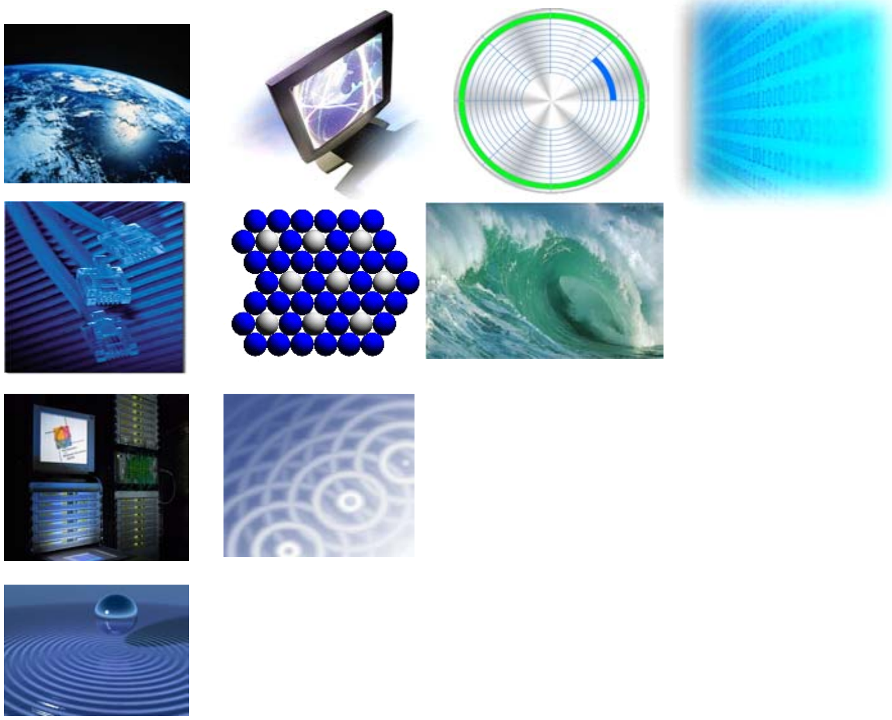
Validation Report

Enterprise Storage Group Lab

Storage Knowledge • Business Results

ESG Lab Report™



Intransa IP5000™

intransa

A product validation study
by
ESG Lab

Fall 2003

Authors:
Tony Asaro Brian Garrett

Table of Contents

Executive Summary

Company Overview

Intransa, a startup based in San Jose, CA, was founded in September 2000. Intransa launched their mid-range iSCSI storage subsystem, the IP5000 in June 2003. The IP5000 is an iSCSI storage subsystem that has a distributed clustered architecture allowing for scalable growth and fault tolerance. Simplicity is not normally inherent with iSCSI storage subsystems but Intransa has focused on making the management of the IP5000 easy and intuitive.

Today's Challenges

In many companies (large and small) the cost and complexity of installing a Fibre Channel (FC) SAN is not justified for some or all of their applications:

- For many applications that run on $3,000 to $5,000 Intel-based servers the cost of the FC SAN connectivity, including the HBAs, FC switches, and fiber cabling is not justified.
- Data is growing at such a rapid rate and the price of the Enterprise-class SAN storage is so expensive that it does not make economic sense to manage anything but mission-critical and priority applications.
- The IT staff in charge of managing the Windows servers may not wish to lose control of their storage by moving their data into the data center, which is managed by the storage IT staff.
- Small companies do not have the staff, experience or budget to deploy even a modest FC SAN.

The benefits of deploying a SAN include reducing the time and the in-direct cost associated with managing islands of storage. Backup and recovery becomes easier and more reliable; there are fewer systems that need to be maintained; customers can manage their data more efficiently with different tools; and data can be more easily shared.

Various elements are converging to enable the success of IP SANs:

1. The iSCSI protocol is mature.
2. Native Microsoft and Linux iSCSI host drivers are available.
3. There are a growing number of storage subsystems that support iSCSI
4. TCP/IP off-load engines and iSCSI HBAs are available for CPU-intensive applications
5. 1GB Ethernet is commodity and offers adequate bandwidths; in the near future 10GB will be a commodity and widely deployed providing high-levels of bandwidth at a reasonable cost.

As more customers install IP SANs and it becomes a "comfortable" choice then it will begin to be used in the data center and eventually with high performance and mission-critical data.

Product Overview

ESG Lab categorizes the IP5000 as a mid-range storage subsystem. The IP5000 has a distributed clustered architecture that supports active/active clustering. The IP5000 software includes read caching, write through caching, full volume mirroring and snapshots.

Intransa claims that the following capabilities make IP5000 a unique and compelling solution:

- A distributed clustered architecture:
    - Provides a high level of reliability
    - Will allow you to add controller modules to scale processors, host ports and capacity
- Easy to install – it takes hours and not days
- iSCSI-based SAN

- Price/Performance leader
- Easy to manage

Intransa IP5000 Project

ESG Lab considers IP SANs to be an emerging technology that has the potential to change the storage landscape. Still in its early stages customers must become comfortable with implementing IP SANs in order for them to become commonplace. With this in mind, ESG Lab set out to test the ease of use, data management features and reliability of the IP5000. Performance was a specific area of focus into which ESG Lab wanted to gain insight in order to help customers understand the practicality of installing an iSCSI-based SAN, and in this case using the IP5000.

Summary and Test Results

ESG Lab spent two and a half days running tests and evaluating IP5000 including installation; accessing iSCSI targets; mirroring services; performance; reliability; etc, with the following results:

- The IP5000 was easy to configure and has an excellent management interface
- IP5000 was easy to install – most of the effort is in cabling and Ethernet switch configuration
- Management functions were intuitive and easy using the IP5000 management GUI – StorControl
- Creating and managing iSCSI volumes was straightforward
- Using the Microsoft iSCSI driver to discover and log onto iSCSI targets is easy
- Adding capacity online was effortless on Windows Advanced Server 2000 using disk administrator
- Creating snapshots and mirrors that provide a point-in-time virtual copy of a volume can be performed using the StorControl or the Command Line Interface (CLI)
- Application level performance for Exchange e-mail and streaming media applications is good and competitive with FC-based mid-range storage solutions
- Maximum I/Os per second (56,000) and bandwidth (185 MB) are comparable to mid-range Fiber Channel solutions
- CPU utilization due to TCP/IP overhead was not noticeable during ESG Lab application emulation testing
- The IP5000 was able to survive drive failures, port failures, and controller failures and continued to operate normally

Issues to Be Addressed

The following are recommendations that ESG Lab for Intransa to enhance their products:

- The IP5000 does not currently support SNMP or SMI-S standard management interfaces. Their StorControl management software is a good management tool but customers are looking for standards-based solutions.
- The IP5000 currently supports RAID-1(mirror) and RAID-10 (stripe of mirrors) protection options. ESG Lab recommends that Intransa support RAID-5 support for reduced protection cost (with the industry accepted RAID-5 performance penalty).
- Although installation is a one-time occurrence cabling and Ethernet configuration for the back-end switch are time consuming and complex. Intransa should integrate an Ethernet switch for the backend to simplify the installation process.
- The IP5000 distributed clustered architecture has the potential for excellent performance and capacity scalability as larger numbers of controller and drive enclosures are supported. In Q1 2004, Intransa will support a larger configuration of two Storage Controller Enclosures (four storage controller modules).
- The IP5000 does not support remote mirroring today. Supporting remote mirroring over native IP can cost-effectively enable budget conscious customers to deploy a business continuance solution.

- The iSCSI protocol does not support remote booting of servers today. This is a limitation that is not in the control of Intransa but is instead a technology issue being addressed by iSCSI HBA and operating system vendors.

ESG Lab has discussed all of the above issues with Intransa and all of these items are in their roadmap.

ESG Lab's View[1]

Customers should look at the IP5000 as a mid-range storage subsystem and consider iSCSI an enabling technology that reduces the cost of deploying a SAN. From a price/performance perspective the IP5000 is extremely well-positioned. If customers examine the total cost of ownership, the economics are compelling.

The IP5000 is a competitive mid-range storage controller supporting 4 GB of cache; 10TB of raw capacity; an active/active distributed clustering architecture; snapshots and full volume mirrors. Over time the IP 5000 will potentially eclipse many other mid-range storage subsystems in terms of storage capacity, cache memory, processors and host ports due to the flexibility of their distributed clustered architecture.

End user customers have told ESG Lab they are concerned with the performance of an iSCSI SAN. The issue of performance is an unnecessary distraction. In terms of pure IOPs and bandwidth, ESG Lab found that the IP5000 had better performance numbers than some of the popular FC-based mid-range storage subsystems; was on par with others; and below the upper end of this segment. Additionally, the IP5000 performed very well with workloads simulating a sizeable MS Exchange environment supporting over 9,000 users.

Customers must determine their storage requirements based on various factors. Here are some questions to consider:

1. Is a FC SAN too expensive to justify implementing?
2. Do you have limited resources and require an "easy to manage" SAN?
3. Do your performance requirements fit within mid-range or high-end storage subsystems metrics?
4. Do you have a problem managing your DAS islands or are you content with your current environment?
5. Are you uncomfortable working with a startup company and a product that is relatively new to the market or are you willing to use latest and greatest innovations to improve your environment?

ESG Lab found that the IP5000 is a competitive product that is very easy to manage; has a scalable architecture; provides hardware reliability and good performance compared to other mid-range storage subsystems. All of these attributes combined with the low cost of IP networks make the IP5000 a price/performance leader in its class.

[1] ESG Lab Disclaimer: It should be noted that ESG Lab did not use and is not using IP5000 in a production environment day-in and day-out. ESG Lab reports are not designed as a substitute for formal engineering testing by potential customers. Each customer should do his or her own evaluations. However, ESG Lab believes its tests provide a solid basis for the opinions expressed in this report.

Validation Report

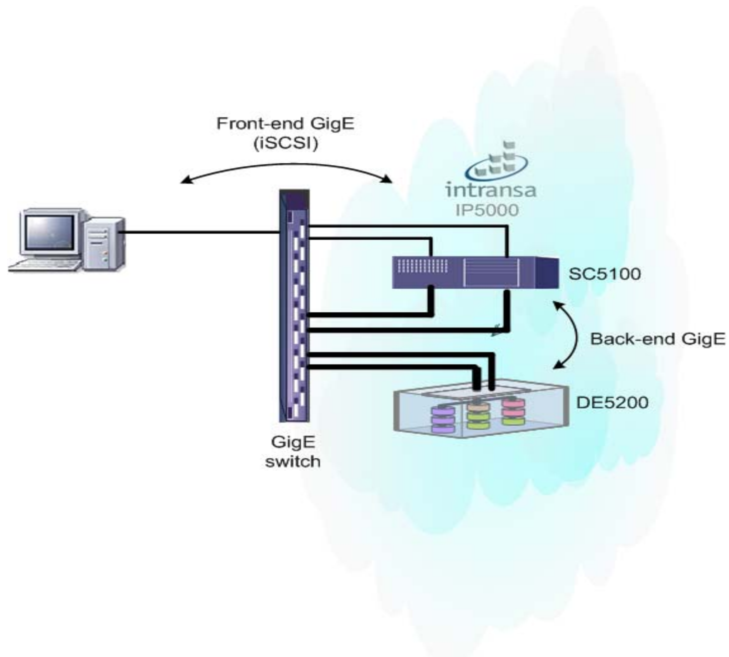
Setting the Stage

The IP5000 is a mid-range storage subsystem that supports iSCSI with an active-active distributed clustered architecture. This design will allow for additional disk controllers to be added to the cluster for scalability of storage capacity and performance in the future.

An IP5000 system consists of one or more SC5100 Storage Control Enclosures and one or more DE5200 Drive Enclosures as depicted in Figure 1. Two physically redundant Storage Controller Modules contained within each SC5100 Enclosure communicate with servers using the iSCSI protocol over Front-end GigE connections.

Today the IP5000, with a basic configuration of one Storage Controller Enclosure (containing two fault tolerant storage controller modules), supports up to 1,000 volumes (iSCSI targets) and up to 128 host connections (iSCSI initiators).

Figure 1: IP 5000



Each DE5200 Drive Enclosure contains up to 16 ATA drives. DE5200 drive enclosures are connected to the SC5100 through Back-end GigE connections running a lightweight protocol defined by Intransa. The SC5100 communicates with each disk drive as an Ethernet attached IP-addressable device. This design provides the following advantages:

- Drives can be added online to increase the size of a single shared logical pool of storage
- Controllers can be added for performance and scalability
- Controllers can take over for each other in the event of a failure for fault tolerance
- Drive spares can be located anywhere in the network

ESG Lab tested installation, configuration, management features, data management functionality, performance and reliability. The next section focuses on performance since customers have expressed a need to obtain more information in this area and is followed by the evaluation of the other aspects of the IP5000.

Performance

Customers are uncertain whether an iSCSI SAN will be able to perform to the levels they need for their applications. The perceived issues are three-fold:

1. The inefficiencies of IP versus the FC protocol
2. FC bandwidth at 2 GB and Ethernet at 1 GB
3. TCP/IP demand on the host CPU and the requirement for TCP/IP Off-Load Engines (TOE)
4. ESG Lab testing shows that an iSCSI SAN can meet the performance requirements of mid-range storage environments:

- ☑ ESG Lab tests show that iSCSI throughput achieves near wire speed
- ☑ For most applications 1 GB is sufficient bandwidth and 10 GB Ethernet will be more widely deployed in the near future.
- ☑ ESG Lab tests have shown that in many cases the impact on host CPU utilization is marginal and a TOE is not required.

The issue is not the IP network but potentially whether the iSCSI storage controller is a bottleneck. ESG Lab tested the performance and scalability of the IP 5000 to determine if it can handle the performance of popular applications as follows:

- Microsoft Exchange workload analysis using Iometer
- Compared IP5000 iSCSI performance to direct attached storage (DAS)
- Audited engineering Exchange testing for comparison to results collected
- Evaluated sequential throughput scaling for use in streaming media and scientific applications
- Interviewed an IP5000 customer who had performed sequential throughput testing
- Measured the IP5000's maximum rate of I/Os per second (IOPS)
- Did an analysis of CPU utilization and audited Intransa TOE/HBA CPU utilization benchmark results

Microsoft Exchange

In 2001 companies circulated more than 12 billion e-mail messages each day.[2] The average size of individual messages is growing annually at a rate of 55 to 210%. The increase in the size of messages, along with the 29% annual growth of e-mail volume has become a challenge for storage administrators. ESG Lab evaluated the performance scalability of the Intransa IP5000 simulating an e-mail server environment.

1. ESG Lab testing was performed using the Iometer synthetic workload generator
2. Workload characteristics of an Exchange server were used to simulate the I/O of:
    a. An Exchange Data Base (EDB) file
    b. The log file (LOG) and the streaming media database (STM) as described in Appendix A

Microsoft guidelines were used to determine the number of "average" Exchange users that the following configurations could support:

[2] Merrill Lynch, The Storage Report: *Customer Perspectives and Industry Evolution*, June 2001

SCSI (2 Drives) - The EDB workload was directed at one SCSI drive while the LOG and STM workloads were directed at a second SCSI drive. Placing the LOG workloads on a drive other than that used for the EBD file is Microsoft recommended best practice.[3]

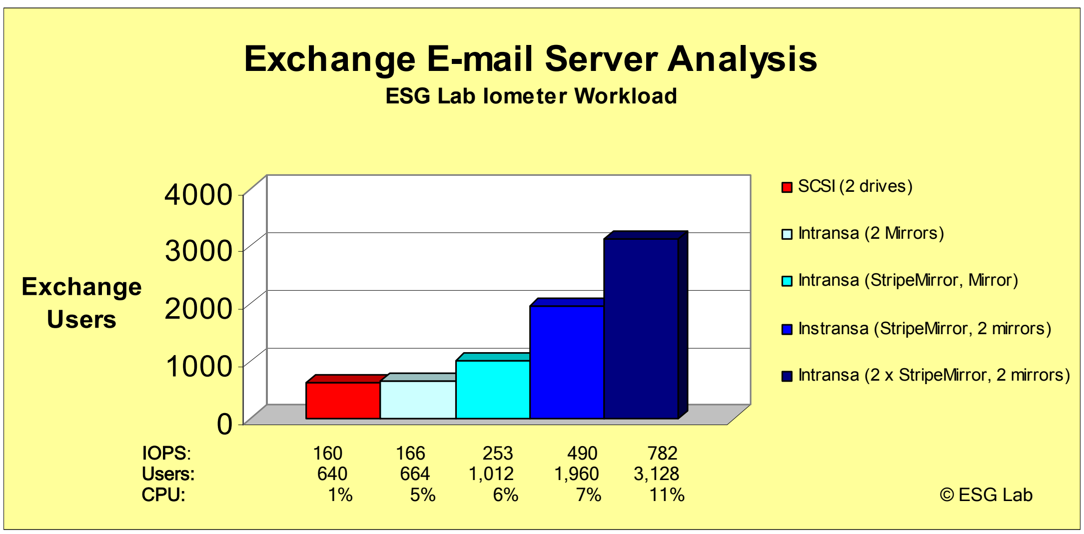
Intransa (2 Mirrors) - Two Intransa mirrored volumes were used. Again, the EDB workload was directed at one volume and the LOG and STM workloads were directed at the second volume.

Intransa (StripeMirror, Mirror) - Two Intransa volumes were used. The first StripeMirror volume (four drive RAID-10) was used for the EDB workload and the LOG and STM workloads were directed at a second mirrored volume.

Intransa (StripeMirror, 2 Mirrors) - Three Intransa volumes were used. The first StripeMirror volume was used for the EDB workload and mirrored volumes were used for the LOG and STM workloads.

Intransa (2 x StripeMirror, 2 Mirrors) – Two PC's accessing an IP5000 through two GigE interfaces were used Each PC accessed three Intransa volumes as described in the Intransa (StripeMirror, 2 Mirrors) test case above.

Graph 1: MS Exchange IOPs. Users. CPU



What the Numbers Mean

The number of I/Os per second reported by Iometer is presented along the bottom of the graph as IOPS. Microsoft guidelines state that an average Exchange user performs approximately 0.25 I/Os per second. ESG Lab used this guideline to calculate the number of simulated Exchange users that each configuration could support. For example, the two SCSI drive configuration performed 160 I/Os per second would support 640 average Exchange users and the IP 5000 with 2 Mirrors would support 664 average Exchange users.

Observations:

- A pair of Intransa iSCSI mirrored volumes performed slightly better than two internal SCSI drives

[3] Microsoft recommends not only isolating LOG files to a different volume, but also using mirroring (RAID-1) for optimal performance and protection instead of RAID-0, RAID-5 or RAID-10.

- Moving the database from a pair of Intransa mirrors towards the best practices of using a striped RAID set for the data base and mirrored volumes for logs and streaming media increases performance dramatically (166 to 253 to 490 IOPS) creating an infrastructure that can support more Exchange users (664 to 1,012 to 1,960 users).

- CPU utilization was low (11% or less) for all tests even though the testing was performed using the built in GigE adapter (not a TOE or iSCSI HBA) in a dual processor Dell 1650 server. The Iometer workload generator used for this test has low CPU requirements. An Exchange Server in a production environment would have a higher CPU utilization.

- Adding a second PC and a second set of files to simulate the deployment of two Exchange servers resulted in the ability to support 3,128 Exchange users.

- The ESG Lab Iometer test methodology was compared to another independent test run at Intransa using a Microsoft approved workload generator and methodology. The ESG Lab Iometer workload was slightly more aggressive than the Microsoft workload generator yielding results that are more conservative than Microsoft results. For example, the third configuration tested supported 1,102 users with the ESG Lab Iometer workload compared to 1,388 average users using the Microsoft workload and methodology.

Intransa internal test results, using the Microsoft workload generator were audited by ESG Lab for larger IP5000 configurations indicating that an Intransa configuration built over two disk enclosures can scale up to 9,120 average Exchange users.

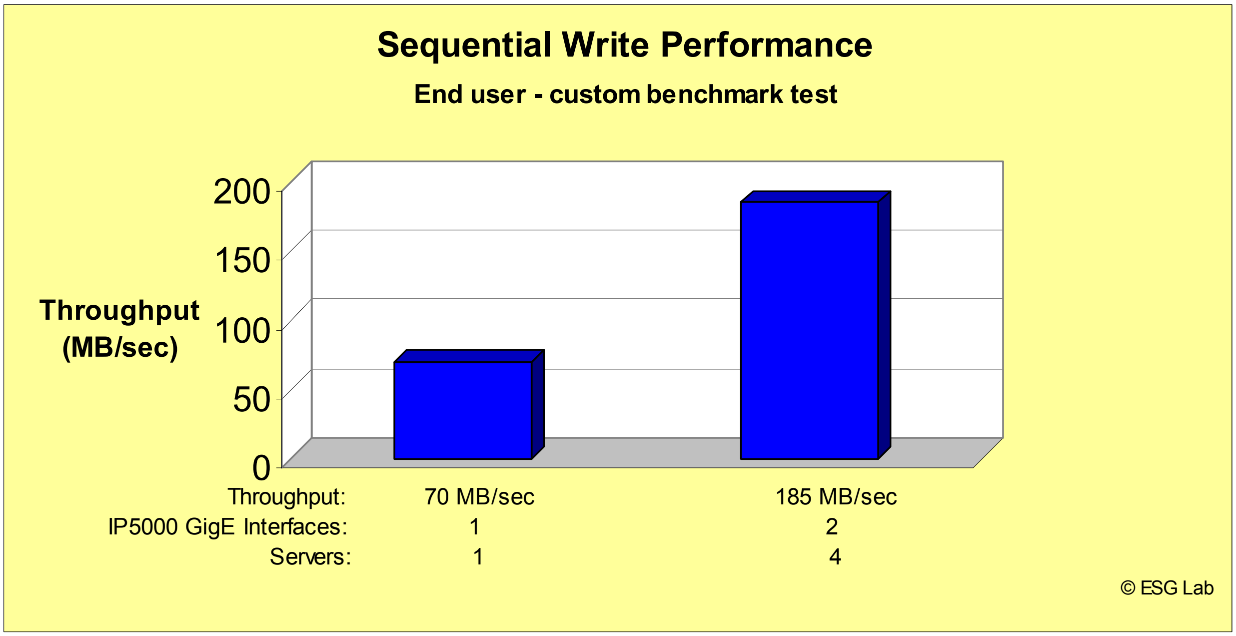
Streaming Media and Large Data Stores

ESG Lab performed tests to measure the throughput of the IP5000. Large block sequential read and write testing was performed from a Dell 1650 server using Iometer to determine the ability to service applications which access large sequential streams of data (e.g. video streaming). Additionally, ESG Lab spoke with an Intransa customer who had performed similar tests with comparable results.

The Intransa customer is responsible for recommending and implementing infrastructure for scientific applications that create and manipulate very large sets of data accessed by a relatively small number of users. The data is collected, analyzed and processed by servers in multiple facilities in different geographies.

A customer using a custom benchmark test evaluated the performance capability of the IP5000 in order to determine whether the low cost of ownership of iSCSI could be used to ease the cost and complexity of their existing FC storage infrastructure. Excerpts from the customer's benchmark results are presented below. Details regarding the benchmark configuration are listed in Appendix A.

Graph 2: End User Sequential Write Performance



What the Numbers Mean

A throughput level of 65 MB/sec was achieved for streams of 256 KB sequential writes from a Dell 2640 server using a single built-in Intel PRO/1000 XT NIC connected to the IP5000 using jumbo frames. Four Dell 2640 servers accessing four StripeMirrored (RAID-10) IP5000 volumes built over 16 ATA drives through two IP5000 GigE interfaces achieved an aggregate throughput level of 185 MB/sec.

Observations:

- 65 MB/sec is a good throughput level for a single target on a single interface given the 128 MB/sec theoretical wire speed limit of GigE.
- Scaling to 185 MB/sec for four streams from four servers over two GigE interfaces with a theoretical wire speed limit of 250 MB/sec is an excellent throughput level.
- CPU utilization for the 185 MB/sec results was 11% using an off-the-shelf GigE adapter. The CPU requirements of the customer benchmark used for testing are low. It is expected that production applications, which are compute intensive, will be deployed using a TOE adapter or an iSCSI HBA.
- Intransa throughput results are comparable to mid-range Fiber Channel controllers. For example, the EMC CX200, supporting up to 30 drives and 2 FC-SW interfaces or 4 FC-AL interfaces, has a published maximum throughput of 200 MB/sec.
- ESG Lab audited the following impressive throughput results performed with four Dell 1650's using built-in GigE adapters accessing eight mirrored volumes in an SC5100 balanced across two storage controller modules and two disk enclosures.

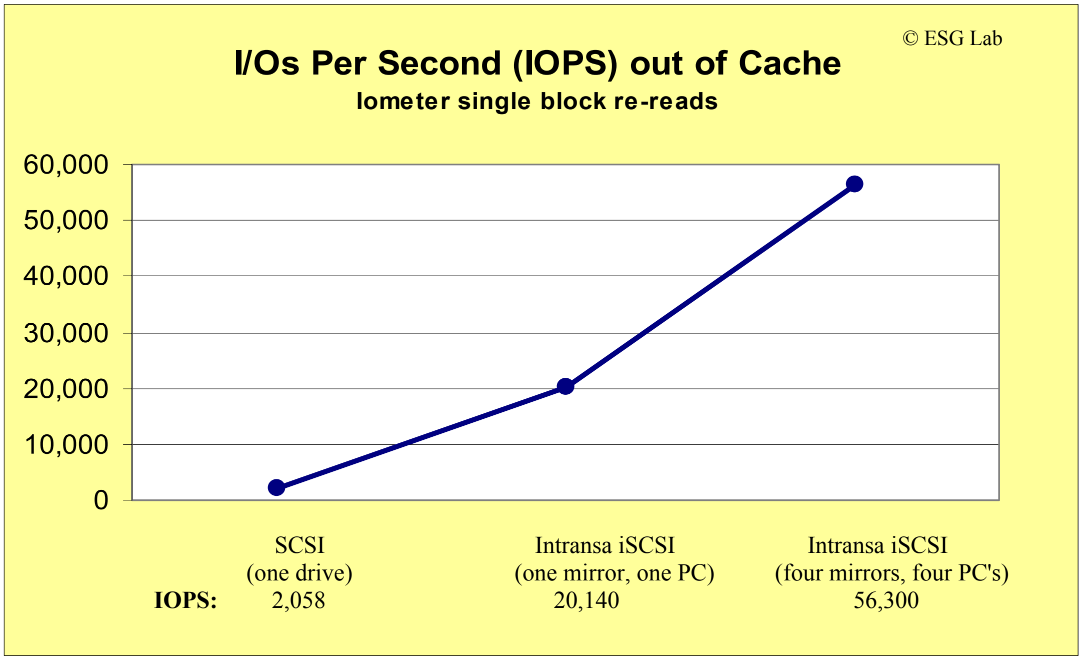
SEQ reads from disk	107 Mb/sec
SEQ READS FROM CACHE	221 MB/sec

There were only two GigE cables between the GigE switch and the IP5000, making the 221 MB/sec (out of a 250 MB/sec theoretical write speed limit) an outstanding result.

IOPS

ESG Lab performed tests to measure the maximum rate of I/Os per second (IOPS) that can be processed by the IP5000. One block reads were performed using Iometer. The range of addresses tested was limited to ensure that all accesses were serviced from cache. For comparison to a SCSI drive all reads were performed to the same block, which was serviced from the SCSI drive's cache.

Graph 3: IOPS out or Cache



Observations:

- IP5000 performance to a single mirror volume of 20,140 IOPS is almost 10 times faster than an internal SCSI drive
- IP5000 IOPS scaling to over 56,000 IOPS over two GigE interfaces from four PC's is an excellent level of performance comparable to midrange FC controllers with two to four FC interfaces. For example, EMC publishes a rate of 40,000 IOPS for the CX200 and 60,000 IOPS for the CX400.

CPU Utilization

During ESG Lab testing and customer benchmark testing, CPU utilization was low and not a concern. However lab testing with a workload generator is not indicative of the CPU requirements of real-world applications, Customers must be aware that CPU utilization should be considered and that application demand will vary.

Intransa is working with their customers to understand the iSCSI CPU utilization issue and the performance characteristics of TOE and iSCSI HBA technology. Intransa recommendations are summarized as follows:

- Off-the-shelf GigE adapters are appropriate for applications that have light disk performance requirements where most processing is done by the application (e.g. Microsoft Exchange).

- TCP I/P Offload Engines are appropriate when other IP services need to be offloaded (e,g, a multi-purpose server doing web, nfs/cifs, and mail).

- iSCSI Host Bus Adapters are appropriate when applications have heavy CPU and disk requirements (e.g. video serving, video editing).

- Benchmark tests have shown that iSCSI HBAs are excellent at reducing server CPU utilization for workloads with a high I/O rate composed of small I/Os (compared to NICs and TOEs).

- Users who are considering iSCSI for use in an old server with an old slow CPU, or a server with high CPU utilization should seriously consider a TOE, HBA or server upgrade.

Why This Matters

The ESG Lab performance testing showed that the IP5000, when compared to FC mid-range subsystems, was above average in raw bandwidth and I/Os out of cache. The IP5000 has a significant price/performance advantage when compared to FC mid-range storage subsystems due to the cost and complexity of implementing a FC SAN.

The IP5000 uses 7,200 RPM ATA-drives, whereas most FC storage subsystems support 10,000-15,000 RPM FC-drives. The FC-drives are faster than the ATA-drives and there will be a marginal performance difference for seek intensive workloads (e.g. heavy response-time critical on-line transaction processing). However, one of the advantages of the Intransa IP5000 is its road map of supporting more and more controllers and drive enclosures enabling customers to scale performance as the environment demands. Today the Intransa IP5000 supports one SC5100 enclosure containing two storage controller modules. ESG Lab tested an alpha version supporting a cluster of two SC5100 enclosures with four storage controller modules. This configuration is targeted for Q1 2004 availability.

The performance characteristics and low cost of ownership of the IP5000 are well suited for:

- Microsoft Exchange
- Streaming media
- Scientific applications with a large data store being processed by a relatively few number of users
- Departmental Microsoft SQL server applications

IP5000 Initial Configuration

Initial IP5000 configuration began by powering up each of the two independent controllers within the physical frame of one SC5100 enclosure. Using a null modem serial cable and a hyper-terminal interface, network settings were provided so that so that each of the storage subsystem components could communicate internally and with other devices in a corporate Ethernet network. Three networks were configured:

1. Host access network for iSCSI traffic
2. Management network for IP5000 configuration and monitoring
3. Storage delivery network for communication between the controller and the disk enclosure

The IP5000 configuration steps are summarized as follows:

1. Standard network settings were defined for physical interfaces
2. Physical interfaces were grouped together into fault tolerant virtual interfaces
3. The two controllers were joined into a fault tolerant cluster (which Intransa refers to as a realm)

The Dell server connected to the IP5000 through a host access network VLAN port on the switch was then powered up and the balance of IP5000 configuration steps were performed using StorControl:

- Configured a logical device (iSCSI target) using simple pull down menus to specify a name for the volume, protection (mirror), capacity in MB, GB, or TB, and placement (e.g. best fit)
- Assigned the volume an initiator address

Summary Conclusions

- Standard network settings are much like setting up a file server, but for three networks with robust fault tolerant options
- Intransa field engineering assists with iSCSI network planning and configuration. The user simply provides IP address ranges.
- Assigning volumes to initiators provides a level of security such that an administrator can restrict volume access to the IP address of a specific server's GigE interface. This is analogous to the LUN masking (a.k.a. access control) functionality provided by FC array vendors.
- Configuring iSCSI volumes using the visually appealing StorControl is easy and intuitive. In particular, the *Volume General View* is a well-arranged screen from which volume operations can be easily launched and observed. In addition, the *Volume Initiator* and *Volume Disk* maps are helpful and visually stunning as they show how volumes are mapped to host GigE interfaces and drives respectively in a matrix fashion.

Using an IP5000 iSCSI Device on Windows 2000

After the Dell server had booted Windows 2000 from an internal SCSI drive, the Microsoft iSCSI Software Initiator driver was downloaded and installed.[4]

The graphical interface of the Win2K driver was used to:

1. Discover devices and IP5000 mirrored device configured previously.
2. Log on to the iSCSI volume (a.k.a. iSCSI targets) and check a box to automatically log on in the future

From this point on, working with an IP5000 iSCSI volume on a Win2K server feels like you are working with a SCSI or FC volume. ESG Lab used the Windows 2000 Disk Administrator Utility to:

1. Re-scan and recognize the iSCSI volume with the Windows 2000 Disk Administrator Utility
2. Quick format the iSCSI volume as a basic disk and assign to drive letter I:
3. Copy a video clip to the I: drive using the Win2K Explorer utility
4. Watch the video using Windows Media Player
5. Check properties of the I: drive and run Win2K disk management utilities (e.g. Disk Defragmenter)

Summary Conclusions

- Approximately one hour after the system had been cabled up, an IP5000 iSCSI volume was accessed as an I: drive
- The difference from a user perspective between dealing with an iSCSI array and a FC array is that instead of dealing with SCSI or FC BIOS to recognize and do initial configuration, the Win2K iSCSI driver is used instead. The Win2K GUI driver is easy to use and only two operations need be mastered: discovering devices and logging on to a device.

[4] The free Microsoft initiator driver supporting Windows 2000 Service Pack 3, Windows Server 2003, Windows XP and Windows XP 64-bit driver is available for download at http://www.microsoft.com/downloads/

- Once the Win2K driver has been used to log on with an iSCSI volume, using an iSCSI device feels like any other SCSI or FC volume.

Online Capacity Growth

One of the biggest complaints that customers have concerning their storage is the ability to grow and manage SCSI and FC LUNs. They would like a simple process that would allow them to add capacity without taking down the system. ESG Lab has identified four aspects to simplifying this process:

- The storage subsystem should support online volume expansion
- The storage subsystem management interface should be easy to use
- The operating system and/or volume manager must support online recognition of added capacity
- The storage subsystem should support online addition of new physical capacity (drives)

Microsoft Windows 2000, 2003 and Novell NetWare support online growth of logical volumes, but Windows XP and most UNIX variants do not. However, UNIX system administrators typically install file systems and applications on virtual volumes created with a volume manager (e.g. VERITAS Volume Manager). An operating system utility is used to discover the new volume and the volume manager is used to grow the virtual volume while online.

The IP5000 supports online volume growth and has developed an easy-to-use StorControl interface. StorControl is intuitive and easy to use supporting a variety of volume expansion methods:

- Percentage to add
- Fixed capacity amount to add
- Drag and Drop from available capacity to a defined volume

ESG Lab tested the ability to increase the size of an existing Win2K drive as follows:

1. Files were copied from the internal SCSI C: drive to an IP5000 iSCSI I: dynamic disk[5]
2. As the I: drive neared 100% capacity ESG Lab added more capacity as follows:
3. StorControl was used to increase the size of the volume
4. The rescan feature of the Windows 2000 disk administrator utility was used to recognize the new capacity and add it to the I: drive.

In addition to expanding volumes while online, ESG tested the IP5000's ability to add disk drives while up and running. A Customer Replaceable Unit (CRU) containing an ATA drive was hot plugged, configured and used in a running system. In addition, ESG Lab witnessed a new drive enclosure with 3.2 TB of raw capacity being cabled and configured into a running system.

Why This Matters

Adding storage capacity to local internal SCSI and SAN-based volumes can be difficult. Often new capacity is installed first and then the data is copied from old volumes to new larger volumes during scheduled system downtime. In some cases the copy is done over the network during off-hours and in a very few cases the volume is backed up to tape and restored onto the new media.

Adding capacity should be a simple and fast procedure. ESG Lab experienced an elegant process that required minimal effort. Storage administrators will find it easy and straightforward to add capacity with the IP5000.

---

[5] Microsoft does not currently have official support for dynamic iSCSI disks. iSCSI basic disks can be expanded using the diskpart utility. The dynamic disk method was chosen for testing in a lab setting for simplicity (GUI instead of cmd line utility) and ease of explanation.

- Adding capacity to an IP5000 to a volume can be performed online
- Adding storage to a volume using StorControl is easy
- Additional capacity is available for use immediately

Snapshots for File Level Restore

ESG Lab used the IP5000 Snapshot feature to create read-only snapshots of a file system within a SpinServer cluster followed by the restoration of a file from one of the snapshots. The procedure below is designed to simulate a customer scenario:

- A storage administrator sets up nightly snapshots of a shared Win2K drive
- Some time later a virus attacks a user's PC corrupting a file on the snapshot protected shared drive
- The snapshot is mounted read-only as another drive letter
- The corrupted file is retrieved from last night's snapshot and copied to the shared drive

Procedure:

1. A document was copied from the C: drive to the I: drive
2. StorControl was used to create a snapshot of the I: drive.
    a. Added the snapshot
    b. Named the snapshot
    c. Marked the snapshot read-only (read/write is also supported)
3. The document was "corrupted" (deleted all but one line of content) and saved on the I: drive
4. Another snapshot was taken
5. Snapshots were accessed as a read-only drive letter
    a. The Win2K iSCSI driver was used to find and log into the snapshot volumes
    b. Win2K Disk Administrator mapped the snapshots as drive letters (J: and H:)
6. Windows Explorer was used to browse the J: and K: read-only snapshot images
7. The "un-corrupted" document was found on J:
8. The snapshot image of the file from before the corruption was copied from the J: snapshot drive to the I: drive

Why This Matters

Creating snapshots on the IP5000 with StorControl was an easy process. Snapshots are a standard feature for network attached storage devices but only in the last year has this feature become commonplace in block-based storage controllers.

- A file or directory tree on an iSCSI volume can be easily and quickly restored from disk (instead of a lengthy restore from tape)
- Multiple snapshots can be created providing greater chance of data integrity and recovery
- Capacity used for snapshots is significantly less than that needed for full volume copies or mirrors
    - Allows user to create more point-in-time backup images
    - Allows a customer to use more disk storage for primary data. Intransa and other industry guidelines estimate that 10% of useable capacity is typical[6]

Local Mirroring Services

ESG Lab used the IP5000 Mirror Volume feature to create and use a mirrored image of a shared Win2K drive. The procedure below is designed to simulate a customer scenario:

---

[6] Small snapshot apertures (i.e. frequent snapshots) and applications with a heavy write/read ratio increase the amount of space consumed by snapshots, which can increase the allocation amount above 10%.

- The data for a mission critical financial application is stored on the shared I: drive
- A new version of the financial application software is available
- New custom reports have been developed
- A test of month-end reporting using the new version of the software with the new reports is desired
- A mirrored image of the I: drive is created
- At month-end, the mirror is broken providing a point-in-time image
- The updated financial software is installed on a spare server
- The spare server maps the Intransa IP5000 mirror image of last month's data as an I: drive
- Month end testing occurs including refinements to the new custom reports
- Performance of the application on the production server is not impacted during testing

The ESG Lab test procedure to accomplish the above:

1. StorControl was used to create a mirror of the I: drive
2. Synchronization progress was noted on StorControl GUI
3. New files were copied to the I: drive
4. The mirror relationship was broken using StorControl[7]
5. The I: drive was deleted and the mirror was accessed as drive F:
    a. The I: drive was deleted using the Win2K Disk Administrator utility
    b. The Win2K iSCSI driver was used to find and log into the mirrored volume
    c. Win2K Disk Administrator mapped the mirror as drive letter I:
6. The hypothetical "month-end" state of the I: drive was noted using Windows Explorer

Why This Matters

The ability to mirror IP5000 iSCSI volumes is a powerful feature enabling a variety of administrative tasks including:

- Mirroring is often used to make copies of data for application staging and testing and performs well since data copies occupy their own drive separate from the primary volume.
- Customers can create a relatively inexpensive disaster recovery solution using the IP5000 and the mirroring feature by "stretching the cluster". The IP5000 supports true active-active clustering that allows customers to physically locate controllers and disk enclosures in dispersed parts of their buildings or campus connected via GigE. Disk enclosure can be "stretched" and located in different facilities with the current release of the IP5000. Support for multiple SC5100 enclosures, planned for 1Q04 and demonstrated for ESG Lab, will permit the entire cluster to be "stretched" and mirrored for high availability and disaster recovery.

Scalability – Processors, Ports, Cache and Capacity

The IP5000 supports a two iSCSI controller configuration (in a single enclosure) with 2 GigE Ports for host level iSCSI traffic access (8 GigE ports in total including ports used for internal communication). The SC5100 enclosure connects to up to three drive enclosures today for a maximum raw capacity of 10 TB.

During ESG Lab testing a two SC5100 configuration with four storage control modules providing a total of 4 GigE ports for host level iSCSI traffic and a maximum capacity of 32 TB, was demonstrated.[8] In the future, Intransa is planning on supporting with storage control modules and beyond. The IP5000 will allow customers to start at a low entry point and grow into larger systems without a huge initial investment.

[7] A minor bug was noticed during this phase of ESG Lab testing. The first mirror of the I: drive had synchronized, and a second mirror was in the process of synchronizing. Attempts to break a mirror were rejected due to a synchronization in progress, even though there was one mirror synchronized and available to break. The problem was reported to Intransa and should be fixed in a future code release. The work-a-round was to simply wait for the second mirror synchronization to complete.

[8] Planned 1Q04 availability

Adding controllers translates into more processing power, cache and GigE bandwidth. The flexible scalability of the IP5000 provides the freedom to tune a configuration for processing, throughput or capacity as needed.

Fault Tolerance

A series of faults were injected during this phase of testing to validate the resiliency of the IP5000. As a Linux attached server ran a write/read compare script, errors were injected culminating in disabling power to an IP5000 drive enclosure and controller.

After building and configuring a system for high availability (e.g. stripe-mirror volumes spread across two drive enclosures), the following errors were injected as background jobs ran without incident:

- Hot unplugged a drive and observed a rebuild starting
- The IP5000 "self-heals" after a drive failure by automatically finding and allocating spare capacity for mirror rebuilds
- Pulled redundant Ethernet cables while online (front iSCSI, back to drive enclosures)
- Powered off a drive enclosure
- Powered off a controller

Why This Matters

- The IP5000 configured in a clustered Highly Available configuration without a single point of failure is a cost effective solutions for use in mission critical applications.
- The self-healing capability of finding unallocated drives for RAID rebuilds makes array management easier (no need to manage hot spares, simply set aside spare capacity for growth AND healing).

Conclusions

- The IP5000 is a competitive mid-range storage subsystem that has an extremely compelling price/performance leveraging standard hardware, ATA drives and Ethernet infrastructure.

- The IP5000 has a compelling distributed clustered architecture that provides reliability and scalability. As the IP5000 supports more storage controller modules it will be able to add storage capacity and performance. This flexibility provides for potential scalability of storage, cache, bandwidth and processing power.

- Since the IP5000 is built on standard hardware it can take advantage of advancements of these technologies including the continuous increase in processing speeds. This strategy helps to make the IP5000 future-proof.

- Intransa designed the IP5000 to be easy to use and typically complex operations have been made simple.

- The IP5000 performs competitively compared to mid-range FC storage subsystems based on ESG Lab testing of IOPs, bandwidth, media streaming and MS Exchange.

Appendix A

The majority of testing was performed using the configuration depicted earlier in Figure 1. Testing culminated with a test of a four storage controller module cluster with eight drive enclosures.

Table 1: Installation and Configuration Specifications

Installation and Configuration	
Intransa IP5000	Single SC enclosure containing two independent controllers running IP5000 code rev 1.1.0. Single DE enclosure, with sixteen 200 GB 7200 RPM ATA drives (3.2 TB raw capacity)
Servers	Dell 1650 , dual PIIIs at 1.26 GHz
Operating System	Windows 2000 Advanced Server, service pack 4
Internal SCSI drives in servers	Seagate 18 GB 10K RPM Ultra 3 SCSI
NIC	Dual built-in Intel NIC Pro/1000 XT
Switch	Dell PowerConnect 5224, Jumbo frames enabled
Exchange Performance Testing	
Workload Generator	Iometer 2003.05.10, 16 IO's per physical device
EDB worker	90% random 4KB I/Os: 73% read, 27% write 7% seq 4K reads 3% seq 4 KB writes
LOG worker	100% sequential 128 KB writes
SDB worker	67% sequential 128 KB reads 27% sequentia128 KB writes
Intransa IP5000	Single SC enclosure containing two independent controllers running IP5000 code rev 1.1.0. Single DE enclosure, with sixteen 200 GB 7200 RPM ATA drives (3.2 TB raw capacity)
Servers	Dell 1650 , dual PIIIs at 1.26 GHz
Operating System	Windows 2000 Advanced Server, service pack 4
Internal SCSI drives in servers	Seagate 18 GB 10K RPM Ultra 3 SCSI
NIC	Intel NIC Pro/1000 XT
Switch	Dell PowerConnect 5224, Jumbo frames enabled
Customer Performance Testing	
Servers	Dell 2640, Dual 2.4 GHz Xeons, 1 GB RAM
Operating System	Red Hat 2.3.20-19.7, Cisco iscsi 3.3.2
Intransa IP5000	Single SC enclosure containing two independent controllers running IP5000 code rev 1.1.0. Single DE enclosure, with sixteen 200 GB 7200 RPM ATA drives (3.2 TB raw capacity) configured as four StripeMirror targets.
NIC	Intel NIC Pro/1000 XT
Switch	Dell PowerConnect 5224, Jumbo frames enabled
TOE/HBA Testing	
Servers	Dell 1650
NIC	Intel NIC Pro/1000 XT
TOE / HBA	Misc.
Operating System	Windows 2000 Advanced Server, service pack 4
Switch	Extreme 11502 Sumitt 5i 1000BT, jumbo frames
Driver	Microsoft iSCSI initiator 1.0

Methodology for Exchange Testing

SCSI drives and Intransa volumes were tested as Physical drives using Iometer. Each disk target was configured for 16 outstanding I/Os. Three workers were used for each test: one for the database access specification, a second for the log, and a third ran the streaming media access specification. The access specifications are defined in Test Configuration table above. Results presented are for a two-minute run with a 5 second warm-up period.