

IP Storage in the IT Mainstream

By: David Dale, NetApp Industry Evangelist, Chair SNIA IP Storage Forum

Introduction

Over the past two years IT organizations have really accelerated their adoption of IP Storage solutions, as the technology entered the IT mainstream. Today, you can find iSCSI-based IP SAN deployments in IT production environments of large, medium and small enterprises around the globe, across all business sectors, including local and national government. The resulting existence of an ever-increasing number of customer references, together with broad platform vendor support is further accelerating the rate of deployment.

This article briefly summarizes the state of the IP Storage market, examines typical deployment scenarios, reviews performance, security and resiliency considerations, and finally looks forward to developments you can expect to see over the next 12 months.

State of the Market

Over the past year we saw enormous growth in the market for IP-based SAN storage technology, as leading server and storage vendors released products, and IT organizations around the world installed IP Storage solutions to consolidate, simplify, bullet proof, and reduce data management costs in their distributed business critical applications environments.

The market for IP Storage technologies is the fastest growing segment of the storage market, growing from an estimated 2500 IP SANs deployed at the end of 2004 to more than 10,000 IP SANs deployed by the end of 2005 – ramping up to more than 22,000 IP SANs by the end of 2006. This growth is expected to continue to accelerate through 2007 and beyond, as iSCSI-based SANs continue to replace direct-attached storage in the Windows market, and as a broad range of robust affordable IP SAN solutions becomes available all host OS environments.

The factors driving this growth have been the continuing need for IT organizations to do more with less – less capital cost, fewer admins per Terabyte, less complexity – coupled with continuing data growth, emerging best practices requiring strict and comprehensive data retention, and the recognition that even tier 2 applications are mission critical in today's 7x24 world. Interestingly, these factors affect enterprises of every size, and most IT managers now recognize that yesterday's direct-attached storage architecture simply cannot meet the needs of today's business environment.

Typical Deployment Scenarios

Initial IP SAN deployments were most often at the departmental level of larger enterprises, particularly in Windows environments comprised of smaller servers where limited admin support, host attach costs and infrastructure complexity have traditionally

inhibited Fibre Channel SAN deployment. This environment is typical of local, state and national government agencies, where IP SAN solutions have become increasingly popular.

iSCSI-based IP SANs have also become extremely popular in large organizations with either distributed resources or widely distributed data. The combination of affordability, an existing high-speed Ethernet infrastructure, and Ethernet-savvy IT administrators, coupled with array-based data protection and data management capabilities such as snapshots, remote copy and asynchronous mirroring, enable these organizations to easily build out efficient multi-location integrated data management environments.

In large enterprises, IP SANs are also popular where multiple separate data centers need to be affordably interconnected and where the management of corporate data needs to be integrated. Many companies have a large robust Fibre Channel SAN environment, but they may still have large numbers of smaller “stranded” servers each with their own internal or direct-attached storage. There are clear advantages in being able to leverage the existing SAN investment if an affordable way to connect these servers into the SAN can be found.

More recently, we have seen the adoption of IP Storage solutions in medium and even in small organizations – once again replacing direct-attached storage and supporting Windows-based business applications. In many cases these applications are mission critical, and data growth concerns, data availability concerns, the need to more efficiently manage the environment with existing admin resources, and limited IT budgets are the major customer concerns. Once again, this is an environment commonly encountered in local and state government.

Although most IP SAN deployments to date have been to replace legacy direct-attached storage, we have over the past year seen a trend towards IP SANs replacing first-generation (1Gb) Fibre Channel SANs. Typically here, the legacy SAN is approaching the end of its lifecycle, the customer needs a more cost-effective solution requiring less admin support, and the applications don’t need the additional bandwidth of an upgraded Fibre Channel SAN solution.

Performance, Security and Resiliency

Despite the growing number of deployments and more than three years of IT experience with the technology, there is a continuing need for education about iSCSI, particularly on the topics of performance, security and resiliency.

Performance

There seems to be a widespread perception, particularly among people with no hands-on experience of iSCSI, that choosing an iSCSI-based solution means you have to compromise on performance. However, the reality is different. iSCSI solutions using standard 1Gigabit Ethernet NICs and the free software initiator that comes with the host operating system provide perfectly acceptable performance for the vast majority of enterprise applications. Disk arrays connected by 2Gb Fibre Channel are not twice as fast as 1Gb iSCSI arrays, and 4Gb Fibre Channel arrays are not four times as fast.

The Enterprise Strategy Group has done extensive comparative performance testing using typical real-world application workloads, and they consistently find that 2Gb

Fibre Channel solutions typically deliver only 5-15% better performance than 1Gb iSCSI solutions using software initiators. For more information, see <http://www.enterprisestrategygroup.com>.

The misconception here is that storage performance is proportional to storage interconnect bandwidth. It's not. A useful analogy would be a car travelling along a road. Adding more lanes to the road doesn't make the car go faster unless there is significant congestion. And there is plenty of bandwidth to spare for most application workloads even at 1Gigabit.

Other factors have a much more profound affect on the performance of the array. The number, type and rotational speed of the disks has the biggest impact – large arrays deliver more performance than small arrays; Fibre Channel drives provide more performance than a similar number of ATA drives. Then the ability of the array to optimize data placement, and its ability to stripe application data over the largest number of disks, also have a significant impact on performance. Interconnect bandwidth is a third-order issue compared with these, and relatively insignificant.

Security

Some storage professionals view Ethernet as a storage interconnect as less secure than Fibre Channel, which is strange given that Ethernet has a plethora of security features and options available (from IP addressing/masking to VLANs to Firewalls to CHAP, SRP, Kerberos, SPKM, and various levels of IPsec protection and encryption) and Fibre Channel has no built-in security features, and fairly limited set of device-specific options (LUN masking, zoning, etc). In addition, iSCSI performs host/target authentication on a per-session basis as part of its standard operation. .

Security however is often simply a matter of strictly adhering to a “best practices” policy. For IP SANs, today's best practices advice is to create a private data center storage network separate from the data communications network – either by using separate physical Gigabit Ethernet switches, or by using VLANs and the existing Gigabit Ethernet infrastructure. This provides the base level of security that most environments require. Then you can make decisions on whether encryption of data in flight or on tape or disk is warranted.

Resiliency

In the SAN world we have become accustomed to building infrastructure with no single point of failure. In particular host I/O paths are always configured for redundancy, with host multipathing drivers ensuring timely failover in the case of path failure. One of the reasons for this requirement is the fact that a Fibre Channel transport has no “retry” facility. If a response to a disk I/O operations is not received before some preset timeout period, the host system assumes the disk is not available, and this usually leads to a graceful or non-graceful application halt.

This requirement isn't so hard and fast in the case of iSCSI, since TCP/IP does support “retries”, and iSCSI initiators can be configured to take advantage of the facility. Consequently, early iSCSI solutions typically used Ethernet “teaming” and this facility rather than the more robust multi-path driver method. More recently, however, OS vendors have started to qualify their entire storage stack for iSCSI, including their native multi-path drivers and host clustering software.

Today, in the Windows environment, you can configure an IP SAN environment which is every bit as resilient as its Fibre Channel SAN equivalent. In addition, the Unix vendors and the Linux community are now releasing native storage stack support for iSCSI, including multipathing and clustering, so equivalent resilience will be available this year for those environments.

What's Next for IP Storage

There are still very large amounts of direct-attached storage in IT environments of all sizes. IP SAN solutions are expected to continue to convert these to networked storage environments. In addition, however, new IP Storage capabilities will further expand the reach of IP Storage solutions. The following are among the more interesting of these new capabilities:

- Complete native iSCSI solutions for UNIX and Linux (including multipathing and host clustering), which will drive iSCSI into the mainstream of departmental UNIX and heterogeneous OS environments.
- Ubiquitous boot-from-SAN solutions driven by the availability of software iSCSI SAN boot from Microsoft and other vendors will drive explosive growth of IP SAN solutions for high-density and blade server environments. There is a great deal of interest in this capability.
- New data-intensive video analysis and forensics applications requiring large amounts of simple, affordable SAN storage will broaden the application base for iSCSI
- The emergence of a defacto standard host interface architecture for next generation TCP/IP Offload Engines and storage systems with 10Gigabit Ethernet connections during 2006 will drive IP SANs in high-performance and large-scale aggregated storage environments

IDC views iSCSI as the fastest-growing area of the storage market today, and they expect a 211% compound annual growth for 2005-2009, with the market reaching \$2.65B in 2009. Given the momentum of iSCSI in the marketplace today, many observers believe that these numbers may be quite conservative.

Conclusion

IP Storage is rapidly becoming a mainstream network storage solution. IP Storage solutions are available from a significant number of vendors to extend and complement existing networked storage environments. The emergence of IP-native SAN solutions using iSCSI means that IT professionals now have SAN storage solutions for a much broader range of environments at a broader range of price points than ever before. Storage consolidation, simplified data protection, affordable disaster recovery, and improved data management are the major focus of IP Storage deployments today – many of which are occurring in government sectors.

The time is now right for all IT executives to take advantage of the new networked storage opportunities available for them, and find out how IP Storage can address some of their most critical data management pain points.

For more information on the IP Storage Forum, visit www.snia.org/ipstorage