Top 10 Storage Technologies

for Systems and Network Professionals

By Vic Dery and Ashish Nadkarni

INTRODUCTION

This article provides an overview of some of the hot topics in the storage industry that all Systems and Networking Industry Professionals need to be familiar with. While some of them are relatively new and their future is unknown, most others have been around in some shape or form and are morphing into commonplace technologies. The order in which these topics are discussed is not indicative of their value. The information provided is also not vendor specific and may or may not be a fetcher of all vendors. As this industry continues to grow, technology that is changing and growing to fill the voids and gaps in the storage industry are always going to be there. As with any of these technologies the proper investigation and research prior to implementation or testing would be best.

IP SAN TECHNOLOGIES

IP SAN Technologies are technologies that use IP as data transport for storage. IP SANs have evolved from the need to use the benefits of a serial transport like IP with the functionality provided by SCSI-3. There are three main IP SAN protocols—iSCSI, Fibre Channel over IP (FCIP) and the Internet Fibre Channel Protocol (iFCP). We will discuss each of these briefly.

iSCSI—iSCSI or the Internet SCSI protocol tunnels SCSI-3 over IP in a manner similar to Fibre Channel Protocol which tunnels SCSI-3 over Fibre Channel. Both Fibre Channel and IP are serial transport media. The growth of iSCSI networks is largely due to Gigabit Ethernet. Before the availability of GigE (as it is popularly known), the only serial transport that could provide speeds fast enough for storage services was Fibre Channel. While Fibre Channel speeds today are still faster than GigE, one of the big disadvantages of the former is the requirement to use special hardware interconnects or adapters (known as host bus adapters) for tunneling SCSI-3. GigE has no such restrictions although special cards known as TCP offload engines and iSCSI Host bus adapters are available. This is largely due to the fact that the TCP/IP stack is well evolved and CPU and memory overheads required by iSCSI are insignificant. iSCSI can be implemented using two methods—native and bridged. In native iSCSI the storage device can communicate iSCSI directly over the network while in bridged mode an appliance performs the translation. The latter method is often deployed in situations where exiting capacity on non-iSCSI storage arrays needs to be used.

FCIP—fibre Channel over IP is a tunneling protocol that allows the extension of Fibre Channel based SANs over geographical distances greater than those supported by native Fibre Channel. This allows the extension of such networks for Business Continuity and Disaster recovery purposes such as replication, remote backups, and datacenter resource consolidation.

iFCP—The applications of iFCP are very similar to FCIP; however the protocol stack is completely different. While FCIP is a tunneling protocol (Fibre Channel frames are encapsulated in IP frames at the source and un-encapsulated at the destination; as a result the Fibre Channel communication end-to-end is maintained), iFCP is a gateway protocol that replaces the transport layer of the Fibre Channel FC-2 layer with an IP network (i.e. Ethernet), but retains the upper layer (FC-4) information, such as FCP (SCSI-3 over Fibre Channel).

STORAGE VIRTUALIZATION

Fabric or network virtualization: Network virtualization has to date remained somewhat of a religious argument between systems and network administrators. To date all of the volume management functions such as mirroring, striping, concatenation etc., are host based. Even though enterprise arrays provide some of these functions, at the end of the day storage presented to the host almost always undergoes some sort of a "Volume management transformation" before it is presented to the application stack in the form of a file system. Even though volume management software has become more advanced than the first generation

versions, some of the limitations posed by them still remain. For example, a systems administrator in a large environment may find it challenging to apply volume manager patches or perform upgrades across multiple hosts. Multiple operating systems make this task even more challenging. While most vendors bundle volume managers with their Operating systems, advanced features often require additional licenses that are cumbersome to manage. Now imagine a situation where all of these functions are managed in the network in a platform agnostic. Now suddenly, all volume management tasks across multiple hosts and/or multiple arrays can be performed in a centralized fashion with little or no intervention on the host side. Fabric virtualization has been somewhat slow to be adopted across the enterprise, but the technology is here to stay.

Virtualization in the array: Array based virtualization is somewhat misleading. Hardware RAID, which most modern arrays are based on, can be thought of as "disk virtualization." However, vendors like Hitachi Data Systems and Sun Microsystems have added a layer of abstraction in their arrays which allows them to virtualize volumes or logical units (LUNs) on a third-party array and present it as native volumes. The HDS TagmaStore Universal Storage Platform for example, allows you to virtualize volumes on an EMC Symmetrix or an EMC CLARiiON array and present it to a host as HDS native volumes. While this feature is very similar to fabric virtualization that we discussed above, it differs from the former because of the fact that native array technologies can be applied to virtualized volumes as well. For example, the HDS USP allows one to replicate EMC volumes to a remote location using TrueCopy, a technique that is ordinarily available in an HDS array only.

BACKUP TO DISK AND VIRTUAL TAPE LIBRARIES

Gone are the days that backups were only performed to tape. Known in the industry as B2D and VTL, this is one of the technologies that most major vendors have started to explore or at least partner for how to grow this technology. The widely accepted usage of these types of technologies is growing because of several factors including, compliance, speed and amount of information being retained. Rather than traditional retention of information being stored on tape, disk arrays are being used as the centralized storage point for retention and recovery. The ability to emulate storage libraries and different tape devices can replace up to 32 libraries with a single solution. These solutions hold numerous terabytes of information in a single array solution. Although we seem to have a perfect solution there are some additional steps and features to insure recovery and management are done correctly.

The CLARiiON Disk Library from EMC and Sepaton are examples of virtual tape libraries. In essence the concept used in a VTL is very simple: It presents a bunch of disks as a bunch of virtual tapes. It allows companies to perform backups much faster than sending them directly to tape and then in the background de-staging to tape can be performed as a cloning operation. Thus the impact to production sites because of backups is minimized. This concept is slightly different from backup to disk where data is written directly to disk either in the form of a mounted file system or as a raw volume. VTLs allow companies to seamlessly integrate disk into existing backup environments with little or no changes.

CONTINUOUS DATA PROTECTION

Data protection by journaling everything the system writes. This allows for a selected point in time rolling back the system. Several areas where CDP is getting the most traction are Database and Exchange recoverability. The ability of the administrator to have more content control and data protection on a real time environment is appealing to application owners. These products will not replace the traditional backups and should not be confused with snap-shot type technologies. The difference between CDP and Snap-shots is that the CDP technology is a constant collection of the information changes in the system. Snap-shots are based on changes from a single or multiple points in time (There are some minor drawbacks depending on the vendor of choice).

NAS OR FILE VIRTUALIZATION

The concept behind NAS virtualization is simple—it is the ability to provide a single global (and virtual) file system across an organization. Anyone who has administered a NAS environment will attest that the single biggest limitation of network file system (NFS or CIFS) is the inability to scale beyond an individual server or appliance, thereby introducing additional costs and complexity when storage capacity needs to be expanded or re-allocated. NAS virtualization appliances address this key shortcoming by presenting a file system that virtualizes individual file servers and appliances behind the scene. Modern file virtualization appliances like Acopia go a step further and present a global namespace that allows administrators to break the static mapping that exists between users and file system resources like files and directories.

SAN FILE SYSTEMS

This concept in the storage technology world is in theory what every multi-O/S Storage Areas Network has been looking for. This would be the ability for multiple operating systems to communicate regardless of the operating system and the server that owns the file and share. If it sounds like CIFS or NFS, it is a similar concept but those protocols are not used to perform this function. As with all good solutions, there is a price for the convenience to make this SAN communicate this way. The SAN speed with the ability to cross communicate within a SAN is being pushed by several vendors. Although this product is not going to be a fit in all Storage Area Networks it will help those challenged by the problems administrators face in sharing disk resources between multiple platforms. Since this technology is fairly nascent, we are presenting a few innovations that may take it to the next level.

Object Store is a concept pioneered by IBM and now a TWG (Technical Working group) at SNIA. It is well on its way to becoming a SNIA standard. Traditional file system access in a SAN is host based—filesystems allow hosts direct access to block devices in the SAN. All file systems hooks such as IO access, file layout, permissions etc. are host based—if a host sends corrupt data to the block device the file system is no longer usable. In this new scenario, these typical activities are moved to the storage device itself. In this new scenario host access is performed through a standard object interface rather than a traditional block-based interface such as SCSI or IDE.

HOST VIRTUALIZATION

Host virtualization is not really a new concept. It is the ability of a system to create virtual platforms, each running an independent instance of the same or different operating system. This operating system uses and shares the same hardware resources as the parent or host OS but functions independently within a set of defined operational

footprints such as CPU and memory resources. Examples of host virtualization are VMWare, Solaris Zones and Microsoft Virtual server/PC. Virtual host platforms allow one to make efficient use of existing processing capacity. They also allow you to provision processing capacity on demand and containerize applications so that they can function within pre-defined boundaries.

WIDE AREA FILE SYSTEMS

WAFS systems are being deployed to ease the network challenges of using file servers in remote offices or environments across the WAN. With the growth of companies in a global market place companies are looking to make ease of collaboration without increasing network problems and adding server purchases. Different vendor products boost methods on how their product works to optimize the WAFS. Optimization across the network is done by utilization of large amounts of cache to retain information so that WAN communication is limited to only transactions not stored. When it does need to send information across the WAN each system or appliance will compress the data prior to sending it across the WAN. The ability to file lock while changes are made is present to keep files from being corrupted, along with using Active Directories to secure the files access.

CONTENT ADDRESSABLE STORAGE

CAS is an object-oriented system coined by EMC in 2002 with the release of Centera product line. CAS is a great way to address fixed information, or information that is written but never changed. Some common examples would be X-rays, individual emails, financial invoices, and just about any piece of data that once it is written would require a new file rather then an update, like a database. Regulations like HIPPA require medical information to be retained for periods outlined by law. Since CAS places unique identifiers to each file, this gives the ability to control the content and information with policies. The CAS design uses a concept known as RAIN or redundant array of independent nodes. Inside a CAS box are multiple independent nodes (1U rack servers) that run a tuned operating system. There are two types of nodes inside a typical CAS environment—storage nodes and access nodes. The role of storage nodes is to store data while access nodes are used to retrieve and deposit data into the system. A CAS system can be accessed in one of two ways—using an API that is integrated with an application like Microsoft Exchange or using a standard interface like a CIFS share. Additional software can be used to control policies such as retention period.

STANDARDS BASED STORAGE MANAGEMENT

Storage management is not without its challenges. These challenges are more pronounced in a large heterogeneous environment. When Fibre Channel Storage area networks became popular, each vendor went about their own way in implementing an interface (or API) to manage its own storage arrays. This was an excellent interface when all storage devices belonged to the same vendor. However, that is not always the case and the problems faced by storage administrators in managing heterogeneous environments became evident. To address these and other inter-operability issues the Storage Networking Industry Association (SNIA) was formed. One of the first charters of SNIA was to create standards for inter-operability and storage management. Thus was born the Storage Management Initiative (SMI). The main goal of this initiative is to develop and standardize interoperable storage management technologies, and to aggressively promote them to the storage, networking, and end user communities. SMI created the SMI-S or Storage Management Interface Specification. It provides the architecture required for fully interoperable Storage Area Networks (SANs). The goal of SMI-S is to provide developers and end users alike with the ability to deploy equipment out of the box that functions in a seamless storage network, populated with a variety of hardware and software vendors, without concern for compatibility.

SMI-S attempts to solve the problem of interoperability by replacing existing disparate models, objects and protocols with common models for each object class, and a common protocol for management interactions. Models and protocols in SMI-S are platform-independent, enabling applications to be developed for any platform, and to be run on different types of platforms to interoperate. SMI-S models are extensible, enabling easy addition of new devices and functionality to the model, and allowing vendor unique extensions for value-add functionality.

Companies like AppIQ with their Storage Authority Suite are aggressively pushing standards based storage management to new levels. After all, Storage is one big happy family.

CONCLUSION

All of these technologies have their place in the current SAN world. The fact is that some of them actually compliment each other and allow for users to push beyond boundaries in their current environment. Depending on your needs, size and requirements, some of these technologies may or may not help. The true test of these products is to look at what you have, define your goals and see which technology will allow you to reach them with reasonable cost, ease of management and best ROI.

REFERENCES:

Acopia (http://www.acopia.com)—NAS Virtualization
ANSI (http://www.ansi.org)—ANSI Standards
AppIQ (http://www.appiq.com)—Standards based Storage Management
Brocade (http://www.brocade.com)—FCIP, Fibre Channel routing
IBM (http://www.ibm.com)—Object Store based File systems, Storage Virtualization
SNIA (http://www.snia.org)—IP SAN Technologies, SMI-S
IETF (http://www.ietf.org)—IP SAN Technologies
EMC (http://www.emc.com)—CAS, Storage Virtualization
Cisco Systems (http://www.cisco.com)—IP SAN Technologies, Storage Virtualization, WAFS
Sun Microsystems (http://www.sun.com)—Solaris Zones
VMWare (http://www.vmware.com)—VMWare
Microsoft (http://www.microsoft.com)—Virtual PC
T11 (http://www.t11.org)—Fibre Channel standards
T10 (http://www.t10.org)—SCSI interfaces

NaSPA members Vic Dery and Ashish Nadkarni work for GlassHouse Technologies.