the Availability Digest

www.availabilitydigest.com

Banks Use Synchronous Replication for Zero RPO

February 2010

Two banks, the Bank of New York and the Fifth Third Bank, have each built highly-resilient, triplexed data center complexes that use a mix of asynchronous and synchronous replication. By doing so, the banks have achieved zero data loss and a recovery time of only two to four hours following any disastrous failure of their production data centers.

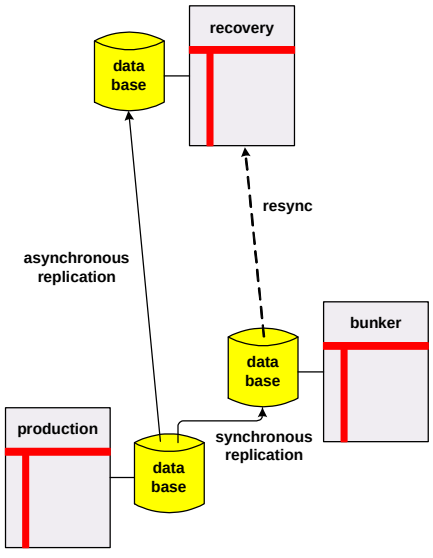
The Problem

Real-time banking services are fundamental to commerce today. If consumers cannot access and manage their cash, and if businesses cannot move funds between their own accounts or those of their vendors and customers, commerce comes to a halt. For this reason, many banks maintain disaster-recovery data centers that can be put into operation should their production data centers be put out of commission by some disastrous event. To make it unlikely that any single disaster will take out both the production and the disaster-recovery data centers, disaster-recovery sites should be located far from the production sites (typically by a few hundred miles).

Equally important is the data upon which the banking applications depend. Transactions lost due to a production-site failure can wreak havoc on the bank's customers. Therefore, it is important to be able to recover all completed transactions following a production-site outage. This implies real-time synchronization of data between the production site and the backup site so that no transaction completes on the production site before it has been safe-stored at the backup site. To prevent performance deterioration caused by the application having to wait for the backup site to store the data, the production and data sites must be located close to each other.

This presents a conundrum of sorts. The sites have to be far apart, and they have to be close together.

The Bank of New York and Fifth Third Bank have solved this problem by combining asynchronous replication to a remote disaster-recovery site with synchronous replication to a third nearby "data bunker" site. Asynchronous replication[1] to the recovery site may lose some data if the production site fails, but synchronous replication[2] to the bunker site guarantees that all data is preserved. On the other hand, the bunker site may be taken out with the



Triplexed Data Centers

[1] Asynchronous Replication Engines, *Availability Digest*; November 2006.
[2] Synchronous Replication, *Availability Digest*; December 2006.

production site (hopefully not) due to a common disaster; but in this extreme case, the recovery site is still available to take over.

The Bank of New York

The Bank of New York is the oldest bank in the U.S., having been founded in 1784. It was the first corporate stock to be traded on the New York Stock Exchange. The Bank operates in 33 countries and is one of the largest security clearing agents in the U.S. It clears 50% of all U.S. government securities.

The Bank's Data-Center Problem

As it entered the 21st century, the Bank maintained several data centers. The production and disaster-recovery sites were in close proximity. The production sites were in lower Manhattan, and the disaster-recovery sites were elsewhere in New York City.

The 9/11 disaster changed all that. The Bank realized that it had to move its IT operations to a less risk-prone area. Furthermore, it had to separate its production and disaster-recovery facilities by a large distance to be immune to a widespread disaster. But it had to do this in such a way that it would lose no data in the event of a production data-center disaster.

The Solution – Triple Data Centers

As a consequence, the Bank decided that it would consolidate its multiple data centers into a single production data center with remote protection.[3] Unlike the common practice of other companies, the Bank did not attempt to locate its IT facilities near its headquarter operations. Rather, after an extensive search, it chose a remote area with a lower risk profile several hundred miles from New York City to locate its production data center. It chose a disaster-recovery site in another state with an equipment configuration that was an exact mirror of the production facility.

The Bank planned to keep the backup data-center's databases in synchronism with the production data center via asynchronous replication. But this meant that some data, up to 60 seconds by the Bank's estimate, might be lost should the production center suffer an outage. To correct this, the Bank planned a third data center near the production site. The third data center would use synchronous replication to keep its database synchronized with that of the production site.

Therefore, should the production site fail, the nearby data center would contain all transactions that were executed up to the time of the outage. No data would be lost. This data center was to be hardened so that it could survive the effects of a disaster that took down the production site. In effect, it was a data bunker.

The data bunker was to be linked to both the production site and the backup site. In the event of a production-site outage, the backup site would quickly bring its database up-to-date by establishing a session with the data bunker and downloading only the data changes that it had missed. Once this was accomplished, the backup site would be put into production with zero data loss.

Construction of the backup data center began in 2002 while normal operations continued. It was constructed with excess space and populated with excess server capacity to provide scalability for the Bank's rapid growth. Once it was completed in 2004, it became the backup site for the existing production data centers.

[3] The Bank of New York: TPC Data Center Consolidation Project, *The Computerworld Honors Program Case Study*; 2006.

The new consolidated production data center came online at the end of 2005 with a contingent of equipment identical to that of the new disaster-recovery data center. Five petabytes of data were migrated to it. Once this was complete, and after exhaustive testing by key customers and business partners, consolidation of the Bank's three original data centers into the new production center was accomplished in ninety days. The initiation of asynchronous replication to the backup site and synchronous replication to the data bunker completed the move. At no point during the move was customer service impacted.

Networking

To achieve the desired reliability of IT operations, the network interconnecting the sites had to be as resilient as the multiple data-center configuration. The Bank worked closely with several communication carriers to ensure carrier diversity and route diversity so that there would be no single point of failure in the network. An extensive broadband wide-area network backbone was established. If the Bank had been a telecommunications carrier, it would have been the ninth largest in the U.S.

Operating the New Data Center Complex

To maintain continuity of operations in the event of an outage, the Bank set up identical help desks and virtual command centers at both the production and backup sites and moved its experienced personnel to these sites. Sufficient additional personnel were hired so that the entire banking operation can be handled from either site.

Though either site can handle 100% of the help-desk load, customer service is split 50/50 between the two sites during normal operations. The entire operations of the data center can be managed equally well from either site. Therefore, should there be a disastrous event of some sort, no personnel have to be moved, a feat that might not otherwise be possible under some conditions following a regional disaster. Uninterrupted customer service can be provided around the clock.

Data-Replication Technology

The Bank chose EMC storage systems and replication software to implement the redundant data network needed to execute its contingency plans. Multiple EMC Symmetrix DMX high-performance storage systems, each capable of managing about one petabyte of data, were installed in each data center.

Synchronization of the diverse databases is accomplished via EMC's Symmetrix Remote Data Facility (SRDF), a software-based data-replication engine that runs on the Symmetrix storage systems. The SRDF/Star option provides asynchronous replication to the disaster-recovery data center and synchronous replication to the data-center bunker from the same source volumes. It facilitates rapid resynchronization and failover in the event of a source-system failure.

The Result

The triply-redundant data-center configuration implemented by the Bank of New York[4] meets an RPO specification (recovery point objective) of zero data loss and an RTO specification (recovery time objective) of two to four hours. This compares to several hours of lost data and a recovery time of sixteen to twenty hours with the old tape backup method.

[4] The Bank of New York is now The Bank of New York Mellon following its 2007 merger with Mellon Corporation.

The new contingency configuration provides resiliency to man-made and natural disasters such as the 9/11 terrorist attack, Hurricane Katrina, and the 2003 U.S. East Coast blackout. It is fully compliant with the regulations outlined by the Federal Reserve, the Office of the Currency, and the SEC (Securities and Exchange Commission) in their white paper "Interagency Paper on Sound Practices to Strengthen the Resilience of the U.S. Financial System," which reflects the lessons learned from 9/11.

Meeting an RPO of zero and an RTO of two to four hours with a geographically-distributed system ensures that the Bank's services to its customers will continue in the face of any disaster.

Fifth Third Bank

Meanwhile, another bank, Fifth Third Bank, faced the same problems as the Bank of New York. Fifth Third Bank has almost 1,200 banking centers across ten states and manages $100 billion in assets. Handling millions of dollars of banking transactions per day, its business and retail customers depend upon the accuracy and availability of their banking data.

Following the 9/11 terrorist attack, the Bank had to comply with new government mandates for information protection and recovery. Consequently, it felt that it had to improve the resiliency of its IT systems. It also saw an enhanced opportunity to market its services to new customers if it had a strong resiliency story. Furthermore, if the system were structured properly, testing of new applications and application modifications could be done with current data rather than weeks-old stale data based on magnetic-tape backups.

The Fifth Third Bank Solution

The Bank implemented a comprehensive business-continuity, backup, recovery, and archiving solution[5] similar to that described above for the Bank of New York. It built new production and disaster-recovery data centers that were 370 miles apart. This gives the Bank protection from a large disaster that might take out its production data center. The disaster-recovery data center is kept in synchronism with the production data center via asynchronous replication.

However, the Bank estimated that it might lose up to 30 seconds of data following a production-site outage because of the asynchronous-replication solution. To solve this problem, the Bank constructed a third data center closer to the production center to act as a second up-to-date data repository. The third data center is kept synchronized via the synchronous replication of data from the production site. This data center is far enough away from the production site to ensure that it will unlikely be taken out by a common disaster but close enough so that application performance is not affected.

In the Fifth Third Bank's solution, failover can be either to the nearby site or to the remote disaster-recovery site. The site chosen to become the new production site maintains synchronism with the other site via asynchronous data replication during the outage of the original production site.

The Bank also chose EMC Symmetrix DMX storage systems. EMC's SRDF/Star is used to provide the dual synchronous/asynchronous replication. Being an early adopter of dual replication, the Bank worked closely with EMC to refine the SRDF product. Enhancements included more efficient cache configurations, better monitoring and reporting capabilities, and fuller utilization of the available bandwidth.

[5] Fifth Third Bank: Business continuity, backup, recovery, and archiving, *The Computerworld Honors Program Case Study*; 2007.

Rapid recovery is only possible if the data is consistent and is synchronized across all mainframes and open systems. To ensure this, the Bank uses SRDF's Multi-Session Consistency feature to ensure that all databases are synchronized to the same point in time should a recovery be necessary.

The Result

The Fifth Third Bank's triplexed data centers give it extreme resiliency in the face of any man-made or natural disaster. Following an outage of the production data center due to such a disaster, the Bank will account for all completed transactions at the time of the outage and will be able to recover its services within two to four hours. This compares to a previous recovery time of sixteen to twenty hours from magnetic tape as well as hours of lost transactions.

Interestingly, application managers were slow to accept the new system. They were very comfortable with their existing tape backup procedures upon which they had depended for years. However, with time and education, this reluctance has been resolved.

Fifth Third Didn't Stop Here

Fifth Third Bank is a leader in providing its customers with high- and continuous availability of its services. In addition to the triplexed data centers described above, it also runs applications on an active/active NonStop system.[6] This system has nodes in Florida and Michigan. The transaction load is normally split between the two nodes, and their databases are kept synchronized via asynchronous replication. Should a node fail, recovery is in seconds since all that is required is to reroute all transactions to the surviving node, which is known to be good since it is currently processing transactions.

In fact, the Bank uses this capability to switch all load to the Michigan site when a hurricane approaches its Florida facility.

Summary

The implementation of a triplexed data-center architecture using a mix of asynchronous and synchronous replication is fairly new. By choosing this route, an enterprise can "have its cake and eat it, too." It can ensure that no common disaster will prevent it from offering its IT services to its customers and partners while at the same time ensuring that no data will be lost.

The solutions described above use EMC's SRDF/Star replication facility. Other products can be used as well. For instance, HP's NonStop systems can be configured in this way using remote mirroring to protect against data loss (what HP calls its Zero Lost Transaction, or ZLT, option). A third distant disaster-recovery site can then be synchronized via the HP NonStop RDF (Remote Database Facility) asynchronous replication engine or via several other asynchronous replication engines available from third parties such Gravic, Oracle, and NTI.

It is expected that a new technology called *coordinated commits*[7] will be introduced in the next year. Coordinated commits provide synchronous replication by asynchronously replicating changes and synchronizing the target database only at transaction-commit time. This technology will substantially remove the distance limitation on synchronous replication, allowing zero data loss and rapid recovery times to be achieved with only two systems. However, even with this technology, a third node in the application network is useful to ensure that there is no single point of failure during the time that one node is taken down for maintenance or upgrades.

[6] Major Bank Uses Active/Active to Avoid Hurricanes, *Availability Digest*; October 2007.
[7] HP's NonStop Synchronous Gateway, *Availability Digest*; June 2009.