Extending SANs Using Cisco MDS 9000 24/10-Port SAN Extension Modules

Cisco MDS 9000 Family IP storage services extend the Fibre Channel SAN reach by using open-standard, Fibre Channel over IP (FCIP)-based technology. MDS switches using this technology can connect geographically separated data center SANs to gain highly available data replication and backup for disaster recovery and business continuity requirements of an enterprise network.

The Cisco® MDS 9000 24/10-Port SAN Extension Module helps extend the SANs across longer distances using Fibre Channel over an existing IP network using FCIP. This solution boosts the performance across the WAN, offering greater flexibility on existing MDS 9700 platforms.

Introduction

Enterprises usually build two or more geographically dispersed data centers with SANs in each of them to achieve high availability, load sharing, and data replication and backup. Interconnection of these data centers is a critical consideration, with active-active configuration using high bandwidth and efficient throughput. With the introduction of the Cisco MDS 24/10-Port SAN Extension Module, companies can now build such data centers in one-to-one and/or one-to-many relationships with active-active or active-backup configuration for a business continuity with disaster recovery model. This paper addresses some of the technical details, customer use cases, and important things to consider during any such design using this module (Figures 1 and 2). Technical specifications are listed in Table 1.
Figure 1. Cisco MDS 9000 Series 24/10 SAN Extension Module

Figure 2. Cisco MDS 9000 Series 24/10 SAN Extension Module Front Panel Port View

Table 1. Technical Specifications

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interfaces</td>
<td>• 24x 2/4/8/10/16G line-rate FC ports</td>
</tr>
<tr>
<td></td>
<td>• 8x 1GE/10GE IPS ports</td>
</tr>
<tr>
<td></td>
<td>• 2x 40GE IPS ports</td>
</tr>
<tr>
<td>Tunnels</td>
<td>3 FCIP tunnels per interface</td>
</tr>
<tr>
<td>FCIP options</td>
<td>Compression, Write Acceleration, Tape Acceleration, IPSec Encryption*</td>
</tr>
<tr>
<td>FC port groups</td>
<td>Six port groups, 4 ports per group</td>
</tr>
<tr>
<td>IPS port groups</td>
<td>Two, Group 1: IPS ports 1–4 and 9; Group 2: IPS ports 5–8 and 10</td>
</tr>
</tbody>
</table>

* Future support

Licensing
There is no additional license requirement to use the FC, FCIP, compression, and acceleration functions of the 24/10 SAN Extension module.

IPS Interface Configurations
This module can function in any one of the configuration modes listed in Table 2 based on the IPS mode requirement for IPS ports. Figure 3 shows the front panel port numbering.

Table 2. Configuration Modes

<table>
<thead>
<tr>
<th>IPS Interface Options</th>
<th>Total Active Ports</th>
<th>Processor-1</th>
<th>Processor-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>8</td>
<td>4x 1GE</td>
<td>4x 1GE</td>
</tr>
<tr>
<td>2</td>
<td>8</td>
<td>4x 10GE</td>
<td>4x 10GE</td>
</tr>
<tr>
<td>3</td>
<td>8</td>
<td>4x 10GE</td>
<td>4x 1GE</td>
</tr>
<tr>
<td>4</td>
<td>8</td>
<td>4x 1GE</td>
<td>4x 10GE</td>
</tr>
</tbody>
</table>
IPS Interface Options | Total Active Ports | Processor-1 | Processor-2
--- | --- | --- | ---
5 | 5 | 1x 40GE | 4x 1GE
6 | 5 | 1x 40GE | 4x 10GE
7 | 5 | 4x 1GE | 1x 40GE
8 | 5 | 4x 10GE | 1x 40GE
9 | 2 | 1x 40GE | 1x 40GE

*Future function*

**Figure 3.** Front Panel Port Numbering

---

**Customer Use Cases**

The MDS 9000 Series 24/10 SAN Extension module is ideally used between two data centers, which are geographically apart. Using this SAN extension module, two data centers can be connected over IP for replication and/or backup purposes. The two data centers can be configured for business continuity or disaster recovery mode. Following are some of the use cases using this module in the MDS 9700 chassis.

**Use case 1: Interconnecting Cisco 24/10 SAN extension on both sites.**

This use case describes how to connect the new 24/10 SAN Extension module on both data centers, primary and secondary, to get higher throughput and performance using a director class chassis on both sites. The total throughput can reach up to 80 Gbps across any distance using IP over WAN. This module can be part of any MDS 9700 chassis. This use case provides the robust performance and reliability needed for a disaster recovery site where high availability and replication helps transfer the workloads to a backup primary site quickly (Figure 4).

**Figure 4.** 24/10 SAN Extension Modules Installed on Both Sides
Use case 2: Interoperability and Compatibility

This use case describes how the new 24/10 SAN Extension module easily integrates with an older Cisco MDS FCIP solution. This module can create a 10G tunnel (using two 5G FCIP tunnels on the same 10GE interface) with Cisco MDS 9250i and 1G FCIP tunnels with Cisco MDS 9222i, an 18/4 MSM module, and SSN16 modules (Figure 5).

**Figure 5.** Interoperating 24/10 SAN Extension Module with Older MDS FCIP Solutions

Use case 3: Hybrid SAN extension over FC and FCIP

In this use case, a 24/10 SAN Extension module is used to create FC Inter-Switch Links (ISLs) over a long-distance dense wavelength-division multiplexing (DWDM) network and FCIP tunnels over an IP network across a longer distance. An MDS 9700 24/10 SAN Extension module connects to an MDS 9250i and MDS 9222i at the remote data centers (Figure 6).

**Figure 6.** Disaster Recovery and Business Continuity Using Multihop Replication
Use case 4: Write acceleration using FCIP

FCIP tunnels are created between multiple MDS 9700 24/10 SAN Extension modules. Each data center has multiple MDS 9700s with a 24/10 SAN Extension module installed. These modules are port channeled to multiple WAN routers to provide redundancy and reliability. Load balancing over a port channel can be based on SrcID/DestID or SrcID/DestID/OXID. Virtual SANs (VSANs) can have one port channel between them across FCIP switches (Figure 7).

**Figure 7.** Write Acceleration Using 24/10 SAN Extension Modules

Use case 5: Tape acceleration using FCIP

The MDS 9700 24/10 SAN Extension module can also be used for tape acceleration between multiple data centers across a WAN using high-speed FCIP links. Load balancing can be based on SrcID, DestID, or OXID. Each WAN link is carrying multiple FCIP tunnels and the link is selected based on the FSPF cost. Fabric Shortest Path First (FSPF) is the standard path selection protocol used by Fibre Channel fabrics. The higher the FSPF cost, the higher the priority for packet path selection. Port channels or ECMP links are not allowed between switches (Figure 8).

**Figure 8.** Tape Acceleration Using 24/10 SAN Extension Modules
FCIP Concepts

**FCIP** is a tunneling protocol used to extend Fibre Channel networks distributed across longer distance over a WAN, MAN, or LAN using IP. FCIP consists of one or more TCP connections between two sites, one connection for data frames and a second connection for control frames. A FCIP link is created among two FCIP switches using virtual E (VE) port (similar to E port in Fibre Channel). The VE ports on each end create a virtual Fibre Channel ISL link by initiating a virtual ISL and initiating VE ports on both sides. All of the Fibre Channel communications for domain management, VSAN, zones, etc. are identical to normal Fibre Channel behavior. Following are some of the basic terms we use while configuring FCIP on these modules:

- **GE interface:** This interface is the physical interface on a MDS 9500 or MDS 9222i FCIP module or chassis that is used to carry FCIP traffic.
- **IPS interface:** This interface is the physical Interface on a Cisco MDS 9700 or MDS 9250i FCIP module or chassis that is used to carry FCIP traffic.
- **FCIP profile:** The FCIP profile is bound to an IPS or GE interface, the local endpoint on both sides to this profile. The FCIP profile contains important parameters such as bandwidth, TCP options, etc.
- **TCP connections:** Although this setting is within the FCIP profile, it is a critical setting within the profile. Every TCP connection can be initiated in two modes: active and passive. By default, the active mode is enabled to actively attempt an IP connection. With the passive mode, the switch does not initiate a TCP connection but waits for the peer to connect to it. To activate the FCIP link, one of the TCP connections has to be in active mode; otherwise the FCIP tunnel will not come up.
- **FCIP interface:** This interface is the logical FCIP interface (VE port) endpoint; it contains remote peer information, optional FCIP and ISL features (compression, acceleration, port channels, VSANs, etc.).

Steps to Configure Basic FCIP Tunnel

To configure an FCIP link, follow these steps on both switches:

1. Configure the IPS (or GE) interface (IP address/mask, MTU size).
2. Create an FCIP profile (minimum/maximum bandwidth, round-trip time (RTT) value, TCP port number, and number of TCP connections), and then associate the Gigabit Ethernet interface IP address to the profile.
3. Create an FCIP interface by associating a FCIP profile to the interface.
4. Configure the remote peer IP address and TCP port number to the FCIP interface.
5. Enable optional advanced FCIP features (see next section).
6. Enable the interface.

Advanced FCIP Features

Advanced FCIP features include:

- **FCIP compression:** This feature allows the FCIP traffic across two Cisco MDS switches to be compressed, reducing the amount of bandwidth that is required to transmit the data. This feature has multiple modes to configure the compression level and mechanism to be applied to the traffic. The Cisco MDS 24/10 SAN Extension Modules, MDS 9250i, MDS 9500, MSM-18/4, and SSN-16 modules support auto, mode1, and mode2 compression modes. Mode1 provides high compression over high-speed links with low latency. Mode 2 provides moderate compression over low-bandwidth links with higher latency. Mode3 compression mode is deprecated in Cisco MDS NX-OS Release 5.0(1a) and later. The default auto mode selects either mode1 or mode2 depending on the traffic and latency conditions.
The auto mode (default) selects the appropriate compression scheme based on the card type and bandwidth of the link (the bandwidth configured in the FCIP profile). Cisco MDS switches use a hardware engine to compress the traffic. By default, this feature is disabled.

- **Write Acceleration:** FCIP Write Acceleration (WA) across a WAN significantly improves the application write performance across the storage area traffic between two data centers. This FCIP WA feature reduces the delay introduced by the round-trip delay and can improve performance up to 50 percent for application write performance. The write acceleration feature needs to be enabled on both sides of the link.

  The Small Computer System Interface (SCSI) Write sequence is a two-round-trip protocol. First the SCSI Write is sent by the initiator to the target, which returns a transfer ready. Then the initiator sends the actual data to the target. This process requires two round trips across the data-center-interconnect (DCI) network, adding latency to the sequence. Using FCIP Write Acceleration, the SCSI Write command is acknowledged by the local FCIP switch, allowing the initiator to start transferring the data without having to wait for the command and response to traverse the network. FCIP Write Acceleration can be enabled for multiple FCIP tunnels if the tunnels are part of a dynamic PortChannel configured with channel mode active.

- **Tape Acceleration:** Tape storage devices read and write data sequentially. The FCIP Tape Acceleration (TA) features improve the read and write speed by optimizing the SCSI READ and SCSI WRITE operations. FCIP Tape Acceleration proxy, the SCSI transfer ready, and SCSI write commands on the local switch provide accelerated data transfer to tape devices over the DCI network. In other words, Cisco MDS switches will act as buffer for tape storage devices and transfer data when it receives transfer-ready signal from tape device.

  In tape acceleration for write operations, the backup server issues write operations to a drive in the tape library. The local Cisco MDS switch proxies the transfer ready to signal the host to start sending data on behalf of the remote tape drive. After receiving all the data, the local Cisco MDS switch proxies the successful completion of the SCSI WRITE operation, again on behalf of the remote tape library. This response allows the host to start the next SCSI WRITE operation. This proxy method results in more data being sent over the FCIP tunnel in the same time period compared to the time taken to send data without proxying.

  At the remote side (at the tape end of the FCIP tunnel), the other Cisco MDS switch buffers the command and data it has received. It then acts as a backup server to the tape drive by listening to a transfer ready from the tape drive before forwarding the data. The proxy method improves the performance on WAN links.

  The Cisco NX-OS maintains write data integrity by allowing the WRITE FILEMARKS operation to complete end-to-end without proxying. This mechanism provides reliable data delivery to the remote tape drives using TCP/IP over the WAN. The WRITE FILEMARKS operation signals the synchronization of the buffer data with the tape library data. While tape media errors are returned to backup servers for error handling, tape busy errors are retried automatically by the Cisco NX-OS Software.

  In tape acceleration for writes, after a certain amount of data has been buffered at the remote Cisco MDS switch, the write operations from the host are flow controlled by the local Cisco MDS switch by not proxying the transfer ready. On completion of a write operation when some data buffers are freed, the local Cisco MDS switch resumes the proxying. Likewise, in tape acceleration for reads, after a certain amount of data has been buffered at the local Cisco MDS switch, the read operations to the tape drive are flow controlled by the remote Cisco MDS switch by not issuing any further reads. On completion of a read operation, when some data buffers are freed, the remote Cisco MDS switch resumes issuing reads.
The default flow-control buffering uses the automatic option. This option accounts for the WAN latencies and the speed of the tape to provide optimum performance. The flow-control buffer size range is 64 bytes to 12 MB.

Best Practices

To achieve best performance between two sites connected using a Cisco MDS 24/10 SAN Extension module on one side and the same module or MDS 9250i, MSM 18/4, or SSN16 on the remote side, the following points should be considered. These settings involve parameters about maximum bandwidth allocation, minimum bandwidth allocation, round-trip time value, and number of TCP connections per tunnel.

- Set maximum bandwidth equal to available or allocated WAN bandwidth.
- Set minimum bandwidth to 80 percent of maximum-bandwidth setting (mentioned previously).
- Set the RTT (round-trip time) value equal to measured value using either the `ips-measure-rtt` command-line interface (CLI) command (preferred) or the ping CLI command.
- Set tcp-connections = 2 if maximum-bandwidth setting is less than 1G.
- Set tcp-connections = 5 if maximum-bandwidth setting is greater than 1G.
- Set IPS (or GE) interfaces to use 2500-byte Jumbo Frames.

Sample Configuration Examples

1. Creating FCIP tunnel between MDS 9700 24/10 SAN extension modules (Figure 9)

**Figure 9.** Tunnel between MDS 9700 24/10 SAN Extension Modules

In Figure 9, the MDS 9700 24/10 FCIP module is connected on both sides to WAN routers using 1/10 GE IPS ports. The WAN bandwidth is 10G between two sites, with the RTT (round-trip time) value measured up as 10ms. The configuration will look like the following:

```plaintext
On site A:
fcip profile 11
ip address 10.10.100.1

tcp max-bandwidth-mbps 10000 min-available-bandwidth-mbps 8000 round-trip-time-ms 10
interface fcip 11
use profile 11
peer-info ipaddr 10.10.100.2
tcp-connections 5
interface ips 1/1
ipaddress 10.10.100.1 255.255.255.0
switchport mtu 2500
```

On site B:
fcip profile 11
ip address 10.10.100.2

tcp max-bandwidth-mbps 10000 min-available-bandwidth-mbps 8000 round-trip-time-ms 10
interface fcip 11
use profile 11
peer-info ipaddr 10.10.100.1
tcp-connections 5
interface ips 1/1
ipaddress 10.10.100.2 255.255.255.0
switchport mtu 2500
On site B:
fcip profile 11
ip address 10.10.100.2

tcp max-bandwidth-mbps 10000 min-available-bandwidth-mbps 8000 round-trip-time-ms 10
interface fcip 11
use profile 11
peer-info ipaddr 10.10.100.1

tcp-connections 5
interface ips 1/1
ipaddress 10.10.100.2 255.255.255.0

switchport mtu 2500

2. Creating FCIP Tunnel between MDS 9700 24/10 SAN Extension Module and MDS 9250i

As per above topology, MDS 9700 24/10 SAN Extension module is connected to MDS 9250i at remote site B.

The WAN bandwidth is 10G between two sites with 15ms rtt value. Here is the sample configuration:

Site A:
fcip profile 12
ip address 10.10.100.1

tcp max-bandwidth-mbps 5000 min-available-bandwidth-mbps 4000 round-trip-time-ms 15
Port 3226
fcip profile 11
ip address 10.10.100.1

tcp max-bandwidth-mbps 5000 min-available-bandwidth-mbps 4000 round-trip-time-ms 15
interface fcip 12
use profile 12
peer-info ipaddr 10.10.100.2 port 3226

tcp-connections 5
interface fcip 11
use profile 11
peer-info ipaddr 10.10.100.2

tcp-connections 5
interface ips 1/1
ipaddress 10.10.100.1 255.255.255.0

switchport mtu 2500
Site B:
fcip profile 12
ip address 10.10.100.2
tcp max-bandwidth-mbps 5000 min-available-bandwidth-mbps 4000 round-trip-time-ms 15
Port 3226
fcip profile 11
ip address 10.10.100.2
tcp max-bandwidth-mbps 5000 min-available-bandwidth-mbps 4000 round-trip-time-ms 15
interface fcip 12
use profile 12
peer-info ipaddr 10.10.100.1 port 3226
tcp-connections 5
interface fcip 11
use profile 11
peer-info ipaddr 10.10.100.1
tcp-connections 5
interface ips 1/1
ipaddress 10.10.100.2 255.255.255.0
switchport mtu 2500

3. Creating FCIP Tunnel between MDS 9700 24/10 SAN Extension Module and MDS 9222i/18/4 MSM Module or SSN 16 Module

Above topology is showing FCIP connection between MDS 9700 24/10 FCIP module with MDS 9500/MDS 9222i using 18/4 MSM module/SSN 16 modules at remote site. The WAN bandwidth is 1G between two sites with 8ms measured rtt value.

Site A:
fcip profile 11
ip address 10.10.100.1
tcp max-bandwidth-mbps 1000 min-available-bandwidth-mbps 800 round-trip-time-ms 8
interface fcip 11
use profile 11
peer-info ipaddr 10.10.100.2
tcp-connections 2
interface ips 1/1
ipaddress 10.10.100.1 255.255.255.0
switchport mtu 2500
Site B:
fcip profile 11
ip address 10.10.100.2
tcp max-bandwidth-mps 1000 min-available-bandwidth-mps 800 round-trip-time-ms 8
interface fcip 11
use profile 11
peer-info ipaddr 10.10.100.1
tcp-connections 2
interface ips 1/1
ipaddress 10.10.100.2 255.255.255.0
switchport mtu 2500

Basic Troubleshooting
Here are few basic connectivity tests to verify FCIP configuration.

Ping: Use the ping or the ping ipv6 command to perform a basic check of host reachability and network connectivity. e.g. switch# ping 1.1.1.12. Successful ping indicates connectivity between hosts.

Traceroute: If ping fails, use the traceroute or the traceroute ipv6 command to determine where connectivity is failing. e.g. switch# traceroute 11.8.85.12, it will show up last successful host it is able to reach.

Show ip route: Show ip route or the show ipv6 route command can be used to verify the static route to the remote device. e.g. switch # show ip route.

Clear arp: clear ips arp or clear ipv6 neighbor command to clear the Address Resolution Protocol (ARP) or neighbor cache to verify that the activity is the most current. E.g. switch# clear ips arp interface gigabitethernet 4/7
arp clear successful message after this command indicates arp has been cleared from interface.

Show ips arp: Show ips arp or the show ips ipv6 neighbors command to verify the hardware address for the remote device. e.g. switch# show ips arp interface gigabitethernet 4/7.

Show interface: Use the show interface command to verify that the local IPS interface (Gigabit Ethernet interface) is up.

FAQs
- Minimum NX-OS required for Cisco MDS 9700 24/10 SAN Extension module is 7.3(0)DY(1).
- Cisco MDS 9700 24/10 SAN Extension module requires NX-OS 7.3 DY release train.
- NX-OS 7.3(x)D1(1) releases do not support this module.
- Minimum required DCNM version is DCNM 10.1(1).
- FCIP WA can be configured across multiple FCIP tunnels provided they are part of same port channel group configured in active mode.
- FCIP TA feature cannot be used across FSPF equal cost path as response and command can take different paths during the flow exchange. If FSPF is used, it should be configured in such a way that every scsi write /read command should take the same route between initiator and target.
- Time stamp control should not be configured on FCIP interface using FCIP WA/TA feature.
- During Software upgrade, 24/10 SAN Extension module all of the traffic will be disrupted across all IPS ports.
● If a switch has multiple FCIP modules, all of the FCIP modules will be upgraded in sequence with at least 5 minutes delay between the upgrade process.

● IPS ports support IPV4 and IPv6.

● MTU on any IPS port can be configured from 576 to 9000 bytes with default set to 1500 bytes. An MTU value of 2500 bytes is recommended to avoid segmenting FC frames. Changing the MTU size is disruptive for all FCIP links.

● The port Channel configuration should be static port channel, not the 802.3ad protocol.

● The interface values for PortChannel can only be Ports 1-4 and 5-8.

● To get maximum performance from 24/10 SAN Extension module, set maximum bandwidth as 10Gbps, minimum bandwidth to 8Gbps with MTU size of 2500 and tcp-connections per interface to 5.

● To achieve maximum performance, make sure to match the bandwidth parameters, MTU size, number of tcp tunnels and tcp connections on both sides.

Additional Resources

Cisco MDS 9700 Series

- Cisco MDS 9700 Series Multilayer Directors
- Cisco MDS 9718 Multilayer Director
- Cisco MDS 9710 Multilayer Director
- Cisco MDS 9706 Multilayer Director
- Cisco MDS 9000 24/10-Port SAN Extension Module for Cisco MDS 9700 Series Multilayer Directors
- Cisco MDS 9700 At-a-Glance Documents
- Cisco MDS 9000 Configuration guides
- Cisco MDS 9700 Data sheets
- Cisco MDS 9700 Whitepapers
- Cisco MDS 9000 Interoperability guide