Storage Appliances: Information Appliances for Data Storage

ATTO Technology, Inc.
155 CrossPoint Parkway
Amherst, NY 14068
716-691-1999
www.attotech.com

SECTION 1 *ABSTRACT*

Information Appliances have become increasingly popular in both large- and small-scale networks. These devices are generally inexpensive and limited to performing one or two essential functions of a network or system.

A similar type of device is emerging for storage applications. Perhaps best described as a Storage Appliance, these devices will perform storage-related activities usually found on a general-purpose server such as tape backup, fail over, security, or license control. The advantages to this approach are increased scalability and performance. Storage Appliances will also lead to storage becoming more independent from file servers, file systems, or operating systems.

SECTION 2 *WHAT IS A STORAGE APPLIANCE?*

It is important to define what a Storage Appliances is and isn't. As is common for new technology, it is easier to define what a Storage Appliance is not, rather than what it is. *It is not a server in the traditional sense*. While it is possible to use Storage Appliances in a client-server environment or have them act the part of a server in a system, their use is not dependent on this particular architecture. *It is also not a general-purpose computing device*. On the contrary, it is specific in terms of its tasks and functionality. Storage Appliances serve only one function or set of functions. It is not a Network Appliance per se. While it can be argued that a Storage Appliance uses a network (Fibre Channel, Ethernet) to communicate with other devices, it is not required that it do so. It is truly oriented toward storage functions and not networking, Internet, or general application tasks.

Just what *is* a Storage Appliance then? A Storage Appliance can best be described as a standalone-computing device whose purpose is to assume storage related functions currently found on a general-purpose server. Storage Appliances could perform backups, handle license management or storage management, implement a security model, and perform file and volume access control for the storage. It is possible that a Storage Appliance will even provide an OS neutral interface to the storage file system. This would then create an *inherent* cross platform file capability, a function currently available only through server-based add-on software. By providing this function in the storage or storage sub-system infrastructure, storage would become completely independent of the operating system.

Take this idea a step further, where application level interfaces – such as database access – are implemented in a Storage Appliance and you have specialty storage devices that can work in any environment for any platform. This feature would be especially useful in distributed computing architectures such as Jini environments.

SECTION 3

WHY DO WE NEED STORAGE APPLIANCES?

A product is only useful if it solves some particular problem for a customer. While a Storage Appliance could be deployed to overcome many *specific* system constraints, Storage Appliances, as a class, solve many *general* problems for system architects. They are:

- Lack of scalability in storage sub-systems
- Cost, especially Total Cost of Ownership (TCO) for high-end computing systems
- The need for increased reliability
- Limited support for distributed and open systems

Scalability

Several factors inhibit scalability in storage sub-systems. These include:

- Cabling limitations
- Lack of connectivity
- Reliance on the server
- Bandwidth Management
- Lack of a common interface to storage

Fibre Channel addresses many of these issues at the physical layer. Unfortunately, it does nothing to solve these problems at the application layer. It is no different than trying to implement a program written for a single user to many people over a LAN. The physical architecture is there, but the application support is lacking. Other than cabling and access issues, Fibre Channel by itself does not solve the problem. It does however provide the correct environment for applications and application level devices such as a Storage Appliance to provide a full solution.

The Storage Appliances concept addresses these issues by making the storage function independent of the server. By offloading server functions to the storage itself, in the case of an embedded Storage Appliance, or to an independent device that front ends the storage, the storage becomes less reliant on servers, operating systems, or networks. Storage systems can then be deployed in amounts and configurations that suit the system and not the server. This allows the storage system to be more easily managed and scaled. This is a natural requirement for distributed systems where scalability and independence are linked directly.

Cost

One of the more important reasons for the creation of Storage Appliances is to reduce costs in the storage area network and overall system. Besides collateral cost benefits derived from the ability to create systems in a new and more efficient way, the Storage Appliance brings direct cost savings to the system.

The first way is in the device itself. Since these are specialized devices that only need to be built to do one thing, they can be built much less expensively than a general-purpose server that must be capable of supporting several applications. Less expensive processors and operating systems that are tailored to a small number of applications are just two ways these devices can be made less costly than a standard Windows NT server is.

The second way and easily the more important, is the Storage Appliance's ability to reduce a system's Total Cost of Ownership (TCO). Increases in TCO are a major focus of the IT industry and devices such as the Storage Appliance are capable of brining them in line.

Aside from a lower initial capital cost, the Storage Appliance reduces TCO by being a simpler device. This reduces the amount of maintenance necessary to keep it up and running. In general, if a Storage Appliance fails, you remove it and replace it. Since the software is embedded in the device, updates are quick to internal flash ROM memory. Configuration is easy and simple since limited functionality means limited configuration options. Many require little management or maintenance past initial configuration, which reduces the need to add additional systems personnel to manage these devices.

This is a norm for Network Appliances but new to storage devices where maintenance and management costs often quickly outstrip initial acquisition costs.

Reliability

A simpler device is a more reliable one. Fewer things can go wrong than with a server running a high-end operating system. The more reliable the devices, the more reliable the system. As a single purpose device, a Storage Appliance will have fewer instances where outside influences can disrupt its operation. In a typical NT or UNIX server, the variety of hardware, network attachments, applications and utility software, and the operating system itself all represent points of failure. They also represent parts of the server that can be changed and manipulated by humans, which can introduce failure. The closed box nature of the Storage Appliance makes it much less likely that human intervention will result in failure of the device.

Support for Open and Heterogeneous Systems

Storage Appliances are SAN devices and easily designed around industry standard protocols. This makes them naturally OS independent and well supportive of open systems. By moving functions that are normally housed in an OS dependent server, Storage Appliances make those functions accessible to the entire SAN rather than those devices capable of interacting with a specific operating system.

Storage Appliances also support open and heterogeneous solutions by providing proxy devices for more closed systems. Many legacy storage devices are tied to specific server technology by the need to accommodate OS specific file systems on the disks or tapes. A Storage Appliance can provide a proxy for the networked storage that allows different operating systems to interact seamlessly with whatever file system is on the storage itself.

Currently, many storage product vendors are embedding servers, usually NT servers, into their products. This is typical of Network Attached Storage. While this is a popular approach, and appropriate in many instances, it ties these products directly to a single operating system. This does not support truly open and heterogeneous architectures. A Storage Appliance proxy device, on the other hand, would use no particular OS and instead force reliance on standard protocols such as IP or SCSI.

The Storage Appliance Relationship to the SAN

Storage Appliances are an essential part of the SAN. They provide control points, security, and distributed functioning to the storage network at low cost. These devices are the “R2-D2” of the SAN. Like the little ‘droid in “Star Wars,” the Storage Appliance will perform a variety of specialized housekeeping, interface, and data movement activities in the SAN, keeping it stable and efficient.

The Storage Appliance increases the value of storage in the SAN because it off-loads server storage-oriented functions to the storage device itself. Just as network computing devices relieve the LAN from reliance on the server; the Storage Appliance will relieve the SAN from reliance on the server and make data storage an independent sub-system.

SECTION 4

STORAGE APPLIANCES ARCHITECTURE REQUIREMENTS

Many different Storage Appliance architectures are possible. Most designs and architectures will be application bound in some manner. However, some general elements, such as a hardware platform coupled with system and application software, are likely to be required of a Storage Appliance device regardless of the application.

Hardware

In thinking of what a Storage Appliance hardware platform might look like, it is important to keep in mind that this is an *Appliance*. This places the following platform requirements on the hardware:

- High speed I/O Ports
- Support for SAN environments
- Large amounts of internal bandwidth; a high-speed international bus
- Large amounts of fast memory for buffering and application use
- An internal processor

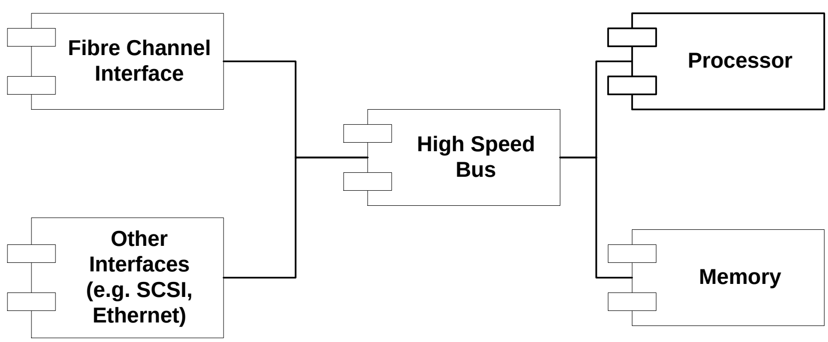

FIGURE 1. Storage Appliance Architecture

Owing to the need for high-speed I/O and support for networked storage environments, there are few I/O interface choices outside of Fibre Channel. The internal processor is necessary to host the OS and Application software that makes this a Storage Appliance in the first place. In other words, this cannot be a passive or "dumb" device.

Software

Software can vary greatly depending on the type of Storage Appliance that is being designed. Some obvious elements of the software architecture are:

- An embedded operating system or virtual machine
- Application specific programming
- High speed I/O interfaces

The Java VM is an example of a type of embedded OS or VM that may be useful for many applications. A real-time OS such as VxWorks is another possibility. The choice of OS will depend on the application that is to be housed. What is important is that the OS is fast and flexible. What it can't be is "fat." A general purpose OS would provide too much functionality, be too slow, and defeat the purpose of the Storage Appliance.

SECTION 5

WEAKNESSES OF THE INFORMATION APPLIANCE APPROACH FOR STORAGE

As with all technology and approaches to technical problems, the Storage Appliance concept has some plusses and minuses. Overall, the pros outweigh the cons, but it is important to keep in mind the weaknesses of this approach when designing storage networks.

One core weakness is that the management of many small independent devices can be challenging. In a small SAN, this is not much of a concern. However, in a large distributed environment, managing these devices can become difficult. Luckily, many of these devices will ship with support for network management software. Already we are seeing the emergence of software from providers such as Computer Associates and Tivoli that address SAN management.

Another drawback to this approach is that they are essentially closed boxes. This means that you are tied directly to the manufacturer of the product. Unlike a server-based system, there will be few instances where independent software will be available for these units.

SECTION 6 *STORAGE APPLIANCE IDEAS*

While the range of storage related information appliances is limited only by what is needed in a SAN and imagination, it is easy to conceive of several important devices immediately including:

- A Copy and Data Movement Device
- Volume, File, and LUN Management
- An OS Independent File System
- A Jini/Jiro Proxy Devices for Legacy Storage

The Copy Device

One of the first Storage Appliances is the data mover or copy device. This product has already been implemented, usually by manufacturers of SCSI-to-Fibre Channel bridge products. This type of device does the actual work of moving data between different storage devices within a Fibre Channel SAN. The first application for this has been serverless backup. At this time, software external to the copy device is needed to tell the copy device where to move what data and to insure that a current catalog of the data is up to date. In time, the control functions will move to the copy device itself making it a full-fledged data mover.

Volume, File, and LUN Management

SANs represent a unique problem for the system administrator – shared access to storage resources. Generally, there is no protection built in to Fibre Channel or most operating systems to regulate read and write access to files and storage volumes. This often leads to corrupted data and overwritten files. Currently the only solution is workstation and server-based file and volume management software. A volume manager would take on the function of traffic cop for files and volumes, regulating read and write access in real time to disks and tapes.

Similar to the File Manager is a LUN Manager. Operating at a lower level, a LUN Manager Storage Appliance would provide policy-based access to specific LUNs in a SCSI-based (SCSI-FCP) SAN. It might also include the ability to hide LUNs from specific client applications, pool LUNs into groups of virtual LUNs, and other low-level storage functions.

An OS Independent File System Device

One of the most important uses for Storage Appliances technology will be in implementing file systems independent of the operating systems with which it wishes to interact. In this scenario, a high-performance file system will be desirable for accessing on-line storage or use of a single file system will be beneficial in a heterogeneous environment. In both cases, the file system deployed is divorced from the operating systems available on the system's servers. Currently, cross OS access to a single file system is achieved with server-based software. This limits the deployment to known and supported file systems, as opposed to new file systems, and requires extensive management of the software, which often takes up significant server resources. Another popular approach is to use an OS neutral file system such as NFS or the new General Purpose Files System that is being researched. Again, these suffer from the need to have OS dependent software reside on the server.

The Storage Appliance-oriented approach would place the file system interfaces inside a File System Storage Appliance. In this way, a standard protocol such as IP could be used to get data from the device, which in turn would translate from an on-board file system. In addition to what Network Attached Storage products do now, the Storage Appliance could also translate into the native OS File Systems while providing a different file system on the storage.

A Jini/Jiro Proxy Devices for Legacy Storage

Jini, the distributed network software model from Sun Microsystems, has the potential to be a major force in computing. The base requirements for participation in the Jini network, or Djinn, are that the device support IP and have a Java virtual machine with the appropriate Jini classes. Jiro, Sun's Distributed Storage Management architecture based on Jini also has the same requirements.

Unfortunately, virtually no legacy storage devices support these requirements. Yet, with IP over Fibre Channel a reality, and quickly becoming more common, it is inevitable that system architects will want to deploy this technology over Fibre Channel SANs. The solution is a Jini proxy device.

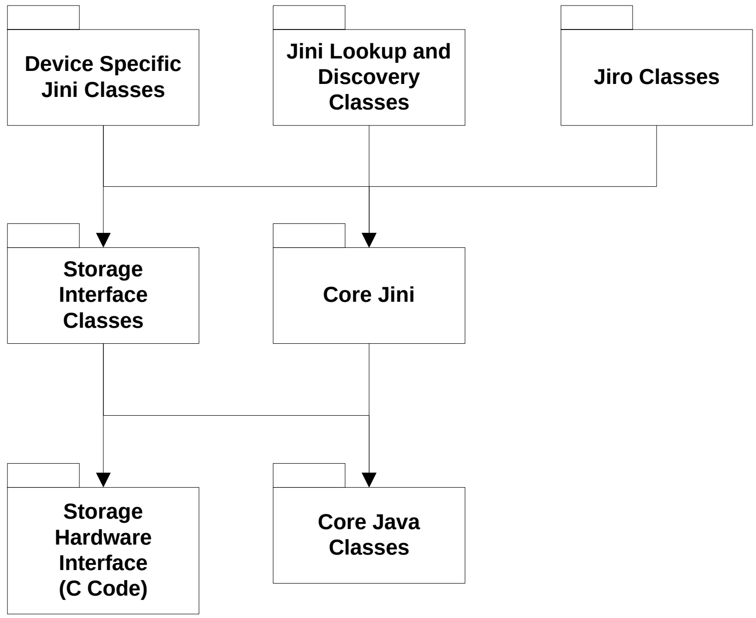

FIGURE 2. **Jiro-Jini Proxy device software architecture**

The Jini proxy device and it's extension, the Jiro proxy device, will provide the resources and interface between current Fibre Channel or SCSI-based storage, and the Jini software network. It will allow current storage to be attached to the Jini Djinn by providing the Java virtual machine and embedded Jini classes as well as specific, device-type dependent classes. The Jiro classes and additional device-specific classes would also be added to allow the storage device to support Jiro-based resource management of the devices. This can be achieved by attaching the storage to a server, which in turn would act as a Jini "file server." A Storage Appliance with an embedded JVM, however, could perform this function more efficiently, more reliably and at a lower cost then the server-based solution.

SECTION 7 *CONCLUSION*

Storage Appliances represent an approach that is more scalable, more reliable, and less costly than server-based storage solutions. It allows storage sub-systems to act independent of any particular OS or server and the rest of the system. This approach will result in better storage sub-systems, especially better SANs, with a lower total cost of ownership of that system.