How Cisco IT Achieves Subminute Recovery Point Objective

Fibre Channel over IP solution overcomes latency inherent in long-distance replication.

Cisco IT Case Study / Storage Networking / SAN for Business Continuity: This case study describes Cisco IT's internal use of Fibre Channel over IP (FCIP) with EMC SRDF (Symmetrix Remote Site Data Facility) replication within the Cisco network, a leading-edge enterprise environment that is one of the largest and most complex in the world. Cisco customers can draw on Cisco IT's real-world experience in this area to help support similar enterprise needs for database replication.

> "FCIP creates huge benefits for business continuity at Cisco. It improved RTO [recovery time objective], reduced our RPO [recovery point objective] to under 60 seconds, simplified data mobility by enabling us to offload database replication from the hosts to the storage arrays, and enabled us to centralize SAN management."
>
> **– Rich Harper, ERP Storage Architect for Cisco IT's Enterprise Storage Services group.**

BACKGROUND

Cisco Systems® maintains production ERP databases in San Jose, California and a disaster recovery (DR) and development site in Research Triangle Park (RTP), North Carolina. The production databases are mission-critical: "Loss of an ERP database has an immediate business impact, including inability to ship product, take orders, or provide effective customer support," says Rich Harper, ERP storage architect for Cisco IT's Enterprise Storage Services group.

Companies measure the effectiveness of database failover using two metrics: recovery point objective (RPO) and recovery time objective (RTO). RPO indicates how up-to-date the database is when restored to service, while RTO refers to how long it takes to bring the DR database online with the best possible RPO. A perfect RPO of zero seconds would require synchronous replication between sites. Given the 3000-mile distance between sites, Cisco strives for an RPO of less than one minute from the production database's current state.

There are usually trade offs between RPO and RTO in a viable DR architecture. For replication distances of 100 miles or less, the best RPO and RTO combination, zero seconds and a fail-over measured in seconds, can only be achieved through synchronous replication of the entire database over high-bandwidth, low-latency interconnects such as dense-wave-division multiplexing (DWDM) and dark fibre. But for distances over 100 miles, the latency that the speed of light introduces—approximately one millisecond per 100 miles—begins to affect performance for highly active, transaction processing databases. Some types of companies, such as financial institutions, choose synchronous replication over short distances because their paramount goal is zero data loss. Other companies, including Cisco Systems, accept nominal data loss in exchange for better database performance and the arguably better disaster protection that comes from using a DR site that is farther away.

CHALLENGES

Latency Over 3000 Miles

In its SAN environments, Cisco has long used EMC SRDF for local database replication, primarily for daily backups. When it became possible to use SRDF/S synchronously over FCIP, however, Cisco IT concluded that this approach did not meet the company's business needs for replication between production databases in San Jose and the DR site in RTP. The reason is that SRDF/S in synchronous mode—the only SRDF/S mode that enables an RPO of zero seconds and ensures data consistency during replication—depends on a sequential process. This process streams write I/Os one at a time. That is, when a write I/O occurs on the production or primary frame, the process first transfers the changed track containing the I/O to the remote frame, where it updates the copy. Only then does the primary frame acknowledge the I/O to the host so that the next write I/O can be processed. "Because of the latency incurred when commands and data travel between the two frames, the farther apart the frames are located, the longer it takes to acknowledge the I/O back on the host," says Harper.

And latency adds up quickly. Even if the primary and secondary arrays are in the same room as the host, the latency from the host to primary frame to secondary frame and back to the host is approximately four milliseconds. Over fibre channel, every hundred miles adds approximately one millisecond of latency due to the speed of light, which amounts to about 30 milliseconds for the 3000 miles between San Jose and RTP. More latency results from the roundtrip handshake and acknowledgement process for each I/O activity. With all factors considered, the round-trip time (RTT) latency for transferring a single write I/O to RTP is 61 milliseconds. "That might not sound like a long time, but users notice the slowdown of the database when I/O latency reaches about eight milliseconds," says Harper. "With 61 milliseconds of latency, a transaction that usually would take ten seconds, for example, might take two minutes—far outside the scope that our users would tolerate."

Electronic Journaling

To support disaster recovery, Cisco uses an application-based asynchronous process called electronic journaling. In Oracle electronic journaling, transactions that arrive at the production database are first stored in a redo log before being committed to the actual database volumes. Individual transactions continue to fill the redo log. When the redo log fills up, the redo log pointer is switched to an empty log and the Oracle archive copies the previous log to create an archive log. This archive log, in turn, is copied over the WAN to RTP using host-based scripts that Cisco developed internally. Once the archive log arrives at the DR site, it is applied to the standby (DR) database after a built-in delay of 45 to 60 minutes, to protect against logical data corruption. The sequential archive logs contain a map for all changes occurring in the production database between two points in time.

To set up this electronic journaling DR process, Cisco IT performs the following steps:

Copy the production database from backup over the WAN with host-based scripts using a UNIX "dd", in a copy process that can take 24 to 48 hours.

In parallel with step 1, transfer archive logs from production to the remote DR host for storage.

After the database is completely copied, apply the large backlog of archive logs that has accumulated during the database copy to the DR host. This step can take an additional 24 to 48 hours.

The host-based replication method for DR databases and logs posed several drawbacks for Cisco:

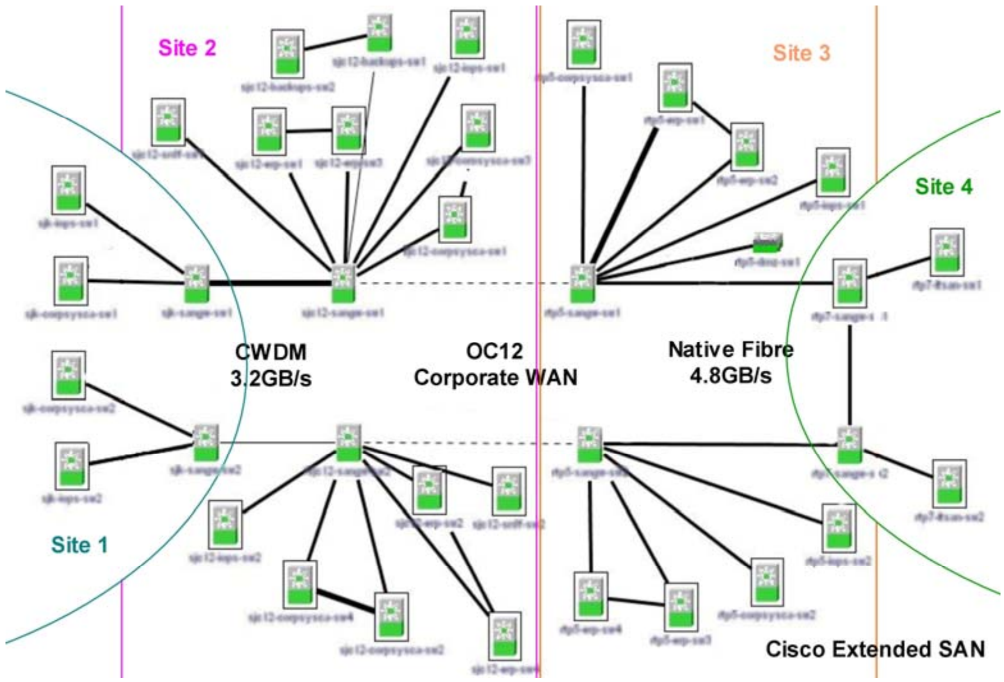
- High management overhead—Cisco IT had to manage host-based, archive-replication scripts for each database on each host. These scripts compressed and cleaned up archive file systems when the remote hosts acknowledged receipt of the logs.
- Use of host CPU resources—Production host resources were used to perform compression as well as to perform a network "push" to the remote DR host. Furthermore, Cisco IT maintained a dedicated host or

hardware partition at the DR site to apply the logs to the mounted standby database.

- Shared host resources for DR and development—To avoid purchasing and maintaining standby hosts capable of handling a failed-over production database that would normally be underutilized in standby mode, Cisco IT created a small host partition to apply archive logs to the DR standby instance and used the remaining larger host partition for development. In the event of a fail-over, Cisco IT would reverse the partition sizes so that the larger partition was available for production. However, changing the partition sizes took time, increasing the RTO.
- Log management—If an archive log at the DR site was lost before being applied to the standby instance but after having been purged from the production site, the DR database became useless because it could not be advanced in time past the missing log. Therefore, Cisco IT had to initiate a full, host-based copy of the entire production database over the WAN. Another problem with log management arose during periods of high database activity, when the production database might generate more than 100 logs per hour. Unable to handle this volume, the log transfer mechanism fell further and further behind. Managing the build up of unapplied logs before they overflowed the DR host's file system was becoming a full-time job. "Eventually, the back up would be more than a day behind or a log would be lost, which is unacceptable for disaster recovery," says Harper. In this situation, Cisco IT had to once again refresh the entire database over the WAN.
- Potential for getting too far behind to catch up—As databases became larger and more active, a new challenge arose. "Suppose a database had to be recopied over the WAN to refresh the DR database," says Harper. "In the time it took to replicate the database using host-based scripts, so many logs accumulated that we couldn't catch up for several days, if at all. At that point, we faced the same log-management challenge and the possible need to copy the database again. We might never catch up."
- Outage required during failback—If Cisco failed over the production database to the DR site, the process of failing back required an extended outage while the entire database was replicated back to the primary site.
- Continuing need for daily backups—The DR process did not replace the need for daily backups at the primary site.
- Potential for logical database corruption—A risk of real-time or near-time replication schemes is that logical data corruption will propagate to the DR site, rendering both databases useless. The closer in time that the two databases are replicated, the greater the exposure.

Cisco IT evaluated potential solutions. A dedicated storage IP network that spanned a distance of 3000 miles and provided the bandwidth needed for the best RPO and RTO would be prohibitively costly—and also inconsistent with Cisco's commitment to shared networks. Therefore, the new solution would have to share bandwidth on the corporate WAN without affecting other critical data streams. Any proposed solutions could not degrade production performance because Cisco ERP databases are very sensitive to I/O latency. Finally, Cisco needs the ability to roll forward or backward the database at the DR site. This meant that, regardless of the solution, Cisco would need to transfer archive logs between sites.

Figure 1. Cisco Fabric Manager View of Director-Class Switches. A dual redundant fabric comprises approximately 5000 ports and spans 3000 miles and four data centers.



SOLUTION

Unifying SANs By Connecting Remote Sites With FCIP

The Cisco SAN architecture includes two Cisco dedicated MDS-9000 Director Switches in each data center—one for each fabric—to act as SAN gateways. All other switches in the data center tie back into this core. “This design works well for companies that need to install special services blades, such as IP Storage Services (IPS) blades, into the gateway switches or to tie metro sites together with single mode fibre, CWDM [coarse-wave division multiplexing] or DWDM,” says Harper.

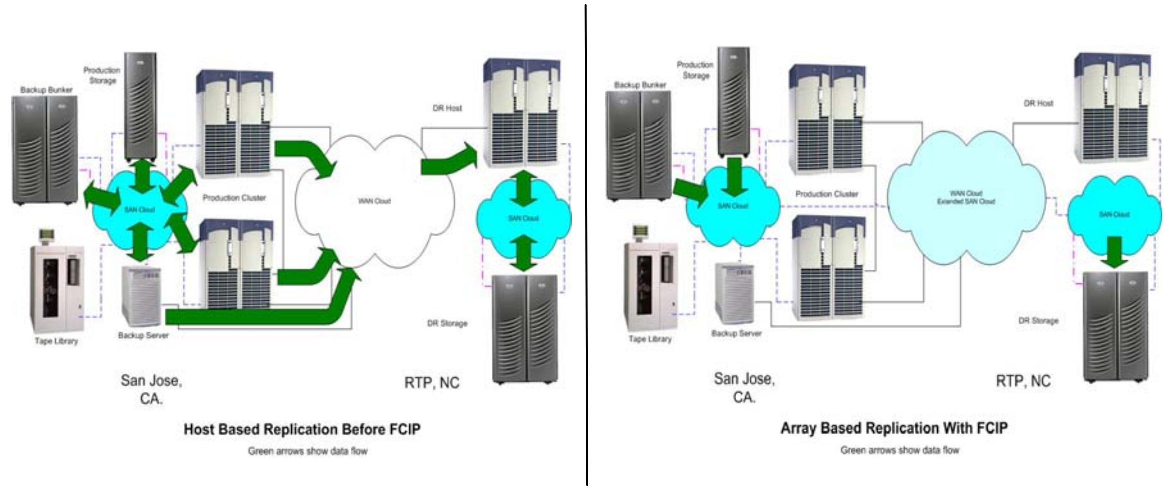
As the first step in upgrading its replication architecture, Cisco IT installed FCIP blades in the Cisco MDS-9000 Director Switches and configured FCIP tunnels. Cisco performed its initial testing using IP Storage Services (IPS)-8 modules, and then later upgraded to multiprotocol services (MPS)-14/2 modules to take advantage of their hardware compression features.

Deploying FCIP created an integrated, monolithic view of the IT data center’s SAN and enabling centralized management of a SAN spanning four data centers and 3000 miles (Figure 1). Now Cisco IT can manage elements in the SAN such as Cisco MDS-9000 Director Switches, VSANs, zonesets and zones from a single view.

“The Cisco philosophy is that voice, video, and data can travel over the same converged IP network,” says Harper. “SAN traffic carrying database transactions to and from the storage arrays is just data—lots of it.”

By interconnecting the storage arrays’ remote data facility (RDF) ports via FCIP, Cisco IT enabled replication between arrays on demand, and without overburdening host resources (Figure 2).

Figure 2. Comparing the Data Flow With Host-Based Replication (Without FCIP) and Array-Based Replication (With FCIP)

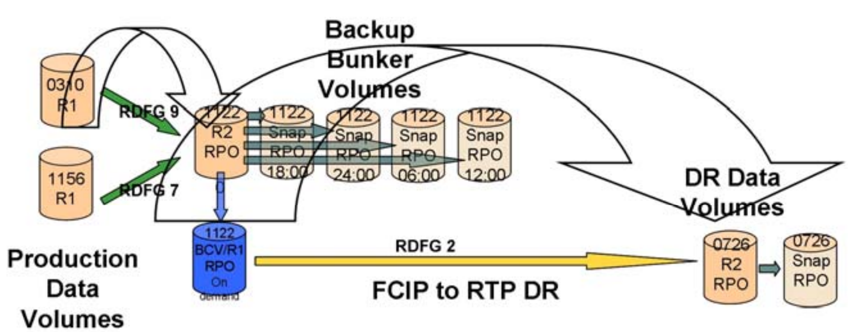


"If Cisco does not maintain incremental relationships between the production and DR databases, failback is a painful process that requires extended downtime," says Harper. "And without the ability to failback in a reasonable period of time, Cisco could only make the decision to fail-over in extreme circumstances. This limits fail-over testing, which can result in a lack of preparedness to properly execute a DR plan."

Achieving an Incremental Relationship

A limitation of host-based replication is that it does not allow Cisco to make incremental copies of only those changes since the last refresh. Cisco initially set out to resolve this issue by using EMC SRDF/S (traditional SRDF) over FCIP. The goal: maintain an incremental relationship between the production backups and DR databases but not try to keep them synchronized. With this approach, if a DR database fell too far behind in its application of logs, or if a log was lost, Cisco could incrementally refresh the entire DR database from the last production disk backup—a process that would take hours, not days. To accomplish this, Cisco connected the two sites' SANs with FCIP and then used an SRDF "multi-hop" approach from the backup bunker frame to avoid affecting production performance (Figure 3).

Figure 3. Multi-Hop Example: RDF Data is Replicated From Production to DR Site over FCIP



Using EMC SRDF/S, Cisco could copy data directly from the production backup frame to the disaster recovery frame over FCIP every six to twenty-four hours. The frame transmitted only those changes that had occurred since the last backup. Typically, Cisco ERP databases change less than ten percent in any 24-hour period. Therefore, by transmitting only the write-folded[1] changes between refresh cycles, Cisco IT avoided the problem of accumulating

[1] Write-folding refers to sending only the last write of multiple writes to the same track.

huge archive backlogs while refreshing a host-based DR database from the backup host. “In our pilot with SRDF/S over FCIP, we were able to fail over and later fail back within a two-hour window, with minimum business impact,” says Harper.

Traditional SRDF worked well for Cisco with a few exceptions, according to Harper. “On the positive side, the only incremental cost was adding an additional multi-hop device in the backup bunker [shown in blue in Figure 3.] This device, known as a Business Continuance Volume (BCV)/R1, ensured that the replication process did not directly affect production. The approach allowed Cisco IT to maintain snapshot copies off the R2s[2]. The drawback of using multi-hop devices is that it complicated backups and required an additional storage copy, with its associated cost.

Harper notes that a limitation of the using a traditional SRDF/S approach to make an asynchronous copy is that any interruption to the replication process means that the DR database will not be consistent with the production database. Therefore, it will not be able to restart.

Achieving an Asynchronously Updated, Consistent Relationship Without Latency

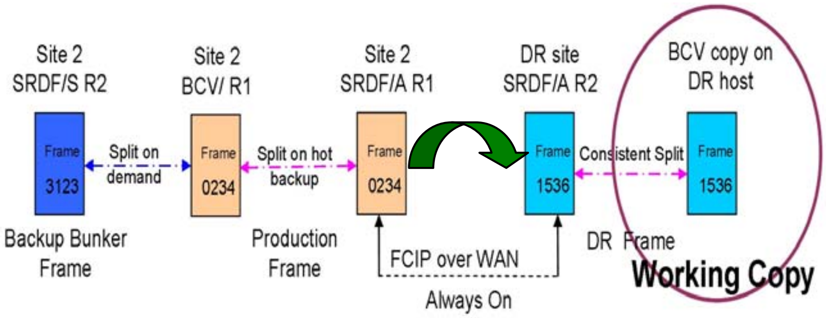
To overcome the limitations of SRDF/S for long-distance database replication, Cisco conducted a test pilot using EMC SRDF/A (asynchronous), which can sustain a long-distance, consistent production image between databases and also avoids latency to the production instance. “Because it’s designed to overcome the challenges of long-distance replication, EMC SRDF/A works over a long-distance infrastructure like FCIP,” says Harper. If an SRDF/A image is interrupted during its refresh cycle, no more than two cycles are lost—a total of 60 seconds in Cisco’s testing—and the remainder of the DR database is immediately restartable.

Unlike SRDF/S, SRDF/A does not depend on synchronous communication to maintain consistency. Rather, it buffers the write I/Os for a specified time period—30 seconds in Cisco’s case—and then transfers this buffered set of write I/Os across the WAN while SRDF/A begins collecting a new set. At the DR frame, the sets (or cycles) are not applied until the frame receives the entire 30-second cycle. The combination of this approach and the write-consistency techniques found in EMC’s Enginuity Consistency Assist (ECA) product ensures that data on the remote site remains consistent and restartable, no matter what interruptions might occur outside of the remote storage array. “Not only does SRDF/A sustain a constantly updated incremental relationship between the production frame and DR frame, it provides some of today’s best possible RPO and RTO targets over this distance,” says Harper.

The pilot showed that SRDF/A frame-to-frame replication of the entire database would avoid the need for Cisco to use host-applied archive logs. Instead, the two frames remain consistent by communicating changes outside the Oracle process. Seconds after a change occurs on the production frame, it is transferred and applied at the track level to the disaster recovery frame without any involvement from the host. The pilot also showed that by using the frame-based approach, Cisco could dedicate all host resources at RTP to development until they were needed for failover, freeing up the partition previously reserved for applying archive logs to the mounted standby database. Therefore, Cisco could dedicate the entire host to bringing up the failed-over database when required, which decreases RTO.

[2] R1 and R2 are EMC terms, and the “R” stands for “resource.” R1s are primary devices or production databases, and R2s are replication targets, such as backup or DR databases. In failover mode, R1s and R2s can switch roles.

Figure 4. Multi-Hop to Local Backup Bunker, and SRDF/A to Remote DR frame over FCIP. The green arrow marks the normal SRDF/A data flow.



Cisco's SRDF/A pilot had some limitations because the EMC Enginuity 5671 array micro-code was not yet internally qualified for production databases at the time testing took place. This prevented Cisco from using concurrent SRDF/A to maintain the current production backup strategy. Instead, Cisco used a multi-hop device to perform backups. The production R1 volumes were directly associated via FCIP to the remote DR database volumes, which became R2s (Figure 4). "This approach complicated backups but it provides a working solution until we can use the Enginuity 5671 micro-code in the production environment and eliminate the multi-hop devices," says Harper.

Cisco's goals for the SRDF/A pilot were to:

- Test the solution's manageability
- Measure WAN impact
- Verify DR consistency under various test conditions.
- Determine the solution's advantages, disadvantages, and viability across multiple DR environments.

"SRDF/A performed very well during the pilot and successfully maintained a remote conistent copy that was restartable[3] under all the failing test senarios that we tried," says Harper. The remote copy was recoverable—that is, archive logs could be applied against it—only when the production database was placed in hot backup mode and the remote BCVs split after sixty seconds. BCVs need to be used because the SRDF/A R2s do not support snapshots. When Cisco IT split off the remote BCV and used it as the valadation copy, performing tests or database copies did not interrupt the replication process of SRDF/A. .

For the pilot, Cisco used SRDF/A over FCIP to transfer archive logs, redo logs, and database volumes. "In effect, this meant that all database changes were replicated three times," says Harper. In fact, the standard archive log shipment process, which Cisco IT did not change in response to the fact that the original non-SRDF/A replicated DR database instance operated in parallel—resulted in a total bandwidth consumption during the pilot of 3.5 times the amount of the change. Even after FCIP compression, approximately 1.75 times the bandwidth required for write changes was being transferred over the WAN. "Even at 1.5 times the bandwidth, this solution could not scale across the number of multiple DR environments we wanted to replicate," says Harper. "We would need to increase the size of the corporate WAN to OC-48, a very expensive proposition considering the distance between the two sites."

Therefore, Cisco IT postponed plans for additional roll-outs of a SRDF/A-based, fully-replicated databases until it could resolve the bandwidth issue. However, Cisco IT realized it could immediately use SRDF/A in other ways to improve RPO at the DR site. The objective would be to improve the process for transferring archive logs between

[3] Restartable in this sense implies that the database can be brought up with Oracle performing crash recovery. The database cannot be brought up in standby mode, nor can archive logs be applied, but it can be opened as a valid database.

sites. Although the existing method, an internally-developed script, was reliable, it could fall behind when processing and shipping logs between sites. Backlogs were especially likely to occur during periods of heavy loads, when the database activity might generate one hundred 500-MB logs per hour.

Transferring Archive Logs

During the pilot, Cisco transferred archive logs using the same SRDF/A mechanism used to transfer redo logs and data files. “It became apparent that even though the corporate WAN could not sustain a full roll-out of SRDF/A across all the desired DR database environments, we could immediately benefit by using SRDF/A instead of our internally-developed scripts to transfer the archive logs,” says Harper. With SRDF/A over FCIP, no more than one or two logs are in transit, and therefore unavailable, at any given time (Figure 5). In this approach, a BCV copy of the SRDF/A R2s containing the archive log file system is split using the `-consistent` option on a defined interval (hourly), and then the volume is activated, file system checked (fsck’ed), and mounted on the DR host. The logs can then be copied to an archive file system, or soft linked directly so that they are available for the application to apply to the standby database.

Cisco IT will decide whether to replace this method of transfering archive logs with Oracle Data Guard after it tests Oracle Data Guard 10g.

Achieving the Best Possible RPO Without Replicating the Full Database

Though a near-time transfer method for replicating the archive logs improved Cisco’s RPO, it still left an RPO gap. That is, transactions in the redo logs that have already been commited to the databases’s raw volumes can represent seconds or several minutes’ worth of data, depending on the size of the log and number of transactions processed. In Oracle, transactions are written to the redo logs before being commited to the data volumes of the database. To achieve the best RPO possible with an SRDF/A solution, only the redo logs and archive logs need to be continually replicated—an approach that Cisco IT adopted (Figure 6). “In the SRDF/A approach, in which only the archive and redo logs are continuously replicated, RTO suffers somewhat because logs still have to be applied at the DR site prior to a database fail-over,” says Harper. “However, a benefit of the approach is added protection against logical data corruption at the remote site.”

Figure 5. Comparing the Transfer Flow for Archive Logs and Redo Logs

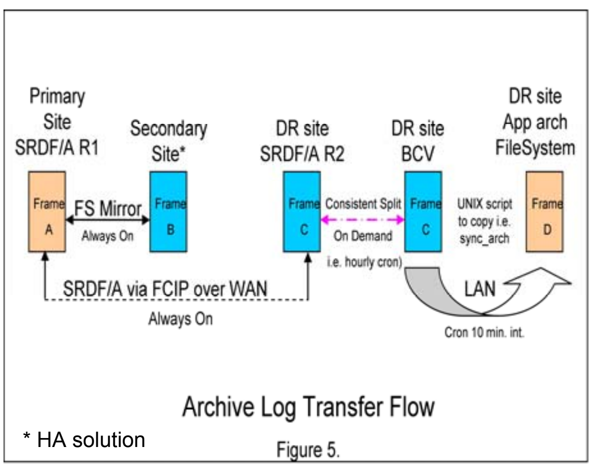
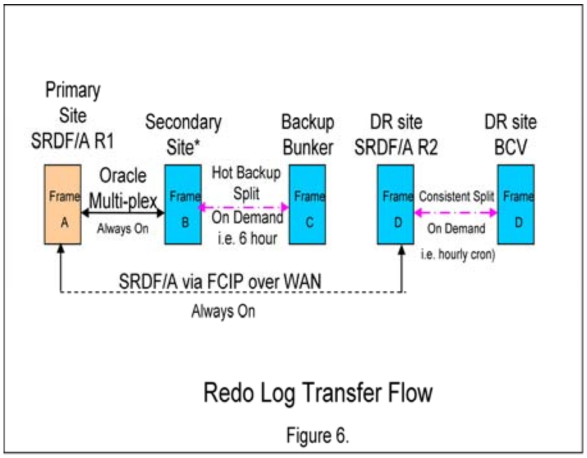


Figure 6. Comparing the Transfer Flow for Archive Logs and Redo Logs



RESULTS

Cisco experienced the following results from the pilot with SRDF/A over FCIP:

- Greatly reduced RPO—Cisco achieved sub-minute RPO over a distance of 3000 miles.
- Greatly reduced RTO—By replicating the entire database using SRDF/A over FCIP, Cisco could achieve an RTO measured in minutes instead of hours.
- High general transfer rates—Using SRDF, Cisco achieved frame-to-frame transfer rates of 50MB/second, and higher rates were possible, according to Harper.
- Incremental relationship between production and DR—Using both SRDF/S and SRDF/A, Cisco maintained an incremental relationship between the production and DR databases, enabling fail-back without an extended downtime.
- Faster DR WAN refreshes—Using an incremental frame-to-frame refresh, Cisco IT can refresh a 4TB database over the WAN in approximately two hours, with one day of accumulated changes, compared to 24 hours using a host-to-host transfer.
- Simplified SAN management—By tying the two remote sites together, Cisco gained a single SAN view of four data centers. This enables Cisco to manage the entire SAN from any switch, using Cisco Fabric Manager, a Web-based management application for Cisco MDS 9000 Family switches.
- More effective WAN utilization—Used for transferring data volumes, archive logs, and redo logs, the SRDF/A replication solution across the entire DR infrastructure would not scale when using Cisco's OC-12 corporate WAN. By using SRDF/A over FCIP and transferring both archive logs and redo logs, Cisco has achieved the same RPO as it would have by also transferring the data volumes. Cisco achieved an average compression ratio of 2:1, using mode-1 compression. The Cisco IPS-14/2 blades provide hardware-based compression and can compress at the 2:1 ratio up to full line rate.
- Reduction of host resources—Frame-to-frame transfers avoid the need for host resources to perform the replication, and using SRDF/A to maintain database data volumes eliminates the need to assign a DR host to apply logs in standby mode.
- No production impact—SRDF/A did not affect production or introduce additional latency, and neither did setting up a multi-hop using SRDF/S. Cisco did require downtime at the outset of the pilot to move the database to the SRDF/A architecture.

Harper summarizes, "The SRDF/A over FCIP architecture provides greater flexibility and allows better RPO and RTO than traditional electronic journaling. However, it also adds more complexity, significantly increases the WAN load, and requires at least one additional BCV copy at the DR site. Companies considering SRDF/A over FCIP should weigh the relative advantages and disadvantages in their business environment before implementing. For Cisco, achieving RPO of less than one minute outweighs any disadvantages."

LESSONS LEARNED

Harper and his team intend to apply the following lessons learned to future replication projects at Cisco:

- Unless the WAN is running near full capacity, assigning storage traffic to the scavenger class of service (CoS) does not affect the performance of SRDF/A or FCIP. The Cisco WAN provides three CoS: voice, standard (e-mail and network management activities), and scavenger class. Storage traffic is considered scavenger class because it is not affected by slight delays. "FCIP is more aggressive than normal IP traffic, and if it has to retry, it bounces back more quickly," Harper explains. Using scavenger class for storage traffic helps ensure that capacity is used for time-sensitive applications such as IP telephony.
- The write-acceleration (WA) feature on Cisco MDS 9000 switches increases the transfer rate with traditional SRDF/S by halving the latency of a write I/O. However, it does not benefit transfer rates with SRDF/A because SRDF/A does not wait the one RTT for the remote storage to send back XFER_RDY, which the WA feature saves by spoofing.
- Numerous factors can affect transfer rates, including minor configuration settings on the storage arrays and FCIP tunnels. For the EMC DMX frames running SRDF/S (not SRDF/A), the default setting of the DA_COPY parameter internal to the frame does not allow filling a 3000-mile pipe using only one RA group. However, increasing this parameter from the default of 28 hex to 50 hex increased the transfer rate on the Cisco network from approximately 21 MB/s to approximately 50 MB/s. The formula for calculating the maximum transmission rate of SRDF/S through the EMC RAs is:
- (DA_COPY parameter value) x (track size) / (RTT x 2)
- When using the WA feature, substitute `RTT` for `RTT x 2`.
- SRDF over FCIP attempts to load-balance itself, and therefore reduces its operating speed to match the slowest side of the fabric. "Any factor that can affect the transfer rate of one fabric in a dual-redundant fabric using traditional SRDF will typically affect the other as well," says Harper. Identifying the side of the fabric with the problem can present a challenge.
- Transferring only the archive logs and redo logs—not data volumes—maximizes RPO at a minor cost to RTO. For Cisco, this is an acceptable trade-off.
- It is possible to achieve incremental relationships between the production and data recovery databases with traditional SRDF/S and FCIP. "Even if a company cannot afford to use SRDF/A to keep databases within 30 seconds of each other, which was Cisco's goal, it might want to investigate using SRDF to achieve some degree of incremental relationship," says Harper. "The ability to update databases over FCIP once a day or once every six hours, for example, is valuable for most companies."
- With an SRDF/A solution, protection against logical data corruption is especially important if the DR site is used to restore a corrupted database. At Cisco, the backup database is co-located on the same host with production. Therefore, Cisco maintains multiple RPOs on a backup bunker, using both SRDF/S R2s and snapshots, that can be used to restore the production database. For further protection against logical corruption that might render the DR database ineffective, Cisco could establish and split the BCV copy at the DR site with the SRDF/A R2s on a given automated cycle, such as daily.

NEXT STEPS

Cisco IT plans to use a combination of strategies to further improve its ability to meet business goals for replication using existing resources. For example, when the EMC Enginuity 5671 frame micro-code is fully qualified, Cisco IT will establish concurrent SRDF/A relationships: one between production databases and the remote DR, and another between production databases and a backup bunker. The link between production and DR will not be kept in an established state, as it was in the pilot, but rather will be established and split on a defined time cycle. This is the same approach that Cisco uses for the backup bunker copies—that is, the database is synchronized only one hour before to the backup target time and then split after the database is put into hot backup mode (see Figure 7). Using an extended cycle time for the DR database updates extends the advantages of write folding—which refers to sending only the last write of multiple writes to the same track—to minimize the amount of data transferred across the Cisco WAN.

Figure 7. Data Volume Transfer Flow

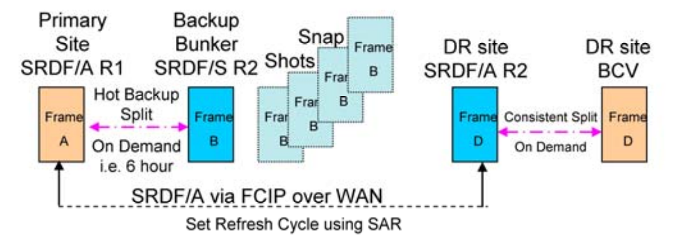


Figure 7.

Using SRDF/A in this architecture is important for three reasons:

- As the remote disks asynchronously synchronize, the latency does not affect the protect volumes.
- By using the production volumes as the source, Cisco can avoid the cost of a multi-hop device.
- Fail-back replication can take place directly to the production volumes and not through a two-stage refresh via the multi-hop devices .

By staggering the refresh from production to DR of the most critical DR databases, and by transferring the redo logs and archive logs all the time (Figures 5 and 6), Cisco can achieve the best possible RPO while still maintaining an incremental relationship between production and DR. “This approach provides all the incremental benefits of a reasonable fail-back plan as well as the ability to incrementally refresh the DR database on demand,” says Harper. In this hybrid approach, Cisco would still maintain a standby database applying logs to minimize RTO on the protected BCV copy of the R2s. The logs would be discarded after application. In this way, Cisco can prevent the database volume refresh process over the WAN from interrupting the standby database’s log application. If the standby copy ever falls too far behind, or a log is lost, the protected BCVs can be merged with the R2s and synchronized. Cisco might automate this standby synchronization cycle to occur daily, ensuring that DR never falls behind in log application by more than 24 hours.

The refresh cycle time of the DR data volumes over the WAN—for example, hourly or every six hours—and the RTO target might make it possible to avoid mounting a standby database or applying archive logs at all except in failover situations. This would create several advantages for Cisco, including using the DR hosts entirely for development except during failover, and eliminating management of the standby databases and logs for application. The downside to this approach, however, is that for every refresh, Cisco will need to place the production database into hot backup mode before splitting the remote R2s after they have been fully synchronized via SRDF/A. The remote BCV copy plays an important role in this solution by preserving the last RPO while the R2s are being updated. Another, minor disadvantage of this server-less approach, compared to applying archive logs to the standby database, is that it does not provide constant validation, which ordinarily increases Cisco IT's confidence in the integrity of the standby instance.

Additional goals for replication at Cisco include:

- Begin replicating the Amsterdam NL ERP databases in RTP. When this occurs, Cisco IT will have acquired a single enterprise view of the entire IT SAN, and centralized management.
- Use FCIP to tie engineering SANs to the IT SANs so that Cisco IT can centrally view and manage the entire SAN infrastructure.
- Utilize FCIP as a fail-over path for campus inter-building SAN extensions, such as CWDM and native dark fiber.
- Implement FCIP and SRDF/A for other mission-critical ERP databases as Cisco increases its WAN capacity between RTP and San Jose. "The business decision for any company is whether the increased disaster recovery capabilities are worth the cost of a WAN capacity upgrade," says Harper. "Disaster recovery capabilities are rarely needed, but when they are, companies are very glad they invested in the infrastructure."

Harper summarizes, "FCIP creates huge benefits for business continuity at Cisco. It improved RTO, reduced our RPO [recovery point objective] to under 60 seconds, simplified data mobility by enabling us to offload database replication from the hosts to the storage arrays, and enabled us to centralize SAN management."

FOR MORE INFORMATION

To read the entire case study or for additional Cisco IT case studies on a variety of business solutions, visit Cisco on Cisco: Inside Cisco IT www.cisco.com/go/ciscoit

NOTE

This publication describes how Cisco has benefited from the deployment of its own products. Many factors may have contributed to the results and benefits described; Cisco does not guarantee comparable results elsewhere.

CISCO PROVIDES THIS PUBLICATION AS IS WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING THE IMPLIED WARRANTIES OF MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE.

Some jurisdictions do not allow disclaimer of express or implied warranties, therefore this disclaimer may not apply to you.

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.**