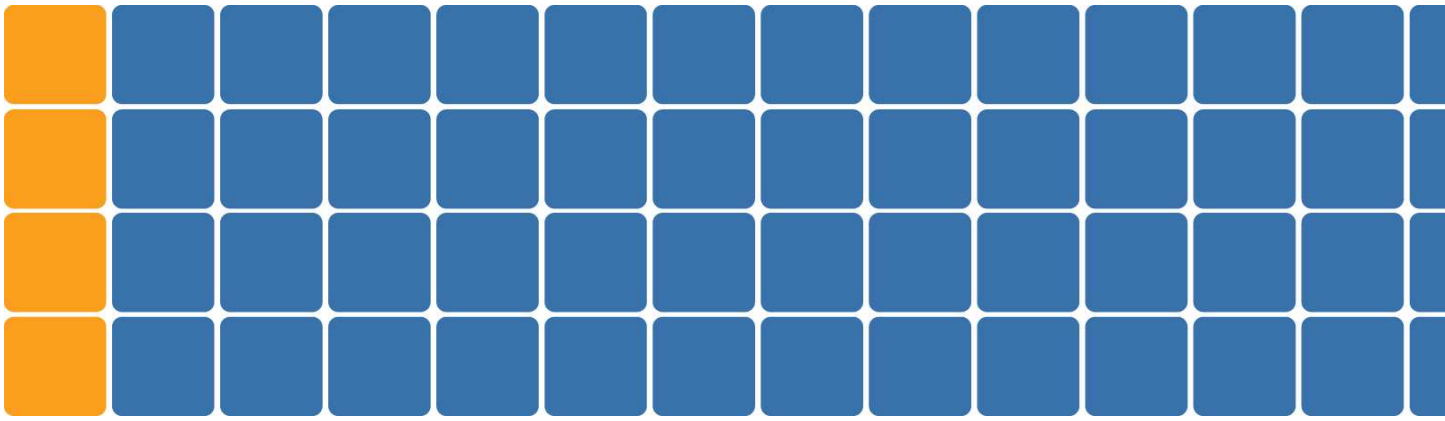
VMWare & CELEROS SANs



By Stephen Rieks
Syracuse University
02/20/07

VMWare & Celeros SANs

Objective

This document provides an overall understanding of how Storage Area Networks function and should help the reader make an informed decision concerning their implementation. This white paper contains functional information on Celeros XL Series SANs (Storage Area Networks), an introduction to their features and functionality, and use with VMWare ESX, GSX, and Windows 2003 server.

Executive Summary

In today's growing organizations, it seems as if everyone is being asked to do more with less staff, less training, and less time to get the job done. This is especially true in the world of information technology (IT). From the helpdesk staff member to the executive director, everyone feels the pinch. As a result, we are being forced to find, and sometimes create, out of the box solutions for problems on a daily basis.

With the increased amount of information and data becoming necessary for an organization to function, the importance of possessing a stable storage environment is critical. The costs of unscheduled downtime can run into tens of thousands of dollars per hour. Finding and matching technologies that best use capital and other limited resources is not always an easy task. Many IT departments have responded to growth by adding servers and storage systems dedicated to specific applications, business functions, and customers. SANs have emerged as the best solution for advanced storage requirements. SANs have been found to minimize some problems faced by IT staff and provide a more manageable, scalable and efficient method of storing, archiving, and replicating information and provide advanced functionality, as well as positive Return on Investment (ROI) vs. a Direct Attached Storage (DAS) environment.

There are an increasing number of SAN technologies on the market. However, all are not created equal. To realize the positive benefits of SAN deployment, the choice of vendor and SAN implementation partner is as critical as the SAN equipment itself. Centralized, efficient management, intelligent SAN services, a truly robust and flexible platform, and 365 x 24 x 7 service and support are requisites for next-generation SAN technology. The optimal SAN solution should have a robust, high performance architecture that creates new opportunities and alternatives while protecting resource investments from unexpected turns in the economic environment and changes in market adoption of new technology.

After reviewing many SAN solutions on the market, we selected the Celeros XL series line for its advanced features, good performance, onsite warranty, and attractive price. Utilizing the Celeros iSCSI (Internet Small Computer System Interface)SAN platform and VMWare ESX and GSX server, we tested various configurations with the Celeros XL series SAN including creating, expanding, and deleting volumes using both Qlogic iSCSI HBAs (Host Bus Adapters) and typical gig-e network cards using VMWare ESX, GSX, and Windows 2003 R2 server. Testing scenarios also included successfully remote booting VMWare ESX and 2003 server from the Celeros SAN. Finally, we devised and built a flexible, cost effective, and highly scalable application and storage environment that was the perfect solution for our IT needs.

Storage Basics

Organizations require effective ways to store and maintain growing quantities of data. Today's technology offers three basic storage options: Direct Attached Storage (DAS), Network Attached Storage (NAS) and Storage Area Networks (SANs).

© 2007 Celeros

Direct Attached Storage (DAS)
In its simplest form, DAS consists of a disk drive, or multiple disk drives, attached directly to a server. Data is typically transferred using Small Computer System Interface (SCSI), Intelligent Drive Electronics (IDE), or SATA commands. DAS is very easy to implement, but there are a number of limitations to the DAS approach including high cost of management, distance limitations and limited scalability. In particular, in order to increase storage capacity, an organization must purchase more servers with more drives. DAS is a local drive technology, designed to be attached within a few inches or feet of the server. These are some limitations have driven the need for network storage.

Network Attached Storage (NAS)
NAS is file-based storage architecture with resources attached directly to the LAN. Storage traffic is transmitted through files, the most common means of transferring data over LANs. Since it uses Ethernet technology, NAS resources can be managed by existing IT staff with minimal training in storage management, which may reduce IT costs. Another benefit of NAS is flexibility, since the storage unit(s) can easily be attached to the network. However, this is not a highly scalable option, since storage access is dependent on general network traffic. Increased network traffic may seriously decrease the performance of the LAN, and slow storage access can affect application performance.

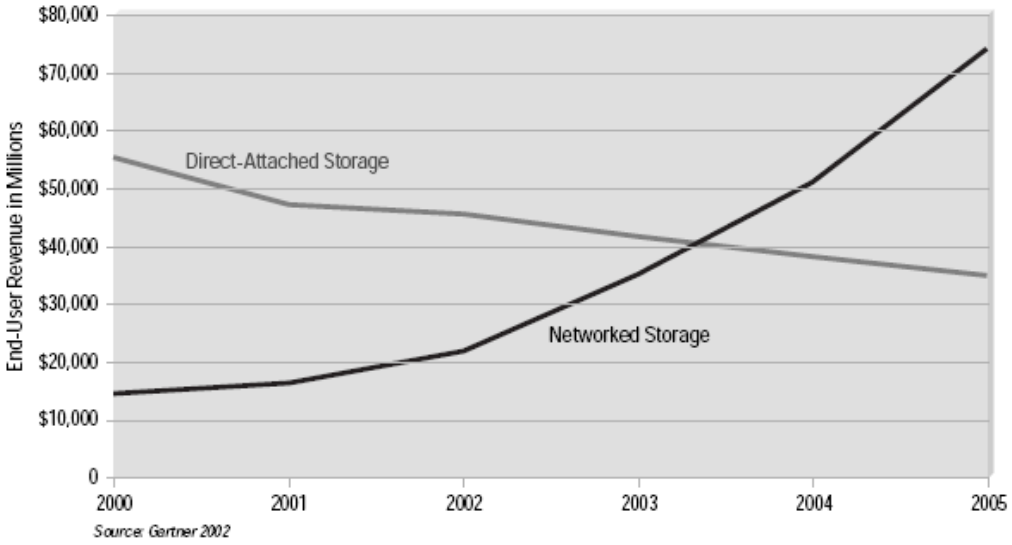
Other shortcoming of NAS technologies includes limited support for various name spaces and impractical methods in which they import, export, and sync's authentication information with your directory service platform. Our testing of a common NAS device on the market does not sync in real-time with our directory service database. It merely imports the list of user and group at some periodic interval that even their tech support staff could not answer. This makes timely file access permission modification impossible for new accounts.

Storage Area Networks (SANs)
SAN is a storage network that provides shared access to storage by multiple computers and servers (Jepsen, 2003). It typically uses an isolated dedicated network that connects servers to storage devices and transports storage traffic without burdening the enterprise LAN. Several factors such as scalability, superior protection for critical business information, and the ability to flexibly manage the available storage space are attractive features. In the absence of storage management tools, servers that host highly demanding IO applications often become overloaded while others remain relatively idle. SANs manage storage strategies can help balance data demands across servers and control storage costs. Growing storage demands can be met by simply reallocating existing storage or by installing more storage and network resources. The centralized data management provided by SANs can also help avoid redundant copies of identical data that consume disk space and can cause reconciliation problems.

Benefits of Networked Storage

The evolution of SANs has progressed to the point where many organizations can evaluate their role in their future storage needs. According to a report from Gartner Dataquest, the market for SAN-attached external storage in 2005 will exceed $22 billion, representing almost three million terabytes of data. The figure below includes all storage-related revenue (disk, software, services, networking equipment, etc.) and illustrates the increasing transition from DAS to SAN storage.

© 2007 Celeros

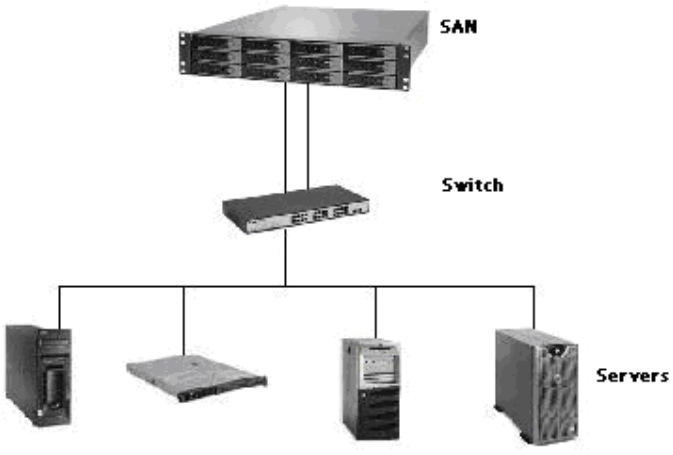


Changing to SAN storage

As small IT organizations feel this pressure, they look towards finding more efficient methods of serving their customers by keeping costs down, while keeping servers/services up and running. Some of the problems experienced were:

- Difficult management of storage at multiple locations. The 'islands of storage' approach makes it difficult for applications running on separate systems to share space.
- Loss of productivity when powering down to add new storage devices.
- Time restraints of scheduled backups for multiple systems.
- Inefficient levels of storage. For example one processor may run out of storage while another processor may have unused storage space that cannot be made available to the processor that requires it.
- Confusion over the various storage technology alternatives, including Fiber Channel, Internet Small Computer Systems Interface (iSCSI), Fiber Channel over IP (FCIP), and Infiniband

The move away from DAS onto an iSCSI storage area network SAN device was imminent. DAS is regarded as the most straight forward method of delivering storage capacity to a server. The storage input/output (I/O) subsystem is directly attached to the server either internally via ribbon cables or externally via external SCSI cables and an external storage array. Using this scheme, only one server may access the storage subsystem. In comparison, the SAN removes the storage I/O subsystem from the server and places it an external array where one or more servers can access the storage.



While DAS storage is a time-tested traditional method of creating storage, it does add considerable costs to a new server installation.

There are also limitations on what tasks can be done on the server and is regarded an inefficient model due to how much of the storage space actually gets used during the server's life. Working with these restrictions were becoming costly and time consuming, hence our decision to move to a SAN environment. Finding a suitable product was a time consuming process. Based on our research we narrowed our search to iSCSI devices because they best matched our business requirements, allocated budget, and I/O requirements.

Some of the vendors we considered were Promise, NEXTSAN, and EqualLogic, and Celeros. Celeros offered the best combination of features, performance, service, and support. We found the Celeros XL series SAN very easy to use. The web interface was straight forward and clear. The features are organized via logical tabs across the top of the screen. After clicking through the tabs and becoming familiar with the web interface, we found ourselves creating volumes and attaching them to targets within minutes.

Choosing the Right Vendor

Some attractive features offered by the Celeros XL series SAN include:

- Dynamic volume expansion. Celeros XL23, allowed us to allocate the amount of space needed by a server at the specific time and as the server required additional storage. It dynamically expanded the SAN target volume and the server's logical volume without ANY downtime.

- Variable bandwidth SAN to SAN replication. The SAN to SAN replication feature of the Celeros XL23 allowed us to modify our disaster recovery (DR) and business continuity (BC) plans. Previous to the SAN, we were creating weekly server images to external drives and used host based mirroring tools such as 'robocopy', Taskmaster, and double-take to sync data from one system to another. The SAN to SAN replication feature allowed us to revamp this process and reduce hardware and licensing costs.

- Creation of multiple snapshots per volume. The volume snapshot feature was an attractive alternative to tape restores but also allowed us to greatly expand our backup window. Unlike other divisions within the campus that have the luxury of an entire weekend to conduct full backups, as a 24x7 operation there is a very limited backup window. The Celeros snapshot feature allowed us to take snapshot our servers: directly mount the snapshot on the backup server; and backup the volumes during normal production times.

These features are generally found on much more expensive iSCSI SANs solutions.

Server Virtualization

As with many small to mid-sized IT departments, we had a number of servers that did not require large amounts of disk space, memory, or high-end CPU's. Utility based servers, patch management servers, remote access workstations, and small database servers were just a few areas in which we believed server virtualization made sense. From the management perspective the use of virtualization has many benefits such as the possibility to dynamically change the size of a virtual volume; move storage capacity from one volume to another or install new physical devices without disrupting system operations. This may be performed by a software component such as Volume Manager, or a hardware component such as a RAID controller.

To create a virtualized environment we purchased a copy of GSX server, 2003 standard edition, a dual processor IBM xSeries 336 server, 4 GB of RAM and a small DAS RAID array to begin our move to virtualization. Using various tools, we either migrated or recreated many of our low powered physical hardware to virtual servers. All works well. We were able to reduce the amount of hardware or metal in the server room which also equated to less power consumption, cooling, rack space, and miscellaneous items such as keyboard video mouse (KVM) switches and server monitors. We found that

our utility servers ran faster due to shared faster hardware. The implementation was seamless and our users were clueless to any backend operations.

One advantage of server virtualization is the ability to create 'template' servers that can be used to quickly create new guest operating system (OS's). Within 30 minutes, we were able bring up a new 2003 server with just a SID, IP, name change, and license purchase compared to the initial process that took 2-3 weeks to purchase new hardware, unboxed, configure, manually install the OS, patch, etc. We were able to run utilities such as PowerQuest's V2i or Symantec's Live State Recovery on a GSX host OS and create a disaster recovery(DR) or business continuity (BC) image of all guest OS'. These guest OS' could then be transported via SAN replication or USB drive transportation to our DR site and launched on another GSX server.

While all worked well, it was discovered that the maintenance cycles of the host and guest systems quickly synchronized and host patch management effect not only the host OS, but also the guest OS'. Other disadvantages also included memory allocation via GSX, and disk I/O from one guest OS affected others. However, from our perspective the pros of virtualization far outweighed the cons.

These shortcomings can be regarded as small and manageable for some static organizations. Further investigations showed that ESX does not suffer some of the problems encountered with GSX. ESX allocates memory to its guest OS' differently than GSX and does not require as much system maintenance. Together with ESX 3.00 support for iSCSI, we began experimenting with ESX with plans to move away from GSX and fully onto ESX. When this was done, we needed to find a way of reproducing the same DR/BC functionality as we had with GSX.

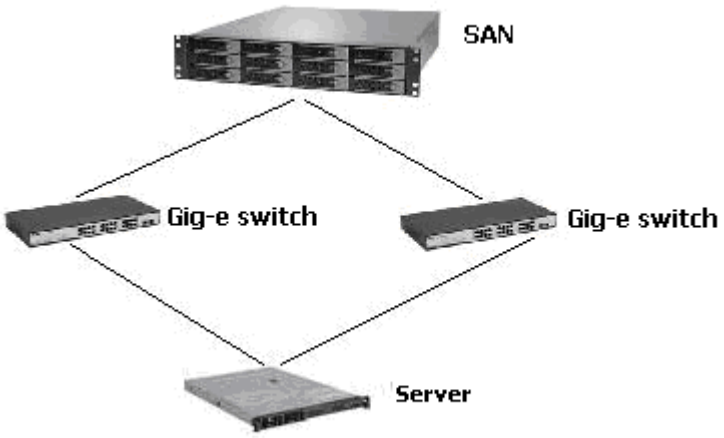
Using Celeros' volume snapshot feature a "volume snapshot" can be either manually created via the browser based SAN management console or scheduled to "snap" automatically. The advantage of SAN based snapshots is that they are independent of the OS. Not only can the data volumes be snapped, but the entire server can be snapped and backed up to tape for BC purposes or replicated off-site.

SAN Performance, Disaster Recovery and Business Continuity

Disaster recovery (DR) and business continuity (BC) are not new problems for IT organizations. Well managed shops have had well tested, well documented DR and BC plans for many years. SANs and server virtualization have only enhanced and added to the methods in which IT shops can protect its valuable systems and data. Such methods include both asynchronous and synchronous data replication solutions. The advantages to iSCSI SANs from vendors such as Celeros is that they can replicate data to a backup SAN via the existing gig-e network without the need to add any additional infrastructure. The Celeros XL23 SAN offers two Gig-E ports that can be split, one for a dedicated iSCSI fabric, and one dedicated to data replication. Using this methodology, one can maintain the high speed transfer rate that Gig-E offers (1000 Megabit / second = 125 Megabyte per second transfer rates) while delivering a near real-time off-site DR / BC copy of your entire environment via slower network links.

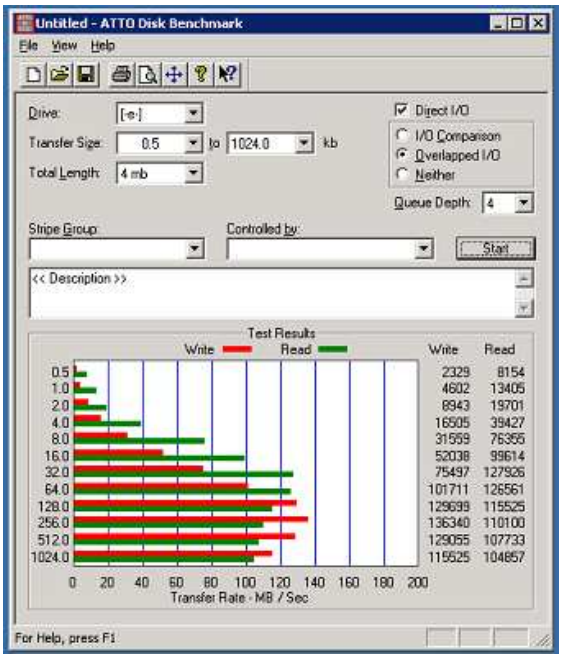
A disaster does not need to be as large as 9/11 or Katrina for your environment to suffer. It can be as small as a localized power outage, a network disruption or a failed system component. One of the things we found comforting about the Celeros XL series SANs is the ability to create a multi-path iSCSI SAN environment using 2003 server. Currently VMWare ESX does not support active-active multi-pathing and iSCSI. With each port of the XL23 connected to a separate gig-e switch, we were able to build a totally redundant iSCSI fabric. One 'glitch' with this configuration was that we needed to purchase a dedicated Qlogic controller. The Microsoft iSCSI initiator, v2.0.2, will only allow you to select a single target IP address when using iSCSI software emulation and as such, you can only connect to the SAN from one channel. Only with hardware HBA such as a Qlogic 4052 were we able to create redundant paths to our SAN. This configuration allowed for full path fault tolerance and gave us a 2 Gigabit connection to our SAN.

© 2007 Celeros



Multi-path Configuration

Performance was good. We used bench32 v.2.3.4. from ATTO technology to conduct basic performance tests. When experimenting with different configurations, we found bench32 a good, quick tool. Seen below are test results from the multi-path environment above.



Basic Performance Tests via Bench32
V2.3.4 from ATTO Technologies.

SAN Configuration

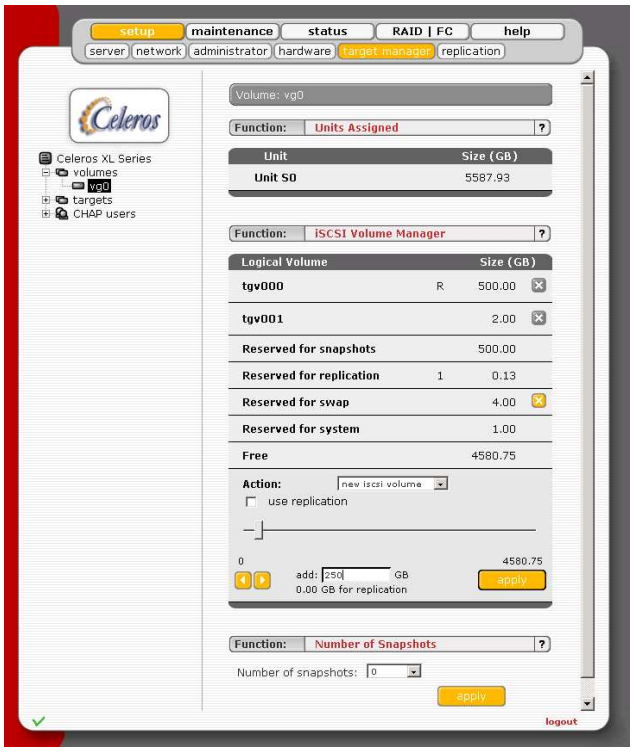
We purchased two identical XL23 SANs, both with 12 750 GB drives, dual dual-core 1.6 GHZ CPU's, dual gig-e network cards, and 1 GB of cache memory on the RAID controller. The SAN was configured with

RAID 6 and two hot spare disks. With 12 drives to work with, we ended up with 5.6 TB of usable storage. See table 1 below

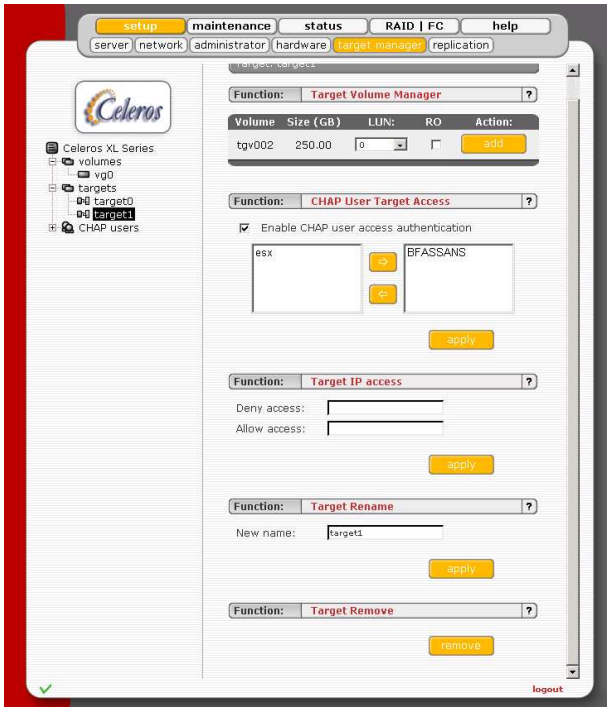
12 Drives total x 750 GB per drive	9 TB total
2 Drives set aside as hot spares	- 1.5 TB
10 Drives remaining, RAID 6 created	5.6 TB RAID Array

Table 1: SAN RAID Array Configuration

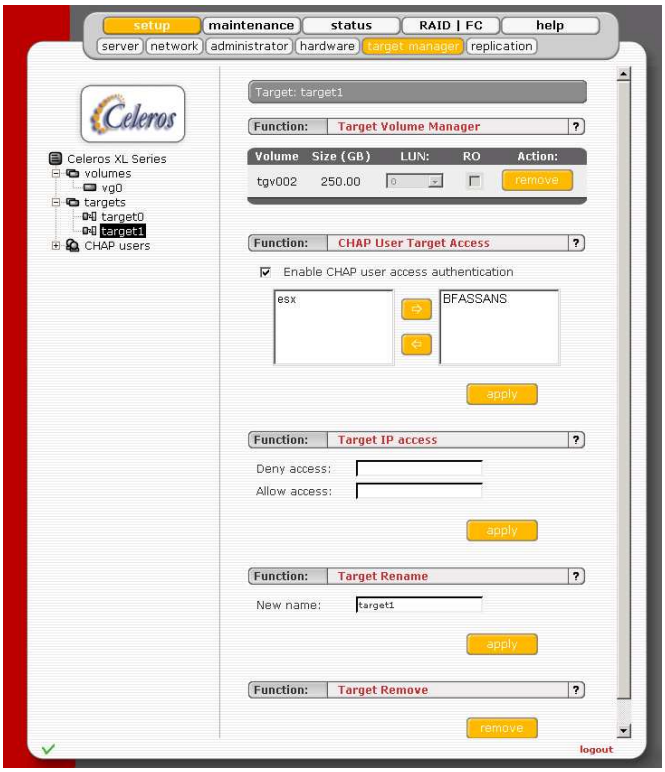
There are 4 basic steps required to carve out SAN space, assign it to a LUN and deliver it to your target server. These steps are shown below.



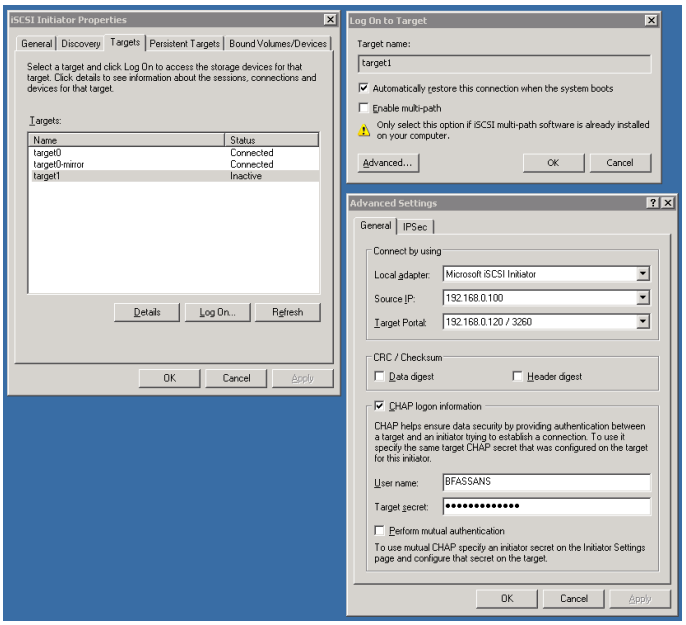
Step 1 (see figure above) is to create a target volume on the SAN. The figure shows a 250GB target volume.



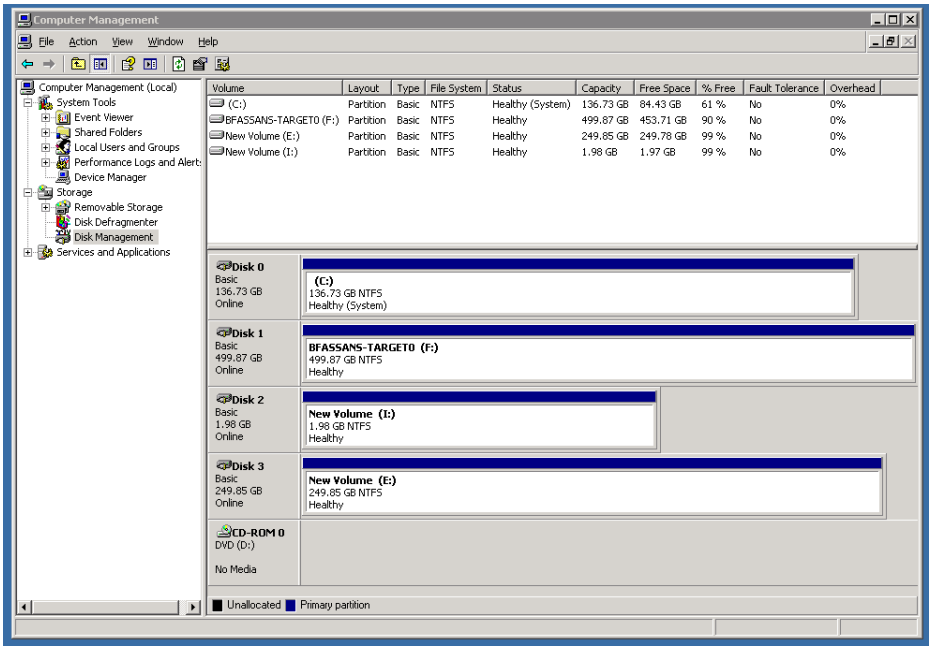
Step 2 (see figure above) is to create a target disk with either Challenge Handshake Authentication Protocol (CHAP) access or IP restrictions. You should use either CHAP access or IP restrictions to limit what servers can see which logical unit number (LUNS).



Step 3 (see figure above) is to attach the target volume (250GB volume created in step 1) to the target disk, step 2. You can allow the target full read/write (RW) access to the target volume or Read-only access depending on how you plan on using it.



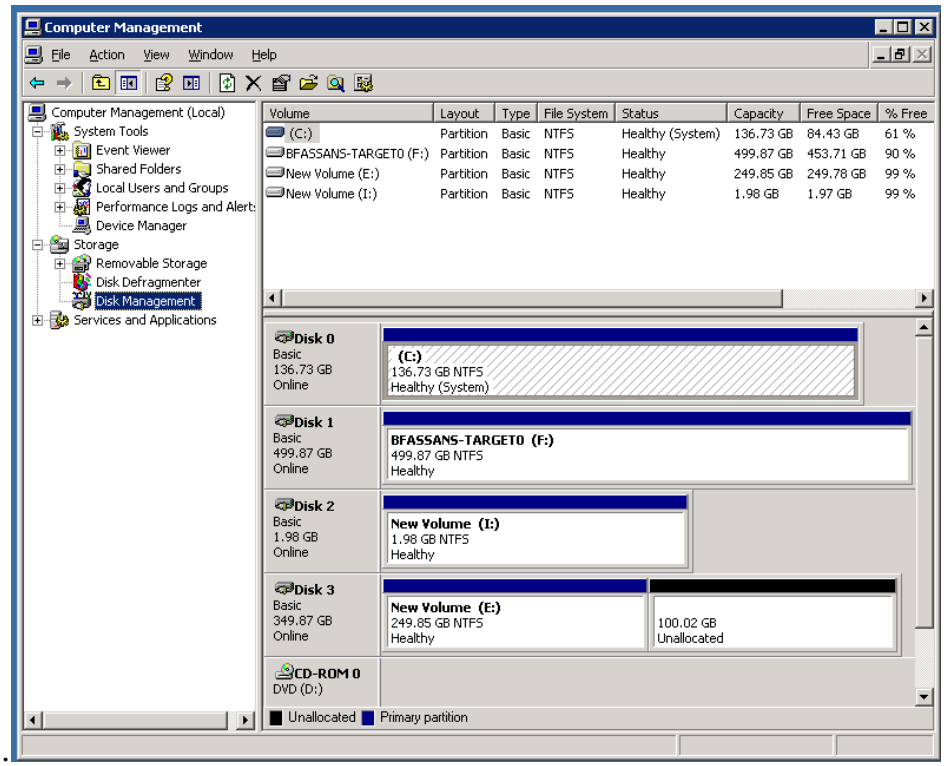
Finally, step 4 (see figure above) is to attach the target volume created in step 3 to your server or workstation. This is dependant on the host OS you are using. In this example, the 250 GB volume (target 1) is connected to a MS 2003 server using MS iSCSI initiator. Press OK in the lower right window to connect to target1.



After connecting to target1, the Disk Management tool was used to initialize and format the 250GB E: drive.

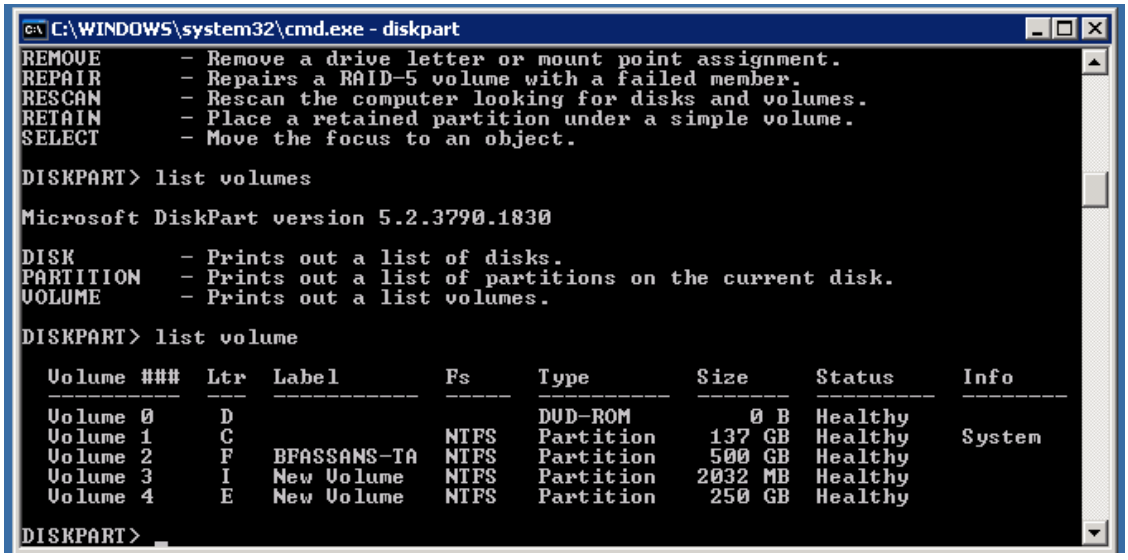
That's it. Now you have a 250 GB volume on our 2003 server. To illustrate how to dynamically expand a volume, we again used a Microsoft 2003 server. The 250GB volume was expanded by 100GB to 350GB using the Celeros web management console.

As soon as this one step was completed, the disk management utility on 2003 server saw the additional 100GB of space.

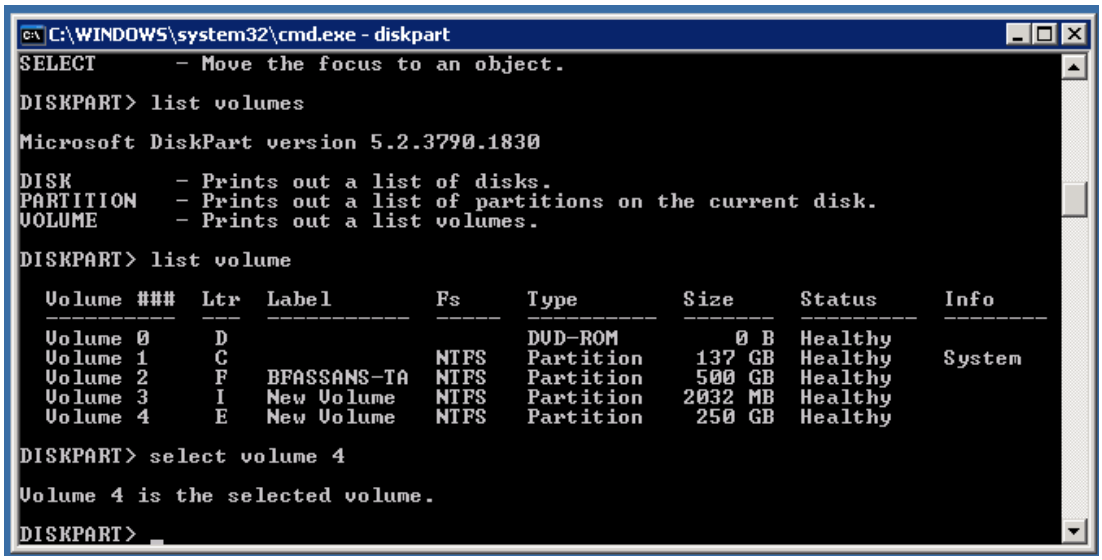


At this point one can elect to do either. Create a new volume from the 100 GB of space or extend the 250GB volume to 350GB. To extend the volume, there are 3 steps that need to be completed.

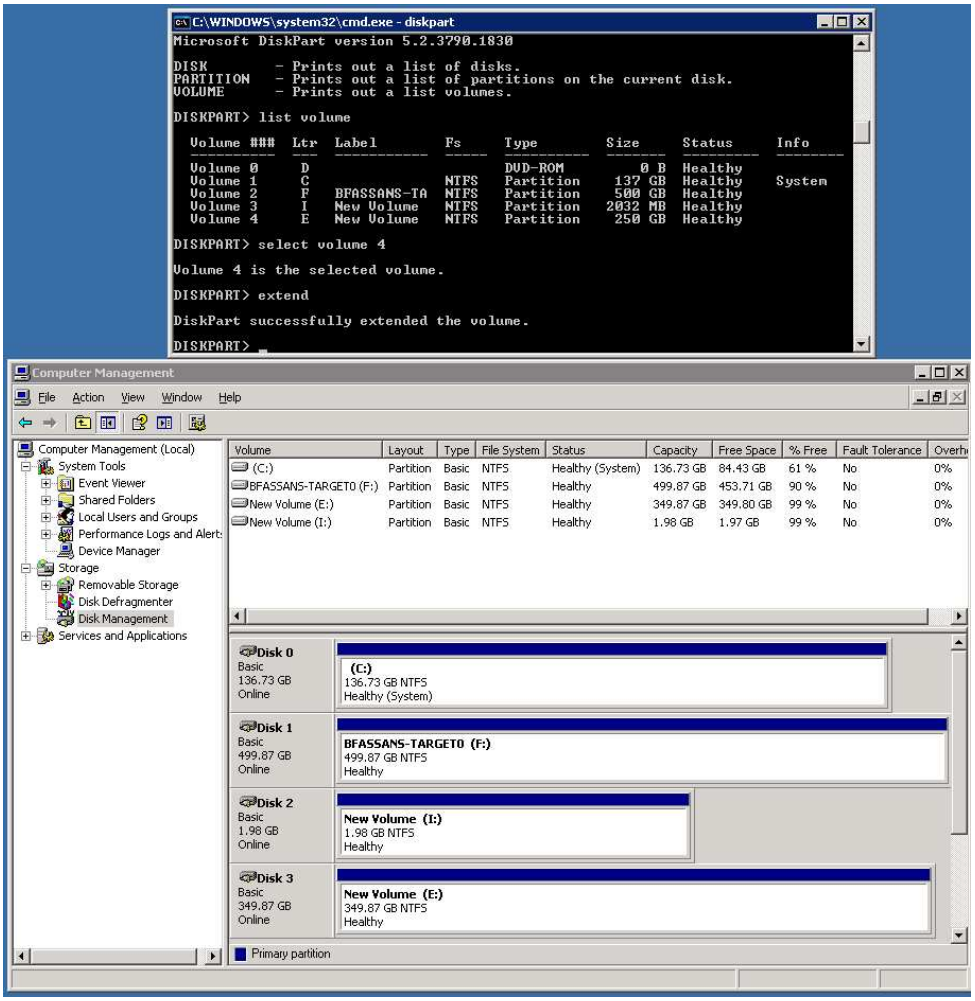
Step 1 is to open a command prompt *(cmd*) and type *diskpart*. This command will launch the diskpart utility and give you a *DISKPART*> prompt. At the prompt, type *list volume*. This will display all volumes in the system as seen below.



Next, select the volume that you wish to expand, in this case volume 4. You select the volume by typing the *select volume 4* at the DISKPART> prompt.

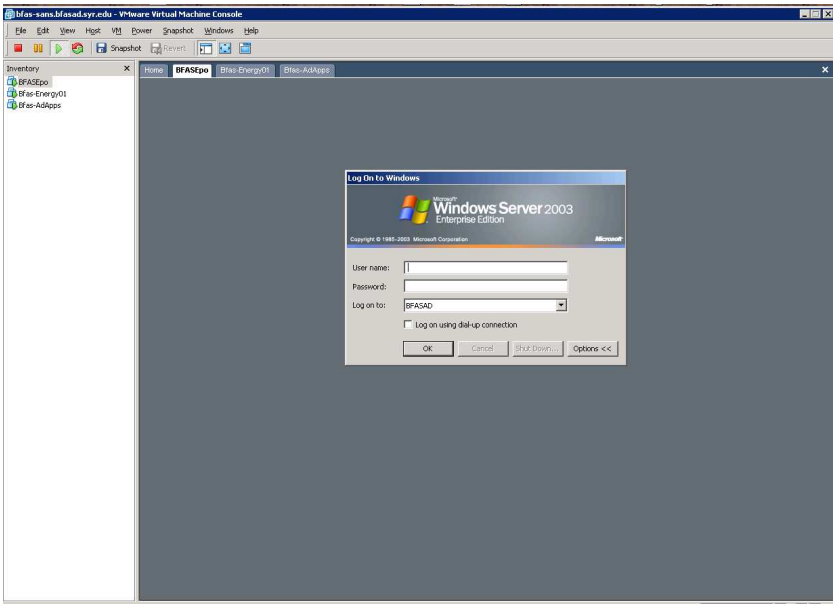


Finally, type: *extend* to extend the volume.



GSX Server / VMWare Server testing

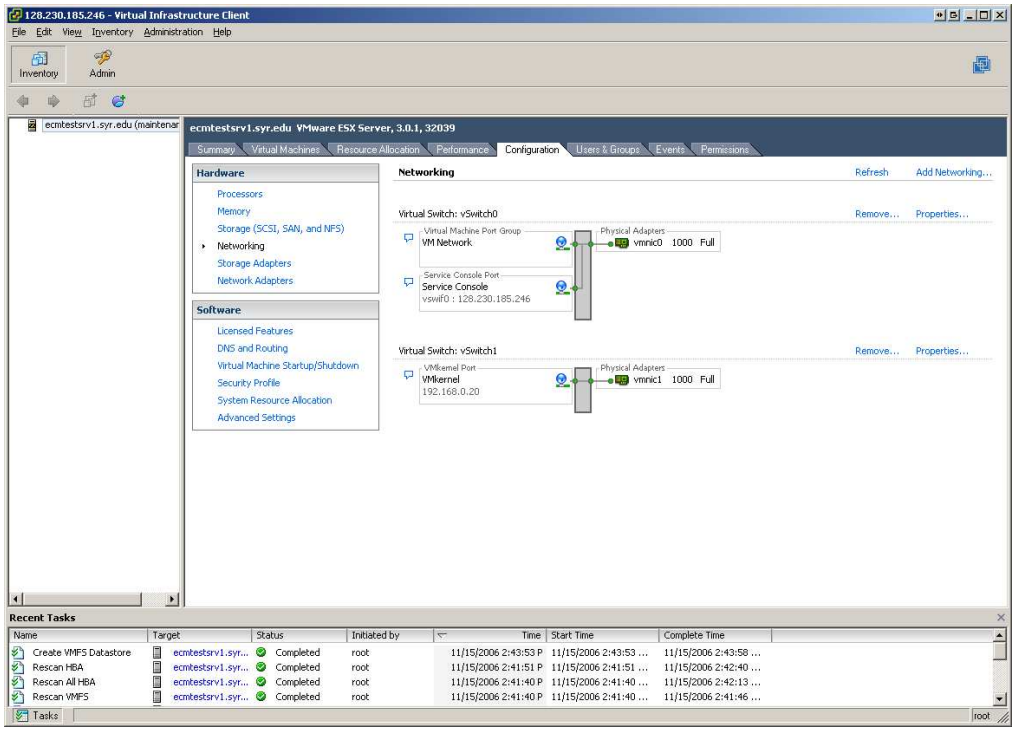
VMWare GSX server and VMWare server run on top of a host operating system. As a result, whatever the host can access in terms of storage is available to GSX. In our testing, we configured 3 virtual servers using GSX server and placed them on the Celeros SAN. Performance was good and everything worked as we wanted. As stated above, we were able to use Live State Recovery and create a host OS disk image that included the 3 guests OS'.



ESX Server Testing

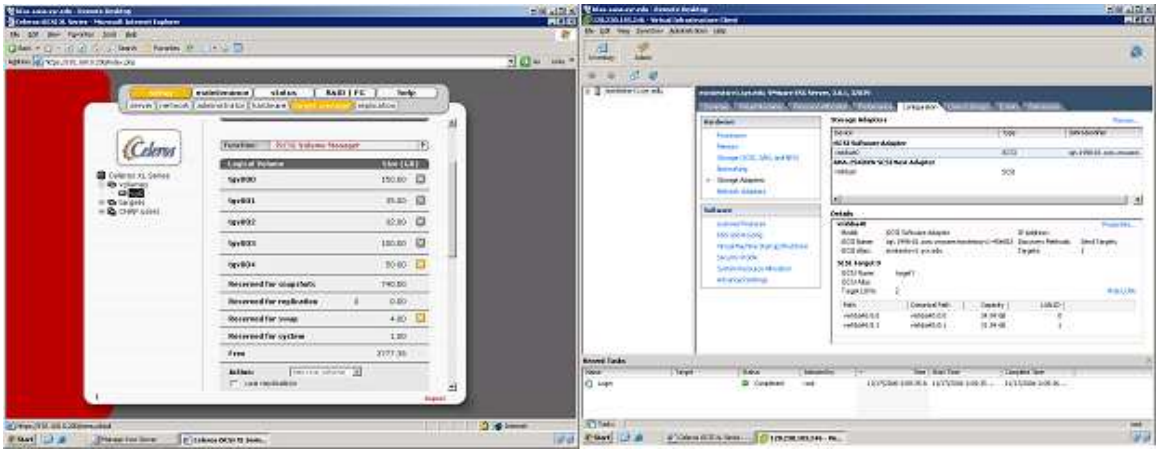
ESX server 3.00 and 3.01 supports iSCSI. One of the server administrators wanted to expand his knowledge of virtualization products by experimenting with ESX. I thought he could bring up ESX with the 4010c and I would work with ESX's software emulation support for iSCSI. (ESX 3.01 fully supports the Qlogic 4010c and recently began supporting Qlogic 4050 and 4052 HBA.) Both server builds were done at the same time. About an hour later, there a knock on my office door and I faced with a smiling server admin. His ESX box was up, and he could see the SAN. He enquired about mine server build. I muttered a few choice words about ESX and setting up iSCSI software emulation, and then got back to work. After a couple more hours, I finally got ESX to see the SAN. Yes, you can connect to an iSCSI SAN via ESX and off-the-shelf gig-e network cards, but we found it much easier to use the Qlogic HBA controllers. I will admit though that I'm not one to read the manuals or attend training sessions. I imagine if I had, things would have gone much smoother.

After my initial issues with connectivity, all went smooth and performance was quite good. Below illustrates my final network configuration using ESX and iSCSI software emulation.



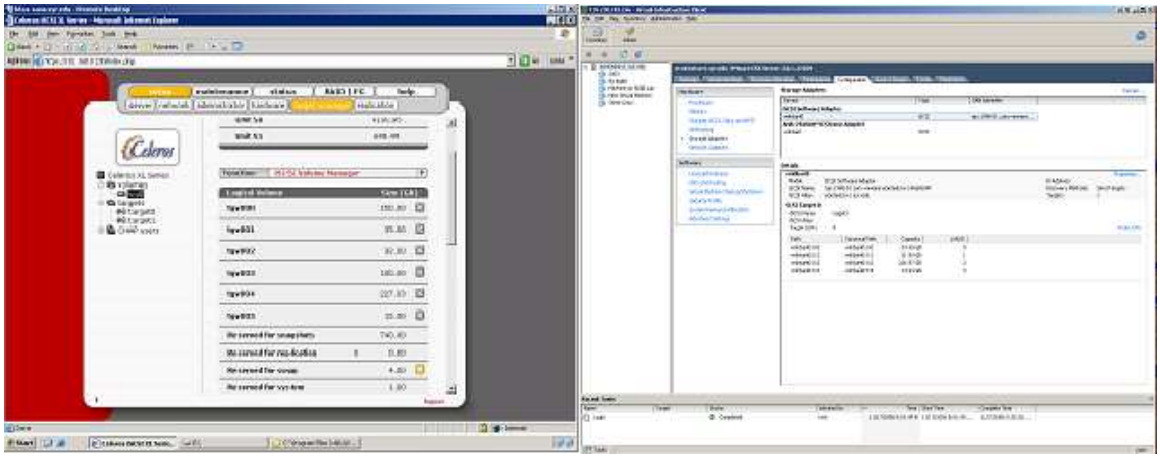
ESX 3.01 and iSCSI software emulation

SAN testing with ESX went smoothly. We didn't experience a single issue with the Celeros SAN and ESX. The only issue focused on our inexperience with ESX but after having to search online for some answers, I was able to get everything working. I created two target volumes, a 32 GB and a 35 GB on the SAN and presented them to ESX. See below



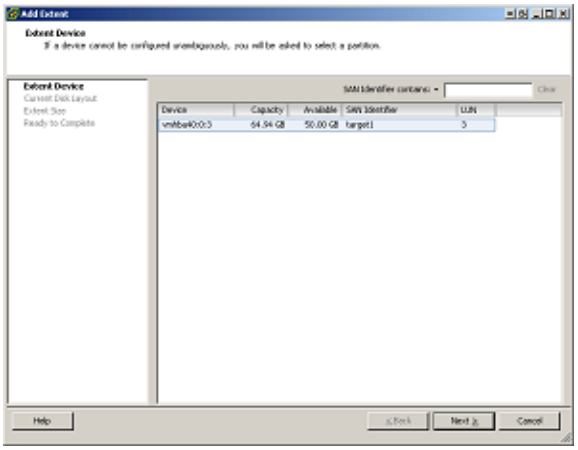
Two target volumes on a
Celeros XL23 presented to ESX 3.01

We thought we'd try utilizing some of the advanced features of the SAN and create and expand a volume. We created a 15 GB volume on the SAN and presented it to ESX as seen below.

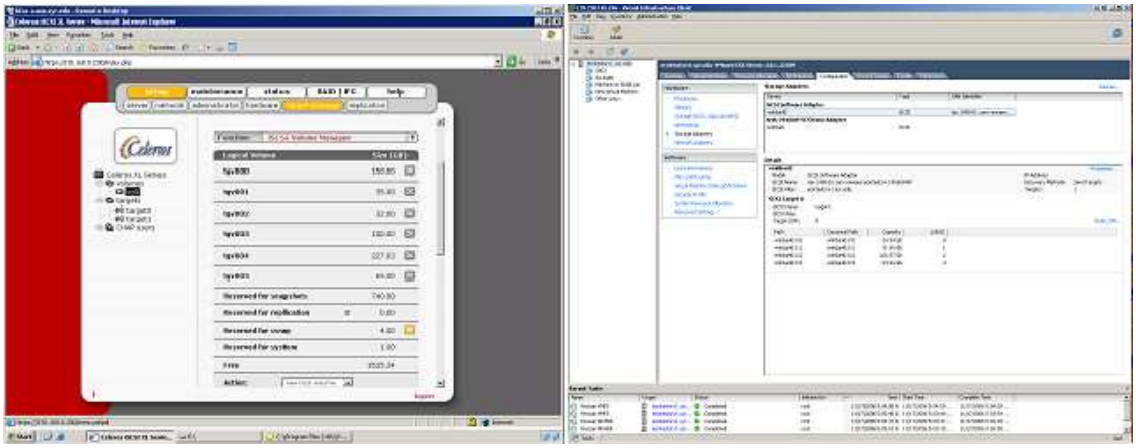


15 GB Target Volume Create - Preparing For Volume Expansion Test

Next, we created a Microsoft 2003 server on this volume and set the C drive to the full 15 GB. We booted the Microsoft 2003 server and verified proper operation. Upon shutdown, I expanded the 15 GB volume to a total of 65 GB using the SAN configuration utility. Next, we connected to the ESX console and began the volume extension process. This involved right clicking on the 15 GB target volume and selecting properties as seen below.



Using the wizard, we were able to property extend the volume to 65 GB. See below.



15GB Volume Expanded to 65GB

Snapshotting was also as easy process. We conducted a manual snapshot of one of the ESX server volumes then added it to a target volume on another ESX server. This process did need to go into the advanced settings on ESX and enable logical volume manager (LVM) resignature to see my snapped volume. We were successfully able to mount the snapped volume created on ESX server 1 and 2.

iSCSI Booting:

Booting from the Celeros XL23 was a straight forward process. We created a target volume on the Celeros SAN, presented it to the IP address of the Qlogic controller and properly configured CHAP access. Again, the Qlogic controller had to be used due to the onboard BIOS on the Qlogic card. To test this functionality, we downloaded the proper Qlogic drivers from qlogic.com and began the Microsoft 2003 server installation. When prompted for additional boot device drivers, I depressed the F6 key and popped my floppy in the drive. The Microsoft 2003 server saw my 4010c driver and proceeded to present my 50 GB volume to the OS. The rest of the server installation was done in a text book fashion. Feeling great about how easy it was to get Microsoft 2003 server to boot from the SAN, we proceeded to test ESX SAN booting. We were not disappointed. The test server booted from my ESX 3.01 evaluation CD, saw the 4010c card, loaded the driver, and continued with the installation.

Booting either Microsoft 2003 server or ESX off a SAN does offer additional advantages vs. local DAS booting. In theory, you can change your server's hardware vendor, transfer your Qlogic HBA to the new hardware, and mount and boot the OS. Yes, you will need to update some hardware drivers, but at least your RAID controller (qlogic card) should be able to boot your OS enough so you can update. The disadvantages are that you need at least one iSCSI HBA with onboard BIOS and not every OS you wish to run may support SAN booting.

Many people that I've communicated with who boot their servers (ESX, Microsoft 2003 server, etc) environment from DAS vs. SAN seem to prefer DAS booting and using the SAN for additional storage. Costs are about the same either way. A pair of 36 GB SCSI and a RAID 1 card will not cost much more than an iSCSI HBA with onboard BIOS and troubleshooting why your server will not boot is, in theory, a simpler, more straight forward process and directly supported by one vendor.

With DAS booting, you again need to be concerned with drive failures, RAID array failures, firmware updates, compatibilities, and additional system monitoring (some of the advantages of using the SAN in the first place). But if you DAS boot, you can use low cost, TCP/IP Offload engine (TOE) offloading gig-e cards from vendors such as Intel for your SAN connectivity.

Performance Tuning

Tuning your environment is very important. One song that I constantly sing to my staff is to "ALWAYS update the BIOS, firmware, drivers, and patches to the latest and greatest version". By doing so, you may circumvent a problem before you even get to it. Test your design in the lab before you put it into production. You can design the best system on paper, but due to issues that you might not be aware of, your design may not deliver optimal performance.

As described above, I created one large array RAID 6 array. I found this the easiest and a more flexible arrangement; but it does not deliver the most IO performance. If you are looking to squeeze every ounce of speed out of your SAN environment, you will want to create multiple RAID arrays and deliver each array to a separate server. This way the IO of one server and the spindles it uses will not affect the others but if you are planning to do this, you will need to plan carefully how you plan on using your SAN. If you need to reconfigure your SAN at a later time, you may need to delete one or more arrays and start again.

If at all possible, you should build a separate iSCSI network for your SAN environment. Yes, you can connect your SAN to the existing gig-e network and mount its drives from another building but you may pay a performance penalty by doing so plus you are competing with normal network traffic. As applications begin using more and more of a gigabyte connection or others begin using the bandwidth for video streaming, MP3 downloading, or whatever, if you elect to put your SAN on a normal communications network, be prepared for the day when you will need to move it. One case in which you may wish to put one leg of your SAN on a corporate network is when you are conducting SAN to SAN replication. With the Celeros XL series SAN, you can limit the amount of bandwidth you allocate for replication. We have limited the amount of bandwidth to 30GBs. We did some research and discovered that most of the time, our existing network has this much free bandwidth between our primary and disaster recovery site. As a result, I was able to use this extra capacity without causing anyone; including myself, any headaches.

In conclusion, the Celeros XL series SAN is an excellent value for the money and offers high end features with a reasonable price tag. Combine it with VMWare ESX, GSX, or virtual server, and you have the major tools you need to build a state-of-the-art server environment without breaking your budget.

About the author

Stephen Rieks is an Associate Director of Systems, Services, and Application Development for the Business, Finance and Administrative Services of Syracuse University. He is an adjunct professor at the School of Information Studies at Syracuse University where he teaches courses on network operating systems administration and configuration. He also owns his own consulting firm. Stephen can be reached via email at sjrieks@syr.edu.

Sources

"2001 Disk Storage Systems Worldwide", Gartner 2002
"Slaying the Storage Beast"; Forrester Research; March 2001
http://www.cisco.com/global/EMEA/ciscoitatwork/pdf/Cisco_IT_Ops_Practices_Storage_Utilization.pdf
http://www.cisco.com/en/US/netsol/ns340/ns394/ns259/ns261/networking_solutions_white_paper09186a00800c464f.shtml
Jepsen, Thomas C. 2003. Distributed storage networks : architecture, protocols and management. Chichester : Wiley.