



ESG Lab Review™

EMC MPFSi: Accelerating Network Attached Storage with iSCSI



**A Product Review
by
ESG Lab
May 2006**

**Authors:
Tony Asaro Brian Garrett**

Table of Contents

Introduction	3
ESG Lab Review	4
<i>Transparent to Users and Applications</i>	5
<i>The MPFSi Performance Advantage</i>	6
<i>MPFSi Performance Scalability</i>	9
ESG Lab Review Highlights	10
ESG Lab’s View	10
Appendix.....	11

ESG Validation Reviews

The goal of the ESG Lab reports is to educate customers about specific storage-related products including storage systems, backup-to-disk solutions, storage management applications, backup/recovery software, storage virtualization platforms, etc. ESG Lab reports are not meant to replace the necessary evaluation process that end user customers should conduct. ESG Lab reports are designed to provide insight to what is compelling about various products and how they can solve customer problems. ESG Lab provides third-party expert perspective based on ESG Lab analysis and interviews with customers using these products in production environments.

All trademark names are property of their respective companies. Information contained in this publication has been obtained by sources The Enterprise Strategy Group (ESG) considers to be reliable but is not warranted by ESG. This publication may contain opinions of ESG, which are subject to change from time to time. This publication is copyrighted by The Enterprise Strategy Group, Inc. Any reproduction or redistribution of this publication, in whole or in part, whether in hard-copy format, electronically, or otherwise to persons not authorized to receive it, without the express consent of the Enterprise Strategy Group, Inc., is in violation of U.S. Copyright law and will be subject to an action for civil damages and, if applicable, criminal prosecution. Should you have any questions, please contact ESG Client Relations at (508) 482.0188.

Introduction

Network Attached Storage Systems (NAS) understand files and metadata and SAN storage systems only understand block data, which has no meaning for users or applications. Users can share, recover, move, and access data more easily with NAS storage. However, because NAS is file aware, it operates at a higher layer and therefore introduces more latency into the read/write process. Because SAN storage operates at a lower layer it is blazingly fast.

For years the question has been: what to use, SAN or NAS? For many companies, the answer is both. While it is not a hard and fast rule, many companies use SAN for their database and e-mail applications and NAS for file storage. However, with the introduction of the EMC Multipath File System (MPFSi), companies can now use SAN and NAS not as discrete functions but as a truly unified solution. MPFSi combines the performance benefits of the iSCSI SAN protocol with the intelligence of NAS using a single Ethernet network.

EMC invented the MPFS protocol over six years ago and named the product HighRoad, which was developed for EMC customers with extreme performance requirements. HighRoad used the Fibre Channel SAN protocol to accelerate NAS data transfers. Due to the additional cost and complexity of running Fibre Channel connections to NAS clients, early adoption was limited to a select few customers who required more performance from a single NAS system than was available at the time.

However, the iSCSI protocol has emerged as an alternative to Fibre Channel for block based storage area networking. EMC was involved heavily in the development of the iSCSI specification and has added iSCSI support to their entire line of SAN and NAS storage systems. Over the past two years, the adoption of iSCSI in production systems in companies of all sizes and industries has grown dramatically. ESG believes that iSCSI is entering the early mainstream phase of market adoption.

EMC announced MPFSi in January 2006. MPFSi uses the iSCSI protocol instead of Fibre Channel to accelerate NAS access. While the combination of FC SAN and NAS requires two networks, iSCSI and NAS can share a single network infrastructure. Therefore ESG believes that MPFSi will ultimately address a much greater market providing SAN level performance with NAS level intelligence over a unified storage network.

Further, EMC has made available the MPFSi client software to the open-source community to encourage the building of applications using the technology. EMC is promoting industry standardization of the MPFSi protocol in the form of an extension to the NFSv4 specification called "parallel NFS".

MPFSi is innovative technology that combines the intelligence and ease of use of NAS with the speed of a SAN with the following benefits:

- Faster than traditional NFS
- Runs over existing Ethernet infrastructure
- Transparent to users and applications
- Windows and Unix client support
- State-less clients which enable mixed NFS and MPFSi access
- Existing Celerra NAS systems can be upgraded to use MPFSi
- Celerra blades support more clients, users and applications than with NFS

ESG Lab reviewed an MPFSi solution at an EMC facility in Southborough Massachusetts. The primary goal of the testing was validation of the performance boost of between two to four times compared to NFS with excellent performance scalability.

ESG Lab Review

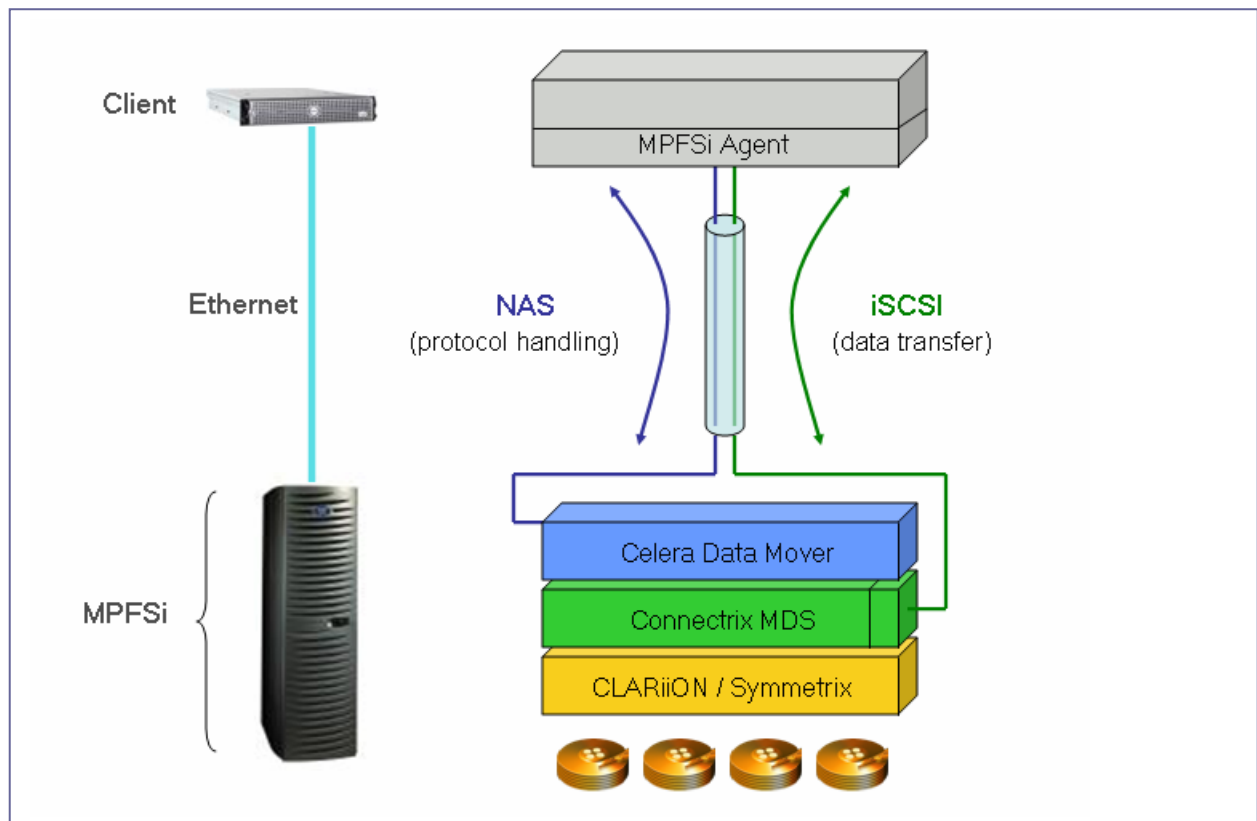
Before we examine the results of the ESG Lab testing, let's take a quick look at how MPFSi accelerates network attached storage with iSCSI. As shown in Figure One, a Windows or UNIX client accesses an EMC MPFSi solution over standard Ethernet wiring and switches. The moving parts of an MPFSi solution are:

1. MPFSi agent software installed on the client
2. A Celerra blade which maintains the file system and handles the NAS protocol
3. A Connectrix MDS storage switch with an IP Storage Service module that handles iSCSI data transfer
4. A CLARiiON or Symmetrix storage system

NAS protocol handling and file system awareness is managed by the Celerra. iSCSI data transfers are handled by an IP Storage Services module in a Connectrix MDS storage switch. IP Storage Services modules, available in four and eight port packages, move data to and from the clients using the iSCSI protocol, and to storage systems using the Fibre Channel protocol.

Blades and a storage system are standard items included in every Celerra NAS solution. Celerra file systems are upgraded to MPFSi by adding agent software and one or more IP Storage Services modules. The most significant cost associated with MPFSi for most customers will be the IP Storage Services Module.

Figure One: MPFSi Accelerates Network Attached Storage with iSCSI

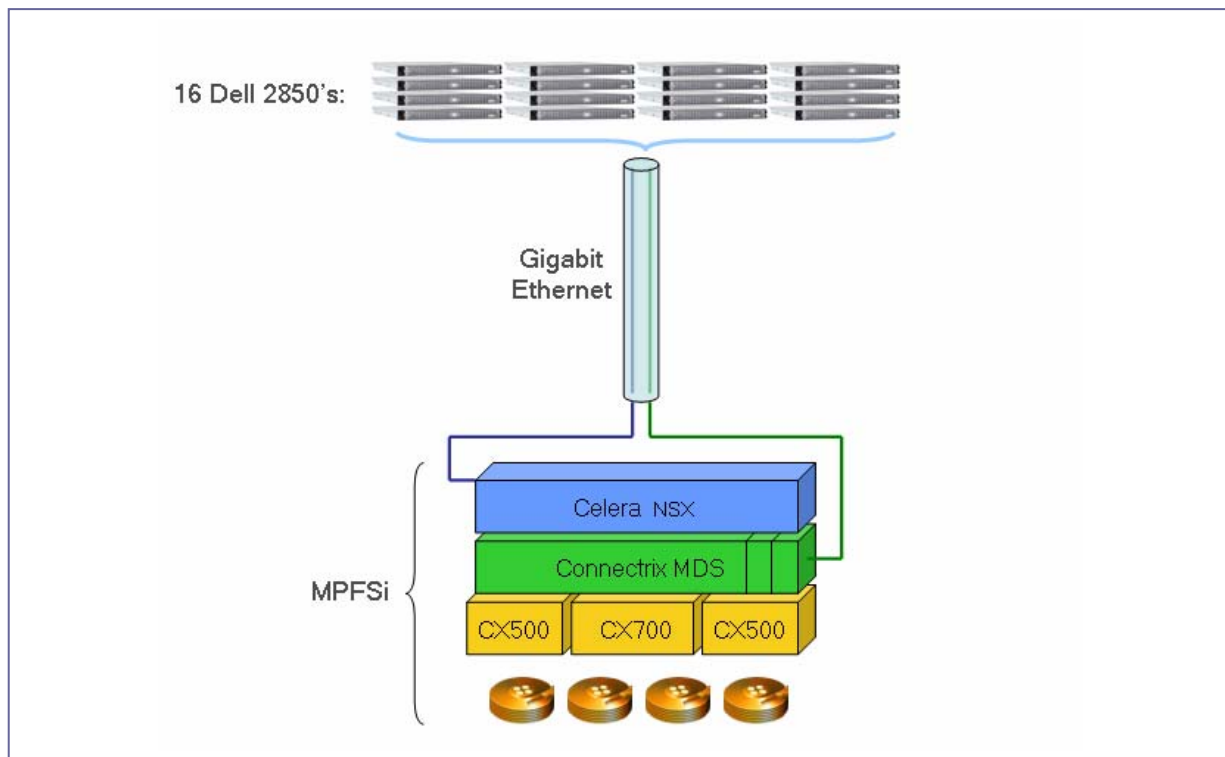


Transparent to Users and Applications

ESG Lab testing began on a pre-wired and configured test bed as shown in Figure Two¹. Sixteen Dell servers running Red Hat Linux were connected over Ethernet to a single Celerra NSX X Blade, two Connectrix MDS IP Storage Service modules, and three storage systems (a CLARiiON CX700 and a pair of CX500's)².

MPFSi agent and iSCSI initiator software was installed on each of the Linux clients so that file systems could be accessed using either the standard NFS protocol or the EMC MPFSi protocol. Switching back and forth between NFS and MPFSi was easily performed many times during testing. An option on the mount command (-mpfs) was used to turn on the MPFSi protocol. A remount of the file system without the option switched the protocol access method back to standard NFS. Operating system utilities and applications worked seamlessly and transparently with MPFSi compared to NFS. The only noticeable difference was the speed of MPFSi which will be examined next.

Figure Two: ESG Lab Test Configuration



Why This Matters

Users and applications can benefit from the performance boost of MPFSi without making any changes. Leveraging existing Ethernet infrastructure, administrators can manage which applications and users are configured for MPFSi. Mixed mode access with some clients using MPFSi and some using traditional NFS is supported so that administrators can get comfortable with the technology on a test client or two. Later when the administrator is ready, a remount of an existing production file system is all that is needed to realize the performance benefits of MPFSi.

¹ Configuration details are documented in the Appendix

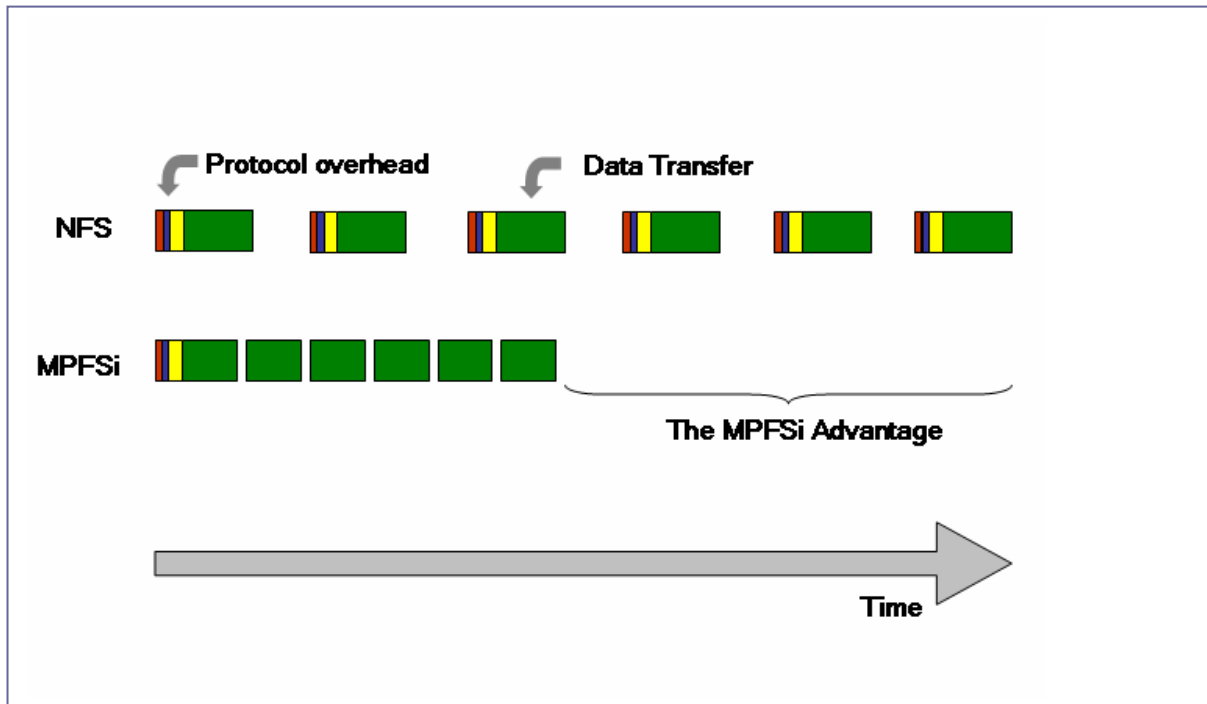
² The configuration is shown in Figure Two and documented in the Appendix.

The MPFSi Performance Advantage

NAS is fundamentally slower than iSCSI due to the chatty nature of high level NAS protocol handling. iSCSI operates over the same Ethernet infrastructure and is a fast and efficient data transfer protocol. The combination of iSCSI for fast data transfers, and NAS for file system protocol awareness is used in the MPFSi protocol to make a NAS system run faster.

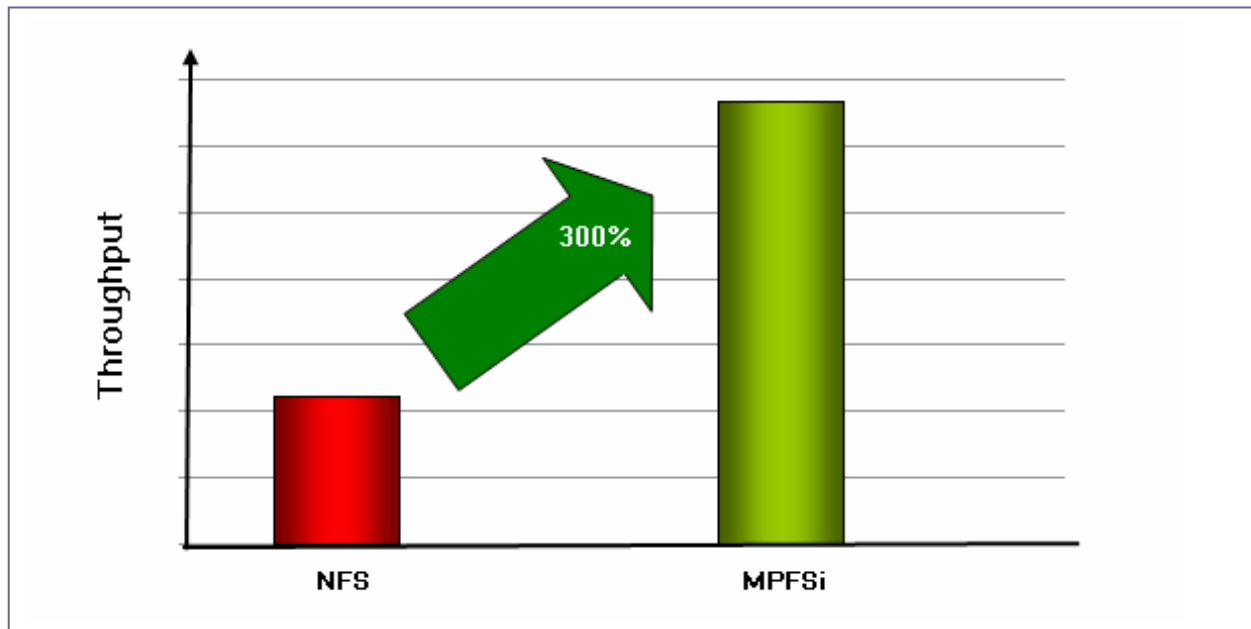
The following diagram shows how this works for a typical file access. The NAS system is burdened by metadata protocol overhead for each data transfer. In comparison, the MPFSi protocol incurs protocol overhead only at the beginning of each file access. After that initial handshake, data transfers run quickly and efficiently over iSCSI. Because MPFSi can transfer more file data over a shorter period of time, MPFSi has a performance advantage over traditional NAS access.

Figure Three: The MPFSi Performance Advantage



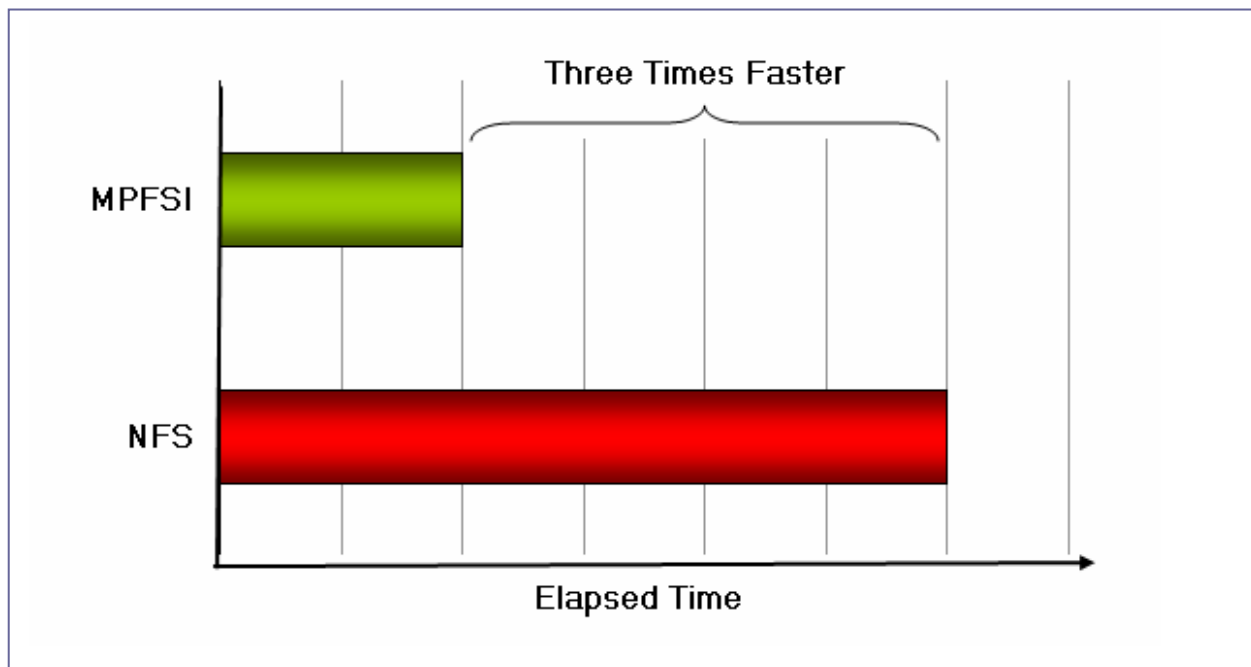
The performance advantage of MPFSi was measured using the industry standard Unix data dump (dd) utility. The utility was used to time how long it took to read files that had been previously created with a test utility. Sixteen Linux clients were used to read 80 MB files in one Megabyte intervals. As shown in the following diagram, the aggregate throughput rate of MPFSi was 300% faster than NFS.

Figure Four: The MPFSi Bandwidth Boost



The results above compare the aggregate throughput capabilities of MPFSi to NFS. These results indicate that many users, applications and processes can do more in parallel with MPFSi than with NFS. Another way of looking at the same performance results is shown below. This is how a user perceives the performance advantage of MPFSi - opening, saving and copying files on a shared network attached drive is three times faster with MPFSi.

Figure Five: MPFSi Moves More Data in Less Time



MPFSi is not only faster than NAS, it is more efficient. NAS protocol handshakes between NAS clients and an NAS servers consumes a lot of CPU cycles. Switching to the MPFSi protocol reduces the burden on NAS clients and the EMC NAS server. As a result, more clients can share the same Celerra NAS system. ESG Lab reviewed the results of EMC testing which indicates that MPFSi delivers a dramatic reduction in CPU utilization as shown in the following tables.

Table One: MPFSi Reduces NAS Server CPU Utilization

File Access Pattern	NFS CPU Utilization	MPFSi CPU Utilization
Reads	40%	3.0%
Writes	58%	1.2%

Table Two: MPFSi Reduces NAS Client CPU Utilization

File Access Pattern	NFS CPU Utilization	MPFSi CPU Utilization
Reads	14%	2.0%
Writes	25%	6.3%

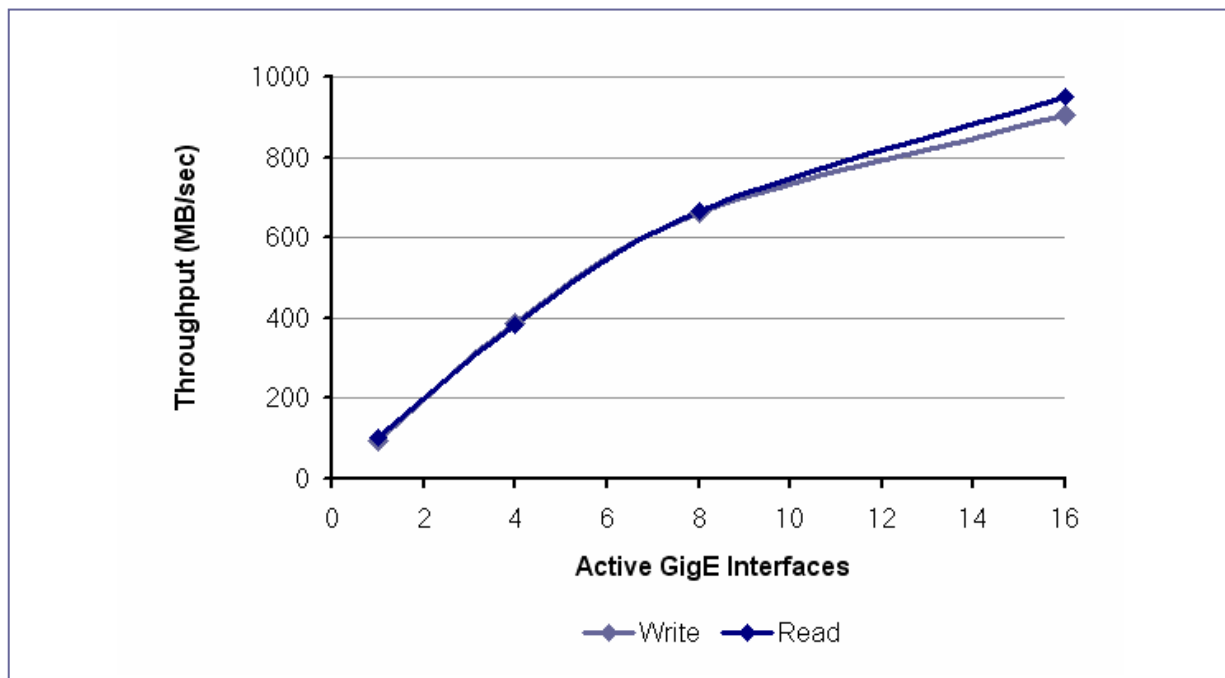
Why This Matters

MPFSi combines the performance benefits and efficiency of the block based iSCSI protocol with the simplicity of NAS using a single Ethernet network. ESG observed significant CPU efficiency gains which can be used to connect more clients to a single Celerra data mover. An MPFSi performance boost of 300% compared to NAS was measured. The bottom line - with MPFSi you can do more, with less data movers, and you can do it faster.

MPFSi Performance Scalability

ESG Lab performed a series of test results with a goal of determining the performance scalability of MPFSi. The EMC fctest utility was run from 16 Linux clients connected to a single Celerra NSX blade and three CLARiiON storage systems as shown previously in Figure Two. The utility was used to create files of various sizes which were then exercised using a variety of access patterns. The performance scalability of reads and writes, over large and small files, using random and sequential data access patterns was reviewed. In all cases, performance scaled in a near-linear fashion as the number of clients and Ethernet interfaces was increased. The following diagram shows the performance scalability of large files accessed sequentially. ESG Lab noticed that near wire speed throughput for large sequential reads and writes over a single Gigabit Ethernet interface scaled well as the number of interfaces using the MPFSi protocol was increased to 16.

Figure Six: MFPSi Performance Scalability



Why This Matters

Companies that rely on NAS storage systems that are shared by a large number of users and applications need predictable performance scalability. The performance scalability limits of legacy NAS systems have forced some companies to deploy many NAS systems to meet this need which increases complexity and cost.

ESG Lab verified that the modular architecture of MPFSi delivers predictable performance scalability from a single system. The number of blades, Connectrix IP storage services modules, storage systems and hard drives can be varied to meet a wide range of performance and capacity requirements.

ESG Lab Review Highlights

- ☑ ESG Lab verified that MPFSi is transparent to users and applications.
- ☑ A mount operation is all that was needed to switch between traditional NAS access and MPFSi.
- ☑ NAS and MPFSi were used to access the same file system.
- ☑ ESG Lab measured a 300 percent MPFSi performance advantage compare to NFS.
- ☑ Near wire speed throughput for large sequential reads and writes over a single GigE interface scaled well as the number of interfaces using the MPFSi protocol was increased to 16.
- ☑ MPFSi significantly reduced CPU utilization compared to NFS. EMC Celera NSX CPU utilization for random reads was reduced from 40% to 3%, and NFS client utilization was reduced from 14% to 2%.

ESG Lab's View

There is brilliance and innovation in MPFSi. MPFSi leverages the ease of use and more importantly the intelligence of NAS with the speed of a SAN - all over the same Ethernet infrastructure. Offloading NAS data transfer handling from an EMC Celerra blade to iSCSI is fast, efficient, and frees up CPU cycles on Celerra. ESG Lab measured an MPFSi performance boost of 300% compared to NFS when accessing the same file system. Freed from the burden of managing data transfers, Celerra data movers can handle more clients, users, applications and processes while delivering near line speed throughput with excellent performance scalability.

Early adopters with extreme NAS performance requirements will be the first to adopt MPFSi. Those users can pave the way for broader market adoption. Industry standardization and the open source availability of the MPFSi client will also be important and EMC is pushing this heavily. ESG believes that the standardization of MPFSi, along with the inevitable commoditization and consolidation of iSCSI processing currently provided by Connectrix MDS IP Storage Services modules, has the potential to drive NAS usage into high performance environments that today can only be served by SANs.

MPFSi has the potential to change the way we do storage networking. ESG believes that the world of storage will not be NAS. And it will not be SAN. It will be an amalgamation of both leveraging the strengths of each.

Appendix

Test Configuration

Clients	16 Dell 2850's, Dual 3 GHz, 4 GB RAM
Client operating system	Red Hat Enterprise Server, version 3
Data Movers	One Celerra NSX, DART version 5.5
Connectrix MDS IP Storage Modules	Two 8 port modules, 16 GigE ports total
Storage Systems	1 CLARiiON CX700, 2 CX500's, Flare 19
Hard Drives	120 FC drives, 10K RPM, 300 GB