1). The examples below explain three scenarios in a three-switch cluster with switches S1 (node ID 1), S2 (node ID 2) and S3 (node ID 3). S1 is the master switch.

1. In a three-switch operational cluster, if switch S3 fails or loses connectivity with the other two switches, then S3 becomes nonoperational. Switches S1 and S2 will form an operational cluster. When S3 comes up again, it will rejoin the cluster.

2. In a three-switch operational cluster, if the master switch S1 fails or loses connectivity with the other two switches, then S1 becomes nonoperational. Switches S2 and S3 will form an operational cluster and S2 will be the master. When S1 comes up again, it will rejoin the cluster. Note that S2 will continue to be the master.

3. If two switches fail, the cluster will become nonoperational.

These examples describe reboots on all switches in the cluster:

⚠ **Caution**   If you perform any configuration change on a cluster, you must save the running configuration to the startup configuration by entering the **copy running-config startup-config** command on all switches before rebooting them. Otherwise, the cluster may not form correctly after the reboot.

1. After a reboot, if all switches come up at about the same time, first a 2-switch cluster will be formed and later the third switch will be added.

    a. If the cluster configurations are the same, S1 (with the lower node ID) will become the master switch and form the 2-switch cluster first; and then add the third switch.

    b. If the cluster configurations are different, the switch that is running the latest configuration will become the master switch and then form a 2-switch cluster; and then add the third switch.

2. After a reboot, if the switches come up one at a time, a 2-switch cluster will be formed after the first two switches are up. Later, when the third switch comes online, it will join the cluster.

    If the third switch happens to be running the latest cluster configuration in the startup configuration (this can happen if you save the running configuration only on this switch but not on the other two), the third switch will not be able to join the cluster.

⚠ **Caution**   It is critical that you save the running configuration on all switches before a reboot.

Four-Switch Cluster Scenarios

The four-switch cluster scenario is very similar to the examples above. The cluster will be operational if the cluster view has at least three switches (N/2 + 1), or if the cluster view has two switches including the switch with the lowest node ID (N/2 with lowest node ID).

In-Service Software Upgrade (ISSU) in a Two-Node Cluster

In-Service Software Upgrade (ISSU) is a comprehensive, transparent software upgrade application that allows you to deploy bug fixes and add new features and services without any disruption to the traffic.

In a cluster comprising of the MDS 9222i Switches as nodes, if the nodes are not able to communicate, then the node having the lowest node identifier (node ID) remains in the cluster while the other node leaves the cluster. However, when an ISSU is performed on a node having the lowest node identifier, a complete loss of the cluster results because both the nodes leave the cluster.

This undesirable situation is addressed in a two-node cluster as follows:

- The upgrading node sends a message to the other node of the intent to leave the cluster. The upgrading node can either be a master node or a slave node.

- The remaining node remains in the cluster and performs the role of the master node if it was a slave node. This node continues to remain in the cluster with the quorum intact.

- After the ISSU is completed and the switches boot up, the upgraded node rejoins the cluster as a slave node.

✎ **Note**   This feature is tied to ISSU logic and no additional commands need to be executed.

Supported Topologies

Cisco IOA supports a single-fabric topology. Multiple modules can be deployed in a Fibre Channel fabric to easily scale-up performance, to enable simplified load balancing, and to increase availability. In a typical configuration, one IOA engine per site is required in each IOA cluster.

IOA clusters include designated backup servers, tape libraries, and one or more MDS switches running Cisco SAN-OS Release 3.2(2c) or alter. One cluster switch must include an IOA engine per site. With easy-to-use provisioning, traffic between any host and tape on the fabric can utilize the IOA services.

Required Cisco IOA engines are included in the following Cisco products:

- Cisco MDS 9000 Family 18/4-port Multiservice Module (MSM-18/4)
- Cisco SSN-16 Module Switch

Single-Fabric Topology

The MSM-18/4 Module can be anywhere in the fabric. Cisco IOA does a one-to-one mapping of the information from the host to the target and forwards the encrypted data to the dedicated HR tape. Cisco IOA also tracks the barcodes on each encrypted tape and associates the barcodes with the host servers.

Encryption and compression services are transparent to the hosts and storage devices. These services are available for devices in any virtual SANs (VSANs) in a physical fabric and can be used without rezoning.

In certain topologies, edge switches are interconnected across the WAN. Plan for deployment at the core and transition of WAN links to core switches for optimal routing.

Cluster Recovery Scenarios

Refer to this section for information on recovery procedures that are used when one or more switches in a Cisco IOA cluster is offline or when you want to change the master switch assignment from one switch to another switch.

This section includes the following topics:

Deleting an Offline Switch from a Cisco IOA Cluster

To delete an offline switch when one or more switches are offline and the master switch is online, use these procedures.

On the offline switch (for example, switch2), shut down the cluster by performing this task:

	Command	Purpose
Step 1	switch# **config t**	Enters configuration mode.
Step 2	switch(config)# **ioa cluster** *ABC* switch(config-ioa-cl)# **shutdown**	Shuts down the ABC cluster on the offline switch.

Note	Repeat the procedure for every offline switch.

On the cluster master switch, delete the offline switch (for example, switch2) by performing this task:

	Command	Purpose
Step 1	switch# **config t**	Enters configuration mode.
Step 2	switch(config)# **ioa cluster** *ABC* switch(config-ioa-cl)# **no node** **switch2**	Deletes switch2 from the ABC cluster configuration. **Note** Repeat this step for every offline switch that was shut down in Step 1.

On the offline switch (switch2), delete the cluster by performing this task:

	Command	Purpose
Step 1	switch# **config t**	Enters configuration mode.
Step 2	switch(config)# **no ioa cluster** *ABC*	Deletes the ABC cluster configuration.

Note	Delete the cluster on every offline switch that was shut down in the first procedure.

Deleting a Cisco IOA Cluster with One or More Offline Switches while the Master Switch is Online

To delete a Cisco IOA cluster that includes one or more offline switches and online master switch, use these procedures.

Caution	Do not remove a cluster master switch from a cluster and then try to revive the cluster on an offline switch. Since the offline switch was not part of the operational cluster, the cluster master may have progressed beyond what is in the offline switch's state. Deleting the cluster master and reviving the cluster on an offline switch can result in stale configuration.

On the offline switch (switch2), shut down the cluster by performing this task:

	Command	Purpose
Step 1	switch# **config t**	Enters configuration mode.
Step 2	switch(config)# **ioa cluster** *ABC* switch(config-ioa-cl)# **shutdown**	Shuts down the ABC cluster on the offline switch

Note	Repeat the procedure for every offline switch.

On the cluster master switch, delete the offline switch (switch2) and then delete the cluster by performing this task:

	Command	Purpose
Step 1	switch# **config t**	Enters configuration mode.
Step 2	switch(config)# **ioa cluster** *ABC* switch(config-ioa-cl)# **no node switch2**	Deletes switch2 from the ABC cluster configuration. **Note** Repeat this step for every offline switch that was shut down in the first procedure.
Step 3	switch(config)# **no ioa cluster** *ABC*	Deletes the ABC cluster configuration.

On the offline switch (switch2), delete the cluster by performing this task:

	Command	Purpose
Step 1	switch# **config t**	Enters configuration mode.
Step 2	switch(config)# **no ioa cluster** *ABC*	Deletes the ABC cluster configuration.

Note   Delete the cluster on every offline switch that was shut down in the first procedure.

Deleting a Cisco IOA Cluster when All Switches Are Offline

To delete a Cisco IOA cluster when the master switch and all other switches are offline, use these procedures.

Note   When all switches are offline, the cluster is offline.

On the offline switch (for example, switch2), shut down the cluster by performing this task:

	Command	Purpose
Step 1	switch# **config t**	Enters configuration mode.
Step 2	switch(config)# **ioa cluster** *ABC* switch(config-ioa-cl)# **shutdown**	Shuts down the ABC cluster on the offline switch.

Note   Repeat this procedure for every offline switch.

On the offline switch (switch2), delete the cluster by using the following commands:

	Command	Purpose
Step 1	switch# **config t**	Enters configuration mode.
Step 2	switch(config)# **no ioa cluster** *ABC*	Deletes the ABC cluster configuration.

Note   Delete the cluster on every offline switch that was shut down in the first procedure.

Reviving a Cisco IOA Cluster

To revive a cluster on the switch that has the latest Cisco IOA configuration version, use these procedures.

This procedure is used to revive a cluster when one or more switches are offline and the cluster is nonoperational (for example, due to a quorum loss). The recovery procedure includes deleting one or more offline switches and then reviving the cluster on the remaining switches.

⚠️

Caution   A Cisco IOA cluster must only be revived on the switch with the latest IOA configuration version as displayed by the **show IOA cluster detail** command. Reviving the cluster on a switch that does not have the highest configuration version can result in stale configuration.

✎

Note   The following procedure assumes that switch1 has the latest IOA configuration version. The steps shown for switch2 should be carried out for every switch that needs to be removed before reviving the cluster.

Step 1: Shut down the cluster on all the nodes in the cluster, by performing this task:

	Command	Purpose
Step 1	`switch# `**`config t`**	Enters configuration mode.
Step 2	`switch(config)# `**`ioa cluster`** `ABC` `switch(config-ioa-cl)# `**`shutdown`**	Shuts down the ABC cluster on the offline switch.

Step 2: Delete the cluster configuration on each node that needs to be removed from the cluster, by performing this task:

	Command	Purpose
Step 1	`switch# `**`config t`**	Enters configuration mode.
Step 2	`switch(config)# `**`no ioa cluster`** `ABC`	Deletes the ABC cluster configuration.

Step 3: For each node that needs to be removed, delete the node configuration from the remaining nodes in the cluster by performing this task:

	Command	Purpose
Step 1	`switch# `**`config t`**	Enters configuration mode.
Step 2	`switch(config)# `**`ioa cluster`** `ABC` `switch(config-ioa-cl# `**`no node`** **`switch2`**	Deletes switch2 from the configuration. **Note** Repeat for every switch that needs to be deleted.

Step 4: Restart the cluster on the remaning nodes, by performing this task:

	Command	Purpose
Step 1	`switch# `**`config t`**	Enters configuration mode.
Step 2	`switch(config)# `**`ioa cluster`** `ABC` `switch(config-ioa-cl)# `**`no shutdown`**	Restarts the cluster on the switch.