White Paper

Using ATA Disk Technology and ADXT™ to Power Next Generation Storage Arrays

Diamond Series Storage Array

ATTO Technology, Inc.
155 CrossPoint Parkway
Amherst, NY 14068
Phone: 716.691.1999
Fax: 716.691.8353
attotech.com
sales@attotech.com

4/6/2004

Contents

I. Introduction

ATTO Technology, Inc. has developed a line of unique, high-performance computer data storage arrays designed to meet the challenging needs of the SAN, NAS and general computer storage markets. The ATTO Diamond Series of storage arrays combine extraordinary density, multiple high-speed interfaces and enterprise class reliability with cost-effective industry standard ATA disk drives to provide a storage product with a significantly better value proposition and the features of the newest Serial ATA Technology. Features like hot plug/hot swap drives, support for command overlap and command queuing, and staggered spin up have been supported since the introduction of the ATTO Diamond Storage Array.

Unlike other storage arrays which use expensive SCSI or Fibre Channel disk drives to achieve performance, the Diamond Series uses lower cost ATA disk drives combined with an intelligent midplane to create a storage array with exceptional price and performance characteristics. The intelligent midplane contains hardware and software which provides ATTO's proprietary **ADXT™** (Aggregated Data Transfer Technology). ADXT™, a switched data management and data movement technology, "aggregates" the performance of individual ATA disk drives together to create a higher performance storage solution. Using this technology, ATTO is able to provide a storage array that delivers faster sustained data transfer rates as well as impressive I/Os per second at a new price point in the storage market.

II. ATA Disk Technology

Today, many Storage Array Networks (SAN), Network Attached Storage (NAS) products and entry servers found within enterprises use ATA drives. Moreover, there are numerous applications such as backup and near-line storage for which ATA drives are proving to be a ubiquitous solution component, not to mention high bandwidth applications such as video surveillance, video manipulation, near-line storage, and high bandwidth back up and recovery procedures that use disks instead of tape.

ATA disk drives originally were known as IDE disk drives, due to the disk interface they contained - Integrated Drive Electronics (IDE). Although IDE began as a low performance, low reliability disk interface, it was continually driven by the demands of the high volume, cost sensitive PC marketplace. Therefore, over the years, IDE was refined and as the drive interface acquired important capabilities required for high reliability and high performance its name was changed to ATA (Advanced Technology Attachment). The ATTO Diamond Storage Array ADXT™ Architecture further enhances the ATA interface and today offers:

- UltraDMA transfer protocol that is similar to high performance SCSI disk protocol
- Double-clocking of data transfers which doubles disk data transfer rates
- UltraDMA protocol which operates at 100 Megabytes per second
- CRC (Cyclic Redundancy Check) code allowing full error detection and reliability of all data and commands
- Multi-threaded I/O support

- Overlapped Command Support which allows commands to be simultaneously active on both drives on the same ATA bus.
- Command Queuing which allows multiple read/write commands to be sent to each drive simultaneously. This reduces command overhead and allows the drive to service the commands in the most efficient manner. This is similar to the SCSI feature of Tagged Command Queuing.
- Higher media transfer rates (up to 240MB/sec)
- A communication protocol and interface with a fundamental lower cost structure than SCSI or FC interfaces

It is important to note that virtually all disk drives from a single manufacturer, regardless of type, share common internal components such as head assemblies, servo mechanical mechanisms and control circuitry. This means the "core" drive technology in SCSI and FC drives is also in ATA drives. As 85% of disk drive production is ATA based, top disk drive manufacturers have improved the design and reliability of their ATA drives to the point that the life expectancy and warranties of these drives often meet those of SCSI and Fibre Channel drives.

ATA disk drives today offer greater capacities than either SCSI or FC disk drives – and at a tremendously lower cost. Advances in platter and head technology will push drive capacity along a very steep curve in the next five years. As for drive speeds, ATA disk drives tend to trend behind SCSI or FC drives however, these drives will march forward to 10K, 15K-etc. over time as well. Today, 10,000 RPM drives are the fastest speed available in ATA.

Based on MTBF numbers there is far less risk with ATA drives today. Statistically the expected life of ATA drives far exceeds the useful life (greater than 5 years). Bearing failures are the main reason disk drives fail over time. Is there a reason to believe that a slower RPM ATA drive will have more bearing failures than a faster RPM SCSI drive?

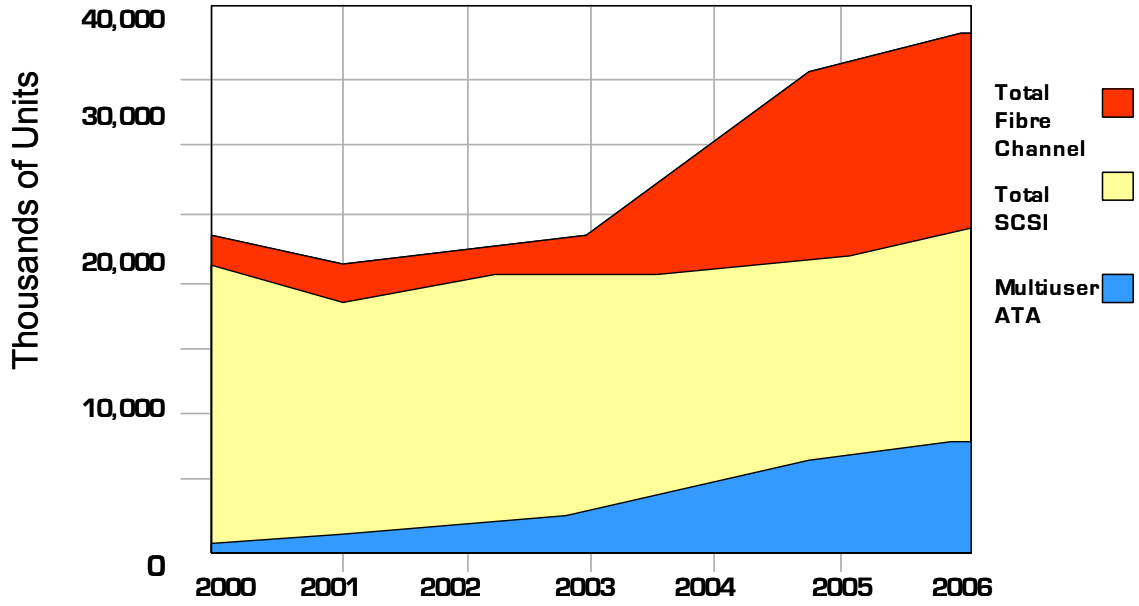
In addition, If you use ATA with RAID for device redundancy, this further improves reliability. For example, applying RAID 5 will improve the AFR (Annual Failure Return) rate by a factor of 3 or greater. The following chart summarizes ATA, SCSI and Fibre Channel disk features:

	ATA	SCSI	Fibre Channel
Connectivity **• Max Devices** **• Cable Length**	**2** **18"**	16 25m	126 / Millions 10 km
Performance **• Bandwidth** **• Bandwidth / drive** **• Multi-threaded I/O**	**100 MB/sec.** **66 MB/sec.** **Yes**	160MB/sec. 10 MB/sec. Yes	200 MB/sec. 1.6 MB/sec. Yes
Max Initiators	**1**	<16	125 / Millions
Topologies	**Bussed**	Bussed	Loop / Fabric
Error Detection	**Yes**	Yes	Yes
I/O Duty Cycle	**20-30%**	60-80%	60-80%
Yearly Power-on Hrs	**8760**	8760	8760
Meantime to Failure	**Up to 1M hours**	>1.2M hours	>1.2M hours
Annual Failure Rate	**< 1%**	< 0.7%	< 0.7%
Capacity	**250GB**	146GB	146GB
Warranty	**Up to 3 years**	5 years typical	5 years typical

This chart clearly indicates that ATA drive technology offers characteristics that ideally suit it for "in the box" usage in high capacity, high performance disk arrays.

From the cost side, while Fibre Channel and SCSI drives tend to be faster, they are significantly more expensive, anywhere from three to eight times, than comparable capacity ATA disk drives. This increased cost can be directly attributed to both technical and basic economic issues including: the limited number of FC and SCSI suppliers (Seagate, Maxtor, Quantum, Fujitsu, Hitachi), much lower shipment volumes and higher interface cost structure. The large cost differential between ATA and SCSI/FC drives is not expected to change over the next three to five years due to the inherent economic and technical differences between the drive types.

With the current demand for ATA disk drives outpacing SCSI by 10:1 (see chart below), it was critical to focus technology improvements on ATA drives to increase reliability and maintain competitiveness. Consequently, for the last half decade, technology now first appears on ATA disks and then moves to other interface technologies. Specific examples include CRC, Dual Edge clocking, GMR heads, and multi-segment CAM assisted caching (Content Addressable Memory cache search accelerators). The final point that needs to be made is this. Shipments of ATA drives are expected to grow rapidly over the next five years with increasing consumer and business demand. In general, ATA disk drives will be the "driver" of all disk storage technology as the costs, volumes and demands of the PC and storage marketplace are the primary focus of the world's largest disk drive manufacturers.



Source : Gartner June 2002

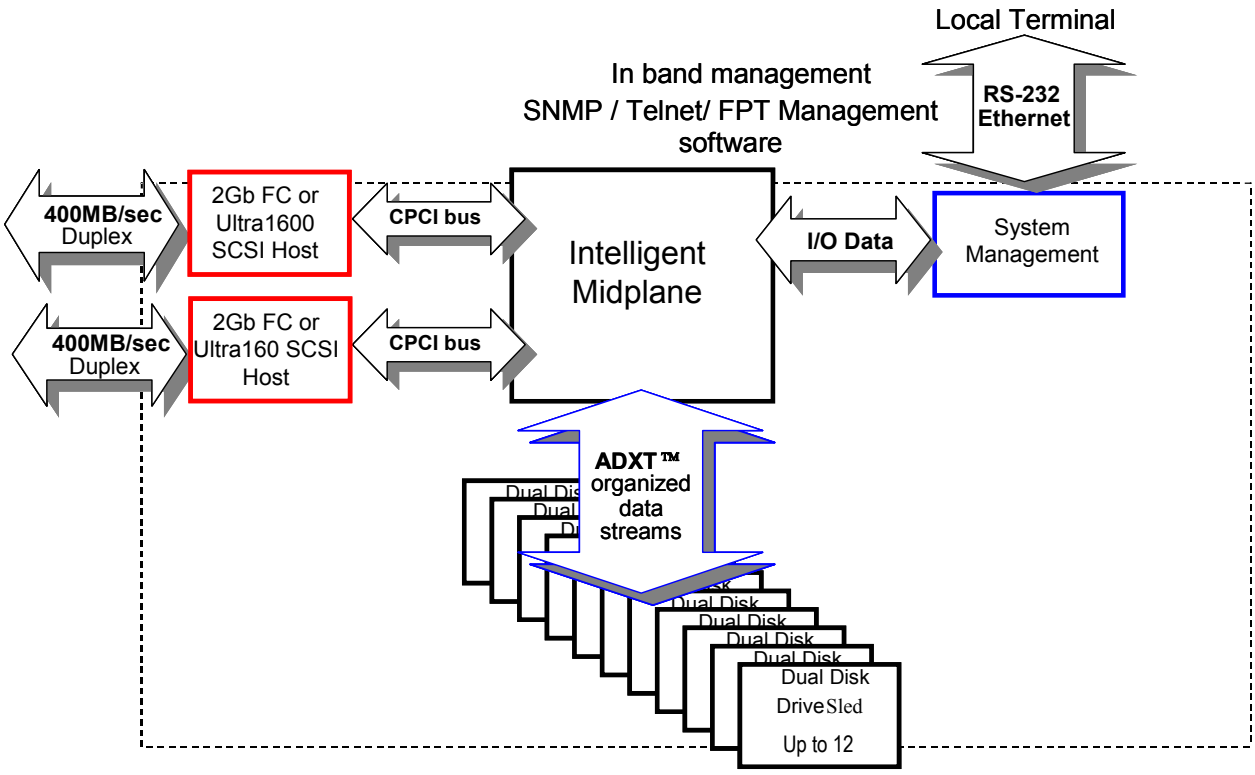
III. ADXT™ Powers ATA to New Levels

The original notion of RAID was to build high capacity, reliable storage subsystems using large numbers of inexpensive disk drives. Thus the original definition of RAID was "Redundant Array of Inexpensive Drives". Over time that definition was morphed into "Redundant Array of Independent Drives" and the notion of an inherent cost advantage in a RAID system was lost. The ATTO Diamond Series takes disk arrays back to their "roots" by providing a high capacity, high performance and highly reliable disk array that uses the merged power of many cost effective ATA disk drives to create these valuable attributes.

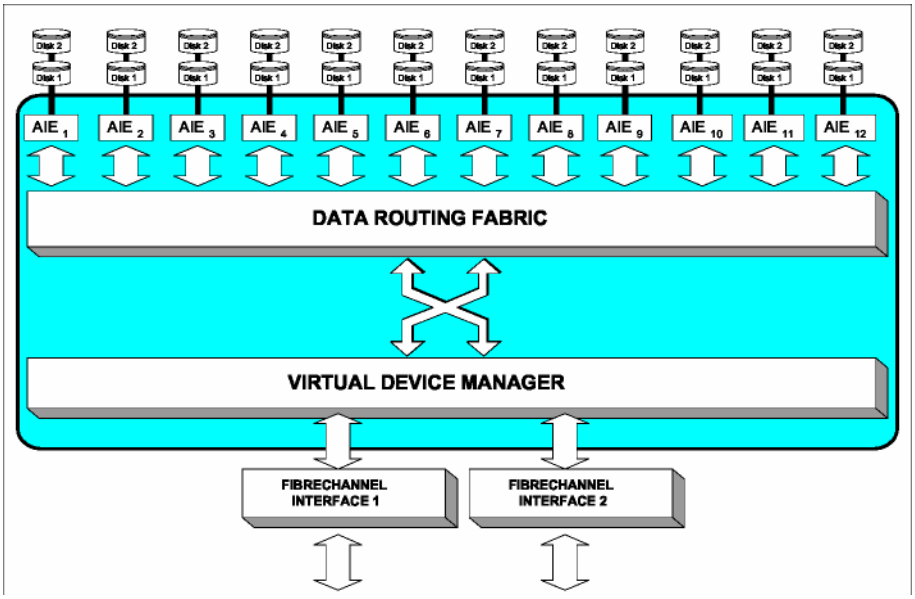
With all of the features that ATA disk drives now offer it is still necessary to add a level of intelligence in the control of these drives to successfully integrate them into high performance disk storage arrays. An individual ATA disk drive is capable of providing 30MB/sec. of sustained data transfer (reads or writes). This rate is adequate for home computer users but far too low for higher performance applications that require 180 to 240MB/sec. of sustained data transfer rates.

ATTO developed an intelligent midplane with ADXT™ to allow the summation or "aggregation" of the data rates from each individual ATA disk drive to create the high data transfer rates needed by SAN's and other high performance computer systems. The Diamond Series of arrays have data transfer rates previously available only in substantially higher priced Fibre Channel or SCSI disk arrays. The sophisticated design of the intelligent midplane also enables features such as server-less backup, advanced error protection, metadata storage techniques, virtualization software, thermal management and advanced enclosure services to be implemented in the architecture as well.

The overall architecture of the Diamond array is comprised of dual host interface cards (HICs), the intelligent midplane, a system management card and up to twelve independent disk drive sleds containing up to 24 ATA disk drives. The data pathways and architecture are:



The heart of the Diamond Series disk storage array is the intelligent midplane. It contains a combination of custom Application Specific Integrated Circuits (ASIC's), processors and proprietary embedded software to make ADXT™ processing technology a reality. It is divided into three main processing sections, which handle the data being read or written to the ATA disk drives from the host interfaces. The Virtual Device Manager (VDM), Data Routing Fabric (DRF) and ATA Interface Engines (AIE) work in concert to smoothly and efficiently organize data streams for storage or retrieval. Schematically, ADXT™ is illustrated as follows:



Virtual Device Manager - The Virtual Device Manager implements virtual SCSI devices and provides capabilities which far exceed the abilities native to the underlying ATA disk drives. Data is accessed through the virtual SCSI devices using a rich implementation of the standard SCSI protocol. The VDM adds capabilities required by advanced storage users such as RAID, data management services, enclosure management services, Zoning and LUN mapping, storage partitioning and other features which serve to increase performance, flexibility, reliability, availability, and management. This layer of software gives the Diamond Series the ability to handle a growing array of user needs, extends product life and reduces the end users total cost of ownership.

Data Routing Fabric - Incoming or outgoing data is routed between the ATA Interface Engines (AIE) and the host interfaces by the custom Data Routing Fabric ASIC. The Data Routing Fabric is a high speed, low latency transfer fabric with over 2 Gigabytes/sec of bandwidth supported by up to 256 Megabytes of ECC memory.

ATA Interface Engine (AIE) - The interface to each pair of drives is through a custom ATA Interface Engine ASIC. The AIE implements the typically software intensive ATA interface completely in silicon. Each AIE contains a dedicated ATA protocol processor to completely automate command and protocol processing for maximum performance. Automated command overlapping and queuing maximizes the performance of multi-threaded I/O. The AIE transfers data using double-clocking technology at a speed of 66 Megabytes per second. The Diamond Series contain 12 AIE's for a parallel transfer rate of 792 Megabytes per second. In addition to guaranteeing data transfer integrity with automatic CRC checking, the command and status transfers are validated using a patent pending technology unique to the AIE.

IV. About the Diamond Storage Array

The Diamond Storage Array is based upon leading-edge technology that is built upon ATTO's proven core competencies in the SCSI and Fibre Channel storage technologies. At the center of ATTO's Diamond Storage Array is its intelligent midplane containing a combination of custom application-specific integrated circuits (ASICs), microprocessors, and proprietary software. Developed by ATTO, this technology comprises patent-pending elements that work together to minimize system overhead and maximize data throughput. ATTO's exclusive Aggregated Data Transfer Technology (ADXT™) aggregates the performance of individual ATA drives to achieve the high data-transfer rates needed for today's Enterprise-level applications. This exclusive combination of features offers the performance of Fibre Channel or SCSI storage arrays with the cost advantages of ATA disk technology. A scalable, high availability storage array such as the Diamond offers the following functional advantages:

Fault Tolerant Design

Great care has been taken in the design of Diamond to ensure that it meets and exceeds the most demanding applications and provides the highest level of data availability as dictated by Enterprise IT environments. Examples of this include:

- Redundant hot swappable, hot pluggable components avoid unnecessary downtime and/or loss of data. Redundant, hot swappable components include the power supplies, blowers and disk drives. Hot pluggable components include the Host Interface Cards.
- Error Correcting Memory, end to end data parity checking and CRC (Cyclic Redundancy Check) to guard again data corruption. ECC memory provides for the highest level of data integrity possible for data transfers in or out of memory and significantly increases the reliability of high speed data access. These special memory components can both error correct as well as inform the host monitoring system of data transfer faults.
- Dual Host Interface Cards provide for redundant pathways to storage ensuring all connected hosts have continuous access to critical data.
- Automatic rebuild of RAID groups in the event of a drive failure will keep the system operational at all times.
- Hot Spares – global hot spares can be assigned to RAID groups in order to ensure continuous operation in the even of a disk drive failure. The hot spare automatically comes on-line and rebuild started if a disk failure is detected

Environmental/System Monitoring

Diamond monitoring provides peace of mind knowing that all operating parameters are within their limits and that any critical condition will be reported to the main service center through the system event manager. Some of these include

- Temperature monitoring – several temperature sensors are located within the Diamond to always ensure proper operating conditions. In the event of an over temperature condition the Diamond will gracefully suspend drive activity until the operating temperature returns to normal.
- Component monitoring – components such as blowers, power supplies and disk drives are constantly monitored for proper operation. In the event of a failure an alert will be generated in several forms including front panel LED's, audible alarm and through remote notification processes via SES and/or SNMP traps.
- Drive/LUN status – This shows which drives and LUN's are active and on-line. This information is accessible through the Command Line Interface as well as the GUI ExpressNAV™.
- RAID rebuild status – Like drive and LUN status, rebuild status/percent complete is information available to the Administrator via the Command Line Interface as well as the GUI Express NAV. Backups will continue even during a rebuild process.
- Status reporting via SNMP and SES – These facilities, typically found only on Enterprise class storage devices, are industry standard methods for reporting system status and operation. Both of these enable easy integration with Corporate IT Management Systems.

Capacity/System Expansion

It is imperative that IT Managers get the most out of their investments which includes the knowledge that their equipment purchases will continue to meet the ever expanding needs for performance, functionality and capacity.

- Scales from six to twenty four drives (1.5TB to 7.2TB). Larger drives can be substituted as drive capacities increase.
- SCSI host interfaces can be later upgraded to Fibre Channel for SAN (Storage Array Network) connectivity. Fibre Channel can also provide up to a 50% performance improvement over SCSI.
- New interface technologies can be added in the future such as iSCSI, Infiniband or U320 SCSI. Each of these is road mapped for Diamond.
- ADXT™ architecture and flash memory allow for upgrades and new features as they become available. For example, a Diamond can be through the combination of new firmware and Management Card. New features recently added to the Diamond product line include hot spares, ECC memory and the ExpressNAV™ GUI.

Component Protection Systems

Often the best defense is a good offense. This analogy can be extended to the Diamond and the attention that has been paid to the most critical operational use cases and the appropriate response to each.

- Soft drive starts – The Diamond array protects against over current conditions by gently spinning up disk drives in a fully populated array configuration. A maximum of six drives are brought on-line/up to speed at one time. This limits the power drain and avoids premature power supply failures.
- Automatic command overlap disengage upon loss of one power supply. In order to conserve power, the Intelligent Midplane firmware will cease command overlap operations to the disk drives should this feature be enabled. This once again limits the power drain and avoids further loss of components until corrective action can be taken.
- High temperature warning system –A sophisticated temperature monitoring algorithm is used to alert a System Administrator of an approaching over-temperature condition. This algorithm examines the temperature readings from multiple locations within the Diamond array and decides if an alert is to be generated.
- Automatic I/O stop at max temperature - If the temperature continues to rise beyond the warning threshold, further action is taken including spinning down drives and generation of an error condition.
- Vibration free drive mounting system – In order to obtain its highest density in a 3U enclosure, the Diamond uses as patent pending drive mounting system which also acts to minimize drive vibration and prolong overall drive life.

System Management

Often it is not feasible to have on-site support services and therefore any new system purchases must be remotely accessible and configurable. Diamond meets these challenges head on.

- In-band via Fibre Channel – A System Administrator has several options for querying the Diamond Storage Array. One method is through in-band via the Fibre Channel interface. A full complement of management functions and status reports are readily accessible.
- Telnet session via Ethernet or Serial ports – Similar to in-band, two other options, Ethernet and Serial Ports, are available to gain access to the Diamond. Again, a full complement of management functions and status reports are readily accessible.
- Flash update firmware – using a Flash Memory System the Diamond allows Administrators to easily update the system firmware which typically includes new features and/or general performance improvements. This process can be done in the field without special tools or training.
- Define/Modify System Configurations – Multiple system configurations are possibly via the Command Line Interface or the ExpressNAV™ GUI. These include setting up Host Interface connections, Serial and Ethernet port, RAID settings (5, 10) Zoning, etc..
- System diagnostics – Upon power-up and through user request, a series of self diagnostics are executed to ascertain the health of the array. If a problem if detected, the sophisticated activity logs will provide valuable insight into the probable cause.

V. Summary

ATA is the dominant disk drive technology today and will be for the foreseeable future. It offers all the cost advantages of a mass produced, consumer driven technology which is rapidly being driven forward. Today ATA also offers all the performance and reliability features needed to create high performance ATA based disk storage arrays.

ATTO Technology views ATA as the perfect disk drive technology for “in-the-box” disk storage array solutions. By themselves however, ATA disk drives cannot achieve the performance needed by enterprise class storage users who demand high bandwidth and IOP’s (Input/Output Operations per Seconds). ADXT™ (Aggregated Data Transfer Technology) with its built-in Virtual Device Manager (VDM), Data Routing Fabric (DRF) and ATA Interface Engines (AIE) smoothly and efficiently organizes data streams for storage or retrieval on ATA disk drives. ADXT™ provides end users with the power and sophisticated data control needed to take moderately priced ATA disk drives combine them in a disk storage array and power them to the performance levels of SCSI or Fibre Channel disk arrays.

ATTO Technology Inc.
155 CrossPoint Pkwy
Amherst, NY 14068
www.attotech.com
716.691.1999