



Configuring SME Cluster Management

DCNM-SAN provides a web browser interface that displays real-time views of your network fabrics and lets you configure the SME with easy-to-use wizards.

This chapter contains information about the SME initial configuration and the tasks that are used to manage SME clusters using DCNM-SAN.

This chapter includes the following topics:

- [Information About SME Cluster Management, on page 1](#)
- [Configuring SME Cluster Management Using the CLI, on page 5](#)
- [Verifying SME Cluster Management Configuration, on page 10](#)
- [Monitoring SME Cluster Management, on page 11](#)
- [Feature History for SME Cluster Management, on page 13](#)

Information About SME Cluster Management

An SME cluster consists of a group of MDS switches running the SME application in a single fabric environment where each switch is a member or node. The cluster infrastructure enables the SME application to offer high availability and load balancing by providing the ability to communicate and coordinate with the other members to maintain a consistent and distributed view of the application's configuration and operational state.

The process of configuring SME on an MDS switch with an installed Cisco MSM-18/4 module, SSN-16 module, or on a Cisco MDS 9222i switch involves a number of configuration tasks that should be followed in chronological order. See the topics in the Before You Begin online help in DCNM-SAN Web Server. Configure SSH and refer to [cisco_sme_getting_started.ditamap#map_FD48D7B73A974D59BE491B1598E630AD](#) and [Configuring SME Interfaces](#) for information about the tasks must be completed before creating an SME cluster.

Cluster Quorum and Master Switch Election

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

Node ID

Every switch in a cluster has a node ID. SME 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-switch cluster, $N/2 + 1$ switches form a cluster quorum.

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

The quorum logic ensures that in the event of cluster partitions, at most 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 describe 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.



Note Because there might be changes in the Master switch, all switches in the cluster need to be configured to handle SNMP configuration, SME roles, user credentials, and SSH. Switches in the cluster should directly communicate with KMC.

Two-Switch Cluster Scenarios

According to the cluster quorum logic [Cluster Quorum, on page 2](#), 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 switch 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 non-operational.

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 non-operational 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.

4. 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.
5. 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.
6. 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.

The examples below 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.

4. 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.
5. 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 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 consisting of the MDS 9222i switches as members, if the switches are not able to communicate, then the switch having the lowest node identifier (node ID) remains in the cluster while the other switch leaves the cluster. However, when an ISSU is performed on a switch having the lowest node identifier, a complete loss of the cluster results because both the switches leave the cluster.

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

- The upgrading switch sends a message to the other switch of the intent to leave the cluster. The upgrading switch can either be a master switch or a slave switch.
- The remaining switch remains in the cluster and performs the role of the master switch if it was a slave switch. This switch continues to remain in the cluster with the quorum intact.
- After the ISSU is completed and the switches boot up, the upgraded switch rejoins the cluster as a slave switch.



Note This feature is tied to the internals of ISSU logic and no additional commands need to be executed.

Server Clusters

A cluster is group of servers linked together to perform a common task.

Clusters provide the following features:

- High availability—If one server in the cluster goes down, the work assigned to that server is migrated to another server in the cluster.
- Load balancing—Clusters allow work to be distributed across different servers.

Clusters can use the shared model or the nonshared model. The shared model requires distributed lock manager (DLM) to manage concurrent access to shared resources. The nonshared model does not require DLM and as a result, requires less overhead. For example, the MSCS (Microsoft clusters) use the nonshared model. This means that a node owns a resource and another node takes ownership of that resource when the owner node fails.

For more information on Cluster-Quorum, see the [Cluster Quorum, on page 2](#).

Configuring SME Cluster Management Using the CLI

You can configure SME Cluster Management using the CLI. This section includes the following topics:



Note SSH feature must be enabled in all the switches to be a part of a cluster.

Creating the SME Cluster

To create an SME tape cluster, identify the fabrics that you want to include in the cluster and configure the following:

- Automatic volume grouping
- Key Management Center (KMC)
- Target discovery
- Tape groups
- Key-on-tape mode
- Recovery
- Shared key mode
- Shutdown cluster for recovery
- Volume tape groups
- Tape compression

To create an SME disk cluster, identify the fabrics that you want to include in the cluster and you configure the following:

- CKMC
- Target discovery
- Disk groups
- Disk device
- Disk path
- Recovery
- Shutdown cluster for recovery

Creating an SME cluster for tape

You can create an SME cluster for either a tape or a disk.



Caution By default, the cluster is capable for SME tapes. However, when you enter the cluster-capability disk command, this cluster can be used only for the disk devices.

To create an SME cluster for tape, follow these steps:

Step 1 switch# **configure terminal**

Enters configuration mode.

Step 2 switch(config)# **sme cluster** *clustername1*

Specifies the cluster name and enters SME cluster configuration submode. A cluster name can include a maximum of 32 characters.

Step 3 switch(config-sme-cl)# **fabric f1**

Adds fabric f1 to the cluster.

Example



Caution You must enable the cluster-capability disk command before adding the first SME interface.

Creating an SME cluster for disk

To create an SME cluster for disk, follow these steps:

Before you begin

Before creating disk clusters, ensure FC-Redirect version 2 is enabled on all switches that are part of the disk cluster. To verify the FC_Redirect version level, enter the following command. The expected output for configuration mode is Mode V2.

```
switch# show fc-redirect configs
Configuration Mode      = MODE_V2
```



Note All switches in the fabric, where SME disk clusters are configured, cannot have FC-Redirect version 1.

Step 1 switch# **configure terminal**

Enters configuration mode.

Step 2 switch(config)# **sme cluster clustername1**

Specifies the cluster name and enters SME cluster configuration submode. A cluster name can include a maximum of 32 characters.

Step 3 switch(config-sme-cl)# **cluster-capability disk**

Defines the SME cluster capabilities for SME Disk.

Step 4 switch(config-sme-cl)# **fabric f1**

Adds fabric f1 to the cluster.

Step 5 switch(config-sme-cl)# **fabric f2**

Adds fabric f2 to the cluster.

Note For SME Disk, you can add up to two fabrics.

Example



Caution For the switches that are in the same fabric, the fabric membership configured in the CLI should be same.

Enabling and Disabling Clustering

The first step in the process of configuring SME is to enable clustering.

To enable or disable the cluster, follow these steps:

-
- Step 1** switch# **configure terminal**
Enters configuration mode.
- Step 2** switch(config)# **feature cluster**
Enables clustering.
- Step 3** switch(config)# **no feature cluster**
Disables clustering.
-

Enabling and Disabling SME Service

SME services must be enabled to take advantage of the SME encryption and security features. After enabling the SME cluster, the second step in the process of configuring SME is to enable the SME service.

To enable the SME service, follow these steps:

-
- Step 1** switch# **configure terminal**
Enters configuration mode.
- Step 2** switch(config)# **feature sme**
Enables SME features.
- Step 3** switch(config)# **nofeature sme**
Disables SME features.
-

Setting the SME Cluster Security Level

There are three levels of security: Basic, Standard, and Advanced. Standard and Advanced security levels require smart cards.

Table 1: Master Key Security Levels

Security Level	Definition
Basic	The master key is stored in a file and encrypted with a password. To retrieve the master key, you need access to the file and the password.
Standard	Standard security requires one smart card. When you create a cluster and the master key is generated, you are asked for the smart card. The master key is then written to the smart card. To retrieve the master key, you need the smart card and the smart card pin.
Advanced	Advanced security requires five smart cards. When you create a cluster and select Advanced security mode, you designate the number of smart cards (two or three of five smart cards or two of three smart cards) that are required to recover the master key when data needs to be retrieved. For recovery, a quorum of cards is required: two of three, two of five, or three of five. For example, if you specify two of five smart cards, then you will need two of the five smart cards to recover the master key. Each smart card is owned by a SME Recovery Officer. Note The larger the number of required smart cards, the greater the security. However, if smart cards are lost or are damaged, the number of available smart cards are reduced that could be used to recover the master key.

To set the SME cluster security level, follow these steps:

-
- Step 1** `switch# configure terminal`
Enters configuration mode.
- Step 2** `switch(config)# sme cluster clustername1`
Specifies the cluster and enters SME cluster configuration submode.
- Step 3** `switch(config-sme-cl)# security-mode basic`
Sets the cluster security level to Basic.
-

Example



Note The CLI is not supported for enabling standard or advanced security mode. Basic mode is also supported through DCNM-SAN Web Client.

Setting Up the SME Administrator and Recovery Office Roles

To set up the SME Administrator, SME Storage Administrator, SME KMC Administrator, and SME Recovery Officer, follow this step:

Command	Purpose
switch# setup sme	Sets up the four security roles.

For more information, see [Creating and Assigning SME Roles Using the CLI](#).



Note When you are accessing SME from Cisco DCNM for the first time, you will be asked to choose the Key Management role for the given DCNM. See the [Configuring Key Management Operations](#) section for more information.

Select the Disk Signature Mode check box to create signature mode clusters.



Note You must download the Master Key file to activate the cluster. If you close the window before downloading the file, navigate to the cluster details page to download the Master Key file and finish the cluster setup.



Note When an error occurs while storing shares on the cards, the cluster should be deleted and recreated.



Note When an error occurs while storing shares on the cards, the cluster should be deleted and recreated.

Verifying SME Cluster Management Configuration

To display SME Cluster Management configuration information, perform one of the following tasks:

Command	Purpose
show sme	Displays a specific cluster configuration, internal information, and transport information.
show sme cluster	Displays additional cluster information.
show sme cluster key	Displays information about the cluster key database.
show sme cluster node	Displays information about a local or remote switch.
show sme cluster recovery officer	Displays information about a specific Recovery Officer or for all the Recovery Officers for a specific cluster.

For detailed information about the fields in the output from these commands, refer to the *Cisco MDS 9000 Family NX-OS Command Reference*.

Monitoring SME Cluster Management

This section covers the following topics:

Viewing SME Cluster Details Using the CLI

This section covers the following topics:

Viewing SME Cluster, Internal, and Transport Information

To verify SME cluster configurations, you can use the **show sme** command to view a specific cluster configuration, internal information, and transport information.

A sample output of the **show sme cluster** command follows:

```
switch# show sme cluster clusternam1
SME Cluster is clusternam1
Cluster ID is 2e:00:00:05:30:01:ad:f4
Cluster is Operational
Cluster is Not Shutdown
Cluster config version is 27
Security mode is basic
Cluster status is online
Total Nodes are 1
Recovery Scheme is 1 out of 1
Fabric[0] is f1
CKMC server has not been provisioned
Master Key GUID is 8c57a8d82d2098ee-3b27-6c2b116a950e, Version: 0
Shared Key Mode is Enabled
Auto Vol Group is Not Enabled
```

Viewing SME Cluster Details

Additional cluster information can be displayed with the **show sme cluster** command. Use this command to show the following:

- SME cluster details
- SME cluster interface information
- Hosts and targets in the cluster
- SME cluster key database
- Cluster node
- SME cluster Recovery Officer information
- Summary of the SME cluster information
- Tapes in a cluster
- Tape volume group information
- Disk group in a cluster
- Disks in a cluster
- SME role configuration

Sample outputs of the **show sme cluster** command follow:

```
switch# show sme cluster clusternam1 ?
  detail      Show sme cluster detail
  interface   Show sme cluster interface
  it-nexus    Show it-nexuses in the cluster
  key         Show sme cluster key database
  node        Show sme cluster node
  recovery    Show sme cluster recovery officer information
  summary     Show sme cluster summary
  tape        Show tapes in the cluster
  tape-bkgrp  Show crypto tape backup group information
  |           Output modifiers.
  >          Output Redirection.
  <cr>       Carriage return.
switch# show sme cluster clusternam1 interface
Interface sme4/1 belongs to local switch
Status is up
```

```
switch# show sme cluster clusternam1 interface it-nexus
```

Host WWN Target WWN	VSAN	Status	Switch	Interface
10:00:00:00:c9:4e:19:ed, 2f:ff:00:06:2b:10:c2:e2	4093	online	switch	sme4/1

Viewing Cluster Key Information

Use the **show sme cluster key** command to view information about the cluster key database.

A sample output of the **show sme cluster key** command for SME tape is as follows:

```
switch# show sme cluster clusternam1 key database
Key Type is tape volumegroup shared key
  GUID is 3b6295e111de8a93-e3f9-e4ae372b1626
  Cluster is clusternam1, Tape backup group is HR1
  Tape volumegroup is Default
Key Type is tape volumegroup wrap key
  GUID is 3e9ef70e0185bb3c-ad12-c4e489069634
  Cluster is clusternam1, Tape backup group is HR1
  Tape volumegroup is Default
Key Type is master key
  GUID is 8c57a8d82d2098ee-3b27-6c2b116a950e
  Cluster is clusternam1, Master Key Version is 0
```

A sample output of the **show sme cluster key** command for SME disk is as follows:

```
switch# show sme cluster clusternam1 key database
Key Type is disk key GUID is aa8c86a783c8a0d9-34ba9cf3af0a17af Cluster is C_SSL, Crypto
disk group is DG Crypto disk is Disk0
Key Type is master key GUID is fc66b503982e816d-a68eba9850f29450 Cluster is C_SSL, Master
Key Version is 0
```

Viewing Cluster Node Information

Use the **show sme cluster node** command to view information about a local or remote switch.

A sample output of the **show sme cluster node** command follows:

```
switch# show sme cluster clusternam1 node
Node switch is local switch
Node ID is 1
Status is online
Node is the master switch
Fabric is f1
```

Viewing Recovery Officer Information

You can view information about a specific Recover Officer or for all Recovery Officers for a specific cluster.

```
switch# show sme cluster clusternam1 recovery officer
Recovery Officer 1 is set
  Master Key Version is 0
  Recovery Share Version is 0
  Recovery Share Index is 1
  Recovery Scheme is 1 out of 1
  Recovery Officer Label is
  Recovery share protected by a password
Key Type is master key share
  Cluster is clusternam1, Master Key Version is 0
  Recovery Share Version is 0, Share Index is 1
```

```
switch# show sme cluster clusternam1 summary
-----
Cluster ID Security Mode Status
-----
clusternam1 2e:00:00:05:30:01:ad:f4 basic online
```

Feature History for SME Cluster Management

The below table lists the release history for this feature.

Table 2: Feature History for SME Cluster Management

Feature Name	Releases	Feature Information
Software change	5.2(1)	In Release 5.2(1), Fabric Manager is changed to DCNM for SAN (DCNM-SAN).
	4.1(1c)	In Release 4.1(1b) and later, the MDS SAN-OS software is changed to MDS NX-OS software. The earlier releases are unchanged and all references are retained.
High availability KMC server	4.1(3)	High availability KMC can be configured by using a primary and secondary servers. In 4.1(3), HA settings are available on the Key Manager Settings page. The primary and secondary servers can be chosen during cluster creation. The primary and secondary server settings can be modified in the Cluster detail page.
Host names are accepted as server addresses	4.1(3)	You can enter IP addresses or host names for the servers.
Target-based load balancing	3.3(1c)	Clustering offers target-based load balancing of SME services.

Feature Name	Releases	Feature Information
Transport settings	3.3(1c)	Allows users to enable or disable transport settings for SME.