NetworkAppliance®
The evolution of storage.™

Windows® 2000 and Unified Storage

by Dan Morgan, Manager, Database Performance Engineering,
Buddy Taylor, Database Performance Engineer,
Ricky Stout, Database Performance Engineer,
Mark Hayakawa, Technical Marketing Engineer,
Antonio Robinson, Technical Marketing Engineer,
Network Appliance, Inc.
August, 2003 | TR 3273

TECHNICAL REPORT

Network Appliance, a pioneer and industry leader in data storage technology, helps organizations understand and meet complex technical challenges with advanced storage solutions and global data management strategies.

Table of Contents

Abstract

The purpose behind this publication is to demonstrate the advantages of using Network Appliance™ technologies to consolidate an enterprise's direct-attached storage environment. The primary focus of this paper is to demonstrate the performance gains that can be experienced by consolidating your direct-attached storage environments onto a central pool of storage that can be shared across the enterprise. All performance work was generated using real functional workloads.

1. Introduction

The paper has been written using a fictitious company called the Acme Medical Supply Company. A small to medium-sized medical supply distribution operation specializing in servicing doctors' offices throughout North America. They have several main offices throughout the country and a large number of remote warehouses used for distributing medical supply products. The company receives the majority of its orders from a Web-based order entry system. Acme has a number of data processing requirements and is struggling in its current environment due to the exceptional growth it has been experiencing.

To best simulate a real world implementation,we will use real workloads in all environments when comparing the different storage technologies. This paper outlines thecurrent Acme environment and demonstrates the problems it is experiencing. A unified storage configuration will then be proposed to demonstrate a solution that will help it meet its growth needs.

2. Current Environment

Acme has three primary data processing environments that are the center of its current operations. The first is an e-mail system running Lotus Domino. This is the only e-mail system the company uses and is critical to the company's success.

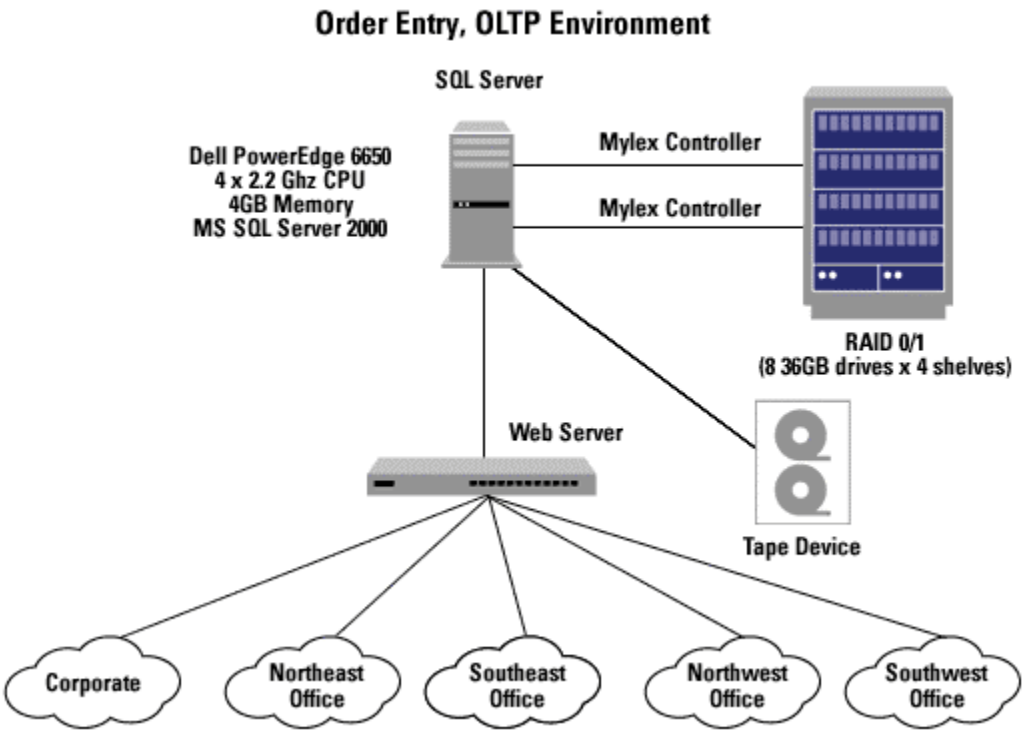
There is also an Internet-based order entry system used to process orders taken directly from the various doctors' offices across the country. This particular system uses a number of warehouses where inventories are kept, and distributions are made from these various facilities. This OLTP environment,

based on the Microsoft® SQL Server 2000 database, is the lifeblood for the company and is the one environment that has been experiencing the fastest growth across the company. Since this application is critical to the company's success, end-user response time for this application is extremely important.

A new addition to the company's data processing environment is a data mining solution. As the company continued to grow, it became much more important to know where it was successful and where it needed to improve and refocus its efforts. Acme decided to implement a SAS environment for data mining to analyze its OLTP transactions in order to make more informed decisions about the company's business. The system utilization will be low at first, however, there are plans to add more users over time.

The problem with this current environment is in the areas of performance and growth. As the company started to grow, the IT department quickly added system (as opposed to people) resources for the various parts of the business. Each server has its own pool of disks on which to store the various data. As the databases grow and as the need for processing more transactions grows, so does the need for a higher-performing storage solution. There are times when both the SAS environment and OLTP systems experience spikes in their workloads due to various issues. The processing load from the Lotus Domino environment remains fairly constant day to day.

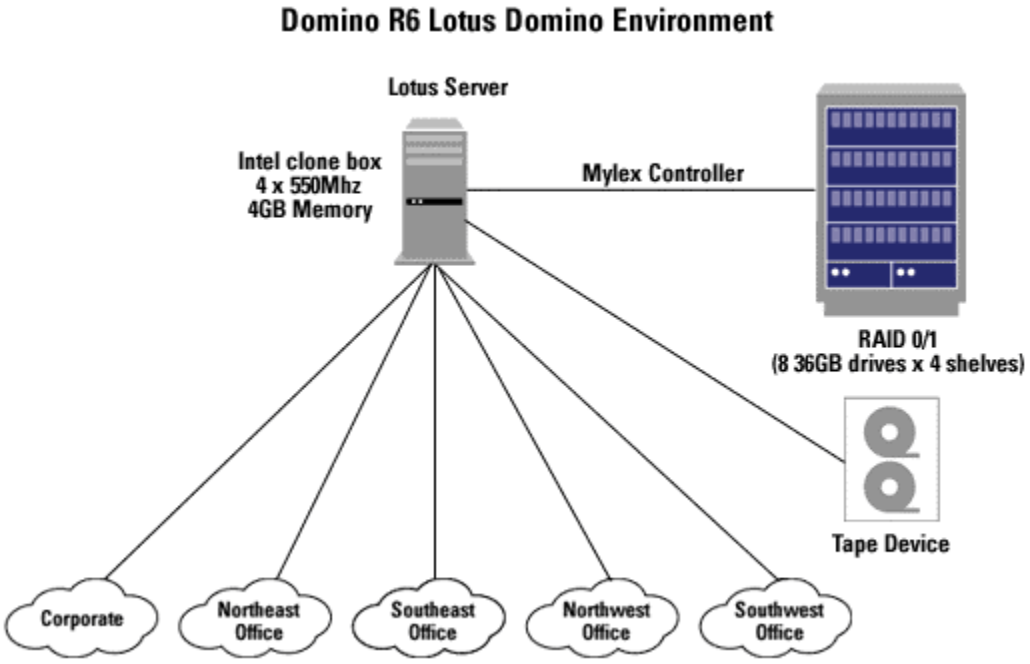
Let's review Acme's current solution for all three deployments.



The OLTP configuration uses twenty-eight 36GB disks for the database. The 28 disks are configured into two RAID 0/1 LUNs with 14 disks each. This leaves 4 disks available as hot spares for this environment. The two LUNs, as configured, yielded 476GB of available disk space for the database.

To simulate the workload for this environment, we set up, configured, and ran a TPC-C-like benchmark to simulate an order entry application. The database files for the actual database were spread equally across both LUNs. The database logs were placed on their own volume. The current database consumes approximately 60% of the available storage. We used a separate two-CPU machine for each of the regions defined above. Each regional office sent OLTP transactions to the global Web server. The Web server then sent all of the OLTP requests to the database server.

While this entire configuration was assembled and running in our database lab, this configuration could have been geographically dispersed and run to support a real production environment.

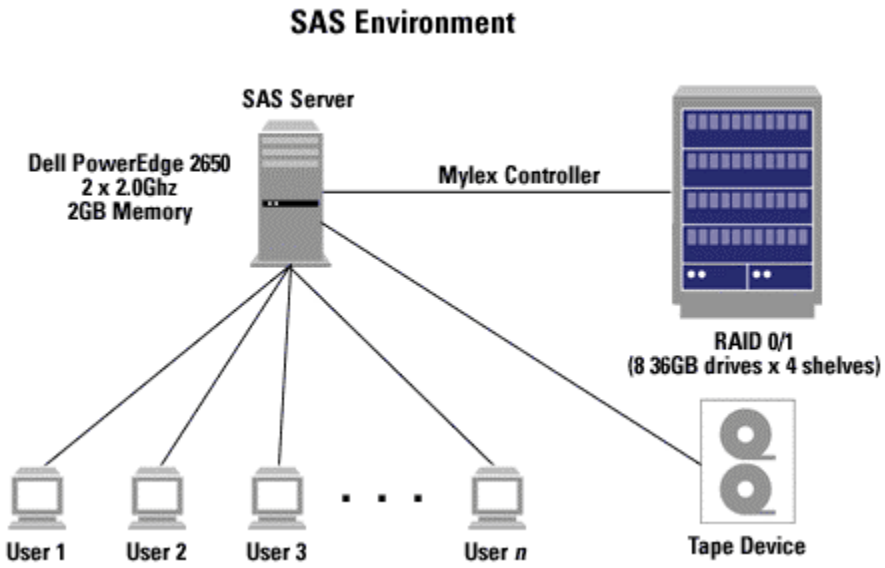


The Lotus Domino configuration uses twenty-eight 36GB disks for the users' mailboxes. The 28 disks were configured into two RAID 0/1 LUNs with 14 disks each. This leaves 4 disks available as hot spares for this environment. The two LUNs, as configured, yielded 476GB of available disk space for the database.

To simulate an actual workload for this environment, we set up and configured the Notesbench R5Mail benchmark. The Notesbench benchmark is an industry-standard benchmark used to measure the performance of Lotus Domino environments.

The users' mailboxes were spread equally across both LUNs. We used an average size of 120MB for each user's mailbox. Therefore, if this environment were to grow to 4,000 active users, it would run out of disk space in the direct-attached environment. The users were simulated in the regional offices by using dual CPU boxes to drive the load generated from the Notesbench benchmark.

Again, while this entire configuration was assembled and running in our database lab, this configuration could have been geographically dispersed and run to support a real production environment.



The SAS configuration also uses twenty-eight 36GB disks for the SAS input data and temporary sort areas. The 28 disks were configured into two RAID 0/1 LUNs with 14 disks each. This leaves 4 disks available as hot spares for this environment. The two LUNs, as configured yielded 476GBs of available disk space for the database. This amount of disk space is overkill for this application. The amount of used space per query varies from query to query. Most queries typically consume around 7GB of space per user for this environment. So 10 users would use only 70GB out of the 476GB of available storage.

This particular application is very I/O intensive, as table scans coupled with data sorts and the creation of temporary files place a much higher throughput requirement on the disk subsystem. Extra spindle throughput capacity was required so the queries would complete in a respectable amount of time.

To simulate an actual workload, we created a SAS query that would read a large input file, nearly 7GB in size. The query had a where clause that would select and create a smaller subset of data. This interim data would then get sorted by seven different columns before being placed into the ending-results file. For ease of setup and execution, each user executed the same query while maintaining separate temporary sort files and end-result files.

The SAS input data and sort areas for each user were spread equally across both LUNs.

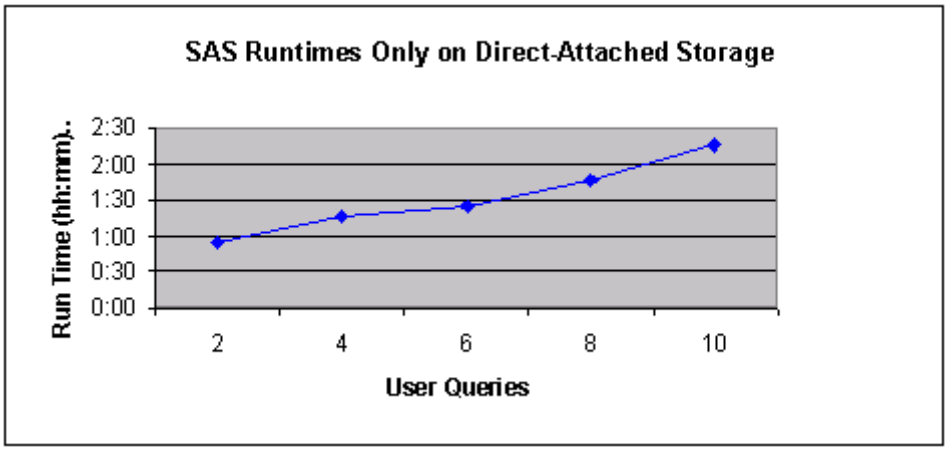
3. Current Performance Issues

Sudden Increases in SAS Queries

Let's step back and review the various obstacles Acme faces within its current environment. First, the company must perform quarterly audits of its books. During this time, auditors from an outside consulting firm are hired on an hourly basis to come in and audit Acme's records. During the audit it is necessary to run various queries to generate supporting data for the audit. It is during these times that many more queries are executed against the companies' data. The queries not only perform vast numbers of table scans but there are also numerous sort requests that require writing interim results to disk. Standard data sets are created from batch jobs every evening. These data files are then used as

input to the queries during the audit.

The graph below reflects the current performance, in the direct-attached environment, of various queries as more and more users submit jobs during this audit.

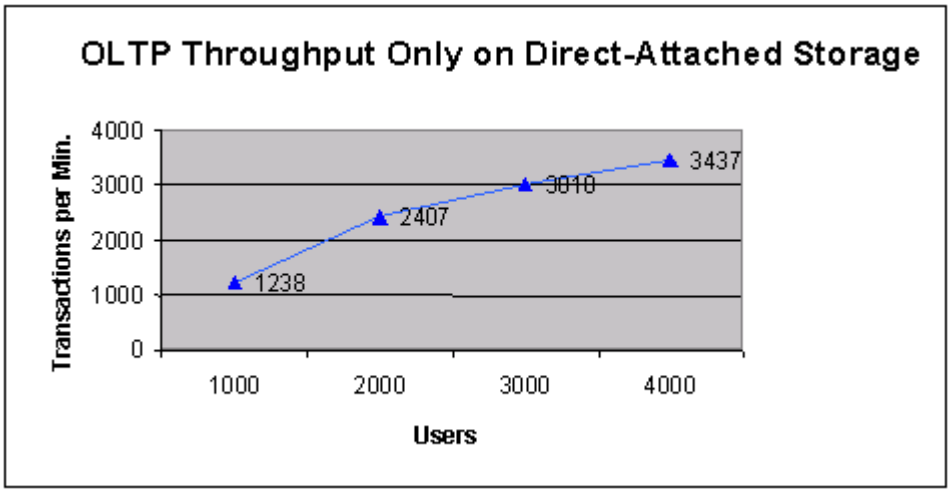


As you can see it can take more than two hours to generate these queries at the high end of the number of requests submitted. The direct-attached environment performs fairly well in the two- to six-query range. However, when more than six requests are submitted, there is a sharp increase in the amount of time needed to complete those queries. This can get quite expensive when our fictitious company is paying its auditors on an hourly basis for the audit.

Sudden Increases in OLTP Transactions

Secondly, there are times when the OLTP environment hits peaks in the order entry system. During peak processing times the disks become saturated with I/O, and response times increase. The OLTP environment has been the fastest-growing environment to date. Acme started this configuration with 14 disks and has constantly had to add additional space as both space and performance demands have increased.

The graph below shows what happens to the throughput during the peak processing times for the order entry system. Data on response times for this environment will be displayed later in this report.



Acme typically sees around 1,000 users online to its order entry system daily. When peaks over 2,000 occur, throughput is poor and does not scale very well. The current direct-attached environment becomes quite spindle-bound during these peak processing times and limits the number of orders that can be placed.

Lotus Domino Issue

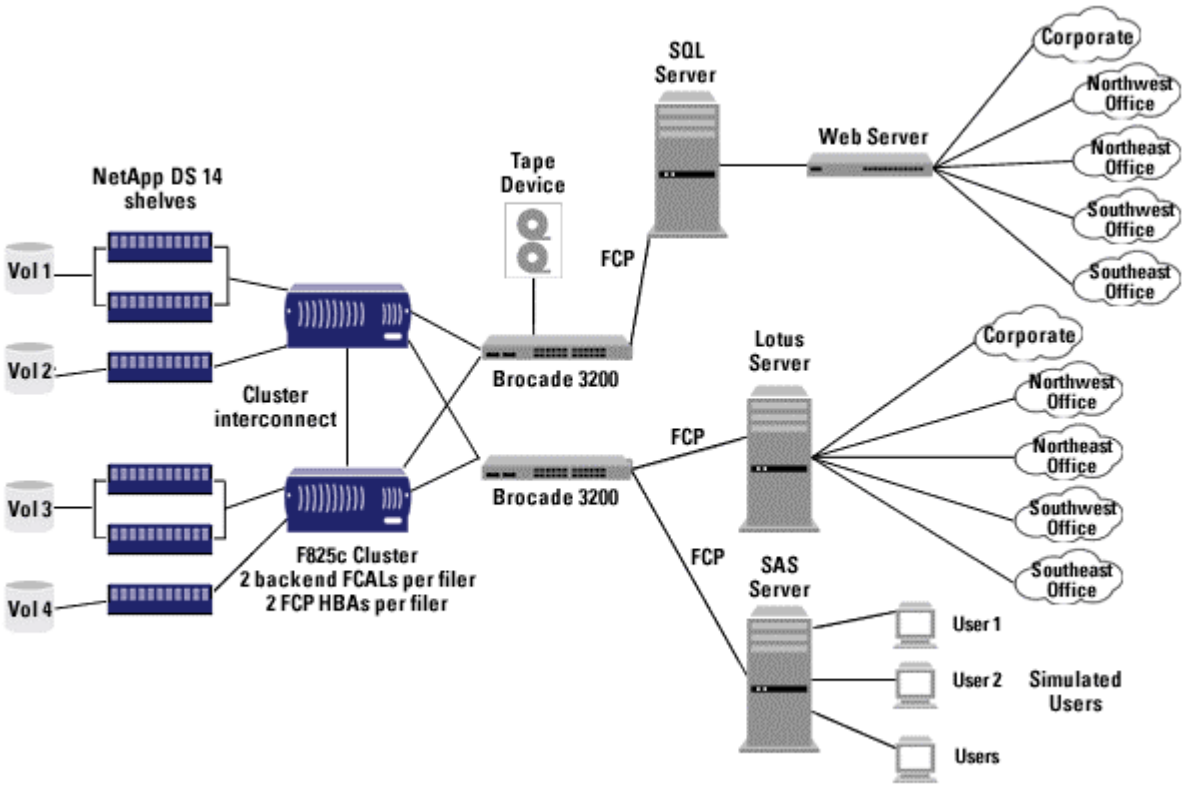
The current hardware configuration seems to be performing well for the e-mail system. The most important issue that faces this environment is the constant need to monitor mailbox space requirements. As users are added over time and as the current users' mailboxes grow, the direct-attached disks need to be constantly monitored for overall space requirements.

The direct-attached solution has a finite amount of space available for the Lotus Domino environment, and adding space when needed is time-consuming and difficult. Acme needed an easier solution for adding space on-the-fly.

4. Unified Storage Proposal

As you can see, there are various limitations in front of Acme that are difficult and complex to work around. It is also very time-consuming for the IT staff to constantly monitor and maintain storage across all these different environments. With the current growth demands being placed on these systems, it's just a matter of time before something breaks. An easier solution was required.

So a proposal was put together to consolidate all of the company's storage requirements onto a Network Appliance unified storage solution. Here is the layout for Acme's proposal.



This proposal removes the need for direct-attached storage from all three systems. This approach helps to consolidate the storage needs into one pool of storage to be shared across the enterprise. There are several advantages to this approach:

1. There is now one pool of storage to manage instead of the three distinct pools of storage from the original environment.
2. By having all disks available to all three data processing environments, you now have the total performance capacity available across the enterprise.
3. Adding storage to one central pool of storage is easier than adding storage to three distinct environments.
4. The centralized storage makes it easier to allocated space to parts of the business that need it most.
5. Spare disks allocated to maintain a level of high availability can now be shared across the enterprise instead of being dedicated to each work environment.
6. If future growth requires more processing capacity, simple head swaps can be performed within minutes to help your storage requirements scale with your company's growth.

The original direct-attached environment used 84 disks for the various databases to service the needs of the three distinct systems. There were also 12 disks across the three configurations dedicated as hot

spares, giving us a total of 96 disks used across all the direct-attached environments.

In the direct-attached scenario, it was difficult and expensive to have spare disks available since spares are needed for every controller on every server. In the consolidation proposal, spare disks are available across the entire disk subsystem, requiring fewer spares overall.

The unified storage proposal uses 76 disks for the various databases that can be shared across all three environments. Files for each environment were spread in a round robin fashion across the four volumes that were created from these 76 disks. A total of 4 disks, 2 per head, were dedicated as hot spares to be shared across all three environments. There were also 2 disks per head that were used for the Data ONTAP™ software. There were 84 disks total in the Network Appliance F825C configuration.

5. Storage Consolidation Performance

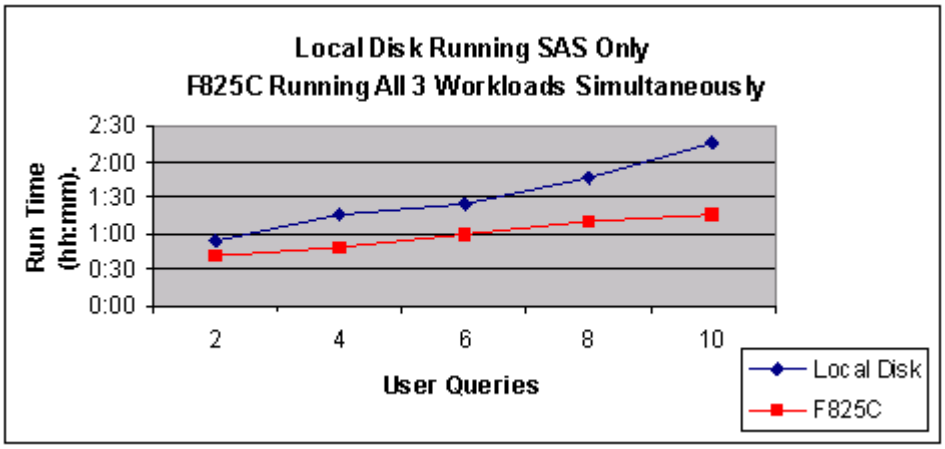
So let's implement this proposal using the customers' workloads. All three environments-order entry, SAS, and Lotus Domino-were arranged so all their data was centrally located on an F825C configuration. The workloads were then executed to simulate an average day's worth of activity.

Note: An average day typically has about two active SAS user queries, around 1,000 Lotus Domino users, and about 1,000 online order entry users. These are the numbers that will be used when running what's called the current workload. Therefore the current workload is defined as running these three applications simultaneously with the user counts as defined.

SAS Peak Processing

Once the current workload was running on the unified storage configuration for all three environments, internal audit time came along for Acme. Auditors were brought on-site to perform a quarterly audit and they needed to perform a number of additional queries above and beyond the current workload.

The graph below demonstrates the performance of those SAS queries while the consolidated configuration was also running the order entry and Lotus Domino users simultaneously under the current workload.



This graph contains a line for the same query times that we experienced during our previous audit

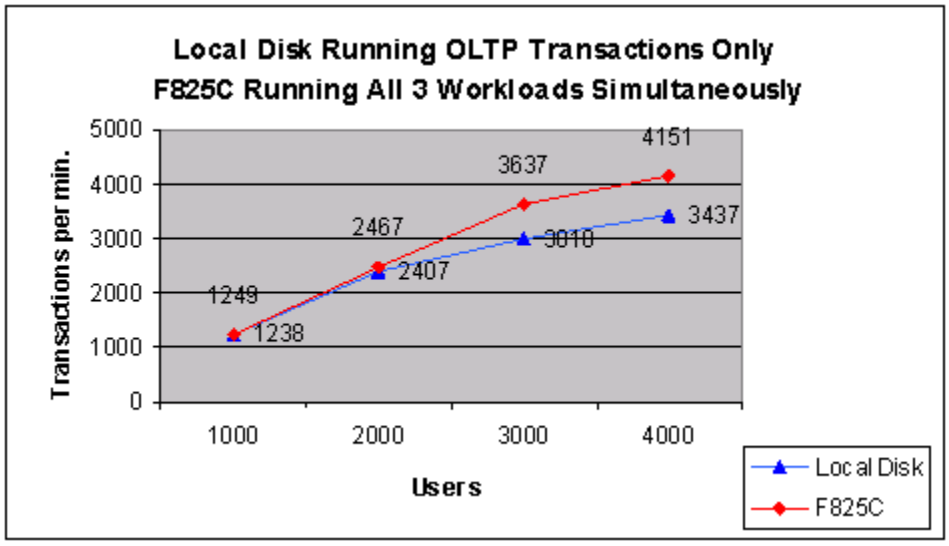
when we were using the direct-attached solution. Now that we've consolidated our environments, all of the disks available in this consolidated pool can be used for all workloads. The run times for the queries complete in less time on average for all the queries. As more and more simultaneous queries are placed during the audit, the consolidate solution utilized the performance of all the disks available in the pool to handle the peak processing demands. Remember that besides scaling the SAS queries, the unified storage solution is also running the OLTP and Lotus Domino workloads at the same time.

Since the unified storage LUNs can span a larger number of disks, it is only necessary for us to dedicate as much storage as needed for the SAS queries as is necessary to complete the jobs. As mentioned in an earlier section in this paper, this environment was forced to use more spindles due to performance requirements and not capacity needs. Therefore, we only configured 80GB of disk space for the 10 users. The direct-attached environment used 476GB of storage and still ran the queries slower.

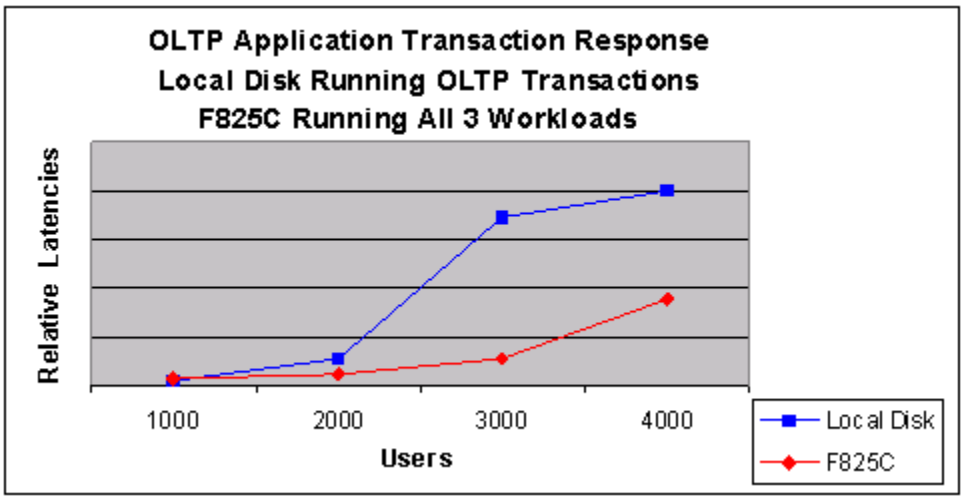
Almost all queries can now be completed in about an hour. The direct-attached environment had most queries taking between one and two hours to complete.

OLTP Peak Processing

The OLTP order entry environment is the most critical application for Acme. Let's see what happens when peak OLTP demands are made in this environment while running the current workload.



As you can see, the direct-attached and consolidated solutions scale about the same at the low end of the user counts. The performance at the low end is currently acceptable in the direct-attached environment. The problem lies with peak processing demands. As more and more users log in to place orders, the overall number of transactions per minute needs to scale too. Another problem with the direct-attached environment lies in the area of end-user response times too. See the following graph for a representation of those user times as more and more users log in.

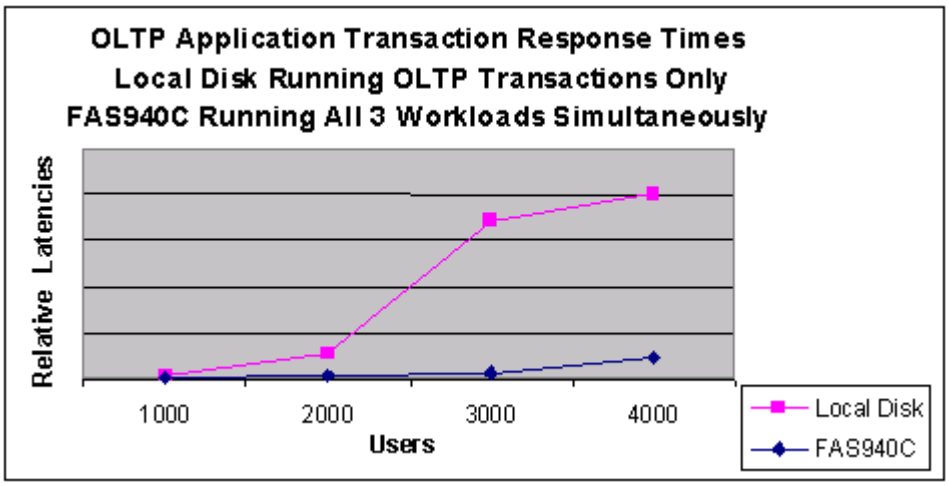


As more users log in to the direct-attached environment the response times go through the roof. The unified storage configuration, however, has acceptable scaling in this particular environment. The environment can handle requests for 4,000 users, but Acme made it clear that adequate response times are required in the long run.

OLTP Future Growth

Acme also feels they need a three- to five-year growth path plan in place. They are concerned with response times in the OLTP environment above the 3,000-user counts. The growth plans for Acme show the ability for the company to grow quickly to 3,000 OLTP users; however, end-user response times are critical and assurances of adequate response times must be demonstrated.

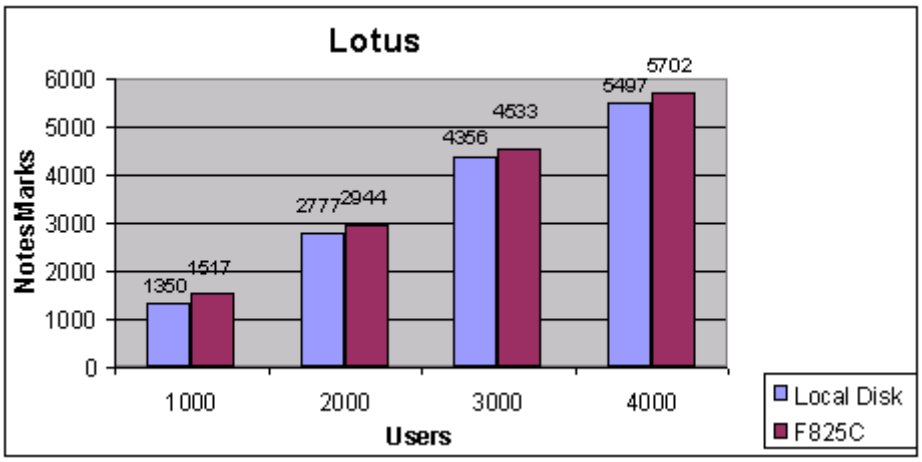
So another experiment was assembled and tested using the same workloads except on a FAS940C configuration. The chart below demonstrates the expected performance behavior using this higher-end unified storage solution with the same number of physical disks. Remember this environment was also running the SAS and Lotus Domino workloads at the same time, while the local disks were running just the OLTP application.



As you can see, the FAS940C is more than able to handle all three workloads and scale the OLTP users well beyond the 3,000-user levels. If the enterprise were to need this extra processing capacity, a head swap could be completed within minutes. It can continue to use its existing disks as they are configured. There is no need to move or migrate any data at any time.

Lotus Domino Environment

The e-mail application does not witness as dramatic an impact on performance since there is a constant flow of e-mail traffic across a fairly static user community. The following graph represents an increase in the number of Lotus Domino users while running the current workload for the OLTP application and SAS applications.



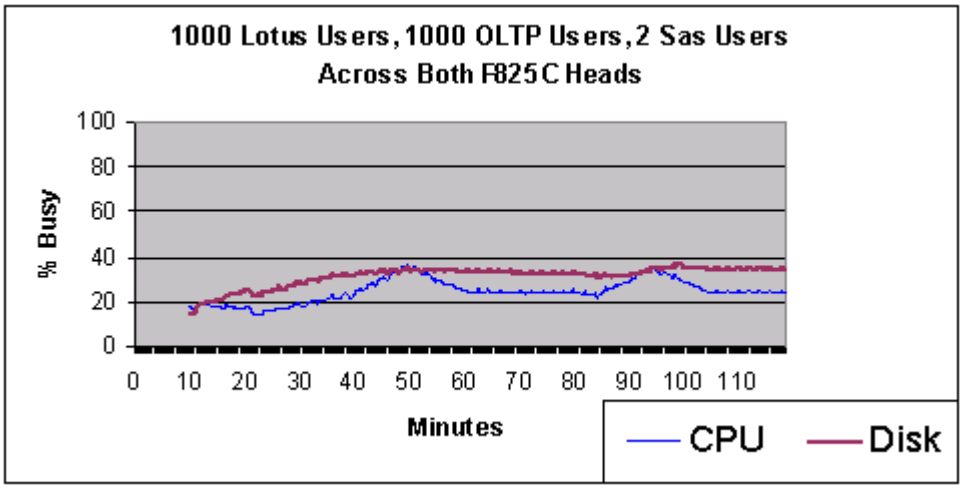
Obviously, in the real world, you would not typically see a dramatic increase in the number of users on a day-to-day basis. The direct-attached and unified storage solutions both scale extremely well given the current configurations. As the graph demonstrates, the unified storage solution consistently outperforms the direct-attached environment. The higher the NotesMarks rating the better.

The Lotus Domino users see excellent response times of less than one second from today's 1,000 users all the way up to 4,000 users in this environment. The processing times for the SAS queries and the OLTP application proceed at their normal pace, basically unaffected by the increase in Lotus Domino traffic.

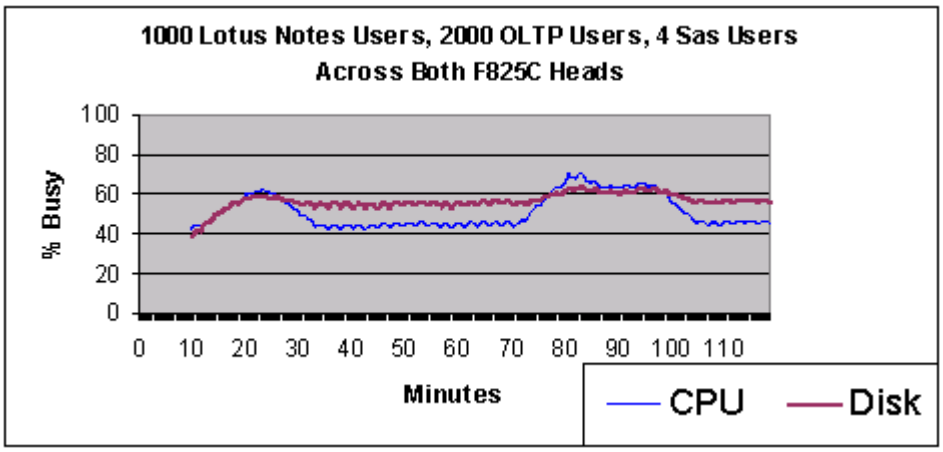
We used an approximate mailbox size of 120MB for each user. At this size, and using 4,000 users, the direct-attached environment would run out of storage capacity. The F825C solution will make it much easier to allocate free space to the e-mail system as needs arise.

6. F825C Performance and Capacity

The above data has demonstrated the clear performance advantages to using a unified storage solution in this particular environment. The next question is, "Just how much head room is there left in this configuration?" The following graphs reflect the CPU consumption and spindle utilizations measured during a time when the unified storage solution was running the current workload.



Here is a picture of the unified storage solutions CPU and disk utilizations during the course of running the current workload as described earlier. Since the load was quite equally balanced, these graphs were generated using the average of both heads in the cluster configuration. The average for both the CPU utilizations and the disk drive utilizations never exceeded 40% while running the current workload. So what would the unified storage solution look like during higher OLTP and higher SAS query demands while running the standard Lotus Domino workload?



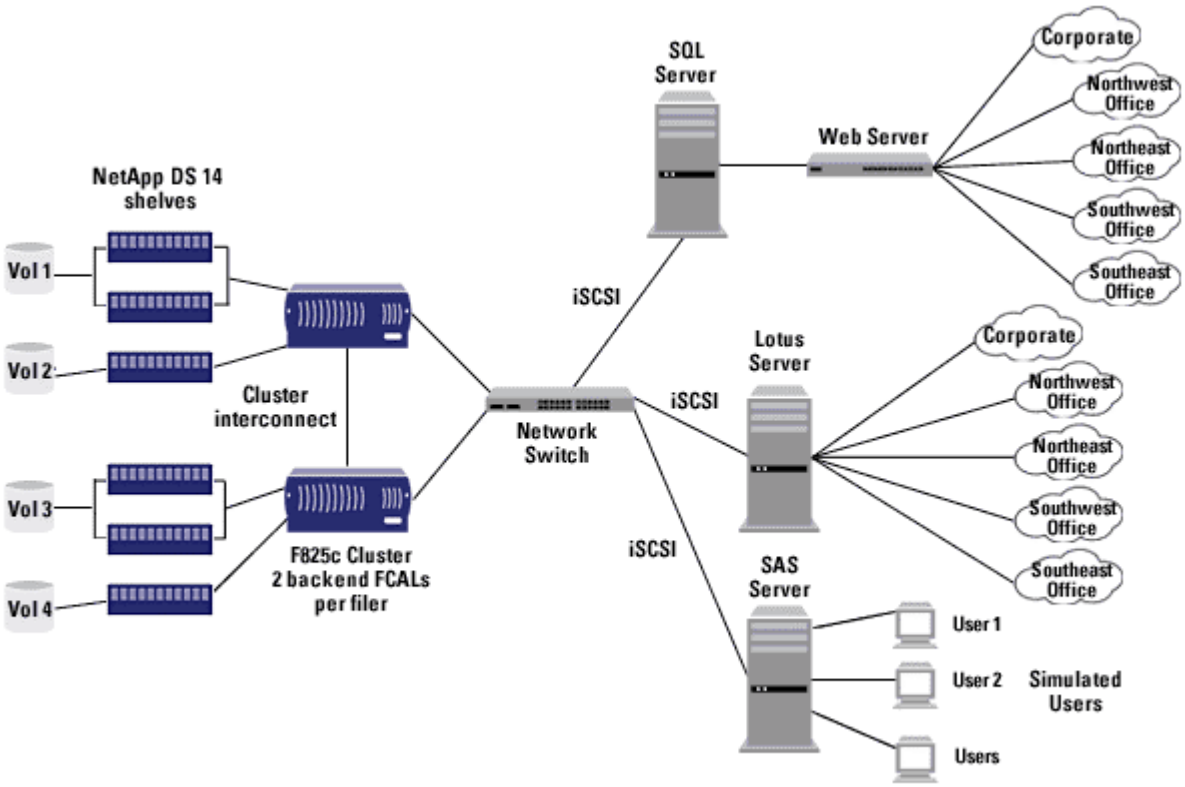
The Lotus Domino users remained constant at 1,000 users while there was an increase to 2,000 OLTP users and 4 SAS users. During this time, the average unified storage CPU utilization spiked to near 75% but averaged in the mid 40% to 50% range. The disk drive utilizations spiked to just above 60% and averaged around 58% during the near two-hour window.

The graphs do show fluctuations in CPU and disk utilizations during the course of the two-hour experiments. This was primarily due to the different phases of the SAS queries that were executed. The queries start out as very read-intensive as the input files are read. Once the where clause was satisfied and the reading was complete, the data was then sorted and written back out to disk.

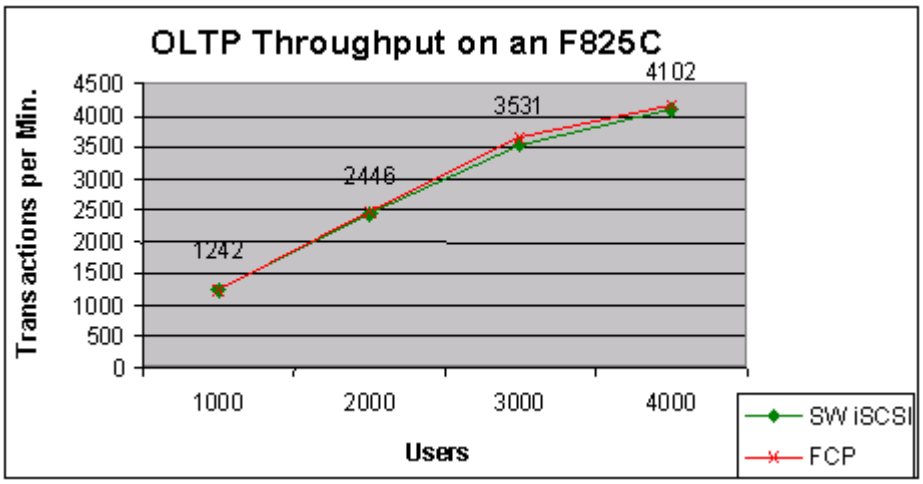
It should also be noted that the F825C configuration described here contained an extra 60GB of disk storage that could be allocated to any of the environments, while using eight fewer disk drives than the direct-attached environment. This was possible due to the ability to create LUNs that would span more disk drives and therefore offer higher performance while utilizing less disk drive capacity.

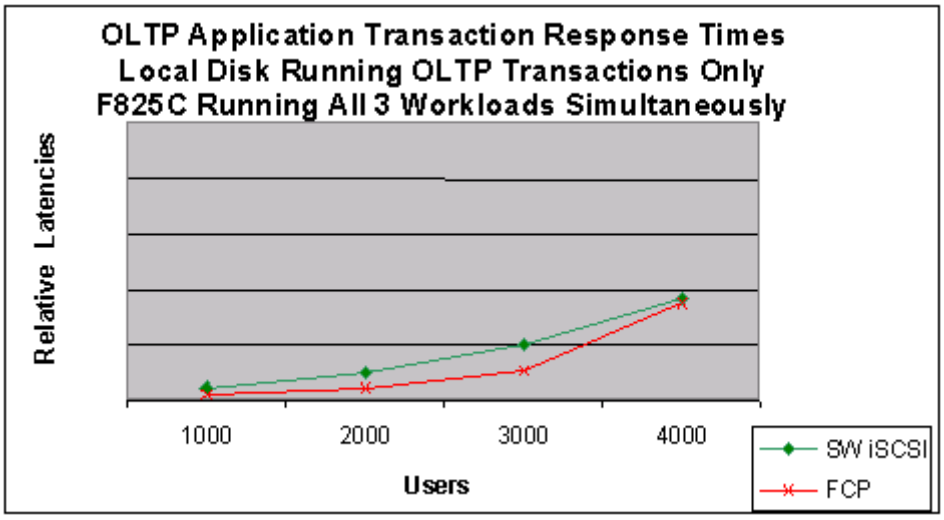
7. What about iSCSI Instead of Fibre Channel Attachment?

For those customers who would like to avoid the cost and complexity of a Fibre Channel infrastructure and like the idea of a unified storage solution, what would the performance look like if you introduced iSCSI into the configuration? The picture below gives you an idea of what a Microsoft iSCSI software initiator implementation would look like in this environment.



In the interest of keeping the overall content to a minimum in this paper, we've only included the OLTP graphs here for your reference. These graphs show what happens when we increase the OLTP demands while running the SAS and Lotus current workload. Similar performance was measured when we increased the workload in SAS and Lotus Domino environments.





As you can see, the performance levels between Fibre Channel-attached and iSCSI software initiator-attached environments are similar. Both environments scale well through all levels of performance. The iSCSI software initiator has slightly higher response times when compared to Fibre Channel.

The iSCSI software initiator implementation will give you similar performance levels at a reduced overall cost.

8. Conclusion

Having a dedicated, centralized, unified storage solution can provide a company with many advantages. Not only can you distribute the entire performance capacity across all parts of the business, you can do this in less time and with less maintenance. When you add capacity to the central storage pool, you add it to the entire enterprise, not to just one system.

A central pool of disks also allows an enterprise to ride through peak processing demands as the needs arise.

Direct-attached storage and/or Fibre Channel-attached storage are common deployments in major enterprises today. However, there are a number of businesses that just cannot afford a major Fibre Channel infrastructure. Not only is Fibre Channel costly to deploy, you also need the human resources with the technical expertise to configure and maintain that environment. iSCSI is a much more cost-effective approach to deploying a storage infrastructure. You can use your current Ethernet infrastructure and utilize the current human resources available to you while achieving similar levels of performance.

NetworkAppliance®
The evolution of storage™

Network Appliance, Inc.
495 East Java Drive
Sunnyvale, CA 94089
www.netapp.com