InfiniBand Commands

This chapter documents the following commands:

ib sm db-sync

To configure the database synchronize feature between the master subnet manager and one or more standby (slave) subnet managers, enter the **ib sm db-sync** command in Global Configuration mode. To disable database synchronization features, use the **no** form of this command.

Note With database sync enabled on all chassis, only the chassis running the master SM will accept partition configuration from the user.

ib sm db-sync subnet-prefix *prefix* {**enable** | **max-backup-sms** *max* | **session-timeout** *timeout* | **poll-interval** *interval* | **cold-sync-timeout** *cs-timeout* | **cold-sync-limit** *cs-limit* | **cold-sync-period** *cs-period* | **new-session-delay** *delay* | **resync-interval** *resync*}

no ib sm db-sync subnet-prefix prefix {**enable** | **max-backup-sms** | **session-timeout** | **poll-interval** | **cold-sync-timeout** | **cold-sync-limit** | **cold-sync-period** | **new-session-delay** | **resync-interval**}

Syntax Description

subnet prefix	Specifies the subnet prefix of the IB subnet on which you want to configure performance monitoring.
prefix	Subnet prefix of the IB subnet on which you want to configure performance monitoring
enable	Enables database synchronization on your IB fabric.
max-backup-sms	Specifies the maximum number of backup subnet managers that will synchronize with the master SM. **Note** Although we offer this configuration option, the master SM currently only supports one standby.
max	Maximum number of backup subnet managers that will synchronize with the master SM. This value defaults to 1.
session-timeout	Specifies the interval, in seconds, during which a synchronization session status MAD packet must arrive at the master SM to maintain synchronization. This value should be greater than the poll-interval value.
timeout	Timeout interval, in seconds. This value defaults to 10 seconds.
poll-interval	Interval at which the master SM polls an active slave SM to verify synchronization.
interval	Poll interval, in seconds. This value defaults to 3 seconds.
cold-sync-timeout	Allots a maximum amount of time in which to perform a cold sync. During the cold sync, the master SM copies all out-of-sync tables to the standby.
cs-timeout	Cold sync interval, in seconds. This value defaults to 10 seconds.
cold-sync-limit	Specifies the maximum number of cold syncs that may take place during the cold sync period. This value defaults to 2.
cs-limit	Maximum number of cold syncs per cold sync period (integer).
cold-sync-period	Specifies the length of the interval during which cold syncs may occur.
cs-period	Duration, in seconds, of the cold sync period. This value defaults to 900 seconds.

new-session-delay	Specifies the amount of time that the master SM waits before it attempts to initiate a synchronization session with a new SM.
delay	Delay length, in seconds. This value defaults to 120 seconds.
resync-interval	Specifies the interval at which the master SM sends a resynchronization request to all active sync sessions.
resync	Resynchronization interval, in seconds. This value defaults to 3600 seconds.

Defaults

Databases synchronize by default. Use the **disable** keyword to prevent synchronizing SM databases. For attribute-specific defaults, refer to the syntax description.

Command Modes

Global Configuration (config) mode.

Usage Guidelines

Platform Availability:

Cisco SFS 3001, Cisco SFS 7000, Cisco SFS 7008, Cisco SFS 3012, Cicso 4x InfiniBand Switch Module for IBM BladeCenter

Privilege Level:

InfiniBand read-write access

Synchronize the database of the master subnet manager with one or more standby subnet managers to retain all database information in the event of a failover.

Note If you make configuration changes to the master SM and then save the configuration, verify that the master and backup have synchronized, then save the configuration on the backup as well.

Examples

The following example enables database synchronization on the IB fabric:

```
SFS-7000P(config)# ib sm db-sync subnet-prefix fe:80:00:00:00:00:00:00 enable
```

Related Commands

show ib sm db-sync

ib pm

To configure performance monitoring, enter the **ib pm** command in Global Configuration mode.

ib pm subnet-prefix *prefix* {**connection** {**monitor** | **reset-counter** | **test**} **src-lid** *source-LID* **dst-lid** *destination-LID* | **polling-period** *seconds* | **port** {**counter** | **monitor node-guid** *GUID* **port-num** *num* | **reset-counter** [**node-guid** *GUID* [**port-num** *num*]]}| **start-delay** *delay* | **state** {**disable** | **enable** | **enable-topspin-switches** | **enable-all**}| **threshold** {**excess-buf-overruns** | **link-downs** | **link-recovery-errors** | **local-link-errors** | **rcv-constrnt-errors** | **rcv-errors** | **rcv-rate** | **rcv-rem-phy-errors** | **rcv-sw-relay-errors** | **symbol-errors** | **vl15-droppeds** | **xmit-constrnt-errors** | **xmit-discards** | **xmit-rate**} *int*}

Syntax Description

subnet-prefix	Specifies the subnet prefix of the IB subnet on which you want to configure performance monitoring.
prefix	Subnet prefix of the IB subnet on which you want to configure performance monitoring
connection	Specifies a connection-level action. Designates a connection that you want to monitor, reset, or test. You specify the connection with the **src-lid** and **dst-lid** arguments.
monitor	Configures monitoring of the port or connection.
reset-counter	Resets the performance monitoring counter(s).
test	Starts a connection test.
src-lid	Specifies the source Local Identifier (LID) of the connection.
source-LID	Source Local Identifier (LID) of the connection.
dst-lid	Specifies the destination Local Identifier (LID) of the connection.
destination-LID	Destination Local Identifier (LID) of the connection.
polling-period	Interval at which monitoring polls occur.
seconds	Interval at which monitoring polls occur, in seconds.
port	Specifies a port-level action. Designates a port you want to monitor or reset. Specify the port with the **node-guid** and **port-num** arguments.
counter	Enables the IB PM port counter feature.
node-guid	Specifies the GUID of the node that contains the port that you want to monitor.
GUID	GUID of the node that contains the port that you want to monitor.
port-num	Specifies the port number to monitor.
num	Port number to monitor.
start-delay	Delay time before performance monitoring starts after being enabled.
delay	Delay time before starting performance monitoring, in seconds.
state	Configures the state of performance monitoring.
disable	Disables monitoring.
enable	Enables monitoring.
enable-topspin-switches	Enables monitoring on all Server Switches in the subnet.
enable-all	Enables monitoring on all ports in the subnet.

threshold	Configures threshold values.
excess-buf-overruns	Configures the threshold for the number of “excess buffer overrun” errors.
link-downs	Configures the threshold for the number of “link down” errors.
link-recovery-errors	Configures the threshold for the number of “link recovery” errors.
local-link-errors	Configures the threshold for the number of “local link integrity” errors.
rcv-constrnt-errors	Configures the threshold for the number of “receive constraint” errors.
rcv-errors	Configures the threshold for the number of “receive” errors.
rcv-rate	Configures receive rate thresholds.
rcv-rem-phy-errors	Configures the threshold for the number of “receive remote physical” errors.
rcv-sw-relay-errors	Configures the threshold for the number of “receive remote relay” errors.
symbol-errors	Configures the threshold for the number of “symbol” errors.
vl15-droppeds	Configures the threshold for the number of “vl15 dropped” events.
xmit-constrnt-errors	Configures the threshold for the number of “transmit constraint” errors.
xmit-discards	Configures the threshold for the number of “transmit discard” errors.
xmit-rate	Configures transmit rate thresholds.
int	Threshold value (integer).

Defaults

Performance monitoring is disabled by default.

Command Modes

Global Configuration (config) mode.

Usage Guidelines

Platform Availability

Cisco SFS 3001, Cisco SFS 7000, Cisco SFS 7008, Cisco SFS 3012, Cicso 4x InfiniBand Switch Module for IBM BladeCenter

Privilege Level:

InfiniBand read-write access

Use performance manager to do the following:

- View IB port counters.
- Test connectivity between two IB ports (test a connection).
- Monitor any/all IB ports for errors, generating SNMP traps and log messages when user-defined thresholds are exceeded.

To monitor IB ports for errors, follow these steps:

- Configure error thresholds.
- (Optional) Configure specific ports and/or connections to monitor.
- (Optional) Configure new start-delay and/or polling-period values.
- Start performance monitoring.

* Either use the **show ib pm** commands to check for errors, or wait for SNMP traps or log messages to be generated by your Server Switch.'

Examples

The following example configures a symbol-errors threshold of 3:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 threshold symbol-errors 3
```

The following example configures a link-downs threshold of 1:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 threshold link-downs 1
```

The following example configures a polling period of 5 seconds:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 polling-period 5
```

The following example configures a start delay of 0 seconds:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 start-delay 0
```

The following example starts performance monitoring on all IB ports:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 state enable-all
```

The following example stops performance monitoring:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 state disable
```

The following example starts performance monitoring on only the specific connections and ports configured by the user:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 state enable
```

The following example configures a specific connection from LID 3 to LID 7 to monitor:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 connection monitor src-lid
3 dst-lid 7
```

The following example configures a specific port to monitor:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 port monitor node-guid
00:05:ad:00:00:01:34:e0 port-num 3
```

The following example resets the counters on all ports:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 port reset-counter
```

The following example resets the counters on a specific port:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 port reset-counter
node-guid 00:05:ad:00:00:01:34:e0 port-num 3
```

The following example resets the counters on all ports on the connection from LID 3 to LID 7:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 connection reset-counter
src-lid 3 dst-lid 7
```

The following example initiates a connection test from LID 3 to LID 7:

```
SFS-7000P(config)# ib pm subnet-prefix fe:80:00:00:00:00:00:00 connection test src-lid 3
dst-lid 7'
```

Related Commands

show ib pm config
show ib pm connection counter
show ib pm connection counter
show ib pm port counter
show ib pm port monitor
show ib pm threshold

ib sm

To administer the subnet manager (SM) on your Server Switch for everything except multicast, and to create and populate partitions, enter the **ib sm** command in Global Configuration mode. To undo configurations and partitions, use the **no** form of this command. Enter this command without arguments to add a subnet manager with default values.

ib sm subnet-prefix *prefix* [**p_key** *pkey* [**partition-member** *partiton member port full member | limited member*] **| priority** *sm-priority* **[sm-key** *key* **| lid-mask-control** *lmc*] **| response-timeout** *timeout* **| sm-key** ***key*** **| sweep-interval** *interval* **| lid-mask-control** ***LMC*** **| master-poll-intval** *mp-interval* **| master-poll-retries** *retries* **| max-active-sms** *SMs* **| ca-link-hoqlife** *life* **| sw-link-hoqlife** *life* **| | switch-life-time** *life* **| max-hops** *1-64* **| mad-retries** *retries* **| node-timeout** *seconds* **| response-timeout** ***milliseconds*** **| sa-mad-queue-deoth** *size*]

no ib sm subnet-prefix *prefix* [**p_key** *pkey* **[partition-member** ***partition-member port***] **| priority** *sm-priority* **| response-timeout** ***timeout*** **| sweep-interval** ***interval*** **| lid-mask-control** ***LMC*** **| master-poll-intval** ***mp-interval*** **| master-poll-retries** ***retries*** **| max-active-sms** ***SMs***]

Syntax Description

subnet-prefix	Specifies the subnet prefix of the subnet manager.
prefix	Subnet prefix of the subnet manager. You may enter any prefix, but we recommend that you enter **fe:80:00:00:00:00:00:00** to indicate a locally administered subnet.
p_key	(Optional) Creates a partition and optionally assigns members to the partition, or assigns a partition key to a multicast group. **Note** With database sync enabled on all chassis, only the chassis running the master SM will accept partition configuration from the user.
pkey	(Optional) Partition identifier, in *##:##* format.
partition-member	(Optional) Node guid for the partition member.
port	(Optional) Port number of the partition-member.
full member | *limited member*	(Optional) Type of partition membership.
priority	(Optional) Assigns a priority level to the subnet manager. Because multiple subnet managers can run on the system and other SMs may run in your IB network, the priority attribute identifies the master SM.
sm-priority	(Optional) Integer value that represents the subnet manager priority level. The higher the integer, the higher the priority.
sm-key	(Optional) Assigns a subnet management key to a new subnet manager. **Note** We recommend that you do not create additional subnet managers. A subnet manager resides on your Server Switch from the moment you boot.
key	(Optional) 64-bit subnet management key.
timeout	(Optional) Maximum amount of time, in milliseconds, that the SM waits for a response after it sends a packet to a port. The *timeout* variable defaults to 400 milliseconds.

sweep-interval	(Optional) Specifies how frequently the SM queries the InfiniBand fabric for network changes.
interval	(Optional) Frequency, in seconds, at which the SM queries the InfiniBand fabric for network changes.
lid-mask-control	(Optional) Assigns the number of path bits present in the base LID to each channel adapter port. Increasing the LMC value increases the number of LIDs assigned to each port to increase the number of potential paths to reach each port. This value defaults to 0.
LMC	(Optional) Number of path bits.
master-poll-interval	(Optional) Specifies the interval at which the slave SM polls the master to see if it still runs.
mp-interval	(Optional) Poll interval, in seconds. This value defaults to 3 seconds.
master-poll-retries	(Optional) Specifies the number of unanswered polls that cause the slave to identify the master as dead.
retries	(Optional) Number of unanswered polls (integer). This value defaults to 2.
max-active-sms	(Optional) Specifies the maximum number of standby SMs that the master supports. This value defaults to 0, which indicates unlimited SMs.
SMs	(Optional) Number of standby SMs that the master supports (integer).
switch-life-time	(Optional) Specifies the packet lifetime inside a Server Switch.
sw-link-hoqlife	(Optional) Specifies the packet lifetime at the head-of-queue of a switch port.
ca-link-hoqlife	(Optional) Specifies the lifetime of a packet at the head-of-queue of the host port.
life	(Optional) lifetime interval (0 - 20). The interval is a function of microseconds.
max-hops *integer*	(Optional) Configure maximum length path for SM to examine for routing.
integer	(Optional) Specifies the number of hops. Range is from 0 to 64. Default is 64.
mad-retries *retries*	Number of times the SM will retry sending a MAD after not receiving a response. The value range is 0 - 100; the default value is 5.
node-timeout *seconds*	Minimum amount of time in seconds that a HCA may be unresponsive before the SM will remove it from the IB fabric. The value range is 1 - 2000 seconds; the default value is 10 seconds.
response-timeout ***milliseconds***	(Optional) Maximum amount of time in milliseconds that the SM waits for a response before resending a MAD. The value range is 100-5000 milliseconds; the default value is 200 milliseconds.
wait-report-response <*true* | *false*>	Determines whether SM waits to receive ReportResponse MADs once. If boolean value is false, SM sends Report MADs once. If true, the SM continues to send Report MADs until either the ReportResponse MAD is received or the maximum number of Report MADs are sent. Default is false.
sa-mad-queue-depth *size*	Size of the SA's internal queue for receiving MADs. The value range is 256 - 1024; the default value is 256.

Defaults

Table 4-1 ib sm subnet-prefix Command Defaults

Variable	Default
sm-key	00:00:00:00:00:00:00:00
priority	10
sweep-interval	10 seconds
response-timeout	400 microseconds
max-hops	64
mad-retries	5
node-timeout	10
response-timeout	200
wait-report-response	false
sa-mad-queue-depth	256

Command Modes

Global Configuration (config) mode.

Usage Guidelines

Platform Availability:

Cisco SFS 3001, Cisco SFS 7000, Cisco SFS 7008, Cisco SFS 3012, Cicso 4x InfiniBand Switch Module for IBM BladeCenter

Privilege Level:

General read-write user.

The subnet manager:

- Discovers the subnet topology and dynamically updates it at a specified sweep interval that you specify with the *interval* variable.
- Assigns the local identifiers (LIDs), global identifier (GID) subnet prefix, and partition keys for each HCA port.
- Assigns the LIDs, GID subnet prefix, and forwarding databases for each switch on the subnet.
- Maintains the end-node and service databases of the subnet, providing a GUID to LID/GID resolution service as well as a services directory.

One subnet manager administers the InfiniBand fabric. All InfiniBand hosts run on this one subnet. The subnet manager loads upon bootup.

Each node in the fabric has a subnet management agent (SMA) to shuttle communication requests between the node and the subnet manager. Communication between the subnet manager and the subnet management agent uses the common management datagram (MAD) message structure.

Only multicast and p_Key configuration attributes are synchronized between master and standby SMs. If other attributes are changed, they must be manually issued at the standby SMs as well.

If, in the future, there is a change in the location of standby SMs, run the command **show config** to list all of the configuration changes previously made at the master SM. Then replay the configuration changes at the new standby SMs.

Regarding Partitions:

Partitions are created, and then ports are added to those partitions to enforce isolation.

Examples

The following example defines a subnet manager, or redefines the existing subnet manager, with the specified priority, sm-key, response-timeout, and sweep-interval configurations:

```
SFS-7000P(config)# ib sm subnet-prefix fe:80:00:00:00:00:00:00 priority 10 sm-key
00:00:00:00:00:00:00:00 response-timeout 2000 sweep-interval 10
```

The following example removes a specified subnet manager:

```
SFS-7000P(config)# no ib sm subnet-prefix fe:80:00:00:00:00:00:00
```

The following example resets the response-timeout value for the specified subnet manager back to its default value:

```
SFS-7000P(config)# no ib sm subnet-prefix fe:80:00:00:00:00:00:00 response-timeout
```

The following example creates a partition and adds a member:

```
SFS-7000P(config)# ib sm subnet-prefix fe:80:00:00:00:00:00:00 p_key 00:02
partition-member 00:00:2c:90:01:1a:c8:00 3 full-member
```

Related Commands

ib-agent
ib sm multicast
ib sm db-sync
show ib sm configuration

ib sm multicast

To administer the subnet manager (SM) multicast on your Server Switch, and to create and populate partitions, enter the **ib sm multicast** command in Global Configuration mode. To undo configurations and partitions, use the **no** form of this command. Enter this command without arguments to add a subnet manager with default values.

ib sm subnet-prefix ***prefix*** **[multicast {mgid** ***GID-address*** **[mtu** ***MTU-value*****] [p_key** ***pkey*****] |[q_key** ***qkey*****] [rate** ***GBPS*****] [sl** ***service-level*****] | ipoib p_key** ***pkey*** **[mtu** ***MTU-value*****] [q_key** ***qkey*****] [rate** ***GBPS*****] [scope {link-local | site-local | org-local | global}]**

no ib sm subnet-prefix ***prefix*** **[multicast {mgid** ***GID-address*** **[mtu** ***MTU-value*****] [p_key** ***pkey*****] |[q_key** ***qkey*****] [rate** ***GBPS*****] [sl** ***service-level*****] | ipoib p_key** ***pkey*** **[mtu** ***MTU-value*****] [q_key***qkey*] [**rate** *GBPS*] [**scope** {**link-local | site-local | org-local | global**}]

Syntax Description

subnet-prefix	Specifies the subnet prefix of the subnet manager.
prefix	Subnet prefix of the subnet manager. You may enter any prefix, but we recommend that you enter **fe:80:00:00:00:00:00:00** to indicate a locally administered subnet.
multicast	Creates a multicast group.
mgid	Specifies the global ID of the multicast group.
GID-address	Global ID of the multicast group.
mtu	(Optional) Specifies the maximum transmission unit of the multicast group.
MTU-value	(Optional) Maximum transmission unit of the multicast group.
q_key	(Optional) Specifies the queue key of the multicast group.
qkey	(Optional) Queue key of the multicast group.
rate	(Optional) Specifies the data rate of the multicast group, in Gbps.
GBPS	(Optional) Data rate of the multicast group, in Gbps.
sl	(Optional) Specifies the service level of the multicast group.
service-level	(Optional) Service level of the multicast group.
ipoib	(Optional) Creates an IPoIB broadcast multicast group.
scope	(Optional) Specifies the scope of the broadcast multicast group.
link-local	(Optional) Applies a link-local scope to the broadcast multicast group.
site-local	(Optional) Applies a site-local scope to the broadcast multicast group.
org-local	(Optional) Applies a org-local scope to the broadcast multicast group.
global	(Optional) Applies a global scope to the broadcast multicast group.

Defaults

There are no defaults for this command.

Command Modes

Global Configuration (config) mode.

Usage Guidelines

Platform Availability:

Cisco SFS 3001, Cisco SFS 7000, Cisco SFS 7008, Cisco SFS 3012, Cicso 4x InfiniBand Switch Module for IBM BladeCenter

Privilege Level:

General read-write user.

The subnet manager

- Discovers the subnet topology and dynamically updates it at a specified sweep interval that you specify with the *interval* variable.
- Assigns the local identifiers (LIDs), global identifier (GID) subnet prefix, and partition keys for each HCA port.
- Assigns the LIDs, GID subnet prefix, and forwarding databases for each switch on the subnet.
- Maintains the end-node and service databases of the subnet, providing a GUID to LID/GID resolution service as well as a services directory.

One subnet manager administers the InfiniBand fabric. All InfiniBand hosts run on this one subnet. The subnet manager loads upon bootup.

Each node in the fabric has a subnet management agent (SMA) to shuttle communication requests between the node and the subnet manager. Communication between the subnet manager and the subnet management agent uses the common management datagram (MAD) message structure.

Only multicast and p_Key configuration attributes are synchronized between master and standby SMs. If other attributes are changed, they must be manually issued at the standby SMs as well.

If, in the future, there is a change in the location of standby SMs, run the command **show config** to list all of the configuration changes previously made at the master SM. Then replay the configuration changes at the new standby SMs.

Regarding Partitions:

Partitions are created, and then ports are added to those partitions to enforce isolation.

Examples

The following example creates a multicast group:

```
SFS-7000P(config)# ib sm subnet-prefix fe:80:00:00:00:00:00:00 multicast mgid
fe:80:00:00:00:00:00:00:00:00:00:00:00:00:00:00
```

Related Commands

ib-agent
ib sm db-sync
ib sm
show ib sm configuration

ib-agent

To configure subnet management agent (SMA) node strings, enter the **ib-agent** command in Global Configuration mode.

ib-agent {**channel-adapter** *HCA-port-guid* | **switch** *switch-guid*} **node-string** "string"

Syntax Description

channel-adapter	Specifies that you are changing the node string for an HCA.
HCA-port-guid	GUID of the HCA that you want to identify with a node string.
switch	Specifies that you are changing the node string for a switch.
switch-guid	GUID of the switch that you want to identify with a node string.
node-string	Specifies the node string description.
string	Node string description.

Defaults

This command has no default settings.

Command Modes

Global Configuration (config) mode.

Usage Guidelines

Platform Availability:

Cisco SFS 3001, Cisco SFS 7000, Cisco SFS 7008, Cisco SFS 3012, Cicso 4x InfiniBand Switch Module for IBM BladeCenter

Privilege Level:

Unrestricted and InfiniBand read-write users.

The **ib-agent** command allows a user to modify the node description string displayed by the **show ib-agent** command. By specifying an IB node (either switch or HCA) inside the switch chassis, and providing a string, the user will override the description string for the given node.

Note This command does not affect how the node appears on the IB subnet, and the IB "NodeDescription" string is not modified by this command.

Examples

The following example changes the node string of a channel adapter:

```
SFS-7000P(config)# ib-agent channel-adapter 00:05:ad:00:00:00:13:f7 node-string "primary
HCA"
```

The following example changes the node string of a switch:

```
SFS-7000P(config)# ib-agent switch 00:05:ad:00:00:00:13:da node-string "Switch 0, LID 2"
```

Related Commands

ib sm
show ib sm configuration
show ib-agent summary