

Data Disaster Recovery and Business Continuance

Fibre Channel over IP Design Using the MDS 9000 Family of Multilayer Switches

Introduction

Data distribution, data protection, and business continuance strategies are critical components of today's information-centric businesses. The ability to efficiently replicate critical data on a global scale not only ensures a higher level of protection for valuable corporate information, but also promotes increased utilization of backup resources, reduces the impact of a catastrophic failure at a single site, and lowers total cost of storage ownership. The Cisco® MDS 9000 Family of Multilayer Directors and Fabric Switches leads the industry in terms of scalability, performance, and availability for building today's critical storage-area network (SAN) infrastructures. As enterprises consider the extension of their SAN environments externally over longer distances to deliver business continuance solutions, the Cisco MDS 9000 offers the industry's only integrated solution. Using the IP Storage Services Module (IPS Module) in the Cisco MDS 9000 Family, SAN designers can use the open-standard Fibre Channel over IP (FCIP) protocol to break the distance barrier of current Fibre Channel solutions and enable interconnection of SAN islands over extended distances. Using the Cisco MDS 9000 solution for FCIP, customers are able to take advantage of an integrated solution that offers the required transport for high speed, high distance SAN extension applications along with an integrated management capability through the Cisco Fabric Manager.

This paper provides a thorough overview of the Cisco MDS 9000 Family IPS Module along with design guidance and recommendations on implementing the IPS Module into a storage environment.

Cisco MDS 9000 IP Storage Services Module

The Cisco MDS 9000 IP Storage Services Module is designed to allow for high capacity integrated IP Storage Services including FCIP and iSCSI into the Cisco MDS 9000 Family of Multilayer Directors and Fabric Switches. Each IPS Module includes eight Gigabit Ethernet ports, each capable of simultaneously supporting three FCIP Tunnels and the iSCSI target functionality. This configuration yields a total capability of up to 24 FCIP Tunnels and 8 iSCSI target interfaces at the same time. In addition to this port capacity, the following features are supported:

- 24 virtual ISLs (FCIP Tunnels) and 8 iSCSI physical targets in each IP Storage Services Module
- Gigabit Ethernet line-rate performance for both FCIP and iSCSI
- Full support of value-added MDS 9000 Fibre Channel features such as Virtual SANs (VSAN), High Availability PortChanneling, and VSAN Trunking TE_Port functionality
- Compatibility with FCIP Devices utilizing a Fibre Channel Bridging Port (B_Port)

With the capability to switch to B_Port mode, the Cisco MDS 9000 IP Storage Service Module allows for compatibility with Bridging Port (B_Port) devices such as the Cisco FCIP Port Adapter Module (PA-FC-1G) for the Cisco 7200/7400 Router product line.

Scaling Remote Connectivity with FCIP

One of the key advantages of FCIP for remote connectivity is its ability to extend distances leveraging TCP/IP. However, distance achieved at the expense of performance is an unacceptable tradeoff for IT organizations that demand full utilization of expensive WAN bandwidth. The IPS Module implements **TCP Extensions for High Performance (RFC 1323)** that allow for efficient full-bandwidth operation over greatly increased distances relative to standard TCP. IETF RFC 1323 adds TCP options for performance including the ability to scale the standard TCP window size up to 1 Gigabyte. As the TCP window size widens, the sustained bandwidth rate across a long haul (more latency) TCP connection increases. In the Cisco MDS 9000 IPS Module, the maximum TCP window size is configurable up to 32 Mbytes.

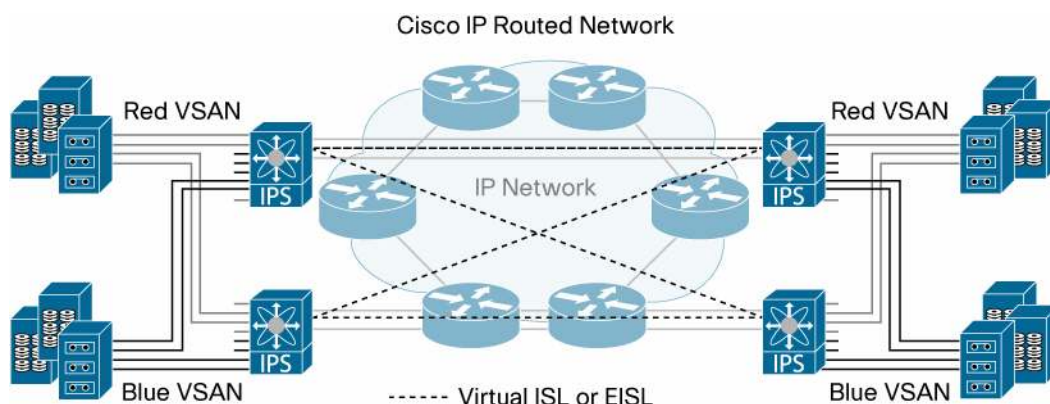
Enhancing SAN Security and Stability

The Cisco MDS 9000 Family VSAN functionality enables the ability to build isolated virtual SAN fabrics within a common physical infrastructure to consolidate costly SAN islands without compromising SAN isolation and stability. VSANs enable more efficient SAN utilization by supporting multiple hardware-based isolated environments within a single physical SAN fabric. Each VSAN can be zoned exactly as one would do with a typical SAN island and maintains its own fabric services for added scalability and resilience. VSANs enable the cost of SAN infrastructure to be shared among more users, while assuring the required segregation and security of traffic is enabled. VSANs also enable independent control and configuration on a VSAN-by-VSAN basis.

The Cisco MDS 9000 IPS Module enables VSANs to be extended across FCIP tunnels. Since the VSAN technology implements a frame-tagging mechanism to identify a frame as belonging to a VSAN, these tagged frames can also be tunneled through FCIP on the IPS Module. Therefore, a user has the ability to extend a virtual ISL and also a virtual Enhanced ISL (EISL) carrying isolated traffic from multiple VSANs.

The ability of carrying multiple VSANs across an FCIP connection enables common wide area services to be shared amongst multiple VSANs while still maintaining isolation. For example, if there are two different instances of remote data replication being used in the data center, one instance could be isolated into one VSAN, and the second instance into the second VSAN. These two VSANs could now be carried across a common FCIP link in a strictly isolated fashion to a remote data center where the VSANs would attach to the respective disk arrays.

Figure 1. Example of an FCIP Deployment Using the Cisco MDS 9000 IPS Module



Best Practices for Designing an FCIP Network

From an IP networking perspective, each FCIP Tunnel requires a pair of static IP addresses and TCP ports with the TCP port number defaulting to 3225. With multiple tunnels extending across a wide area network, it is best practices to have a separate subnet for each FCIP tunnel.

Although it is not an absolute requirement to implement quality of service (QoS), it is highly recommended that QoS be used to ensure high priority for FCIP traffic and to ensure performance degradation resulting from network delays and congestion are kept to an absolute minimum. Although, FCIP timeouts can be tuned to tolerate network anomalies, upper layer applications such as disk replication software might not be as forgiving. Therefore the practices of protecting available FCIP bandwidth at the IP layer can greatly improve the consistent performance of the overall SAN extension solution.

In determining FCIP bandwidth requirements, considerations need to be given as to the distance between remote sites, the required application recovery time, the network fail-over time and the cost of the downtime. With a single instance of synchronous disk replication, typically 15 to 20 Mbytes per second of network bandwidth is required. In data networking terms, OC-3 (155 Mbps) rates are almost a minimum requirement.

However, few storage disaster recovery designs actually mandate synchronous replication over the long haul between active/active drives. More likely than not, local synchronous disk replications are conducted and the mirror copy is usually backed up offline to the remote disks. Thus, bandwidth at or below OC-3 rates should initially be adequate for the typical SAN extension design.

Although SAN extension using FCIP is most often used for remote disk to disk replications over the long haul (greater 100 km), there are other potential uses for this FC encapsulation technology. For organizations that have made significant IP infrastructure investments and wish to leverage some of the investment for storage replication, FCIP can be used without the requirement for additional fiber for native Fibre Channel extension. By interconnecting various isolated SANs within the organization, better disk and tape utilization can be realized. Also, in circumstances where it might be cost prohibitive to implement Metro Optical solutions to extend a Fibre Channel infrastructure, FCIP represents a lower cost alternative. For the very long haul (continental SAN Extension, greater 1000 km), FCIP over a long haul IP infrastructure can also be implemented.

IPS Module Design and Configuration Considerations

In designing an FCIP SAN extension solution using the Cisco MDS 9000 Family IP Storage Services Module, there are several design factors that must be considered. These design factors help determine the achievable distance and performance of an FCIP connection. The following section outlines several tune-able parameters that must be optimized based on the deployment environment and service goals.

Configuring TCP Parameters

As previously mentioned, FCIP uses TCP/IP as its underlying transport protocol. The TCP implementation used within the IPS Module leverages TCP options which enable higher performance as outlined by **IETF RFC 1323—TCP Extensions for High Performance**.

When designing and configuring FCIP tunnels for SAN extension, consideration must be given to such TCP parameters as **Maximum Window Size (MWS)** and path **Maximum Transmission**

Unit (MTU). In addition, various aspects of high availability must be considered and accounted for in the design and configured of the solution.

Configuring TCP **Maximum Window Size**

From a TCP configuration perspective, the TCP Maximum Window Size must be increased as the latency of the network is increased.

This relationship exists to accommodate for the window acknowledgment mechanism that exists in TCP. In TCP, each MWS worth of data must be acknowledged by the other end as having been received intact. Therefore, if the MWS is too small relative to the network latency, the FCIP connection will effectively be throttled due to the latency impact of waiting for TCP acknowledgments from the remote end. However, if the MWS is unnecessarily large, more data will have to be retransmitted (full MWS) if a TCP packet were to be lost or corrupted. Therefore, there is a balance that must be achieved to ensure the full bandwidth potential is usable and still reducing the impact of a lost or corrupted TCP packet in-flight.

For the Cisco MDS 9000 IPS Module, the default MWS is 64 KBytes and can be increased to a maximum of 32 Mbytes. Typically, the MWS calculation is based on the bandwidth of the network (B) and the round-trip-time of the IP packets, represented here as Delay (D). As such the MWS should be greater than the product of the bandwidth (B) of the network and the delay time (D).

$$\text{MWS} > B \times D$$

Where B is the end-to-end bandwidth (The slowest link in the network path) in bits per second and D is the delay or round-trip-time in seconds

For example, if the bandwidth of the network is 155 Mbit/sec (OC3) and the delay is tested to be 10 ms, then the MWS should be at least 194 KBytes.

$$\text{MWS} > 155 \times 10^6 \text{ bit/sec} \times 10 \times 10^{-3} \text{ sec}$$

Converting to bytes:

$$\text{MWS} > 155 \times 10^6 \text{ bit/sec} \times 10 \times 10^{-3} \text{ sec}$$

$$8 \text{ Bits/Bytes}$$

$$= 193,750 \text{ Bytes or } \sim 194 \text{ Kbytes}$$

Latency Calculation

To determine the round trip time or latency, the Ping command within the Cisco MDS 9000 Family can be used. The procedure is to set the target Ping IP address to the IP address of the Gigabit Ethernet port of the peer IPS Module, set the **repeat count** to 10, set the **datagram size** to 2112, and set the **timeout** (in seconds) to 1s. It is recommended to use the average latency expressed in seconds for the above calculation.

The following output shows a sample Ping procedure and its results.

```
w16-excal-1# ping
Target IP address: 5.5.5.2
Repeat count: 10
Datagram size: 2112
Timeout in seconds: 1
Extended commands (y/n): n
PING 5.5.5.2 (5.5.5.2): 2112 data bytes
2120 bytes from 5.5.5.2: icmp_seq=0 ttl=255 time=3.5 ms
2120 bytes from 5.5.5.2: icmp_seq=1 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=2 ttl=255 time=3.4 ms
```

```

2120 bytes from 5.5.5.2: icmp_seq=3 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=4 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=5 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=6 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=7 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=8 ttl=255 time=3.4 ms
2120 bytes from 5.5.5.2: icmp_seq=9 ttl=255 time=3.4 ms
--- 5.5.5.2 ping statistics ---
10 packets transmitted, 10 packets received, 0% packet loss
round-trip min/avg/max = 3.4/3.4/3.5 ms

```

The delay value that would be used from the above example is 3.4 ms.

Configuring Path MTU

By default, **Path Maximum Transmission Unit (PMTU)** discovery is disabled on Cisco MDS 9000 Family switches. When PMTU discovery is disabled, an MTU of 1500 bytes is used for all remote destination IP addresses. The PMTU can be adjusted to take advantage of jumbo frames support, namely Ethernet frames that support a frame size greater than the IEEE standard of 1518 bytes. Using jumbo frames, an FCIP payload can be configured to completely fit a full-size Fibre Channel frame (2148 bytes) thereby eliminating the requirement to fragment the frame into two TCP packets. In order to use jumbo frames support, the entire IP infrastructure between the two connected IPS Modules must support jumbo frames.

In the scenario whereby an FCIP tunnel is built between the Cisco MDS 9000 IPS Module and the FC-PA-1G port adapter module in a Cisco 7200/7400 series router, jumbo frames support is not required on the Cisco MDS 9000 IPS Module as the FC-PA-1G PAM segments FCIP data frames into a maximum MTU of 1500 bytes and does not support jumbo frames.

Configuring Selective Acknowledgment (SACK)

Although TCP is a very robust and adaptable protocol, it has gone through several iterations and additional features have been added to it. Many of these features have enhanced the ability of TCP to perform in longer latency environments such as required for satellite transmission. In these environments where transit delays can be quite large, on the order of 100s of milliseconds, many TCP enhancements have been added to minimize the TCP control traffic and enable its ability to recover faster from any dropped frames.

One such feature that has been added, and is supported by the Cisco MDS 9000 Family IPS Module is **TCP Selective Acknowledgment**. TCP SACK helps TCP connections that are extended over long distances to recover from any sort of frame loss that may occur.

The TCP protocol implements reliability by sending an acknowledgment for received data segments. This is a cumulative acknowledgment scheme where only in-sequence data segments are acknowledged. In case of packet loss, subsequent segments that are successfully received are not acknowledged by the receiver. Because of this behavior, the sender has no way of knowing which segments within the current window of transmission were successfully received after the packet loss. Therefore, the sender is forced to retransmit all segments after the packet loss is detected by a retransmission timeout. This behavior leads to retransmission of segments which were actually successfully received by the receiver thereby wasting network bandwidth. Also, the retransmission timeout causes the congestion window to fall drastically and future transmissions are made at a slower rate than before.

By using the Selective Acknowledgment (SACK) mechanism, a receiver is able to selectively acknowledge segments that were received after a packet loss. The sender, having been notified through the selective acknowledgments, then needs only to retransmit the lost segments. These lost segments or packets are also referred to as "holes" in the data stream.

The Cisco MDS 9000 IPS Module provides the SACK capability as an option that can be enabled on a per-FCIP-profile basis. The SACK option is disabled by default and must be enabled and supported at both ends of the FCIP circuit. Although the SACK option can be turned on with no performance implications, it will only have a noticed benefit for longer, larger bandwidth connections. Hence, in case of lower (bandwidth * delay) networks, there is no significant performance improvement even in case of data loss. In case of larger longer data circuits, the SACK shows a significant improvement in performance in the event of data loss only when compared to non-SACK TCP. In a scenario whereby an FCIP tunnel is created between two Cisco MDS 9000 IPS Modules, it is recommended to enable SACK.

High Availability and FCIP

There are several factors in terms of high availability that must be considered when designing an FCIP solution. High availability design and features must be utilized to ensure that no single point of failure exists in the FCIP service. In addition, it is important to plan how recovery from a failed link or hardware will occur and what the resultant effect will be on the application. The following section discusses some of the high availability design principles and features within the Cisco MDS 9000 Family.

Deploying Redundant FCIP Tunnels

To achieve a higher level of resiliency within an FCIP design, it is possible to configure multiple FCIP tunnels between two data center sites. FCIP tunnels appear as virtual E_Ports (or TE_Ports if carrying VSANs) between the Cisco MDS 9000 switches on the local and remote sites. By adding additional FCIP tunnels, one is effectively adding additional equal-cost paths between the two data centers. These equal-cost paths can be used by the Fibre Channel FSPF routing protocol for load-balancing and recovery purposes.

Although multiple tunnels can be created from a single IPS Module, it is recommended to build parallel tunnels on different ports, modules, and switches to raise the high availability coverage. In the scenario of building two parallel FCIP tunnels between two data centers, it is recommended that the tunnels be deployed in one of the following configurations. The lower numbered configurations are preferred over the higher numbered configurations:

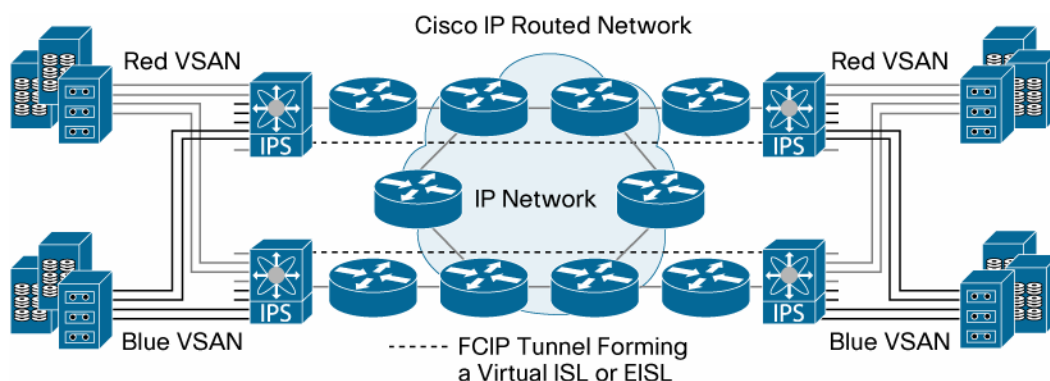
1. **Two tunnels—two IPS Modules—two MDS 9000 switches:** This configuration protects against a port failure, IPS Module failure, switch failure, link failure, and IP service failure. This is the most recommended configuration.
2. **Two tunnels—two IPS Modules—same MDS 9000 switch:** This configuration does protect against a port failure, IPS Module failure, link failure, and IP service failure, but does not protect against a switch failure. It is recommended to use the IPS Module in a Cisco MDS 9500 Director if only using one switch to achieve the desired switch high availability.
3. **Two tunnels—two different ports—one IPS Module:** This configuration does protect against a port failure/link failure and an IP service failure only. It is recommended to use the IPS Module in a Cisco MDS 9500 Director if only using one switch to achieve the desired switch high availability.

4. **Two tunnels—one port—one IPS Module:** This configuration provides minimal incremental high availability capability over a single FCIP tunnel. Although each IPS Module interface supports up to 3 FCIP tunnels, these tunnels should not be used as part of the same virtual ISL or EISL. It is recommended to use at least two different IPS Module ports to create the tunnels as in scenario #3.

It is also very important to ensure that the WAN or MAN connectivity being used to provide the IP transport network is designed in a resilient manner. It is recommended to use redundant router connections and bandwidth circuits for a full high availability solution. Scenario 1 from the list presented above is depicted in Figure 2 along with a suggested redundant WAN configuration. Most SAN designers are often faced with making choices based on factors including available budget and available WAN/MAN services. However, it is recommended to isolate as much of the redundant FCIP tunnel infrastructure as financially possible.

The high availability characteristic of the FCIP implementation can increase when more than 2 tunnels are used between two sites. The following section explains how the Cisco MDS 9000 Family PortChannel feature can be extended across an FCIP infrastructure to further raise the availability of the design.

Figure 2. High Availability FCIP Implementation with Two Parallel Tunnels



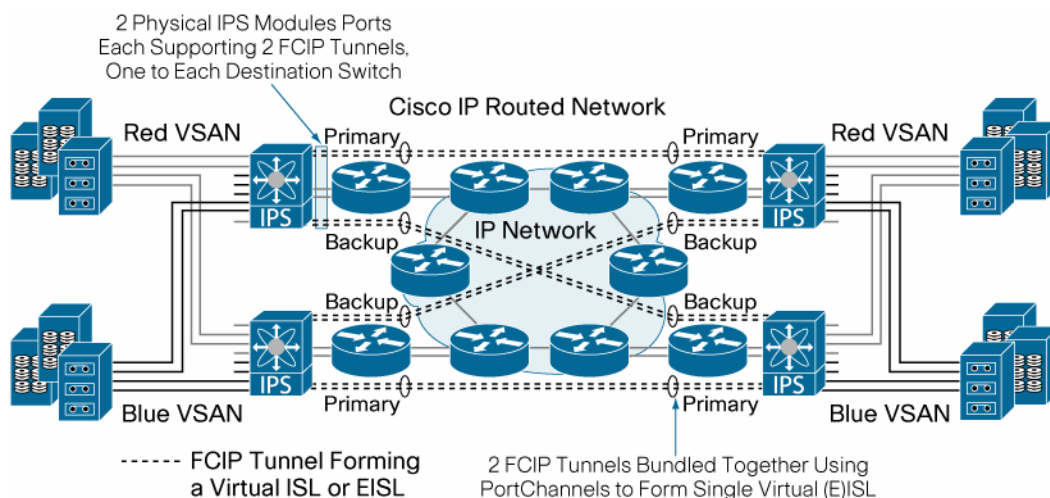
PortChannels and FCIP Tunnels

Since FCIP is merely an extension of a virtual ISL or EISL link between two Cisco MDS 9000 switches, these (E)ISLs can also be bundled together using the Cisco MDS 9000 Family PortChannel feature. The PortChannel feature enables multiple parallel (E)ISLs between two switches to be logically aggregated into one virtual (E)ISL. The PortChannel configuration represents the bundled links as one virtual (E)ISL to the FSPF routing protocol. If a link in the bundle were to fail, the recovery is done at the PortChannel level and not at the FSPF routing level and the results is a sub-second non-disruptive recovery. The PortChannel feature also provides a load balancing capability using one of two load balancing algorithms. The two algorithms enable traffic to be distributed across the bundle based on either the FC_ID pairs (/dst) or the FC_ID pairs along with the exchange IDs (OX_ID/RX_ID). The load balancing capability allows all links within the bundle to be utilized without putting any two frames associated with the same Fibre Channel exchange down two different paths.

Using the PortChannel feature, two FCIP tunnels can be logically combined at the Fibre Channel ISL level to create one virtual (E)ISL. This increases the availability of the extended (E)ISL. However, one requirement for creating a PortChannel is that all the links with the PortChannel must be connected between the same pair of switches. This requirement also exists for bundling of

FCIP tunnels. Therefore, it is still recommended for a fully redundant solution to consist of two Cisco MDS 9000 Directors at each end of the FCIP tunnel. However, since each IPS Module port supports up to 3 FCIP tunnels, it is possible to create a backup set of tunnels to the remote site and weigh them using FSPF to only use if there is a failure of the primary sets of tunnels. Figure 3 shows this configuration. Using the eight defined tunnels, full resiliency is achieved.

Figure 3. Fully Resilient FCIP Design Leveraging PortChannels



VSANs and FCIP Tunnels

As mentioned previously, the VSAN feature within the Cisco MDS 9000 Family enables the ability to create isolated virtual fabrics within a larger physical fabric to provide needed isolation between different departments, applications, and/or environments. The VSAN technology involves the hardware-based addition of a frame tag enabling traffic to be identified and isolated to its particular VSAN. VSAN-tagged frames can also be carried across an FCIP environment using the Cisco MDS 9000 Family IPS Module.

The VSAN technology enhances the manageability and resiliency of FCIP environments as the isolation created within a local data center using VSANs can be extended to remote data centers. This isolation also helps isolate those devices and applications requiring wide-area connectivity from those that don't require such connectivity. By using the VSAN capability, separate virtual fabrics, each with their own Fibre Channel fabric configuration including routing, zoning, name services, and fabric management can be selectively extended across the wide area. Each FCIP tunnel created using the IPS Module can be configured to be a virtual Trunking E_Port (TE_Port) thereby allowing the VSAN tagged traffic to traverse the FCIP tunnel.

For example, in an environment consisting of several SAN islands, each housing a different application require wide-area connectivity to a DR facility, VSANs provide a powerful solution. Each application can be migrated locally to its own VSAN. These VSANs are secure and have their own set of Fibre Channel services and management domains. By extending the VSAN-based services across an FCIP tunnel with VSAN tagging, these virtual fabrics can use the same DR transport facilities while remaining completely isolated. VSANs are also effective for scenarios where multiple disk arrays, each with their own data replication facility, are assigned to separate VSANs and then transported across fewer wide-area FCIP tunnels and still remain completely isolated from a fabric perspective.

VSANs also provide a major benefit of being able to isolate those fabric devices that don't require connectivity across the wide area. For example, in a data center using remote data replication to a DR facility, only the disk array ports require access to the FCIP tunnel. Therefore, by creating a dedicated VSAN on the common SAN infrastructure consisting of only the particular disk array ports, this replication traffic can be carried across the wide area without any interference from other local data center traffic. In addition, local fabric events such as zone set installments, fabric rebuilds, and RSCNs that do not pertain to the disk array ports involved in replication do not get impacted by such events. Also, the control frames in the fabric that are transmitted by these fabric events remain local and do not traverse the wide area.

These are just two of the many benefits and scenarios whereby VSANs can be utilized along with FCIP.

VRRP and FCIP Tunnels

The Virtual Router Redundancy Protocol (VRRP) is used to create seamless fail-over capability between two IP default gateways such as Cisco routers. Using a pair of default gateways, a virtual default IP address and MAC address are created and shared between the two gateways. At any given time, only one gateway owns the IP/MAC addresses. However, should the owner fail, the second gateway would immediately resume the gateway functionality including the same IP/MAC address. Client workstations configured to utilize the virtual IP address as a gateway never lose connectivity to the gateway and do not have to re-learn MAC or IP addresses of the gateway after a failure. The VRRP functionality is also built into the IPS Module providing the same level of availability for IPS Module ports.

Using the VRRP capability, redundant FCIP endpoints can be established between IPS Module ports, IPS Module modules, and even between different MDS 9000 switches equipped with IPS Modules. However, one needs to consider which high availability features to use in which situations. One aspect of VRRP is that it does in fact provide a virtual fail-over point for the IPS Module IP and MAC addresses. However, any failover will result in the need to re-establish the FCIP tunnel as TCP connection state is not kept between different IPS Module ports.

One must consider the recovery capability of an FCIP solution at the IP layer and at the Fibre Channel layer. VRRP is a tool that is used to help aid in the recovery at the IP layer should an FCIP tunnel fail due to a downed FCIP endpoint on an IPS Module module. VRRP helps to provide a redundant endpoint to allow a failed FCIP connection to re-establish. However, this form of recovery is disruptive on the FCIP tunnel and therefore disruptive on the Fibre Channel connectivity if this is the only form of high availability enabled in the solution.

Using redundant FCIP tunnels essentially provides a set of virtual redundant (E)ISLs between two fabrics. This configuration provides protection from the failure of any FCIP tunnel and therefore virtual ISL. Therefore, for an FCIP solution leveraging the Cisco MDS 9000 Family along with the IPS Module, that VRRP be used as a last line of defense against failures. The VRRP feature can always be added as an extra line of protection against FCIP tunnel failure. Using the VRRP solution will allow for recovery of a failed FCIP tunnel.

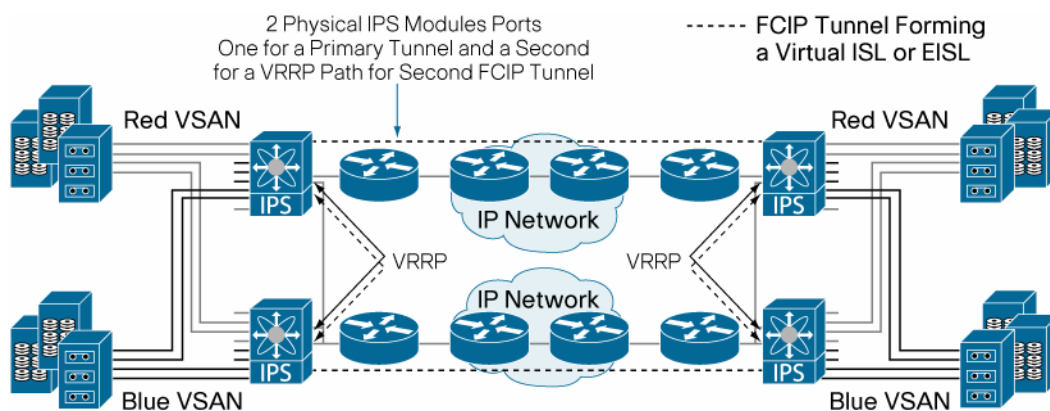
Since each physical Gigabit Ethernet port on the IPS Module is capable of supporting up to 3 FCIP tunnels, it is recommended to create redundant active FCIP tunnels in a fully meshed configuration as shown in Figure 3 rather than simply relying on VRRP for recovery. By building redundant FCIP tunnels, one is protected against failures anywhere in the network and not just at the IP gateway on the IPS Module. In addition, by utilizing multiple FCIP tunnels (virtual (E)ISLs) and the PortChannel (E)ISL aggregation capability when building FCIP tunnels between the two same

Cisco MDS 9000 switches provides for a complete non-disruptive recovery from the failure of a single FCIP tunnel. However, keep in mind that PortChannels cannot be formed across different switches.

Although not in the scope of this white paper, VRRP is heavily used for iSCSI clients accessing the IPS Module for iSCSI services.

However, there is one FCIP configuration that can heavily benefit from the use of VRRP. In the scenario where a full mesh of FCIP tunnels is not possible as shown in Figure 3, VRRP can be utilized to re-establish any failed FCIP tunnels and use all provisioned WAN/MAN bandwidth. For example, if two completely separate WAN or MAN facilities have been provisioned in a dual point-to-point configuration, VRRP can help recover a failed FCIP tunnel due to an offline MDS 9000 switch. This configuration enables the ability to use both provisioned WAN/MAN facilities even with switch or gateway failure.

Figure 4. FCIP Design Leveraging VRRP



Cisco MDS 9000 IP Storage Services—Interoperability

As mentioned previously, Cisco provides an additional solution for FCIP leveraging the Cisco 7200 Series and Cisco 7400 Series routers. The Cisco FCIP Port Adapter Module (PA-FC-1G) for the Cisco 7200/7400 Router product line provides a means of creating a single FCIP tunnel from a router. This FCIP tunnel could terminate at the remote end on another Cisco router, or it could terminate on the Cisco MDS 9000 Family IPS Module.

Although the two product offerings, the Cisco MDS 9000 Family IPS Module and the Cisco 7200/7400 Series Router FCIP PAM, are quite different, a solution involving both products can be highly effective. Consider the scenario where one may wish to create a hub-n-spoke SAN extension network between a central data center and several remote data centers. One may wish to use the Cisco 7200/7400 as an edge FCIP tunnel device in the remote data centers to create one or two FCIP tunnels back to the central site. These tunnels could be used for data replication or remote data access. At the head-end data center, one could use the IPS Module for the Cisco MDS 9000 Family to terminate the multiple tunnels from the multiple remote data centers.

The following section outlines the configuration guidelines for interoperating between the Cisco MDS 9000 Family IPS Module and the Cisco 7200/7400 Series routers with the FCIP PAM.

FCIP Design Considerations—MDS 9000 IPS Module and the Cisco 7200/7400 FCIP PAM

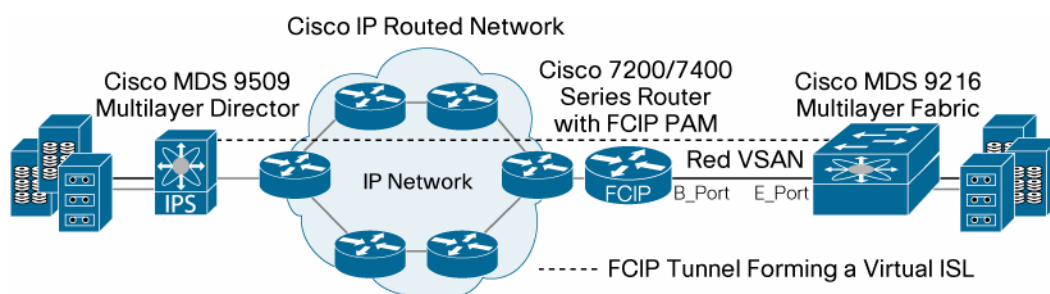
The Cisco MDS 9000 IPS Module capability along with the Cisco FCIP PAM enables several unique FCIP solutions. Using these two complimentary products, SAN designers can build cost-

effective hub-n-spoke FCIP solutions using the FCIP PAM at smaller remote sites. However, there are two design factors that must be considered when combining these two FCIP products.

The first point that must be remembered is that the Cisco 7200/7400 Series router solution with the FCIP PAM only supports transport of a single VSAN. The FCIP PAM is unable to carry the slightly larger frame sizes involved with the addition of a VSAN tag to a frame. However, with the Cisco MDS 9000 IPS Module, a different VSAN can be carried to different data centers terminating on the Cisco 7200/7400 Series router FCIP PAM. As a result, the FCIP link is represented as a virtual ISL with virtual E_Ports on either end as opposed to an EISL link with virtual TE_Ports.

The second design consideration involving the two products is the fact that the Cisco 7200/7400 Series router FCIP PAM only supports a B_Port implementation. This has two ramifications on the FCIP design. First, both ends of the FCIP tunnel must be in B_Port mode. On the Cisco MDS 9000 IPS Module, B_Port mode is a simple configuration and is outlined later in this document. The second factor is that there must be another Fibre Channel switch on the back side of the Cisco FCIP PAM to terminate the ISL link. This switch must be compatible with the Cisco MDS 9000 Family but does not necessarily have to be another Cisco MDS 9000 switch. This switch could be a Cisco SN 5428 Storage Router or a compatible third-party switch.

Figure 5. Combined FCIP Solution Leveraging the Cisco MDS 9000 IPS Module and the Cisco 7200/7400 Series Router FCIP PAM



FCIP Between the MDS 9000 IPS Module and the Cisco 7200/7400 FCIP PAM

Consider the following scenario whereby a centralized SAN fabric requires connectivity to multiple remote SAN islands. This scenario may be the case in which tape and disk resources of a central Data Center are required to be shared with host devices in other remote SANs. In this scenario, it is possible to use the Cisco MDS 9000 Family IPS Module to act as the FCIP hub to multiple remote offices connected over FCIP leveraging the Cisco 7200/7400 Series routers and the FCIP PAM. Figure 5 depicts a simple point-to-point configuration.

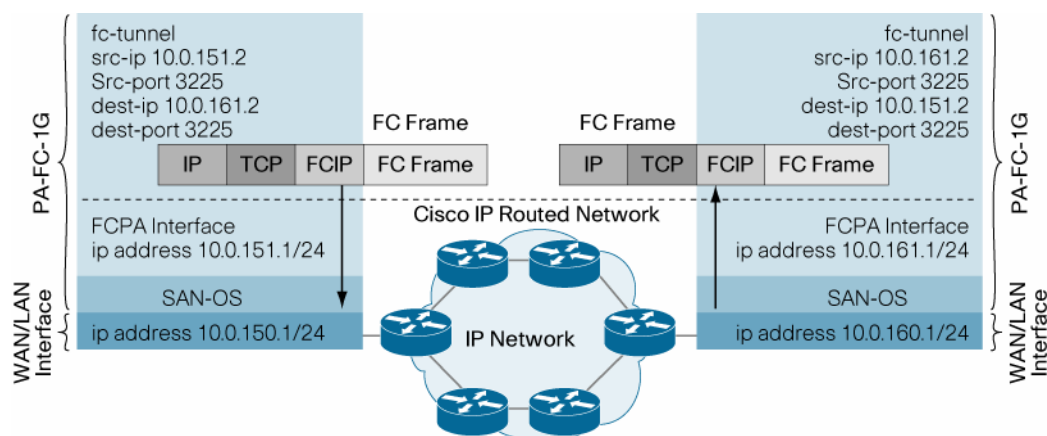
Hub-and-Spoke Topology

A Hub-and-Spoke topology is possible whereby Fibre Channel Switches are attached to a Cisco 7200/7400 Series router with an FCIP PAM at multiple remote data centers. These data centers are then backhauled with FCIP tunnels back to the central data center. With each of the 8 Gigabit Ethernet ports on the Cisco MDS 9000 IPS Module capable of handling up to 3 FCIP tunnels, 24 simultaneous FCIP tunnel are possible. To support the Cisco FCIP PAM, each tunnel created from the Cisco MDS 9000 IPS Module can be configured into a unique VSAN. Environments that can benefit most from this solution are those with existing isolated remote SANs currently disconnected from the central data center. Using the Hub-and-Spoke solution, remote and central resources such as disks and tape libraries can be shared transparently for remote data access or disaster recovery data replication. Figure 6 shows an example involving two remote data centers.

To properly configure ECID on

With the FCIP PAM implementation in the Cisco 7200/7400 Series routers, there is a similar virtual and physical interface model. The FCIP PAM implementation consists of an **FCPA Interface** and an **FCIP Tunnel**. The **FCPA Interface** contains the IP address of the FCPA virtual entity and the **FCIP Tunnel** contains the tunnel information of the source IP address and TCP port number, the destination IP address and TCP port number, and other TCP options such as **Selective Acknowledge** (SACK) and **Maximum Windows Size** (MWS). Figure 8 depicts the FCIP interface model within the FCIP PAM. One will notice, the **FCPA Interface** is in its own IP subnet and therefore must have appropriate IP routing configured to enable IP connectivity to the remote end of the tunnel.

Figure 8. FCIP Model within Cisco 7200/7400 Series Router FCIP PAM



Cisco MDS 9000 IPS Module Terminology

The following definitions pertain to the configuration elements used to establish FCIP tunnels in the Cisco MDS 9000 IPS Module. These elements are depicted graphically in Figure 7.

FCIP Profile—The **FCIP Profile** is used to define an FCIP tunnel instance. It is associated to a particular Gigabit Ethernet port and includes information such as the local IP address to use, TCP Parameters such as time-outs and retransmit times, the Path Maximum Transmission Unit (PMTU), Selective Acknowledgment (SACK), and Maximum Window Size. Since each Gigabit Ethernet port can have up to three tunnels, there can be multiple FCIP Profiles associated to a single Gigabit Ethernet port.

FCIP Interface—The **FCIP Interface** is the actual FCIP tunnel that is enabled when an FCIP Profile is assigned to a Gigabit Ethernet port.

Virtual Expansion Port (VE_Port)—Since each Gigabit Ethernet Port can have up to three virtual (E)ISLs (FCIP tunnels) extended across an IP network, each virtual (E)ISL endpoint is actually represented by a Virtual E_Port as opposed to a standard E_Port.

Bridge Port (B_Port) Mode—For interoperability with FCIP devices that are not actual fabric switches (such as the FCIP PAM), B_Port mode needs to be enabled on the IPS Module. B_Port devices simply pass all information to the attached E_Port on the fabric switch. In the case of the IPS Module, when B_Port mode is enabled, the information is terminated on an internal E_Port. B_Port mode is assigned on a per-FCIP Tunnel basis in the IPS Module configuration dialog.

FCIP PAM (PA-FC-1G) Terminology

The following definitions pertain to the configuration elements used to establish FCIP tunnels in the Cisco 7200/7400 Series router FCIP PAM (PA-FC-1G). These elements are depicted graphically in Figure 8.

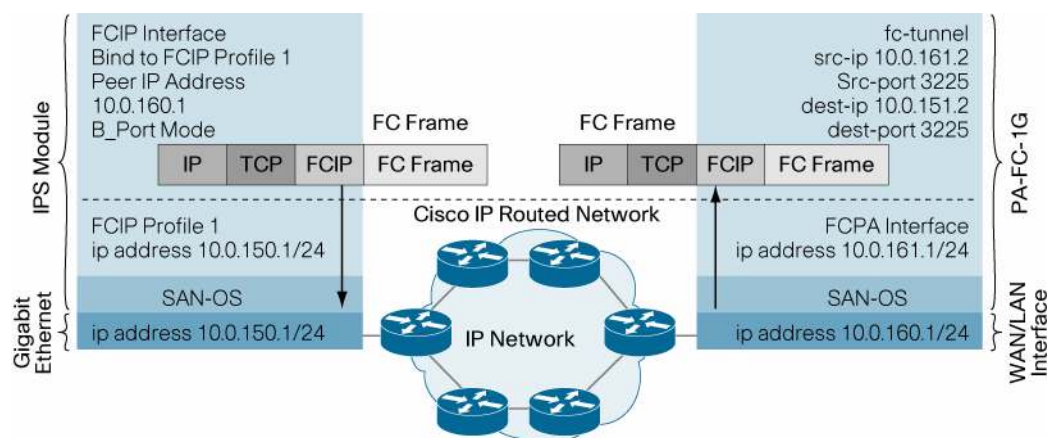
FC-Tunnel—The **FC-Tunnel** is the actual FCIP tunnel that is defined within the PAM. It contains the source IP address and TCP port number and the destination IP address and port number. Consideration must be given to the IP addressing of the FC-Tunnel addresses and their relation to the FCPA Interface addresses. In the case of the FC-Tunnel, the source IP address will be different than the IP address belonging to the FCPA Interface. The two IP addresses for these entities must be in different IP subnets and thus the two configured IP subnetworks must be added to the configured routing protocol to ensure they are reachable from either end of the defined tunnel.

FCPA Interface—The **FCPA Interface** is the FCIP PAM's own physical interface and has its own IP subnet assigned to it.

To configure the Cisco MDS 9000 Series IPS Module to work with the FCIP PAM, the interconnecting FCIP tunnel instance in the IPS Module needs to be set to B_Port Mode. The peer IP address setting in the configured **FCIP Interface** will need to point to the source IP address defined in the opposing FCIP PAM FC-Tunnel.

Likewise for the FCIP PAM configuration, the destination IP address and port needs to be assigned the IPS Module's **FCIP Profile** assigned IP address. In addition, the appropriate routes will also need to be created in the IPS Module to point to the **FC-Tunnel** source IP address. Figure 9 depicts the scenario whereby the Cisco MDS 9000 IPS Module establishes an FCIP tunnel with the Cisco 7200/7400 Series router FCIP PAM. Please consult Appendix B for detailed setup and configuration information.

Figure 9. FCIP Model with the Cisco MDS 9000 Series IPS Module and the Cisco FCIP PAM



Conclusion

The Cisco MDS 9000 Family of Multilayer Directors and Fabric Switches are the industry's first fully integrated multiprotocol storage networking platform. By incorporating the Cisco MDS 9000 IPS Module into the any of the MDS 9000 Multilayer platforms, enterprises can now build enterprise-class storage area networks utilizing cost and distance optimized transports for local and remote SAN connectivity.

In today's world of constant cost cutting along with the need to protect vital data assets, the FCIP solution offers a cost-effective and long-reaching solution for moving critical data to remote data centers and disaster recovery sites. Combined with industry-leading disk subsystem-based data replication solutions, Cisco's MDS 9000 FCIP solutions enable robust data replication within the metro area or over continental distances. Full wire-rate 1 Gbps FCIP transport can be achieved over great distances thanks to Cisco's intelligent merged Fibre Channel and FCIP flow control solutions. Whether the solution requires 20 km or 2000 km of distance, the Cisco MDS 9000 IPS Module can extend SANs with all the performance and required resiliency.

Cisco offers a fully integrated management multiprotocol solution. By combining the management of local data center MDS 9000 Fibre Channel fabrics with the configuration and management of the FCIP infrastructure into the Cisco Fabric Manager, Cisco makes it easy to leverage FCIP for SAN extension solutions.

Appendix A: Disk Array Considerations

Disk Replication Considerations

SAN Extension over an IP network serves as an enabling technology for implementing disk array-based data replication. However, one must determine whether to deploy replication services as asynchronous or synchronous. Consideration must be given to such factors as:

- The distance between the remote sites which impacts the latency of a replication I/O operation. In a synchronous environment, this latency will have a direct impact on the performance of the replicated application.
- The fail-over time of the network should an FCIP tunnel fail. This time will potentially impact the application if the FCIP network is not designed in a redundant manner. As pointed out in this document, a fully redundant FCIP network is the best method to ensure non-disruptive recovery from a failed FCIP tunnel.
- The needed recovery time of the application. Synchronous replicated data is an exact copy of the data whereas asynchronous data may not be an exact copy of the local data depending on the timing and impact of a failure.

However, for remote sites that have latencies that exceed beyond what the recommended MWS can compensate for, or what the application can manage, asynchronous replication becomes the only option.

Asynchronous vs. Synchronous Data Replication

There are two primary forms of data replication that can be performed over remote facilities. Each form has different data 'completeness' and application performance considerations.

In synchronous replication mode, data is copied from the primary site to the secondary site. This data is validated and committed to the secondary site all during the I/O cycle. As such, the reliability and 'completeness' of the remote data is very high. However, the cost of synchronous replication is that the application must wait for the remote committal of the data and an acknowledgment before the application may continue any write I/O. Therefore, as latency is increased between the two sites, the I/O cycle time is also increased and may eventually have an unsustainable impact to application performance.

In asynchronous replication mode, data is also copied from the primary site to the secondary site. However, the data committal process at each end is done separately at the primary and secondary sites. When a write I/O is received at the local end, it is committed to the disk subsystem and then acknowledged to the application host. A second step is then performed whereby the I/O data is copied and committed to the remote end. This separation of tasks allows the local application to continue its I/O processing without waiting for the commit process to occur at the remote site and therefore has less impact on the application. However, there exists a potential for out-of-sync data during a failure condition if the data had not been fully committed to the remote site.

Appendix B: Software Configuration

The following is a sample configuration as depicted in Figure 9 whereby a Cisco MDS 9000 IPS Module establishes an FCIP tunnel to a Cisco 7200/7400 Series router FCIP PAM.

The Cisco MDS 9000 IPS Module consists of the following configuration:

```
w16-excal-1# config terminal
Enter configuration commands, one per line. End with CNTL/Z.
w16-excal-1(config)# int gigabitethernet 2/1
w16-excal-1(config-if)# ip address 10.0.150.1 255.255.255.0
w16-excal-1(config-if)# shut
w16-excal-1(config-if)# no shut
w16-excal-1(config-if)# end
w16-excal-1(config)# fcip profile 1
w16-excal-1(config-profile)# ip address 10.0.150.1
w16-excal-1(config-if)# end
w16-excal-1(config)# int fcip 1
w16-excal-1(config-if)# bind-profile 1
w16-excal-1(config-if)# bport
w16-excal-1(config-if)# peer-info ip-address 10.0.150.2
w16-excal-1(config-if)# end
GigabitEthernet2/1 is up
Hardware is GigabitEthernet, address is 0005.3000.9f36
Internet address is 10.0.150.1/24
MTU 3000 bytes, BW 1000000 Kbit
Port mode is IPS
Speed is 1 Gbps
Beacon is turned off
5 minutes input rate 672 bits/sec, 84 bytes/sec, 0 frames/sec
5 minutes output rate 672 bits/sec, 84 bytes/sec, 0 frames/sec
143985059 packets input, 131549535644 bytes
0 multicast frames, 0 compressed
0 input errors, 0 frame, 0 overrun 0 fifo
143726417 packets output, 131488131638 bytes, 0 underruns
0 output errors, 0 collisions, 0 fifo
0 carrier errors
w16-excal-1# sh int fcip 1
FCIP1 is trunking
Hardware is GigabitEthernet
Port WWN is 20:42:00:05:30:00:6c:de
Peer port WWN is 20:42:00:05:30:00:66:9e
Admin port mode is auto, trunk mode is on
Port mode is TE
vsan is 10
Trunk vsans (allowed active) (1,10,20,30)
Trunk vsans (operational) (1,10,20,30)
Trunk vsans (up) (1,10,20,30)
Trunk vsans (isolated) ()
Trunk vsans (initializing) ()
Using Profile id 10 (interface GigabitEthernet2/1)
```

Peer Information

Peer Internet address is 10.0.150.2 and port is 3225

Special Frame is disabled

Maximum number of TCP connections is 2

Time Stamp is disabled

B-port mode enabled

TCP Connection Information

2 Active TCP connections

Control connection: Local 10.0.150.1:65491, Remote 10.0.150.2:3225

Data connection: Local 10.0.150.1:65492, Remote 10.0.150.2:3225

14 Attempts for active connections, 7 close of connections

TCP Parameters

Path MTU 3000 bytes

Current retransmission timeout is 100 ms

Round trip time: Smoothed 6 ms, Variance: 6

Advertized window: Current: 30 KB, Maximum: 30 KB, Scale: 0

Peer receive window: Current: 30 KB, Maximum: 30 KB, Scale: 0

Congestion window: Current: 63 KB

5 minutes input rate 39 bits/sec, 312 bytes/sec, 0 frames/sec

5 minutes output rate 39 bits/sec, 312 bytes/sec, 0 frames/sec

223153926 frames input, 90785732536 bytes

136299 Class F frames input, 12580812 bytes

223017627 Class 2/3 frames input, 90773151724 bytes

0 Error frames

223677104 frames output, 90891727692 bytes

136301 Class F frames output, 12581508 bytes

223540803 Class 2/3 frames output, 90879146184 bytes

0 Error frames 522 reass frames

The Cisco 7200/7400 Series router FCIP PAM consists of the following configuration:

```
c7200_bottom#config t
Enter configuration commands, one per line. End with CNTL/Z.
c7200_bottom(config)#int fcpa6/0
c7200_bottom(config-if)#ip address 10.0.161.1 255.255.255.0
c7200_bottom(config-if)#no shut
c7200_bottom(config-if)#fc-tunnel bottom
c7200_bottom(config-if-fc-tunnel)#-ip 10.0.161.2
c7200_bottom(config-if-fc-tunnel)#dest-ip 10.0.150.1
c7200_bottom(config-if-fc-tunnel)#-port 3225
c7200_bottom(config-if-fc-tunnel)#dest-port 3225
c7200_bottom(config-if-fc-tunnel)#inservice
c7200_bottom(config-if)#exit
c7200_bottom(config)#exit
c7200_bottom#copy running-config startup-config
c7200_bottom# show fc-tunnel
Interface: Fcpa2/0
FC Tunnel name: bottom
INSERVICE: configured ARP entry: Installed
Source IP: 10.0.161.2
Destination IP: 10.0.150.1
```

Source port: 3225
 Destination port: 3225
 TCP SACK option set
 TCP MWS: 32KB
 TCP KAD: 7200sec
 IP TOS: 0
 MTU: 1500
 MSS: 1440
 SM state: SM_UP_ST
 FC Port Type: B_Port
 FC Port WWN : 100000E0B0FFF2CF
 Switch Port WWN: 200000C0DD00C248
 Switch WWN : 100000C0DD00C248
 FC BB_Credit: 128
 FC RA_TOV: 120000msec
 FC ED_TOV: 60000msec
 FC Link state: UP



Americas Headquarters
 Cisco Systems, Inc.
 170 West Tasman Drive
 San Jose, CA 95134-1706
 USA www.cisco.com
 Tel: 408 526-4000
 800 553-NETS (6387)
 Fax: 408 527-0883

Asia Pacific Headquarters
 Cisco Systems, Inc.
 168 Robinson Road
 #28-01 Capital Tower
 Singapore 068912
www.cisco.com
 Tel: +65 6317 7777
 Fax: +65 6317 7799

Europe Headquarters
 Cisco Systems International BV
 Haarlerbergpark
 Haarlerbergweg 13-19
 1101 CH Amsterdam
 The Netherlands
www-europe.cisco.com
 Tel: +31 0 800 020 0791
 Fax: +31 0 20 357 1100

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

©2007 Cisco Systems, Inc. All rights reserved. CCVP, the Cisco logo, and the Cisco Square Bridge logo are trademarks of Cisco Systems, Inc.; Changing the Way We Work, Live, Play, and Learn is a service mark of Cisco Systems, Inc.; and Access Registrar, Aironet, BPX, Catalyst, CCDA, CCDP, CCIE, CCIP, CCNA, CCNP, CCSP, Cisco, the Cisco Certified Internetwork Expert logo, Cisco IOS, Cisco Press, Cisco Systems, Cisco Systems Capital, the Cisco Systems logo, Cisco Unity, Enterprise/Solver, EtherChannel, EtherFast, EtherSwitch, Fast Step, Follow Me Browsing, FormShare, GigaDrive, HomeLink, Internet Quotient, IOS, iPhone, IP/TV, iQ Expertise, the iQ logo, iQ Net Readiness Scorecard, iQuick Study, LightStream, Linksys, MeetingPlace, MGX, Networking Academy, Network Registrar, PIX, ProConnect, ScriptShare, SMARTnet, StackWise, The Fastest Way to Increase Your Internet Quotient, and TransPath are registered trademarks of Cisco Systems, Inc. and/or its affiliates in the United States and certain other countries.

All other trademarks mentioned in this document or Website are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (0708R)