The Complete Reference

Storage Networks

Part VII

Appendixes

Appendix A

NAS Case Study: The International Image Processing Company

The International Image Processing Company, or IIP, provides a proprietary digital imaging service to academic, educational, and government institutions throughout the United States. Its computer imaging services provide clients with the ability to transfer physical documents, such as historical books, documents, and newspapers, as well as photographic images, such as historical photos, video images, and medical images, into digital media. The company's clients range from major universities to some of the largest archival government institutions.

Despite the scope of its clients, IIP is a small business with less than $5 million in revenues, with plans to grow threefold in five years if it can increase its capacities to bring on more clients. IIP has a small but distributed business organization with headquarters in Texas and major client work sites in New York City, Washington, D.C., and San Francisco. Due to the nature of IIP's work, its IT staff is integrated into the specialized work force that the company employs—given the digital scanning software and processes that are integral to their work. The company also employs two dedicated IT technicians—one for hardware and another responsible for infrastructure software. Additional IT-related activities are handled through the imaging software support and configuration personnel.

The company was challenged by its dependencies on storage, created through its software scanning processes that were increasingly exacerbated by new clients. Daily operations could quickly utilize a terabyte of data storage. Because the imaging scanning product was archived to clients using both optical and tape storage devices, if operations ran short on available storage, the imaging processes would slow and eventually stop until sufficient space was freed up for operations. This slowed billing and eventually cash flow.

The storage situation reached critical levels that started to impact the company's ability to handle new clients and subsequently impacted its planned business growth.

IIP was in the classic predicament of needing a high-performance solution, while being restricted by a minimum budget and a limited IT staff. The evaluation of potential storage solutions ranged from storage area networks (SANs) to IDE RAID solutions. While a network attached storage (NAS) solution would satisfy the size requirement, NAS remained a problematic solution because it also had to be managed remotely at the New York, Washington, D.C., and San Francisco sites.

IIP chose a general-purpose solution, even though its processes were largely proprietary, which was a matter of thoughtful strategy on IIP's part to utilize standard hardware and OS platforms. The NAS solution was chosen for its ability to integrate easily into a small IP environment, while being able to deploy remotely with some level of remote management.

This case study will discuss the IIP storage and related processing problems and challenges in detail. These drove the activities involved in identifying the company's workload and analyzing potential storage solutions. Finally, we discuss how IIP concluded with the decision to implement NAS and the subsequent transition to the new storage model.

The Situation Analysis

IIP developed its storage solution from years of experience in scanning and duplicating photographic images, ranging from military applications to historical documents. The challenges had always been the time required to scan an image, versus the quality required for large clients. This had given way to the amount of space required for each image, given that the IIP client base dealt with millions of images.

The type of work IIP performed was not directed toward the market of flatbed scanners nor even larger scanners in terms of quality and production process requirements. Instead, IIP developed a process and proprietary software that used specialized digital cameras to scan images of various dimensions, physical states, and types—various papers, photographic techniques and types, and so on. The process, coupled with the scanning stations, provided a production-oriented environment where imaging could take place 24/7, if required. The process and software included a fully complete life cycle of imaging capture, correction, and quality assurance before it was placed on a CD or tape for shipment to the client.

Clients of IIP had requirements to digitize documents and photographs to provide a wider distribution and availability of these items through the Internet. Consequently, these clients had become part of the growing movement within both the academic community and public sector to save historical documents. For the most part, these markets are just emerging, given the tremendous amount of material that remains to be scanned and digitized.

IIP's client requirements could run into the 500,000-plus number of images for a single project. That type of work drove the production-oriented environment introduced by IIP in the late 1990s.

IIP Systems Infrastructure

IIP maintained a distributed infrastructure with its headquarters in Texas and field locations in New York, Washington, D.C., and San Francisco. Figure A-1 illustrates the configurations at the headquarters' locations. Here you see the integration of capture stations, process servers, correction/quality assurance workstations, database servers, and archive servers. Each field location is set up in an identical fashion. Each is linked to the headquarters' network and web server through a virtual private network (VPN). E-mail and FTP services are handled in this manner. Architecturally, this setup was designed for future development of remote scan processing and diagnostic imaging services.

The imaging process will describe the storage utilization scenarios and why the process is so data-centric. Figure A-2 depicts the process from start to finish.

1. *Physical images are entered into a database.* The database drives and tracks the processing of images from capture to output on media. In IIP's case, the database is a relational database that tracks the location and status of the scanned image

so that at any one time the image can be located and accessed for specific purposes. The initial step in the process is to have the image scanned into the system. An important note regarding the database: Because of certain restriction and challenges with unstructured data and relation technology, the scanned images are actually stored as files within the system. Therefore, the database is actually quite small, as its main job is tracking the locations and characteristics of the scanned image.

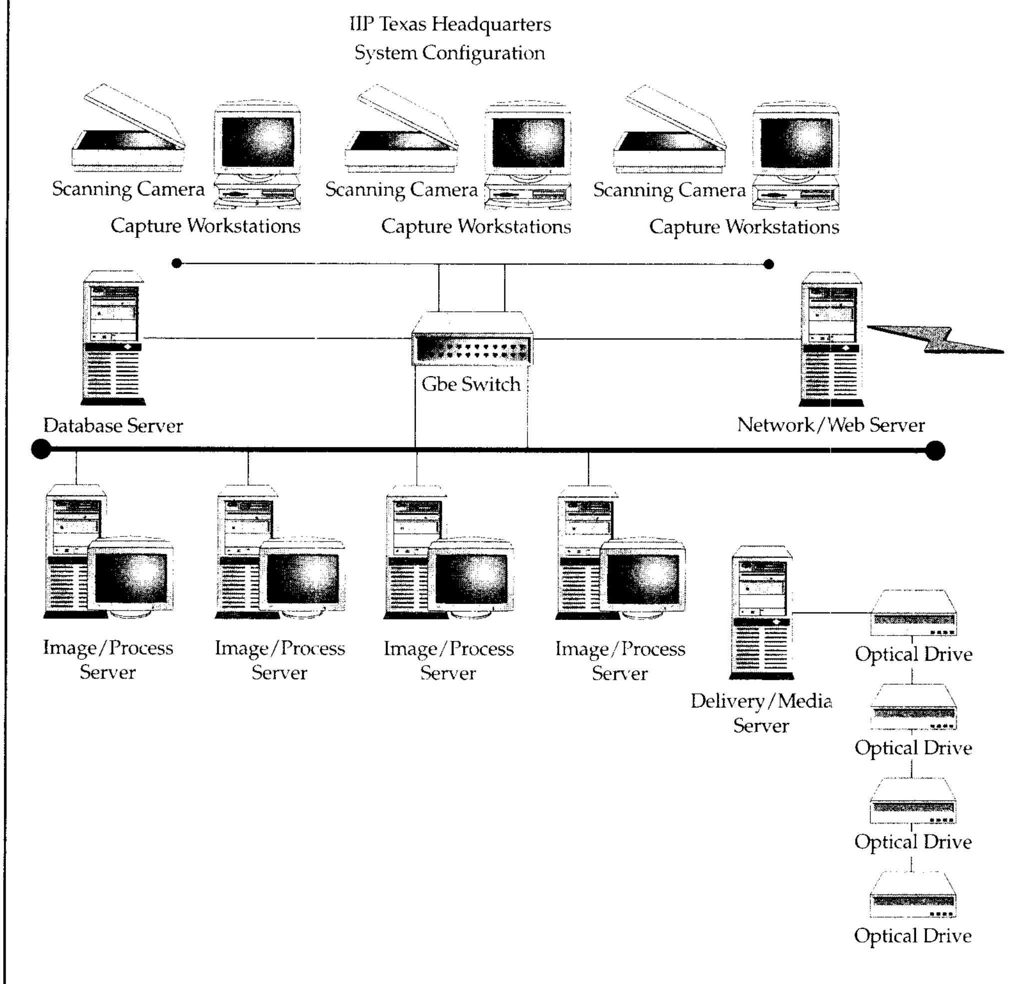

Figure A-1. *IIP systems configuration*

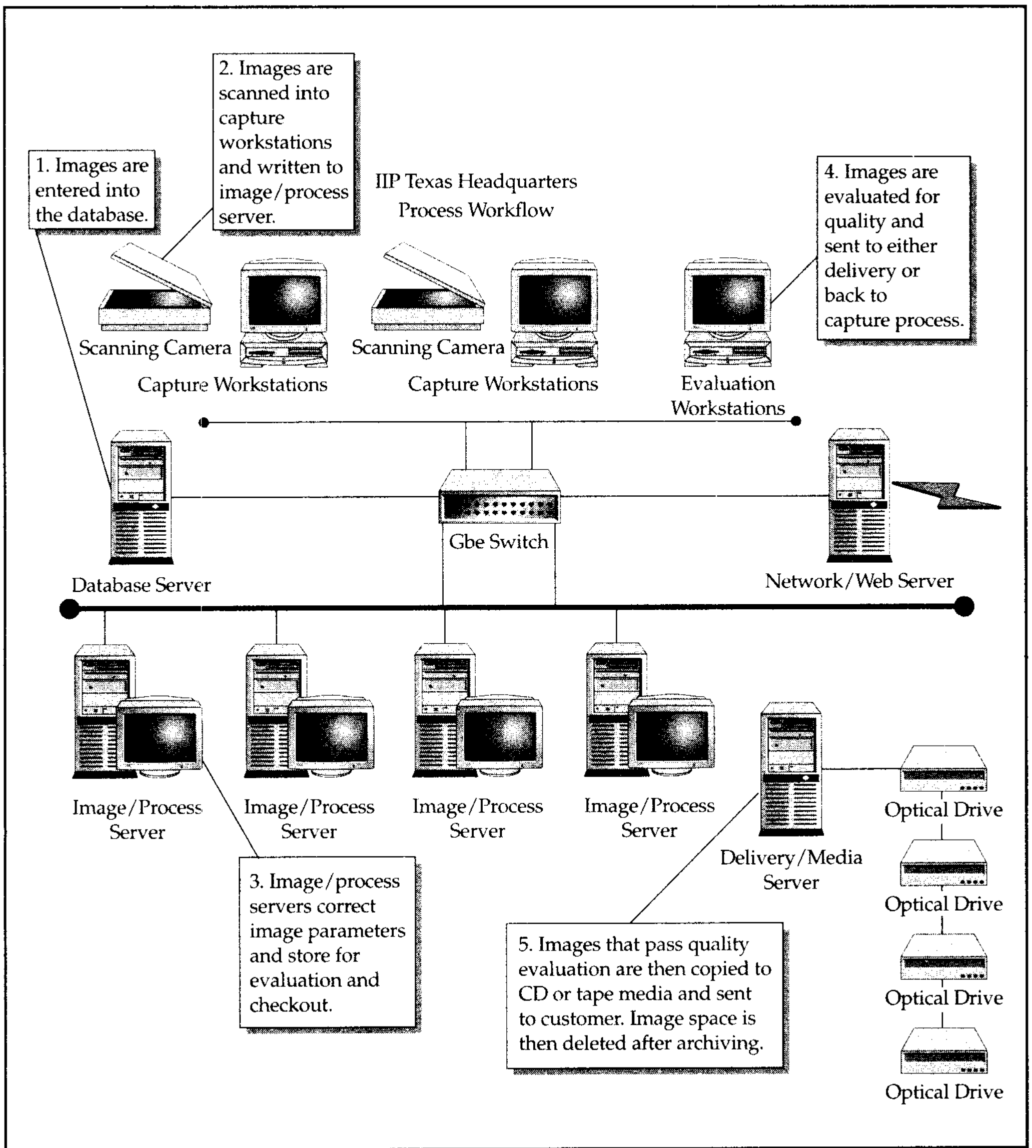

Figure A-2. *IIP imaging application process*

2. *At the capture station, the images are scanned into the system.* Capture stations are high-end workstations that communicate with the specialized digital cameras. A camera scans the image into the capture station; then the capture software,

running on Windows 2000 Professional, writes the scanned image to a network drive on an available process server over the network and moves to the next image. Average images range from 300 to 500MB per raw image.

3. *The process server automatically corrects the image for perspective and positioning and updates the database.* The process creates another set of files for the corrected image. This set of files is smaller because it contains only the changes made to the image. The image is then ready to be evaluated and checked out.
4. *The image evaluators view each image for correct quality of preset parameters, as discussed with clients.* If the images meet the quality criteria, they are sent to checkout for delivery to media servers. However, they may need manual correction, which is performed at the evaluation workstations or sent back for rescanning, as required. This is all tracked by the database.
5. *The delivery point is the media servers where the completed images are batched together and written to optical media, CDs, or tape, per the requirements of the client.* At this point, a final verification is performed, and if passed, the storage becomes available for more scanned images on the process servers.

Problem Detail

What IIP had not foreseen were the systems infrastructure requirements for this type of production work. This required the calculation of both processing cycles and, most of all, the amount of storage space that would be needed on a daily operational basis. Because IIP is a small business, it had resisted a formal capacity plan and had relied on its ability to respond quickly when additional capacities were needed. That meant the additional servers were purchased on an "as-needed" basis, with most of the hardware being "do-it-yourself" built chassis and motherboard configurations.

With this orientation to developing and maintaining the hardware portion of the systems infrastructure, the storage challenges were met with larger and higher speed internal IDE disks. This gave rise to additional server installations that were needed to handle the post scan image processing (see Figure A-2, Steps 2 through 4). This then prompted the acquisition of dedicated media servers to write out the client images using CD or tape media. This is the archival system on the backside of the process (see Figure A-2, Step 5—the delivery process). This meant that a faster network was necessary to speed the transmission of scanned raw files to the process servers, and ultimately it placed the problem back at the storage infrastructure once again as the image scans overtook the capacities on the servers.

A stopgap effort was a move to IDE RAID to provide adequate storage for the process servers. This was largely driven by the "do-it-yourself" mode of the hardware and severe limitations of budgets constraints. Although the IDE RAID facilitated a quick fix, IIP's flexibility in providing reliability and backup protection was problematic. In many cases, the volatility of the data movement over the period of one week could easily surpass more than five terabytes running through a single process server. As the tremendous write activities continued, the IDE drives generally failed twice a month, with minimal success running data protection with RAID level 1. However, the space it provided offered a brief interlude to the space problems that shut down the process entirely.

Given that additional business was coming in with more restrictive time constraints for completion, IIP concluded that a longer term solution had to be found.

The Search and Evaluation

The IIP hardware and software IT personnel researched the solution along with assistance and input from the imaging software specialist. They found that a SAN was a valid consideration, since it appeared to be the choice of others working with unstructured data such as video and audio projects. However, they found that the imaging system, although proprietary by its methodology, used open and commodity levels of hardware and operating environments and was further open to additional solutions that integrated well into the small business environment. Another alternative was to move to larger process servers with external SCSI drive arrays, to scale up in both process and storage power. Yet another alternative was the consideration of a NAS solution, which would integrate easily with the existing network, would use file systems, and would have the capacity they needed.

Estimating the Workload

The IT personnel, working with an outside consultant, used the guidelines mentioned in Chapter 17 in identifying the company workloads. They further moved to understand the types of configuration needed for all three possible alternatives. First looking at the SAN configuration, followed by the larger server with external RAID, and finally the potential NAS configuration. The results are summarized in the following sections.

Workload Identification Looking at a year's history of scanning images, the IIP team concluded that the workload was complex and data-centric, and it fit somewhere between online transaction processing (OLTP) and data warehousing. The workload encompassed OLTP characteristics when scanning the image and then transmitting the write transaction to the process server. Although developed as a synchronous process, it was recently changed to an asynchronous process to facilitate greater throughput at the capture station. However, this still required a sequential write process at the process server as each image was scanned.

On average, the image scans were 300MB in size. The calculation of 300 images per shift × three capture stations working two shifts provided the necessary throughput. It was determined that at least 540GB of free space was needed to accommodate the daily scanning process. This required that the overall storage infrastructure be able to accommodate a 566MB per second throughput rate.

Workload Estimates for SAN Using the guidelines described in Chapter 19, we can quickly calculate that the required components for a SAN could be handled by one, 16-port switch, given that a single point of failure is acceptable for the installation; or it could be handled by two, 8-port switches for some level of redundancy. Three HBA adapters with 2 ports each for redundancy and performance would be required for the process servers. However, not to be overlooked, this configuration will require additional Fibre Channel storage arrays to accommodate and be compatible with the new Fibre Channel storage network. Given that the total capacity of 540GB needs to be available every 24 hours, we can estimate that two storage arrays of 500GB each would provide the necessary capacity with sufficient free space to handle peak utilization as images are processed through the system.

Workload Estimates for Direct Attached Aligning the requirements to new servers, we find that all the process servers would have to be upgraded. This would also require that the storage capacities be carefully aligned with each process server. Even with this alignment, specific workload affinity would have to be observed to utilize the storage effectively. On the other hand, the process server could more easily share storage across the network but would have to reflect some level of duplication for storage requirements to accommodate the total capacity, essentially doubling the entire storage requirement.

This would require, in addition to new servers installed, OS software upgrades, with appropriate maintenance and all the necessary activities of a major system installation. It would result in a normal disruption of service and reliability characterized by new system installations. However, the new servers would have to be configured to handle the I/O throughput of an aggregate of 566MB per second. This would require each server to handle 188MB per second if the workload is evenly distributed, which in most cases will not be the case; however, we will use this for estimating purposes. That relates to a minimum of six Ultra-wide SCSI-3 adapters necessary to handle the sustained rate of 188MB per second. This requires the total storage to be divided among the servers, and subsequently the adapters, and places a limitation of approximately 120GB per LUN. Thus, a more complex management problem in terms of flexibility of reconfiguration based on required storage would be necessary, given that one capture station could generate 180GB of images every 24 hours.

Workload Estimates for NAS Using the worksheet and guidelines in Chapter 19, we can calculate that our workload requirements are definitely within the mid-range NAS device configuration and probably just under the enterprise NAS solutions. Our calculations indicate that the workload requires the following minimum requirements:

- Two network paths
- Eleven data paths

- An additional path for redundancy (calculated using 10 percent special applications category)
- 13.2 total logical paths
- Comparing total logical paths to data paths = 83 percent

Using the quick estimate NAS Sizing Factor Table in Chapter 19, we select mid-range even though our sizing factor is within the enterprise range. This is based on the special application circumstances and because the aggregate data is below a terabyte and would be physically segmented within the aggregate data capacity estimate. In addition, we considered the workload being further characterized by limited users working with an almost dedicated Gbe network.

The NAS solutions also offer the flexibility of storage incremental selection—for example, installing two large NAS servers and one small server, or one large and two medium-sized servers. These solutions also provide the flexibility of RAID processing, network compatibility, and non-disruption to the existing server configurations. In addition, these solutions can be easily configurable to support the scanning projects and mapped as network drives with the same flexibility. They will also provide a closed, yet remotely accessible, solution for the remote network configurations.

One last word on our estimating process: We recognize the characteristics of the small integrated IT staff and the company's lack of any formal capacity planning activities. The process of workload identification and estimates provides this company a level of direction and planning. The result of this exercise has identified that the mid-range NAS devices can meet the company's workload now and within a limited planning period. However, it also provides an insight into future challenges IIP will encounter, as its staff has become aware that it borders on moving into enterprise solutions of either the NAS type or probably a SAN if the budget for infrastructure can support either.

IIP Storage Solutions

IIP chose to acquire a NAS solution for headquarters and each of its remote work sites. The evaluation proved that the requirements for capacity, access, and budget were well suited to the NAS solution.

Given the growth and support required for IIP for the next 12 months, the ability for the organization to absorb a new technology such as Fibre Channel would have been beyond the current budget limitations. In addition, the learning curve mistakes with the SAN could not be afforded with a small company such as IIP, where there is no time, space, or budget for test or beta machines. The IIP IT evaluators determined that SANs would probably be in the company's future as bandwidth catches up to the current network capacity and the need to provide capture stations with at least 100MB-per-second speeds just to keep up with the increased scanning capture operations.

IIP Storage Evaluation Results

The requirements and evaluation are further depicted in Table A-1. The evaluation is based on a score of 1 through 10, with 10 being the highest, or best, score. Vendor selection is specific to the customer and the competitive nature of an acquisition, although it is recommended that a minimum of three bids be requested once workload estimates have been concluded.

Note *We have estimated pricing only as a general reference point. Included in these pricing estimates are total components for system operation, additional software required, and storage array costs. Training for the SAN solution is an add-on, however it's a highly recommended item.*

IIP Storage Acquisition and Implementation

Along with the storage requirements, IIP provided a list of service requirements it had developed in conjunction with an outside consultant. These requirements, along with the capacity, access, and management parameters, drove the decision to go with NAS. Although cost was a factor, the costs of the proposals were all within 10 percent of each other.

Implementation of the NAS devices provided additional activities. Although minimal, these still provided additional time before the devices were used in production. The

Criteria	SAN	Direct Attached	NAS
Meets I/O workload performance and capacity specs	9	5	8
Ease of installation and maintenance	3	4	7
Scalability of storage capacity	7	2	7
Compatibility with existing storage infrastructure	5	6	8
Composite score	24	17	30
System price	$150,000 ($175,000 w/training)	$235,000	$75,000

Table A-1. *Storage Solution Evaluations for IIP* items are summarized in the following list. Note that any of these are reflective of the operating environment and maintenance status of both operating systems and the network.

- **Network Directories Not Recognized on Network** Required additional upgrade to network for compatibility with NAS server network software.
- **Network Directories Being Dropped** Required additional upgrade to Windows 2000 Server software to reflect NAS devices.
- **NAS RAID Not Working** RAID configured improperly, and reconfigured to reflect recovery parameters.
- **NAS Integrated Backup** Used as archival to tape media, however, requires update to database to reflect new delivery device.
- **NAS Slower Performance from Capture Stations** NAS set up to use NFS file protocol, overhead to write much higher than before. Investigating solutions or optimization.

The IIP NAS Solution

The final configuration provides increased storage capacity and enhanced performance. The IIP scanning process is now being upscaled with new customers and imaging projects. In addition, the space is being utilized on an automated basis from new project initiation to deletion after delivery. Figure A-3 illustrates the NAS configurations that support the revised IIP systems infrastructure.

Figure A-3's NAS configuration illustrates the increased capacity for storage of images, but it also provides the foundation for image access through the Internet. This was designed into the system to allow the remote work sites that have identical installations to upgrade to NAS devices. Over and above the increased capacity at all sites will be the potential ability to access images across the IIP storage network infrastructure. In other words, images scanned in New York could be evaluated in Texas, Washington, D.C., San Francisco, or other remote offices as they begin operation. This is an added business efficiency that takes advantage of the imaging expertise across the company without regard to location. It also allows clients eventually to be provided test and quality images through the Internet to further facilitate the delivery and client acceptance process.

In summary, the NAS solution turned out to be a good fit for IIP. Even though its application appeared to have many proprietary processes and software processes, NAS operated within the bounds of leveraging commodity infrastructures for networks, systems, and storage. With the exception of the database server, all the servers within the IIP infrastructure now share the available storage in the NAS devices. The acknowledgment of this value further expands the use of the NAS solution as the company began to configure its web and imaging software code development on the NAS devices.

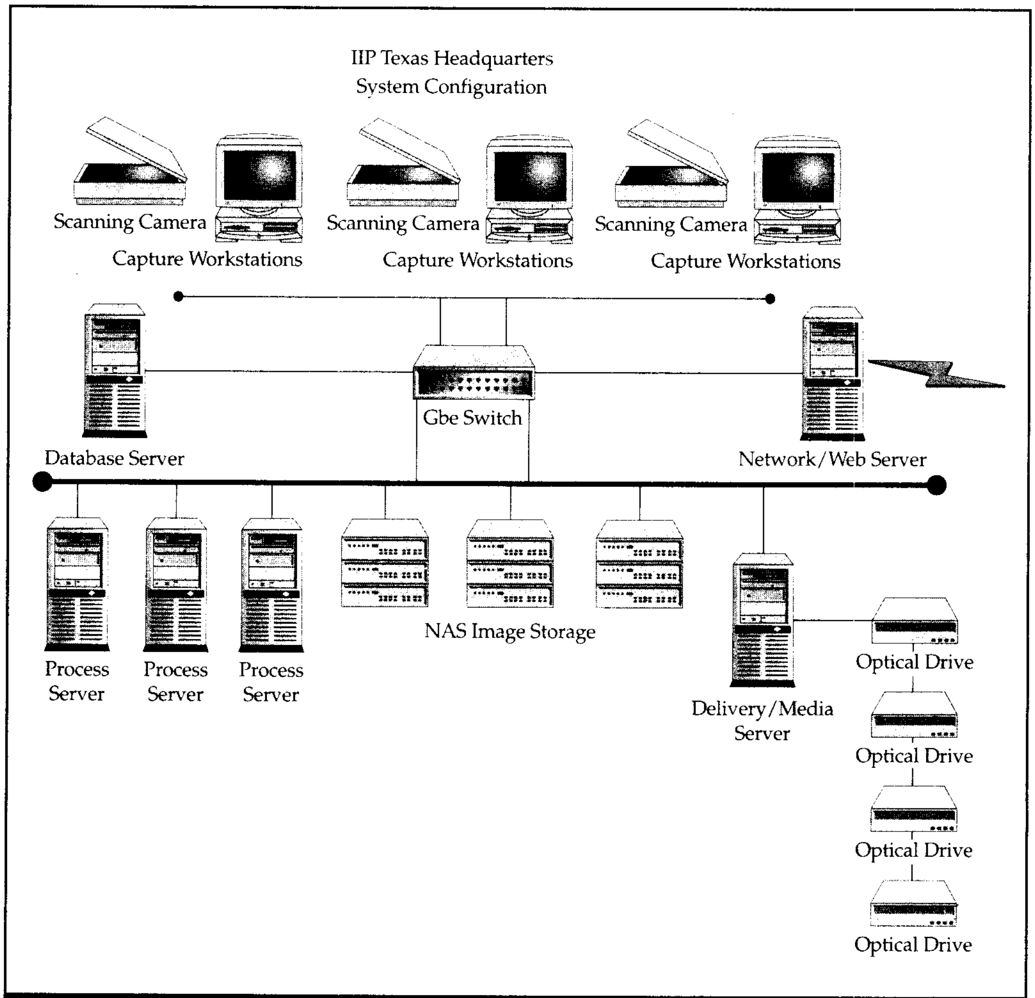

Figure A-3. *IIP NAS storage infrastructure*

The NAS solution for IIP provided a cost-effective solution for a small company, but it also provided the company with the necessary storage resources for expanding its business, which had grown dependent on a storage-centric product. The company's storage infrastructure is now poised to grow into the future with either more sophisticated products such as a SAN, or to continue to scale economically with further uses of NAS devices. Either way, IIP now has the business flexibility to meet the growing dynamics of the imaging business.

Appendix B

SAN Case Study: The Import Auto Industry

The import auto industry provides distribution services for automobiles manufactured in both Europe and Asia. Most successful international automakers have expanded to provide an extensive presence in North America, as most have populated manufacturing plants in the United States, Canada, and Mexico.

The activities that surround the distribution of products within the U.S. have provided a rich datacentric set of activities that has prompted leading-edge data centers in the U.S. Because the automakers' IT organizations' data centers have had to address data-processing activities on an international level, they have consistently been ahead of many other U.S.-based companies in their abilities to handle data on an international basis as well as becoming involved with leading computer technologies. In addition, government regulations exacerbated the amount of the data these companies must collect, analyze, and address.

Although the international aspect of these companies requires that specific additional data center issues be addressed, they must also deal with the basic systems of business in terms of reporting, tracking, and analyzing data. What makes the auto industry unique in many respects is the various sources of revenue that it derives from its products, much of this transparent to the consumer. Key among these activities is the sale of parts products, service contracts, and the increasing amount of technical support provided to dealerships. Many of these are reflective of the increased technological advances within the automobiles themselves. In addition, warranty products and services must be tracked and accounted for. For example, each time a warranty claim is presented and processed within a dealership, the manufacturer pays for the service, parts, and support. Although these activities are not restricted to just the auto industry, it does set the foundation for a set of required information that means that an extremely large amount of receivables be managed by the U.S. distribution companies. This necessitates a synchronized transaction between dealer, distributor, and manufacturer to resolve.

The results of all these operations and interactions make this industry an extremely datacentric and data-sensitive enterprise. The composite collection of IT organizations within the U.S. that represents these companies has become the key user and consumer of high-end computer systems and technologies. However, in many cases these organizations must work in tandem with headquarters data centers in both Europe and Asia. This requires the U.S. data centers to implement or become compatible with a strategy that has been developed outside the United States. Working with the levels of autonomy, each of the U.S. data centers has laid the foundation for the tactical types of technology trends they can justify and implement. This issue has both advanced and stifled some of the most productive trends within these data centers.

This case study discusses the activities surrounding the analysis, justification, and implementation of a storage area network (SAN) to support a major restructuring of key application systems within a major international auto distributor in the United States. The company was advanced in its installation of state-of-the-art mainframes and distributed systems. These systems supported the U.S. headquarters, parts distribution centers, and its extensive network of dealerships. However, the major applications supporting these areas lagged the rest of IT, with largely centralized systems without the necessary distribution of data from the mainframes to enhance or synchronize the data that existed on the distributed systems. Consequently, a consistent and timely analysis of the key areas of the enterprise remained problematic.

This evolved as the company executives mandated more data to analyze portions of the business—key among these are parts, service/warranty, and sales lines of business. They placed hard requirements to the IT executives to facilitate the development of and restructuring of current data warehouse applications into a more responsive set of data marts as well as a corporate data warehouse that provides relationships to the key revenue indicators of the company.

The industry's IT area had been restricted by a static budget for the last two years as the auto business has tried to stabilize its operations and financial positions. IT employs 350 personnel, with an additional 50 to 60 contractors employed at any one time. Within this organization, the technical support area has 35 personnel, with enterprise storage administration being supported by one full-time and one part-time systems programmer. This small contingent works to administrate and support storage operations in conjunction with the assigned mainframe systems programmers and UNIX and Windows systems administrators.

The Situation Analysis

The current application model consists of all critical application systems being run on the mainframes. These are generally OLTP-type applications that provide support to all business functions within the headquarters and distribution centers. These online applications are supported by a heavy batch cycle of processing within a very restricted online and batch-processing model. The dealership network is a separate entity with restricted access to limited sets of inventory and sales data. The applications targeted for redesign and redeployment are some of the few that are operated from the UNIX open systems area. These are the current data warehouses for sales, warranty, service, and parts. They are currently loaded from the operational databases on the mainframe during a nightly batch cycle.

The new application requirements are to restructure each data warehouse and distribute its data into a smaller data mart supporting each of the individual lines of business (LOB). The demands require that the warranty and service data warehouses be consolidated into a larger integrated data warehouse, along with a smaller but integrated data mart for each LOB. In addition, the larger, but longer term requirements are to build a corporate data warehouse to provide key executives with real-time analysis on sales and issues affecting sales. Although this requirement is beyond the scope of this study, the impact will be affected by the storage infrastructure that is ultimately selected by the storage administration team.

Figure B-1 depicts the current systems configuration overview and data capacities as provided by the storage administrator.

The challenge is to develop an infrastructure that will house the information and provide online access to the predefined analysis that will be run against the data. A common set of service levels has been agreed to in the requirements meetings with end users. However, these were approved by the application and database designers

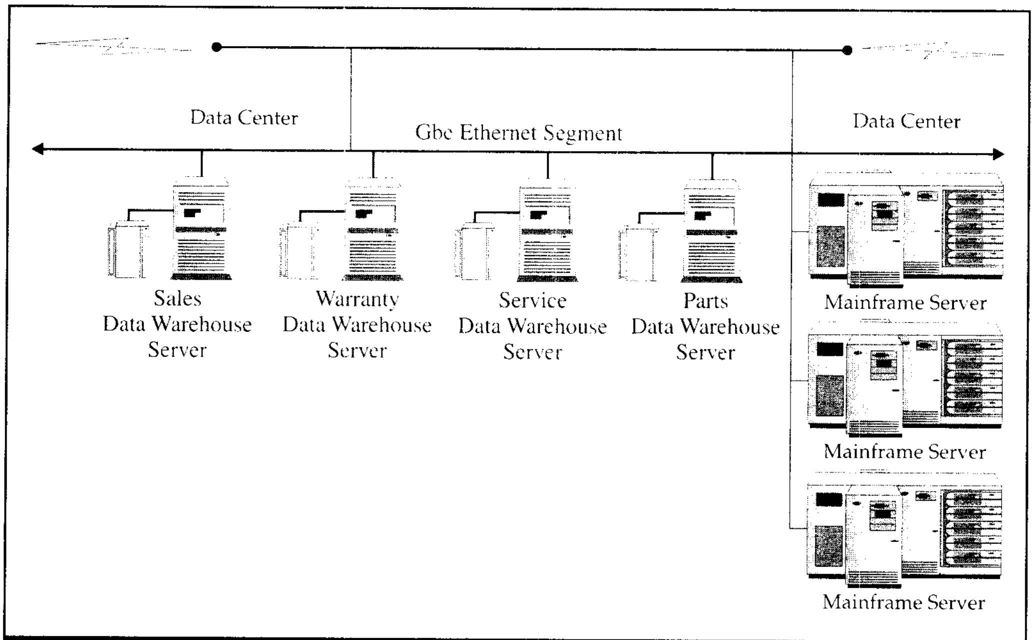

Figure B-1. *An overview of the current systems configuration*

without the concurrence or involvement of the storage admin team, much less the technical support group. Consequently, the storage team's job is to develop the infra.:tructure that will support the predefined capacities, processing requirements, and online service levels. These are summarized in Figure B-2 and Table B-1.

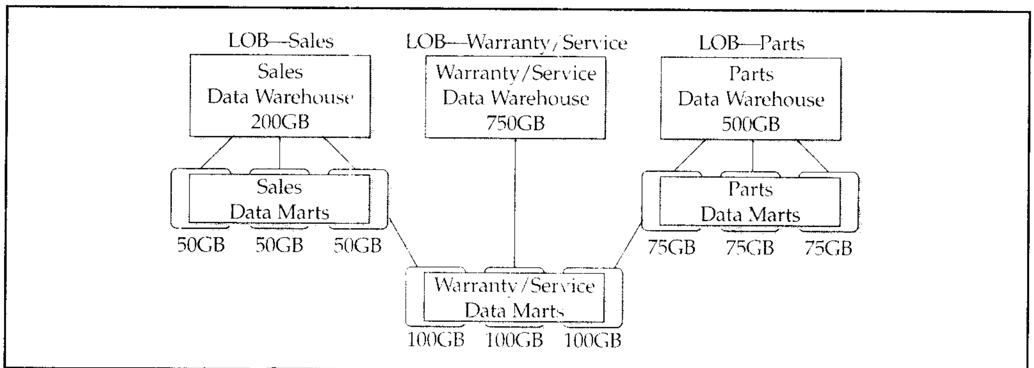

Figure B-2. *Data warehouse/data mart requirements*

Processing Req.	Sales DW/DM	War/Service DW/DM	Parts DW/DM
Age of Data < 24 hours	✓	✓	
Age of Data > 24 hours			✓
Batch Loads	Mon–Fri: nightly	Mon–Fri: nightly	Mon–Sat: nightly
Response Time	< 3 seconds	< 10 seconds	< 10 seconds
Availability	Mon–Fri: 6 A.M.–7 P.M. Sat–Sun: 8 A.M.–5 P.M.	Mon–Sat: 24x7 Sun: as available	Mon–Sat: 6–7 P.M. Sun: as available

Table B-1. *Processing and Service-Level Requirements for Data Warehouse and Data Marts*

Armed with the following information, the storage team had the task of supporting the IT applications area in developing these business systems. Working with the application designers, the timeline was approximately six months, in which the last three were given over to development and beta testing of the new data warehouse configurations.

The Search and Evaluation

The storage team concluded that any solution had to scale to the multi-terabyte level. This was deduced by a quick but careful analysis of current utilization, historical trending of growth, and current estimates. Current utilization of the data warehouse applications was around 750 to 800GB of aggregate data, combining both user and database overhead. This was accommodated through the current installation of 1 terabyte of external SCSI-based storage capacity installed over three UNIX servers (summarized in Figure B-1).

The other quick conclusion was scalability of access given the large amounts of storage; end user access and capability to load and update the databases would be key to the feasibility of any solution that would meet the processing requirements and service levels. The new requirements, as provided by the application design team, called for an additional capacity of 1.125 terabytes, moving the total requirement for the data warehouse and data mart applications to well over 2 terabytes. The team had calculated that the data warehouse storage utilization was growing at 40 percent per year, with no indication that it would decrease.

The storage team had to validate key pieces of information before embarking on its search for solutions. First, it was asked, was the validity of current estimates it received from the application designers for aggregate user data, or did it include overhead estimates for database overhead from the DBAs? Second, were these estimates for day-one installation or have some growth factors been built in? The answers turned out to be good news and bad news, and some new news, which called for even new challenges.

The good news was that the estimates included the overhead as estimated by the DBAs. Consequently, the aggregate data sizing could be used as is. The bad news was that these were day-one estimates. In other words, this was the capacity needed when production started and reflected the actual storage required for the sourcing of the data from the mainframe databases. The new news was that the consolidation of warranty and service data warehouses would continue to use two different database products; therefore, the consolidation would not occur immediately and the two data warehouses would run in tandem until they would be consolidated later in the year. Consequently, the storage team needed to add an additional factor for two databases; doubling the current capacity estimate accomplished this.

Available Alternatives

Considering all the available estimates, and validating the capacities, processing requirements, and services levels, the new data warehouse infrastructure would need to support a 3-terabyte capacity. Further investigation indicated that the user community—although comparatively small, with less than 250 end users—would require an I/O workload of multiple gigabytes per transaction. This also carried over into I/O workload for loading and updating the database, creating an additional multiple-gigabyte workload per transaction. In summary, calculations indicated a total I/O workload of 29,300GB for a 24-hour period. This was reduced to 1200GB per hour and subsequently an aggregate of 20GB per second.

Using an external direct attached storage strategy that would simply enhance the current configurations would require additional servers and additional SCSI storage arrays. A quick estimate revealed a requirement of approximately 50 new servers with the maximum of SCSI adapters for each server to meet the aggregate I/O workload of 20GB per second. An alternative to this scenario would be to replace existing servers with a much larger server machine, such as a UNIX mainframe server, where the machine could be partitioned to run the individual data warehouses with dedicated resources. However, this solution also would require the new storage arrays, although the existing arrays could possibly be used in conjunction with these.

By utilizing the guidelines for estimating SAN capacity (see Chapter 18), it turns out that a SAN configuration will require a total of 240 ports, including calculations for recovery and redundancy factors. Given the type of workload—datacentric (high I/O content), transactional with batch updates, and less than 250 end users—the configuration could be handled by 15, 16-port switches, or 4, 64-port director class switches. However, this solution also requires that new storage arrays be purchased for compatibility and performance within the storage network. This strategy does provide the possibility of utilizing the existing servers.

The storage team, in conjunction with technical support management, chose to evaluate the UNIX mainframe alternative compared with the SAN configuration. This evaluation took place as a mini-capacity planning exercise, since this requirement was addressed from an existing capacity plan that was in place. Consequently, the IT executives, to provide the flexibility to evaluate solutions outside the context of existing budget plans, invoked the end-user non-aggression pact. (Refer to Chapter 23 regarding storage network capacity planning and Chapter 22 for end-user non-aggression pacts.) This exercise revealed an interesting comparison.

Storage Evaluation Results

The UNIX mainframe was a high-end SMP machine that was capable of being partitioned into 2 to 16 logical processing partitions. The SAN configuration was analyzed as 4, 64-port director class switches, with 12 HBAs installed across the existing data warehouse UNIX servers—for example, 4 HBAs each. The storage arrays were configured with 3 terabytes of capacity with specific configurations for each solution—for example, one set for the SCSI-based UNIX mainframe and a different set of arrays for the Fibre Channel–based SAN.

The requirements and evaluation is further depicted in Table B-2. Our evaluation is based on a score of 1 through 10, with 10 being the highest, or best, score. Vendor selection is specific to the customer and the competitive nature of an acquisition. It is recommended that a minimum of three bids be requested once workload estimates have been concluded.

Note *We have estimated pricing only as a general reference point. Included in these pricing estimates are total components for system operation, additional software required, and storage array costs.*

Requirement	SAN	Direct Attached
Meets I/O workload performance and capacity specs	9	8
Ease of installation and maintenance	3	4
Scalability of storage capacity	7	4
Compatibility with existing storage infrastructure	5	8
Composite score	24	24
System price	$700,000	$1,850,000

Table B-2. *Storage Solution Evaluation Table for Data Warehouse/Data Mart Applications*

Solution Acquisition and Implementation

The storage admin team recommended the SAN solution with subsequent approval from IT executives and reluctant concurrence from the application design team. Because this would be the first installation of the storage networking technology for the data center and because of the visibility of the applications, the design team's reluctance was understandable. However, an integrated plan was developed to provide beta support for the application testing of new data warehouse and data mart prototypes.

The decision was based upon both price and performance. Given the increased throughput with the Fibre Channel–based storage arrays, the SAN solution appeared to be one of the most adaptable solutions for this type of high-end throughput application—that is, an aggregate 20GB per second of data. Specifically, the ability to source the data warehouses from the mainframes into the UNIX servers for subsequent preprocessing, loading, and updating was appealing. Moreover, the shadow of the impending corporate data warehouse project provided the pivotal requirement that pushed the decision toward the SAN. The decision makers realized that if the UNIX mainframe strategy was chosen, another would be required to handle the additional load of the corporate data warehouse. Finally, the ability of scaling the SAN or adding another SAN would be more cost-effective, but it would also be more responsive to the application needs in the short and long term.

The installation of the SAN provided additional internal requirements and surprises, but none that were unmanageable or extremely cost intensive. These centered on the new operational characteristics of the SAN, the additional software tools required, and the responsibility for maintaining the SAN software and hardware (from an IT organizational view). Each of these issues was handled within the beta period of installation and subsequent prototype testing of the applications. These issues are summarized in the following section.

SAN Operational Issues

The new SAN required a learning curve and subsequent integration into the existing data center hardware and software processes. The storage admin team and select members of the systems administration team took vendor classes, which provided an overview specific to the vendor product selection and additional guidelines on installation and preinstallation planning.

The installation was somewhat more problematic because of the problems any new infrastructure has in moving into the data center. The issues were centered on three areas: facilities and location; fundamental management processes, such as meeting the existing data center operations rules; and integration into existing server, network, and storage wiring mechanisms.

- **Facilities and Location** This was handled by accommodating more space within the existing server and storage areas. However, given the less restrictive lengths, the storage was initially planned for installation in the mainframe storage area on a different floor. This was scratched because local proximity (for example, server to switch to storage) was better during the initial production cut-over period. However, the flexibility of subsequently moving the storage to another area, given the increased length capability of Fibre Channels, turned out to be an additional plus for the data center facilities planners.

- **Management Processes** Perhaps the most troublesome issue was the operational integration into the data center. Because of the fundamental processes of each SAN director, this required that additional data center operational policies and processes be developed, followed by training and orientation of key operational staff members. The ability to provide configuration information and real-time performance information continues to hamper this aspect of the operation; however, this is becoming more adaptable as the operations of SAN become familiar to both operations and systems/storage admin staff.
- **Integration into Data Center** One of the most difficult activities surrounding the installation and subsequent expansion of the SAN is the integration of wiring complexities and integration into the existing complexities of server and network wiring and switching structures. The ability to implement, manage, and track this process remains quite difficult as more than 200 ports are routed to appropriate points of contact. Although not unlike their Ethernet counterparts, and comparatively small next to the LAN/WAN wiring closets, the critical sensitivity to a lost, unconnected, or inoperative port can affect the highly visible application such as the company's data warehouses. Existing plans call for the SAN infrastructure to be separated from the existing LAN/WAN configurations.

New SAN Software Tools

Once prototype testing was in place, it became apparent that new tools were necessary for managing the storage across the SAN. The requirements specifically came from the need to manage the storage centrally across the supported servers. This required some level of both centralization and drill-down capability for specific array and individual devices access. This was accomplished through the acquisition of new volume management tools and vendor-specific management tools for the storage arrays, both having specific functionality for Fibre Channel–based SANs.

The next level of tool that proved to be more problematic was the backup/recovery tool. Due to the nature of the application, data warehouses generally don't require the stringent restore functions that an OLTP application would need. Therefore, the ability to perform standard backups would impact the nightly update function or prove to have little value to an analysis that uses a 24-month rolling summary of data. Consequently, rebuilding or reloading the data warehouse can be done on a basis that is less time sensitive—for example, performing a rolling backup once a week and probably on the day when the least processing occurs. Given the large amount of data, the specific challenge to relational databases, a snapshot function is being pursued for those database tables that are the most volatile and time consuming to rebuild and reload. In the case of a complete volume or array outage, the snapshot would allow the DBAs to go back to a specific time period and reload and update the database within a minimum amount of time.

IT Organizational Integration

The responsibility to evaluate and justify the SAN initially was given to the storage administration team. The installation and support of the SAN beta testing was accomplished through a team approach, using the expertise of systems admins, application designers, and DBAs, with leadership from the storage admin team. As the SAN became a production entity, the team dissolved into its daily responsibilities and the nebulous nature of some of the SAN components began to affect the total mean time to recovery (MTTR) for any problem scenario.

Consequently, IT management is faced with an interesting dilemma of providing shared responsibility for the SAN or developing a new infrastructure group focused on storage. There are pros and cons for either direction. This study indicates the somewhat conservative nature of the industry to continue to share responsibility across the SAN components. The network, the systems, and the storage staff are all responsible for the SAN or various components of the SAN. This unfortunately results in the increase of the MTTR, since this type of arrangement creates the "it's not my problem" scenario. (Refer to Appendix C for details on how this manifests itself in a storage networking management business case.)

The SAN Solution

The final configuration provides the increased storage capacity and enhanced performance that was expected of the solution. The data warehouse/data mart project was accomplished on time and has proved to be responsive to the I/O workload. Because the estimates were accurate and allowances were built into the configuration for expansion, the storage capacity and access performance is ahead of the I/O workload estimates and business requirements, making both of these execute within service levels.

The current configurations are moving into the planning for the next highly visibly project, the corporate data warehouse. However, this time it has complete concurrence from the application design teams, the DBAs, and system administrators. Further planning is being considered to consolidate additional storage and servers within the data centers. The two outstanding caveats remain the organizational fluidity of responsibility and the continued organizational challenges to effective wiring management. The first continues to elongate any problem associated with the SAN configuration, and the second continues to facilitate the singular failure of port mismanagement, exacerbating the first issue.

The configuration shown in Figure B-3 illustrates a summarization of the data warehouse/data mart application systems. The ability to portray a 200-plus port configuration is beyond the scope of this study. However, it's important to point out the results that the SAN solution provided this company with the ability to manage its business more effectively through the use of multiple technologies. The functionality of the relational database to facilitate the analysis and relationships among business functions has provided a 40 percent increase in the amount of parts sold and a 25 percent cost containment in the area of technical support. More importantly, the applications have been used to reduce the warranty receivables significantly, and that results in an important improvement to the company bottom line.

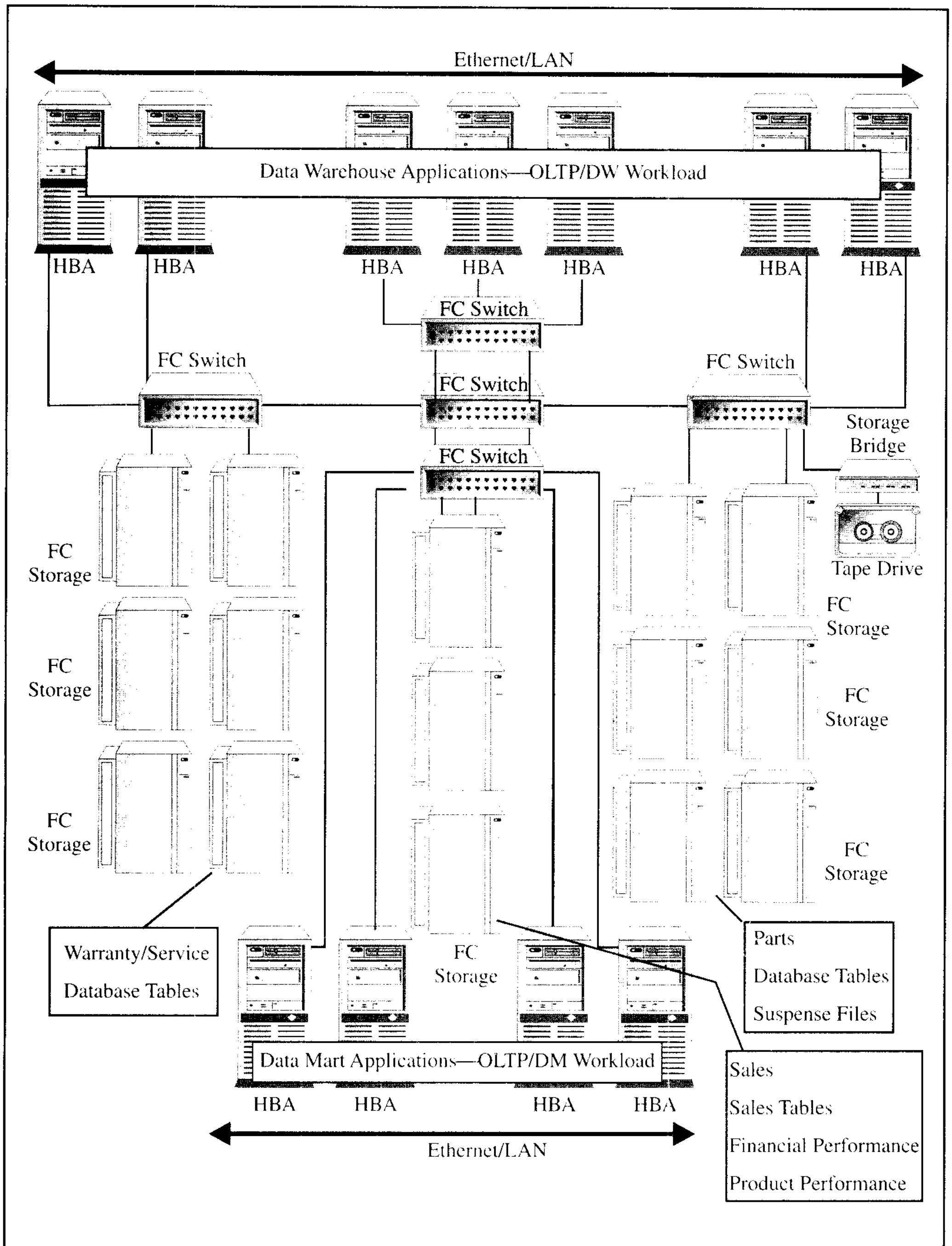

Figure B-3. *SAN storage infrastructure*

The SAN solution provided the infrastructure for these applications. The ability to move and access data at the increased levels of Fibre Channel SAN configurations has provided this company with a significant competitive edge. This case study illustrates the success that storage networking can have with an adequate and reasonable application and I/O workload planning set of activities.

Appendix C

SAN/NAS Management Case Study: The Southwestern CD Company

The most important aspect of business application performance is the ability to access data. Data stored on disk, tape, optical media, or temporary cache, using the traditional direct attach storage models will be less susceptible to performance and operational anomalies than data stored through the network model. Although the location of the data should be irrelevant, if the application cannot get to it or the paths to the data are congested or critically impacted, the performance of the application will degrade—or worse, become unavailable. The ability of an IT staff to locate, identify the cause, and correct these anomalies becomes more problematic as storage configurations move into SAN and NAS configurations.

The key to successfully monitoring and managing the performance of any application starts with the consistent and proactive management of the critical paths to application data. In most cases, this is done with an assortment of vendor- and user-developed tools that support management of storage configurations. In Part VI of this book, we discussed many of the accepted management disciplines that become the foundation for a well-run IT organization, none of which will be more important in this case study than performance management. Having said that, performance management is tied to the necessary yet mundane activities of configuration management and the need for an orderly problem management system of processes and procedures.

In this study you will discover that in the storage networking area, the availability of tools and coordinated systems will be lacking greatly. As a general industry trend, management tools lack the implementation of successful technologies by three to five years. This situation is becoming even more significant and problematic as the integration of infrastructures needs to be taken into account in both small and large configurations. As the chapters on connectivity options for both SAN and NAS point out, the effects and counter-effects of storage networks are experienced on an enterprise-wide basis.

Unfortunately, as this case study will demonstrate, today's storage management industry remains in an often misguided and confused startup mode with point products and integrated storage management architectures and standards useful only to those companies selling industry controversies. However, storage management has become an effective catch-up mechanism for large vendors, especially those in the software management sector, who have more integration challenges than the small startups.

Nevertheless, regardless of vendor size, the current assortments of products are characterized by the following deficiencies, to name only the major ones:

- No correlation of functions among distributed components
- No collection of statistically useful data
- No proactive performance, configuration, or capacity trending
- No identification of "root cause" to problems

All this makes for a challenge when you are faced with installing, managing, and maintaining a storage area network and network attached storage solutions, in addition to the challenge the SAN and NAS vendors have in compensating for these inadequacies. This study will illustrate many of the problems and potential solutions when the Southwestern CD Company deals with its initial SAN experiences.

The Deferred Responsibility Scenario—or Who's Doing What to Which Data?

The Southwestern CD Company distributes all types of entertainment CDs to both wholesalers and retail outlets large and small across North America. From its headquarters in Scottsdale, Arizona, the company ships on average 250,000 CDs a week. By purchasing directly from several manufacturers at optimal prices, the company has achieved revenues of approximately $51 million per annum. Its profit margin is based upon moving product effectively to its retailers to leverage the current entertainment fashion for both music and movies. Because it competes with other, larger distributors, Southwestern must offer better services and a more competitive price structure.

The CD distribution industry is based upon time to market and can rely heavily on technology to facilitate and provide a competitive advantage to the ordering process. CD products are purchased from manufacturers with a return policy that allows the retailers and, subsequently, the distributor to return unsold products. These market dynamics are reflected in the price, generally at the retail level, although the competitive positioning based on availability and amounts can be experienced at the distributor level.

An IT department of 15 people supports the company. Ten permanent personnel and five contract programmers staff the department. They do not have specific storage expertise, but they rely on their systems administrators to handle the operating system infrastructures that include the storage configurations. Their configuration is Microsoft-based with the exception of the database servers, which are UNIX machines running a popular relational database system. They have recently implemented a SAN to consolidate their databases and provide more responsive order-entry and sales-entry applications.

In conjunction with order-entry and sales-entry applications, the company runs an accounting application, a sales analysis application, and corporate e-mail, all from within its Metro Data Area (see Chapter 2). The recent implementation of the SAN has consolidated data from all of these applications within the available storage resources. All applications are now centrally accessible through the SAN.

The company's configuration, business applications, and client network is depicted in Figure C-1.

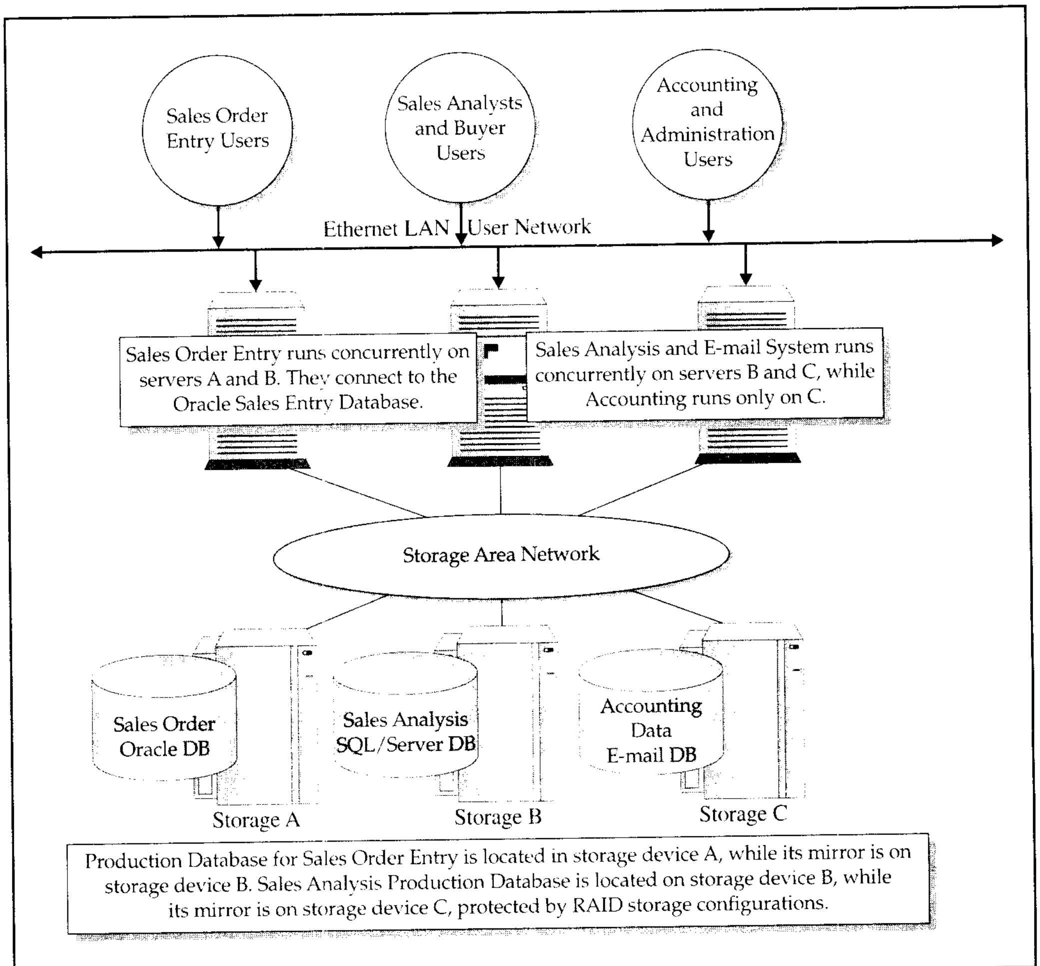

Figure C-1. *Southwestern CD's Metro Data Area*

The Problem

The company's new sales-entry application has experienced periodic slowdowns in its performance. In certain instances, the database has had to be restarted and reorganized to resume operation. Although the rationale behind this has yet to be analyzed or understood, restarting appears to restore performance for a time—all of which takes time away from the sales department, which cannot enter sales data while the system is down. Today's scenario, however, is further complicated. It's the end of the month and the day begins with a handful of frustrated phone calls from sales personnel, who either can't log in to the application or are suffering because the application is running extremely slowly.

Business Situation

At the end of the month, the sales department really needs to hit its numbers so that the orders can be booked for this quarter. Order quotas come from the Southwestern executives, are part of the quarterly goals, and are the basis for the financial incentives for the sales management staff. They rely heavily on the system to be up and performing optimally to enter as many orders as possible. Over and above is the new pricing structure on the hot new CD products they can sell to their largest retailers, to which they are still trying to sell.

Technical Scenario

The IT department is short staffed, because the database programmer recently quit to take a job at a bank. The application database administrator (DBA) now has new and additional responsibility for the systems aspects of the production databases. The periodic slowdowns happened before the SAN installation; however, they were never completely debugged and root causes were never identified. The problems appear to have gotten worse as the SAN went into production. All of the user table data has been transferred to the SAN disk arrays, which was the last activity the recently departed database programmer accomplished.

Problem Dialogue

The following is a summarization of the IT department's events and activities in determining the cause and solution to the problem.

9:45 a.m.—After numerous calls to the Help Desk over the past hour, the IT operations manager fills out a "trouble ticket" and assigns it to the DBA.

10:45 a.m.—Busy implementing the next release of the database while parsing additional database design changes, the DBA, having worked through most of the night, arrives to work late. The DBA initiates the database management program, which provides performance metrics on the database, leading to the following conclusions: the sales-entry database is up and running, and a number of deferred writes are filling up the temporary storage. The DBA decides the problem is an operating system anomaly and reassigns the "trouble ticket" to the Windows administrator.

11:00 a.m.—The Windows administrator, who is also the LAN administrator, is busy implementing a new LAN segment to facilitate the expansion of the accounting department. Consequently, he must be paged when there is no response to the "trouble ticket" within the allotted time.

11:15 a.m.—The Windows/LAN administrator logs in to the sales servers and initiates the OS management software, which provides performance metrics on each of these systems. Two anomalies become apparent: client transactions are being queued at an increasing rate, and this is putting stress on the system as memory and temporary space within the database becomes saturated. System-level paging has increased to an alarming rate. (See Figure C-2.)

11:30 a.m.—After a brief discussion with the DBA, the Windows/LAN administrator and DBA concur that a reboot of the sales system and reinitialization of the databases will clear up the problem. By this time, the entire sales-entry system is down. A message is broadcast via e-mail and voice-mail stating that the sales-entry system will be rebooted during lunch. The VP of sales estimates $500,000 in sales has not been entered. He is not pleased.

12:30 p.m.—The DBA and the Windows/LAN administrator discover that although the servers are rebooting, the database is extremely slow coming up. The Windows Performance Monitor tool now indicates that few I/Os are being serviced from the sales servers. After further discussion, the two hit upon the possibility that some external application is impacting the reboot of the databases.

12:45 p.m.—Inquiries are made and the two discover a large archive job has been running since 8:00 A.M. Initiated by the Sales Analysis department, the job was to dump two years' of sales history from the SQL Server database running on the ADMIN server. Not surprisingly, the sales history database is stored within the same storage subsystem as the sales-entry database. The DBA and the Windows/LAN Administrator (who've ordered in lunch by this point), attempt to contact Sales Analysis management to get permission to stop the archive job and restart it later that night or over the weekend. The entire Sales Analysis management team is out to lunch, of course.

2:00 p.m.—The DBA finally contacts a Sales Analysis manager and negotiates a stop to the archive job, which will restart later that night—but with the mandate that he (the DBA) have the data archived and reorganized for his department by 9 A.M. the next morning. The DBA is not pleased.

3:00 p.m.—The Windows/LAN administrator and DBA terminate the archive job and restart the sales servers and sales-entry databases, all of which come up running at optimum levels. The second message is broadcast via e-mail and voice mail: "System's up. All's well that ends well."

3:30 p.m.—Sales personnel are finally back online and able to enter sales data into the system.

Epilogue

The analyses surrounding activities of problem identification, determination, and correction are indicative of the state of management within storage networking environments. The following epilogue, or post analysis, can provide some insight into the hidden problems that can develop as storage configurations move into a networked setting.

Root Cause

In this case, unexpected I/O activity from a single server impacted the capacity of the I/O subsystem and storage network to the degree that it severely limited the operation of the sales-entry application and access to their data sharing the same network.

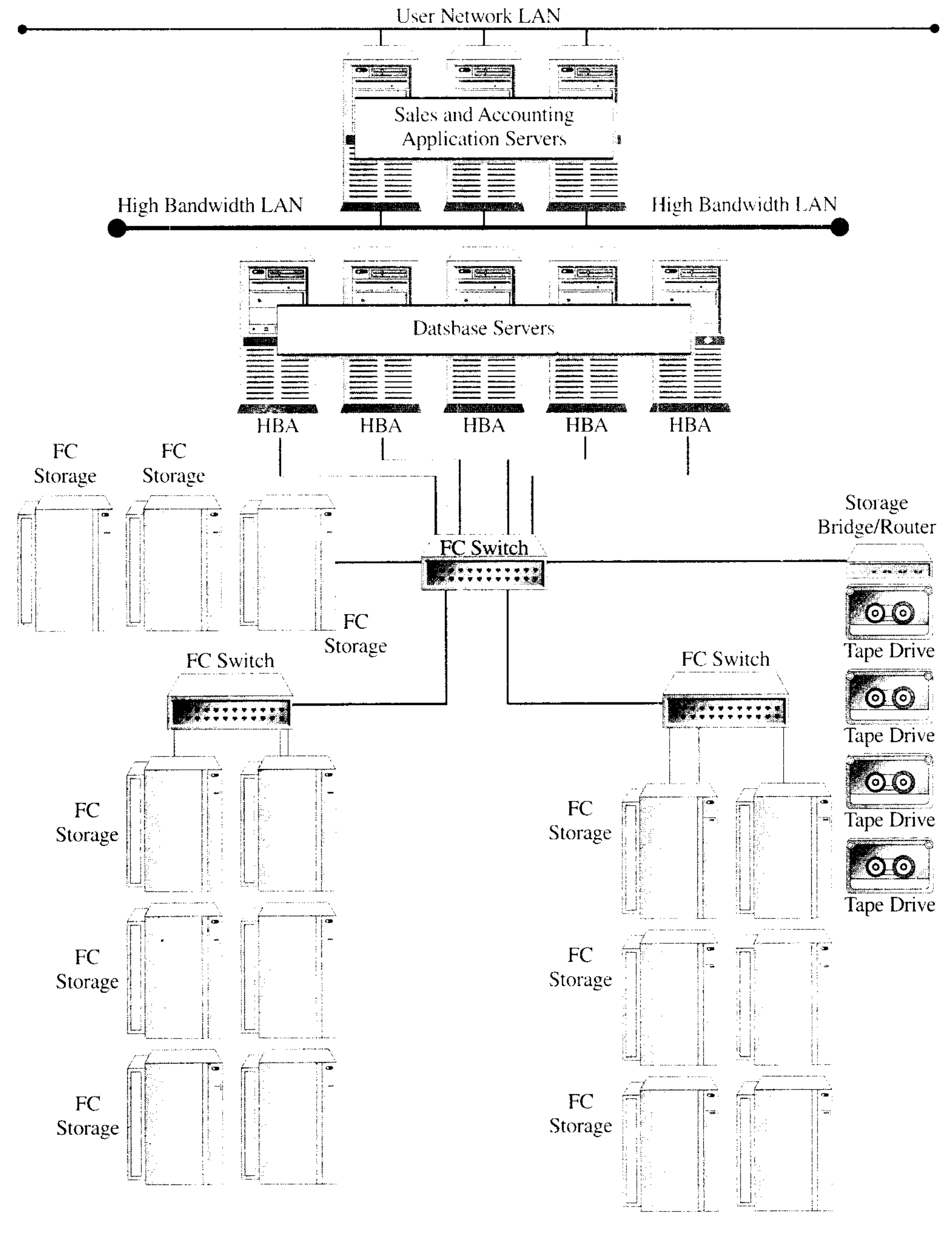

Figure C-2. *Southwestern CD's Systems Configuration*

Solutions

Unscheduled archive jobs were terminated and rescheduled, sales servers were rebooted, sales-entry databases were reinitialized, and recovery programs were run.

Downtime

In effect, the sales-entry application was down from 8:30 A.M. to 3:30 P.M., a total of 7 hours. Due to the sales-history database not being archived and reorganized, it was unavailable for a longer term, some 12 hours.

Effect on Business

As a result, sales was unable to log all of the day's orders, subsequently losing an estimated $1 million in backlog business. Downtime of the sales-entry database also adversely affected other departments, including purchasing, whose buyers were unable to meet order deadlines to record manufacturers, which will increase costs by approximately $2 million for the next month.

Effect on IT

IT management reviews the steps taken to resolve this problem and concludes the following:

- More correlation of information across the systems is needed.
- Greater attention should be given to the "root cause" of a problem.
- No integrated tools are available to coordinate the actions of the Windows/LAN administrator and the Oracle DBA.
- The Windows administrator and the DBA must work toward proactive trending in estimating storage resource requirements.
- Additional storage restrictions must be put in place to prohibit processing intrusion into the Sales production environment.

Evaluation

This scenario can be evaluated in three ways:

- Business rules and policies
- The processing configuration limits and choke points
- What available tools and practices are currently available that are relevant and appropriate to this scenario

Business Rules and Policies

The Southwestern CD Company, like many companies, lacked a complete evaluation and definition of the business applications it was supporting. Prioritizing sales entry, sales analysis, and other workloads would have provided metrics for dealing with many of the timing issues. Defining these rules and policies is the responsibility of the business user and the IT department and should be performed up front, if only on a macro-level (that is, sales-entry must have highest availability of all applications).

Knowledge of Processing Configurations

Not knowing the processing configuration, especially the storage network configuration, was especially harmful in this case. More importantly, knowing and understanding the limitations of the storage network in context to the supported servers masked the root cause of the problem. Finally, and most importantly, was the lack of association between the storage network and related devices to the business application. Had this association been made, the problem would have been identified within 15 minutes and resolved within 30.

Available Tools and Practices

Unfortunately, in this case, any available tools are probably of little use in identifying root causes. The following points provide a quick summary of the value of the major tools currently available:

- **Storage resource management (SRM)** Defensive, reactionary information received after the fact. SRM tools are resource-focused, operate without business policy, and provide no correlation of information across distributed servers; in addition, they have no relationship to the business application.
- **Backup/recovery** Adding new and/or bolstering existing backup/recovery software to the storage network. This is complementary, at best, to resolving the problem. The benefit of faster recovery time is not part of the root cause of the outage. Although necessary, backup/recovery bears no relation to the business application.

Although more software tools are becoming available, they continue to provide disparate, incompatible, and inconsistent levels of information on the storage infrastructure. No single tool provides consistent, proactive management functions that associate business applications with application data. IT management must choose from an assortment of tools that provide only discrete levels of empirical information, ranging from operating system metrics, to database metrics, to I/O and disk metrics. IT users bear the burden of correlating these seemingly unrelated sets of information in an attempt to understand the effects of resources on business applications.

The deficiencies within these tools are compounded by the requirements, costs, and expertise needed to support an increasing set of server platforms, operating systems, and major application subsystems such as relational database management, messaging, and transactional systems.

The following points illustrate only a few of the major deficiencies challenging today's management tools, as well as the inefficiencies surrounding business application availability:

- *No correlation functions among distributed components.* Today's business applications are distributed, which means that events happening on one server can seriously degrade performance throughout the entire system. The ability to correlate important aspects of performance information as it effects the business application is currently unavailable.
- *No proactive trending.* Today's IT managers are expected to drive the bus effectively while monitoring performance through the rearview mirror. Literally all reporting and trending is historical. The information that ultimately reaches the IT user is about incidents that have already occurred and provides little value in determining real-time solutions to business applications that perform poorly.
- *No identification of root cause to problems.* The effects of the conditions stated previously means it is unlikely that the information supplied to the IT user will provide any sort of root cause as to why a business application is down, not to mention any information intended to identify and correct the problem. Most activities address only the symptoms, leading to reoccurring results.

As far as practices go, this case study illustrates a competent and responsive IT department with processes in place for problem notification and assignments. However, the specific expertise and focus of the individuals contributed to the time it took to pinpoint the problem—that is, the database was up but was not processing any transactions.

The following points present alternatives to consider when providing a long-term solution:

- **Reconfiguration/isolation** Adding hardware (servers, switches, and so on)—also known as "throwing money at it." Although a possible long-term solution, this can essentially convolute the problem and put storage isolation as part of the IT solution matrix, thus limiting the inherent benefits of storage networking. Also, reconfiguration/isolation provides no direct association or long-term benefit to the applications.
- **Increased manpower** Hiring more people to monitor and manage the storage infrastructure by business application—also known as "throwing money *and* people at it." This could be a solution, but a highly improbable one, given the monitoring requirements, expertise required, and associated costs. This solution does in fact relate to the business application, but only at a prohibitive cost.

Final Analysis

More powerful servers and storage subsystems are quickly evolving into Metro Data Areas characterized by high data traffic, inevitable congestion, and complex problems relating to data availability. The Internet is also driving the Metro Data Area theory by increasing exponentially the number of data highways leading in and out of these already congested areas.

Business applications are distinguished from other applications through their direct association to the business, their size, and their mission-critical requirements. They are supported within storage systems by practices and products that manage various aspects of maintaining the storage devices and data. Business applications can be enhanced with storage networking technology, although a host of problems remains that will cause difficulties relating to data availability. Currently, no tools directly relate the business application to the storage network.

Through a composite IT scenario, the realities behind operating with a storage networked environment were illustrated in this case, as well as the challenges in problem identification, isolation, and solution within these configurations. Although a SAN configuration was used, a NAS configuration could just as easily be substituted for the same effect. Evaluation of the scenario indicated that best practices and current tools would not minimize or stop these types of availability problems.

Steps to Better Storage Networking Management

As demonstrated in this case study, the association of business applications to computing resources remains for the most part at macro-levels, at best, and not at all within new technologies, such as storage networks. Storage networks, given their disparate components and critical interrelationships, have evolved into an infrastructure category within the data center. As such, they have become a tremendous challenge to manage as a complete entity, much less integrated into the logical set of applications and supported systems. The following provides additional information on what's required for new tools. This is further augmented by best practices for managing storage networks.

The Need for Next-Generation Performance Management Tools

The most critical element of performance for a business application is its availability to its own data. Consequently, a level of performance tools is needed that closely associates the business application with storage networking resources. These tools should provide a level of functionality to perform the following:

- Correlate business application performance information across servers and storage networks
- Provide metrics and monitoring for proactive trending of performance and take appropriate action based upon previously established business rules
- Identify root cause areas, if not components, to IT support personnel and make appropriate recommendations based upon previously established business rules

Such a set of tools would enhance IT management's ability to manage business applications within increasingly complex application and storage networked configurations. These tools would move IT professionals closer to the critical elements of business application performance.

Storage networking best practices to observe include the following.

- *Eliminate unnecessary downtime.* Decrease the negative business impact of storage networking and enhance the inherent abilities of storage networking technologies.
- *Automate performance management tasks.* Decrease the cost to IT management and staff, and maintain specialization in storage networking as implementation and production use increases.
- *Learn/associate business to infrastructure.* Root cause analysis, proactive processing, and management can provide solutions before outages occur.

The following guidelines can provide some visibility into dealing with the Southwestern CD company scenario:

- *Focus maintenance efforts on critical resources—before they go inoperable.* Identify the critical resources through effective configuration management activities and processes. These can help in focusing efforts on current configuration and processing scenarios.
- *Proactively manage growth and change within the storage networked configuration.* As the storage network moves into production, be proactive through an effective problem management and ongoing capacity plan. This will allow the identification of problems and changes to capacity. Further accommodations can be helpful through an effective change management program that will identify and focus on processing anomalies and activities.
- *Gather value oriented performance data.* This will probably be the most intrusive and time-consuming task, given the lack of availability and the programming necessary to gather information from disparate sources such as MIBS, Common Information Model, and vendor statistical collections. This can also provide valuable input to vendors and industry during problem-resolution scenarios and activities.

Appendix D

Glossary

ALU A component of the computer's central processing unit where mathematical calculations are assembled and executed. CPU-intensive devices and applications rely on the ALU heavily, which in turn utilizes temporary storage in locations of CPU registers and system-level cache.

application development life cycle Activities that surround the design, development, and implementation of computer applications. These are generally characterized by logical phases of work that progress in a contiguous fashion. Business justification, external and internal design, coding, testing, and deployment characterize these activities.

archival The activities of moving data from an online storage device to an offline or removable storage device for long-term storage. Archiving data generally results from data that has not been accessed in a specific time period, consequently taking up valuable online disk space. It is distinguished from backup and recovery by being available for reloading into an online system for later access.

backup/recovery Copying data from an online storage device to an offline device for protection against possible data corruption, storage outage, or site disaster. These functions differ from archival activities and functions by making a copy of the data and their time requirements in order to provide a specific data state and other information for recovery purposes.

bandwidth A set of frequencies that designate the band of electronic wavelengths used to transmit signals within a computer system and reflect the transmission capacity of the link. This is measured by the address of the data that can be transmitted and the speed of the signals through various cycle times measured in MHz (megahertz).

blade computing A method of loosely connecting discrete computer systems through a common bus. Blade computing is made up of small, powerful, fully functional, general-purpose computers that are optimized with only CPU, memory, and I/O bus components. Server blades, as they are known, are used in conjunction with an I/O backplane that supports connectivity to a common set of I/O devices, including various storage devices, network devices, and control functions.

block A unit of data used with storage devices and systems that requires a contiguous block of data accessed with each I/O instruction of the system. Block sizes are dependent on both operating system and storage device types, but have direct relationships to the efficiency of the system's I/O operations.

bus A system that provides the interconnection between various devices within computer systems. Generally characterized by a singular and priority-driven operation between two devices within the bus—one transmitting instructions and data, the other receiving them. Various bus mechanisms are used within a computer system such as the PCI bus to interconnect CPU, memory, and I/O adapters, with storage buses such as SCSI providing interconnections to multiple storage devices.

business application A set of computer programs that automates the functions or operations of a business. This is different from utility or management applications that are computer programs which automate or enhance the management of the computer systems. Examples include systems such as backup/recovery and disk defragmentation systems.

capacity plan Activities that identify business elements and translate these into workload requirements. These include the projection and configuration of systems that effectively process the workload requirements. Historically, capacity planning has included processor, memory, and storage capacity as a holistic system. However, storage networking separates the I/O workload from the computer systems and demands a new set of activities that estimate I/O workload but which are integrated back into the supported processor complex.

CD-ROM Storage type that utilizes a polymer material for encoding information which is then subsequently read using a laser light source. CD-ROM has yet to be fully integrated into FC SAN; however, it can be used in some configurations with a SCSI bridge/router. CD-ROM drives are available as NAS devices and can be configured as network storage devices, with certain limitations.

cluster In terms of computing, it's a collection of interconnected CPU resources that provide shared efficiencies, such as workload balancing, reliability, availability, and serviceability to the supporting infrastructure. In terms of storage, it's the smallest addressable unit on a disk drive as used by Windows operating systems. Cluster sizes are dependent on the total storage capacity of the disk—therefore, the larger the disk, the larger the cluster. This has a direct relationship to the efficiency of the system's I/O. Writing a single file to a large cluster uses the entire cluster. Consequently, writing a 1k file to a 32k cluster utilizes the entire space.

Common Internet File System (CIFS) A distributed file system used with Windows operating environments that allows files to participate within networked file systems and network attached storage. CIFS has become an ad hoc standard in enabling Windows clients to participate and share files within networked environments.

controller A device which controls and manages multiple devices that communicate with a computer system. Storage systems utilize controllers extensively to control and manage multiple storage devices that communicate within a bus system.

CPU The central processing unit (CPU) is the mechanism that processes instructions which invoke all other actions within the computer and ultimately produce the results of the end-user application. The CPU is dependent on its basic architecture in how it operates and processes instructions. Two popular architectures are the Reduced Instruction Set (RISC) and Complex Instruction Set (CISC). Each processes instructions differently; however, both utilize temporary locations, such as registers and system-level cache, and initiate I/O operations.

cylinder A measurement of disk geometry, where the association of tracks is measured vertically within the multiple platters of a disk drive. Cylinder allocations can have significant performance benefits for applications that access large amounts of information on a read-only basis.

data mart Related to the data warehouse type of application, the data mart is generally designed around a single subject. Consequently, the scope of the information is reduced, along with its relationships, within a greater set of information. For example, a data warehouse application that supports customer financial accounts will have all customers, their accounts, and their related attributes. However, a data mart providing secondary support may only have information that relates to a subset of this information such as customers in a particular region, or a set of financial products and their status. These subset applications are datacentric, and although they utilize a smaller set of data, they remain challenging given their distributed requirements and use of relational databases.

data warehouse A type of application that allows end users to analyze historical information on a set of related subjects. For example, a data warehouse for a financial application may have information for the past three years on all customers, their accounts, and any related attributes. Therefore, using relation data base technology, an end user can query all customers in a particular set of states that have moved within the last year and have brokerage accounts. These applications are datacentric and can be the largest users of both storage and processing capacity within a data center.

Digital Linear Tape (DLT) A format for magnetic tape that has become an ad hoc standard for open systems. The architecture is based on the DLZ1, a special compression algorithm where data is written on the tape in dozens of straight-line (linear) tracks, usually 128 or 208. Capacities are vendor-dependent but range to >50GB of data when compression is used. A variant of DLT technology, Super DLT, provides capacities to move beyond the 100GB on a single cartridge.

Direct Access File System (DAFs) A file system used specifically with NAS devices, where file access bypasses the normal file location services through the OS and storage controllers and utilizes a new service called virtual interface (VI). This allows files to be transferred through direct-memory transfers, which circumvent the overhead of TCP processing for file access. This, however, requires that a number of new components be implemented, including a new type of network card which enables the VI protocols and NAS devices that process direct memory addressing.

director A class of storage switching devices that are generally defined as having the attributes of >64 ports, additional hardware functions that enable redundancy within the switch, and related software that configures recovery and redundancy features.

disk compression Operations that allow data on a disk to be compressed for efficiency of space utilization. Compression algorithms usually find all spaces and duplicate characters and remove these from the storage data source. These are then replaced when the data is accessed.

disk defragmentation Disks read and write information in a random manner, which is part of their value in accessing data. However, as write operations occur over time, the disk becomes fragmented as the drive writes related data blocks out to available blocks, clusters, or sectors, depending on the type of disk device and operating environment. This fragmentation of the disk, with related data spaced all over the geometry of the drive, begins to cause performance problems since it takes the disk head longer to read the segments of the file, or find available space within a fragmented disk. Defragmentation, sometimes referred to as defragging, moves all the scattered blocks, clusters, or sectors into a contiguous space on the drive, essentially reorganizing the drive so it can operate more efficiently.

embedded database Many computer applications come with a database integrated with each function. Other applications are unattended and utilize databases to collect or store data for its operations, such as an automobile, HVA systems, and aeronautical systems. Each database included in these applications integrates an embedded database. The embedded database is tightly integrated with the application programs and cannot be used for other general-purpose applications. These applications, and moreover their embedded database elements, can be challenging to storage given they have limited performance and allocation parameters. Therefore, their inherent limitations have to be compensated for by the processing infrastructure.

F_Port A FC switch port designated to connect participating nodes, known as N_Ports, to the FC network. These nodes can be devices such as storage arrays and servers.

fabric A Fibre Channel switch, or multiple Fibre Channel switches, that are interconnected in a networked topology to facilitate the physical transmission of FC frames between any two N_Ports. This includes the software operating within the control of a microkernel that facilitates the operation of a FC switch. The software functioning within the control of a microkernel operating system provides the base level services that form the switch fabric network. The fabrics are vendor-dependent and although compliant with FC standards, they have levels of differences in how they implement many of the FC operations.

Fibre Channel (FC) Fibre Channel is a network connectivity standard for a serial I/O bus, with the capability of transmitting data between two N_Ports at 100 MBps. Principally, the X3T11 committee of ANSI governs this standard.

file allocation table Disks have tables of contents so they can access the data stored in a responsive manner. Most operating environments support some type of file allocation table that resides on the disk and stores information regarding the files and physical locations of the data on that disk.

file system File systems are the operating system functions that manage the stored data within the computer. Most file systems allow users to organize their data into a defined naming and hierarchical structure. File systems function in networked environments by permitting other computers to access their data by providing access to segments of the file system. File systems maintain various levels of data integrity by monitoring and managing the processes that allocate and write data to the files. This becomes problematic with storage network environments as multiple systems with their own file systems attempt to manage the files they own—a particularly difficult problem in a SAN where multiple servers may access the same storage array, device, and subsequent file. NAS devices have less of a problem with this since they provide remote file services and offer centralized ownership of files that reside on NAS storage devices.

frame A component of the Fibre Channel standard that describes the method and specifications of the basic transmission packet that encapsulates the user data.

G_Port A port on a FC switch designed to function as an F_Port or an E_Port that connects both nodes. This can be external to FC switches through interswitch links (ISLs), and SCSI-to-FC bridges/routers. This type of universal port is becoming increasingly popular in developing simpler port utilization structures. These ports are vendor-dependent on implementation.

Head and Disk Assembly (HDA) The mechanical assembly that houses the read and write heads for a magnetic disk. The assembly moves mechanically over the spinning platters of the media as it performs its operations. Additional functionality has been moved into the HAD to facilitate greater read/write performance.

InfiniBand An I/O standard that provides a switched fabric network replacing the traditional bus networks of computers. The InfiniBand standard requires that all I/O be connected through an InfiniBand switch with the added benefit that connecting computers together makes clustering a more scalable and efficient process.

iSCSI A new I/O standard that provides for the execution of a SCSI-based I/O operation over a TCP/IP network. This allows a server to execute a block I/O operation to a remote storage device. The new standard provides additional specifications for encapsulating SCSI-based commands into TCP communications and transferring them through an IP packet to corresponding SCSI-based storage devices that are IP addressable and have compatible TCP processing capabilities. Initial implementations have required new NIC cards to support these functions at both the server and storage device level.

JBOD Defined as "just a bunch of disks," this term defines a storage array in which all devices within the array are addressable independent units. These provide additional capacity but do not offer any fault resiliency in the event of an inoperable drive. Partitioning data throughout the disks and providing a layer of virtualization services is generally part of the I/O management of the operating system. These functions are also available through third-party software applications.

Linear Tape Open (LTO) A format for magnetic tape that competes with digital linear tape. LTO uses a linear multichannel bidirectional format but adds an enhanced timing-based servo (a device that automates a process of error correction for a mechanism), hardware data compression, and efficient error correction. Initially developed by a consortium of IBM, HP, and Seagate, it now has two different formats— one for fast data access and another for greater storage capacity. The Accelis format uses 8mm tape on a two-reel cartridge that loads at the mid-point to provide fast data access, specifically for read-intensive applications. The other is the Ultrium format that uses a single reel of half-inch tape to maximize storage capacity, specifically for write-intensive applications, such as archival and backup functions. Ultrium-based products offer a >100GB capacity, while some vendors are approaching the 1TB capacities.

LUN A term used with addressing SCSI devices within a SCSI bus system. A logical unit number (LUN) is defined at the controller level to access, control, and manage the devices within the connected bus. LUNs provide a level of virtualization when communicating with the operating system, and subsequently the application, by shielding the application from the complexities of knowing physical addressing and access information for disk or tape I/O operations.

LUN management The activities, functions, and capabilities that make up the manipulation, control, and assignments of LUNs within SCSI bus systems. These generally take the form of configuration parameters for controller initialization through microcode, but can assume higher levels of functionality, especially when managing multiple SCSI bus systems.

LUN masking Part of LUN management where certain servers access specific SCSI devices within the shared SCSI bus environment, while other servers are blocked from accessing specific LUNs.

micro-kernel A special type of operating system that can be optimized for a particular application. Both SAN and NAS devices use micro-kernels to provide base level operating system services to the SAN switch and NAS server, respectively.

MPP A term used to describe a massive parallel-processing configuration. These configurations are made up of multiple processors (often numbering in the hundreds), sharing no system resources except a high-speed interconnect. They are currently used to address complex scientific, academic, or business analytical processing such as measuring earthquakes, nuclear explosions, and very high-end data warehouse applications, by breaking the application tasks into activities that can be processed in parallel. The processing nodes within MPP systems often have their own storage arrays; however, in some cases, some nodes have a common I/O configuration.

N_Port The FC port that is designated for participating nodes, such as servers and storage arrays. These ports are configured through a Host Bus Adapter (HBA) and used to connect to the FC switch F_Ports.

NVRAM A type of shared memory that is used by multiple system processors. Non-volatile random access memory is a portion of memory that ensures data integrity with the data that is written to the disk as multiple processor I/O operations take place across the same data.

PCI Bus A bus system to connect peripheral devices, which has become a standard within the industry. PCI buses provide the interconnections to the internal and external storage devices through adapter connectivity points, but also provide the interconnections between the CPUs, RAM, and caching systems.

Relational Database Management System (RDBMS) A database system that allows for the organization and access of data using relational operations. RDBMS architectures are developed from tables where entities are defined with each table and attributes are defined as columns or fields within the table. These databases have become the preferred method of storing information for both transactional and analytical systems. RDBMS architectures provide their own file systems and special I/O operations given they are table-driven and have different data access and storage characteristics than file systems.

SCSI A standard term for a type of bus storage architecture used with block-level devices such as disks and tape. The Small Computer Systems Interface (SCSI) defined the standards used with connecting storage through parallel bus architectures. Forming the largest install base of storage devices today, the SCSI command structure has superseded its bus application and is used with both FC protocols and the newly defined iSCSI standard.

service levels Service levels exist on several levels. There are specific agreements with end users that define the level of service the IT department provides. Internal service levels are managed to provide a working measurement of systems and networking performance. Service levels for storage, if they exist, are often integrated into the internal system service levels and are comprised only of capacity measurements for end users. The establishment of meaningful measurements for the storage-networking infrastructure that encompasses storage, systems, and network specifications have yet to be developed.

SMP Symmetrical multiprocessing defines a configuration that provides multiple computing nodes supported by both shared memory and shared I/O configurations. These are generally characterized by multiple CPUs within the same cabinet. The scalability restrictions have long centered on the limitations of the shared I/O structure and related performance problems with storage scalability. SMP configurations are greatly enhanced by the inclusion of storage networks, especially SANs that provide additional bandwidth and speed to the I/O storage subsystem.

snapshot A term used for copying online data at a synchronized time and providing a copy of a data volume, file system, or database for later access. Used in response to user data requiring 24/7 access, snapshot configurations generally require a duplicate volume to be operating so that the actual copy is made from the duplicate volume. This allows uninterrupted backup operations, and provides a consistent snapshot of the data at the time it was executed. Once the snapshot (backup) is taken, the duplicate volume then resynchronizes itself with the primary online volume and continues processing.

T11.3 standard The ANSI Fibre Channel standard specifications. All vendors whose products support Fibre Channel work through this set of specifications. The ANSI Fibre Channel X3T11 Committee governs these.

tape compression A setup of functions and operations that allows data on a tape to be compressed for space efficiency purposes. Compression algorithms usually find all spaces and duplicate characters, and remove these from the storage data source. These are then replaced when the data is accessed.

temporary storage These consist of various locations within a computer system where data is stored temporarily until it can be used by the CPU or written on a storage device. These locations are considered internal when described as system cache, CPU registers, and memory buffers in RAM. They are external when considered as part of a peripheral such as a block device.

track A fundamental measurement of a disk's geometry that spans the circle made by the disk head as it makes a single rotation of the media platter. Track sizes, therefore, are relative to the size and density of the media platter. Tracks are formatted for use with an operating environment by way of the operating system dividing the tracks into clusters or sectors. Tracks aligned vertically with multiple platters make up another measurement called a cylinder.

zones Functions within a FC switch that allow the ports to be segmented into individual areas for node access. Authorized access lists for each port provides login services to the switched fabric defined zones. Zoning can further be determined regarding the attributes of the ports, node login addresses, and login attributes. This allows the SAN to define access to authorized devices such as servers to attached storage devices such as disk and tape arrays.