

InfoWorld

October 3, 2005 ■ Issue 40

GET TECHNOLOGY RIGHT®

TEST CENTER SPECIAL REPORT

iSCSI SANs Unleashed



(EXCERPTED)

BY PAUL VENEZIA

Solid solutions from Adaptec, Celeros, EqualLogic, Intransa, NetApp, and Rasient show that iSCSI is ready to push Fibre Channel off the low ground

WE'VE ALL BEEN HEARING ABOUT THE SIMPLICITY AND low cost of iSCSI for years now — and how iSCSI would topple FC (Fibre Channel) as the storage networking technology of choice for shops moving from DAS to SAN. Yet entry-level SAN systems, such as those from Dell/EMC and Hewlett-Packard, although quick to adopt low-cost SATA drives, have continued to stick with FC interfaces. Even those that have offered iSCSI typically included FC as well. Fibre Channel has remained king, even for small SAN deployments.

That's finally changing. The current crop of iSCSI storage arrays prove that cost-effective SAN storage is definitely mature, leaving FC arrays comfortable only at the high end. For most small infrastructures, the performance of FC simply isn't necessary, and the cost differential is significant.

The combination of SATA hard drives and the iSCSI protocol has finally engendered the new era of storage we've been waiting for, especially for small to midsize infrastructures.

At the true low end of iSCSI SAN implementations, you find that iSCSI-connected servers don't require HBAs to talk to the SAN array — standard Gigabit Ethernet NICs and software iSCSI initiators will provide connectivity. Standard gigabit switches can then be used to create the SAN itself.

Above the iSCSI software initiators come the iSCSI HBAs that off-load iSCSI processing from the server CPU and handle it in hardware on the HBA itself. Working with Gigabit Ethernet switching that supports jumbo frames, this type of iSCSI connection could push more than 100MBps to and from the iSCSI array. Not too shabby for commodity hardware.

The six products I tested show the current spectrum of

TEST CENTER

SPECIAL REPORT

iSCSI solutions available on the market today. The Adaptec, Celeros, and Raslient units represent the lower end in price and functionality, whereas EqualLogic, Intransa, and NetApp bring more enterprise features to their products at additional cost. In addition to providing iSCSI services, these products also present a relatively wide variety of storage features, such as replication, clustering, snapshotting, and supported RAID levels. Adaptec and NetApp even include NFS and CIFS file sharing.

Notably absent is EMC, which expressed interest but ultimately declined to participate. The poor performance turned in by the Dell/EMC AX100i

(infoworld.com/3186) might have something to do with that, but I would have liked to have given the EMC CX300i a run through the lab.

My evaluation focused on configurability, storage management features, and iSCSI performance. I ran a battery of performance tests under both Linux and Windows, using software initiators as well as iSCSI HBAs from Alacritech and QLogic. I tested throughput and I/O using various block sizes, running random split tests (50/50 reads and writes) as a harsh measure of general-purpose performance and streaming read and write tests to get a feel for maximum raw throughput (see test details at infoworld.com/3276).

Because the 4KB tests are relative indicators of many real applications, such as Microsoft Exchange, I've included those results here; the results of tests for 8KB, 32KB, and 256KB block sizes are included in the online version of this

article. The accompanying charts show the numbers gathered from Iometer tests on Red Hat Enterprise Linux 4 with the linux-iscsi software initiator. Given the sector offset issues that cropped up during Windows testing, the Linux results are the best baseline numbers available for all units.

EqualLogic PS200E

EqualLogic has obviously poured much effort into its PS200E. I received two units, each boasting 5TB of raw storage laid across 14 400GB SATA drives.

The PS200E is a no-nonsense array. Instead of a glowing



EqualLogic PS200E

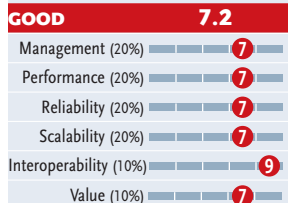
LCD panel or fancy front bezel, it sports a low-key face highlighted by disk access and array health LEDs. The redundant controllers set in the back of the chassis each contain three gigabit NICs with both copper and SFP (Small Form-factor Pluggable) fiber connections. The unit runs NetBSD and boots

quickly to a console-based initial configuration script that provides addressing for the NICs and defines a storage group to assign the controllers. In EqualLogic's view, all storage arrays are grouped into logical units. This abstraction provides a smooth way to cluster arrays for management and redundancy purposes.

After the controllers had been configured on the network, all further administration was handled by the Java-based Web GUI. I found the UI to be well-organized and quite versatile, although I did run into problems related to the JRE (Java Runtime Environment) version on a few workstations.

Adaptec Snap Server 18000

Adaptec snapappliance.com

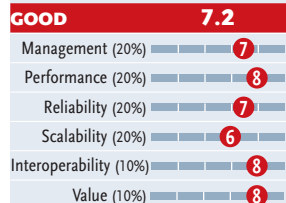


COST: \$10,995 as tested with 2TB raw storage

BOTTOM LINE: Adaptec's Snap Appliance division has been in the low-cost network filer market for a while, providing native NFS, CIFS, and AFP file sharing. The 18000 adds the ability to serve as an iSCSI target. The iSCSI implementation is a bit awkward, however, and is missing features such as selective presentation, fail-over, and redundancy. That said, the 18000 offers a lot of features for not a lot of bucks.

Celeros EzSAN XR23

Celeros celeros.com

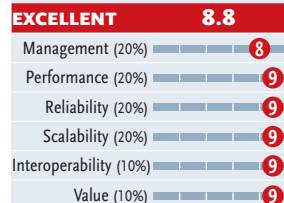


COST: \$7,400 as tested with 3TB of raw storage

BOTTOM LINE: Celeros, a new player in the iSCSI market, combines Wasabi System's Storage Builder OS and commodity hardware in the EzSAN XR23. The result is a very low-priced iSCSI array. Redundancy and snapshot features are absent, but the EzSAN handled well in the lab, proving to be a cheap yet effective way to add lots of storage when redundancy and snapshot features are not necessary.

EqualLogic PS200E

EqualLogic equallogic.com

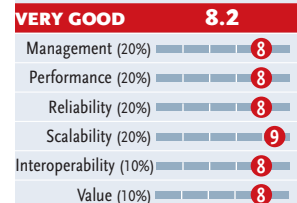


COST: \$51,800 as tested with one PS200E unit with redundant controllers and 5.6TB raw capacity

BOTTOM LINE: EqualLogic's PS200E is a well-designed and superbly executed iSCSI storage array, boasting clean lines, solid management tools, solid snapshot and replication features, and plenty of bandwidth and redundancy. Combine all that with extremely good performance in the lab, and the PS200E is simply an excellent choice for any small or midsize storage network.

Intransa IP5500

Intransa intransa.com



COST: \$61,545, including 7.1TB storage in two disk arrays, two controllers, and SMC gigabit switch

BOTTOM LINE: Intransa's IP5500's all-IP design is quite unique. All communication between the controllers and the disks flows over bonded Gigabit Ethernet NICs rather than Fibre Channel or SCSI. The solution is elegant but needs more up-front I/O: iSCSI bandwidth is limited to one gigabit NIC per controller. Furthermore, adequate Windows performance requires sector offset tweaks for every volume. Intransa is a vendor to watch.

Inside iSCSI SANs

Key differences among our six solutions include redundancy, capacity, and snapshot capabilities. EqualLogic, Intransa, and NetApp offer more enterprise-grade features.

| | iSCSI interfaces (per controller) | Redundancy | Drive technology | Maximum disk size | Maximum raw capacity | Snapshot support | CHAP support | LUN masking | RAID levels | Appliance OS |
|----------------------------------|-----------------------------------|---|------------------|-------------------|---------------------------|------------------|--------------|-------------|-------------|--------------|
| Adaptec Snap Server 18000 | 2x gigabit, copper | Power supplies | SATA | 400GB | 30TB (7 expansion arrays) | Read-only | Yes | No | 5 | Linux |
| Celeros EzSAN XR23 | 2x gigabit, copper | Power supplies | SATA | 400GB | 4.8TB | No | Yes | Yes | 0,1,5,10,50 | NetBSD |
| EqualLogic PS200E | 3x gigabit, copper/SFP | Power supplies, controllers, clustering | SATA | 500GB | 112TB (16 controllers) | Read/write | Yes | Yes | 0,1,5,10 | NetBSD |
| Intransa IP5500 | 1x gigabit, copper | Power supplies, clustering | SATA | 400GB | 51.2TB (4 controllers) | Read/write | Yes | Yes | 0,1,5 | Linux |
| NetApp FAS3020c | 4x gigabit, copper/SFP | Power supplies, clustering | SATA and SCSI | 250GB | 50TB | Read/write | Yes | Yes | 0,1,5,10 | Data ONTAP |
| Rasiliant 4000 | 2x gigabit, copper | Power supplies, controllers | SATA | 400GB | 3.75TB | Read-only | Yes | Yes | 0,1,5,10 | Linux |

I settled on a revision of 1.4.2 that seemed to play nice and created my volumes. As with most iSCSI targets, each volume can be assigned access rights to permit only certain initiators to connect and mount any given volume. The iSCSI standard calls for the use of CHAP, which provides a modicum of initiator authentication, and the PS200E handles that without an issue. Also, initiators can be assigned to LUNs by mapping the initiator IQN to that LUN. Volume presentation is then determined by the requesting initiator IQN. There is no means of grouping or aliasing IQNs, which can get tedious when working with several servers.

I configured the PS200E for performance, running the array at RAID 10 with two hot spares. RAID 10 is a mirror

set of striped arrays, providing better performance than RAID 5 while maintaining redundancy via the mirror. The downside is that only 50 percent of the raw capacity of the filer is usable. But with SATA drives reaching 500GB per spindle, this isn't the constraint it once was.

In performance tests, the PS200E led the field, claiming the highest marks in the raw single-threaded read tests and showing a superlative 101MBps 256KB streaming write throughput with the Alacritech iSCSI accelerator on Windows. Interestingly, the PS200E also responded well to the QLogic HBA, posting the best file creation and deletion times — especially impressive considering the HBA's lack of jumbo frame support. Overall, the EqualLogic PS200E posted the best raw iSCSI performance numbers in the test.

When I built the second PS200E, I initially created a completely separate group for it and configured replication between the arrays. This is extremely simple to do, and the PS200Es will do a block-level synchronization of volumes at scheduled intervals or when manually triggered. The controllers provide no bandwidth shaping, however, so factor that into your plans if you're replicating over WAN links.

After resetting the second PS200E to factory defaults, I joined it to the original array group — again, very simple — and was able to manage both arrays from the group UI. When joined, the two units immediately reallocated volumes between them for better load balancing — a very nice touch. The downside of this is that a failure of one of the group members can affect all the volumes between both units, bringing everything to a halt.

When you add arrays to an EqualLogic storage group, you not only add disk, you add controllers. Each set of controllers in the array has three active Gigabit Ethernet interfaces and uses active load balancing between them to balance server requests. The interfaces on the dormant controller can be used as fail-over interfaces as well. Thus, as you add more disk, you also add more network capacity — another nice feature.

NetApp FAS3020c

Network Appliance netapp.com

VERY GOOD 8.4

| | |
|------------------------|---|
| Management (20%) | 8 |
| Performance (20%) | 8 |
| Reliability (20%) | 9 |
| Scalability (20%) | 9 |
| Interoperability (10%) | 8 |
| Value (10%) | 8 |

COST: \$67,246 as tested with redundant controllers, 14 250GB SATA, and 14 72GB Ultra160 SCSI drives

BOTTOM LINE: The old guard of network storage has still got it. Adding iSCSI support and SATA disks to NetApp's successful filers was an obvious move, and the combination works very well. This test didn't take the filer capabilities of the FAS3020c into account, but they definitely sweeten the pie. The cluster cabling is a bit unwieldy, but fail-over is flawless, performance is solid, and support is stellar.

Rasiliant Rastor 4000

Rasiliant rasiliant.com

VERY GOOD 7.2

| | |
|------------------------|---|
| Management (20%) | 7 |
| Performance (20%) | 7 |
| Reliability (20%) | 8 |
| Scalability (20%) | 6 |
| Interoperability (10%) | 9 |
| Value (10%) | 7 |

COST: \$22,050 as tested with 2.2TB raw storage

BOTTOM LINE: Rasiliant's Rastor 4000 offers a redundant iSCSI SAN array with 15 SATA drive slots in a single 3U chassis, topping out at 3.75TB of raw capacity. The performance proved adequate, a controller failure was handled nicely, and the management UI was functional. However, the physical chassis construction tolerances of the box tested were poor, creating a sense of unease about the overall solution.

TEST CENTER

SPECIAL REPORT

One issue I did have with the PS200E involved the Microsoft iSCSI initiator. In fact, this problem also affected the Intransa IP5500, as the PS200E and IP5500 are the only solutions in the test that use iSCSI redirection to achieve load balancing. At certain times, the iSCSI volumes would simply refuse to mount or would abruptly disconnect during heavy I/O. The crux of the issue was related to the Alacritech iSCSI accelerator cards used on the Windows server in conjunction with Version 1.6 of the Microsoft initiator. Updating the Alacritech drivers to the most recent version and moving to Version 2.0 of the Microsoft initiator resolved these issues.

Overall, the PS200E is a well-designed and well-executed SAN, providing not only a significant bang for the storage buck but also a simple, powerful interface for storage management. Delivering good performance, redundancy, and scalability, this solution is definitely enterprise-ready.

On Target With iSCSI

Given the cost of big SCSI SAN storage today, and the fact that most infrastructures simply don't require the speed and throughput of a Fibre Channel SAN, making the case for iSCSI storage is simple. SATA drives are more than adequate for most e-mail, database, and file storage applications, and so is the 1Gbps iSCSI transport. The low cost of entry, combined with the ease of integration, make the SATA-iSCSI combination a no-brainer when compared with even a stand-alone file server. A rack-mount server with six 147GB SCSI disks will generally cost you more than a low-end iSCSI storage array, and it's a less effective way to provide storage to multiple applications.

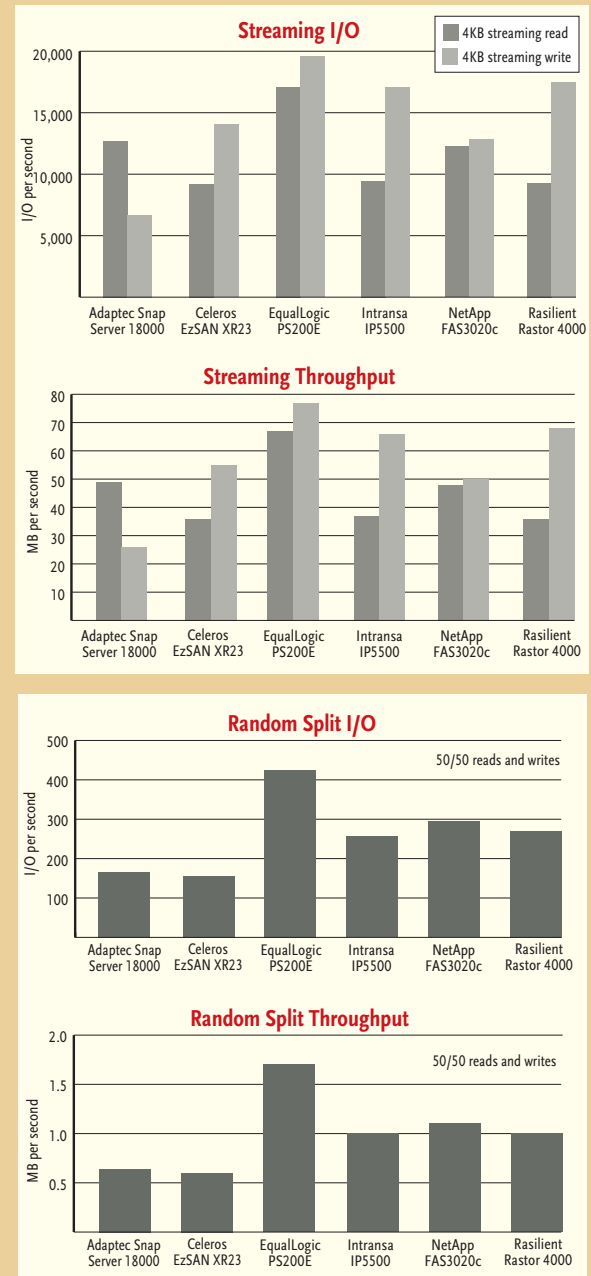
All of the arrays I tested are capable of providing large storage at the center of an infrastructure, but their performance and resilience will differ wildly, depending on the application. For a general-purpose storage array in a mid-size infrastructure, the EqualLogic and NetApp products are excellent choices. Both products are feature-packed and polished. The Intransa solution takes the bronze here, although its capacity, resiliency, and throughput are likewise capable of supporting most applications.

The Adaptec, Celeros, and Rasilient solutions match up well for smaller infrastructures where the dollar needs to go farther. The Snap Server 18000 in particular would function well as a small-office or branch-office storage unit, providing NFS and CIFS file sharing in addition to iSCSI disk-to-disk backups. The EzSAN XR23 and Rastor 4000 provide more native capacity than the Snap Server does, and they're better tuned to provide big volumes to smaller networks. The redundancy in the Rastor 4000 gives it an edge over the EzSAN, albeit at twice the price. ☞

CARY STRENG

Passing the Transaction and Throughput Tests

Results of I/O and throughput tests using 4KB block sizes show that these SANs can meet the demands of transaction-oriented applications and file-intensive workloads. The random split tests, which mix random reads and writes in equal measure, indicate heavily fragmented performance for some applications, while streaming tests highlight maximum capabilities at this block size.



NOTE: Results shown are from Iometer tests run from Red Hat Enterprise Linux 4 using the linux-iscsi software initiator. Test results for 8KB, 32KB, and 256KB block sizes may be found at infoworld.com/3281.



(888) 579-9762 • www.equallogic.com