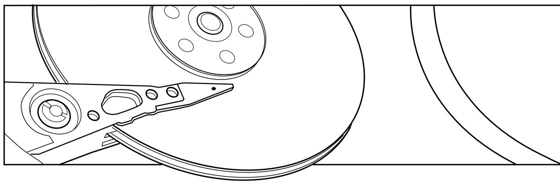

Serial ATA

A Promising New Alternative for Enterprise Storage Applications

➤ Overview/Executive Summary

For desktop and portable computers, the Parallel Advanced Technology Attachment (ATA) interface is currently the most popular protocol for moving data between the hard drive controller and system memory. Designed to provide a "best of both worlds" solution, Serial ATA (SATA) replaces the existing parallel bus with serial links, and adds features designed to increase the performance, reliability and scalability of ATA-based devices while retaining ATA's significant cost advantage.

➤ Background

As an evolutionary replacement for Parallel ATA, a consortium of industry leaders developed a new storage interface specification called Serial ATA 1.0. This new specification was formally introduced in August of 2001. Serial ATA II, a second iteration of the specification for SATA servers and network attached storage, was published in October 2002. It will allow vendors to further address the needs of networked storage segments via enclosure management, backplane signaling, and cabling and performance improvements.

Traditionally, ATA has not supported the requirements of enterprise systems. Traditionally, enterprise storage applications have been dominated by solutions based on the SCSI (Small Computer System Interface) standard, with the more recent entry of Fibre Channel. Systems based on these technologies deliver the high performance and reliability required for mission-critical applications. However, the cost for systems based on SCSI or Fibre Channel can be substantially higher than technologies based on ATA.

Storage vendors quickly realized that Serial ATA has distinct advantages over traditional ATA and parallel SCSI that make it a viable alternative for many enterprise storage applications, enabling storage systems at what current pricing models estimate to be one-third the cost of SCSI-based systems. The result is a cost-effective alternative to SCSI technology for all but the most critical direct-attach and networked storage applications.

➤ Serial ATA Features and Benefits

While SATA represents the next generation of ATA interface technology, it departs from ATA with several significant new features and benefits, summarized below.

Feature	Serial ATA	Parallel ATA	Parallel SCSI
Low-cost implementation	✔	✔	
Point-to-point connectivity	✔		
Cyclical Redundancy Checking (CRC) on commands	✔	(partial)	✔
Hot-plug/hot-swap support	✔		✔
Highly efficient cabling, connectors, backplanes	✔		

Table 1. Hard Drive Interface Features and Benefits Comparison

Point-to-Point Connectivity

Point-to-point connectivity is an important feature that provides significant performance and reliability advantages over the shared connectivity approach employed by both the ATA and SCSI parallel interfaces. Each port on a Serial ATA controller serves just one device; that is, the controller communicates with a given drive only through the port where it is connected (see Figure 1 below).

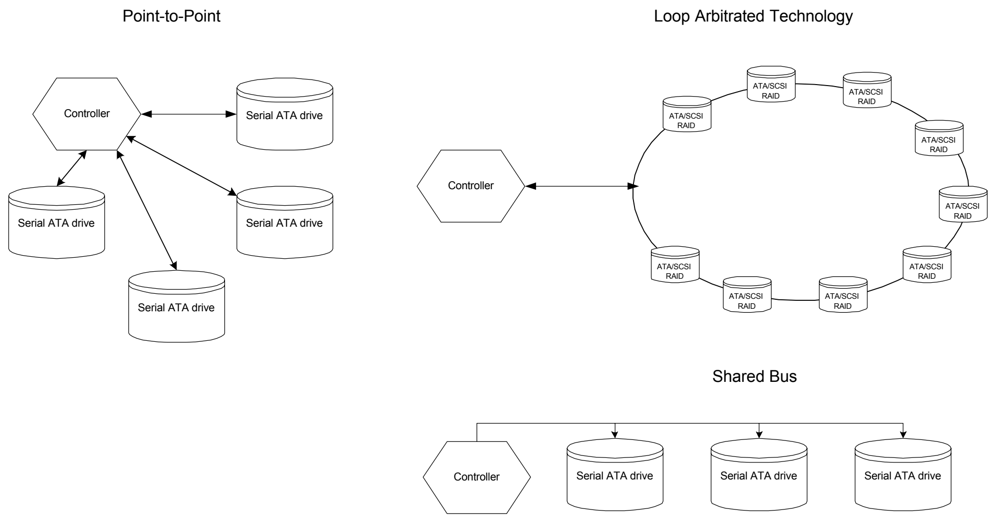

Figure 1. Point-to-Point Versus Shared Connectivity

Because there is no sharing of the bus, each drive can communicate directly with the system at any time. This means that the entire available interface bandwidth is dedicated to each device. The dedicated link approach of point-to-point connectivity eliminates the arbitration delays sometimes associated with shared bus topologies. With a shared bus approach, this overhead increases as drives are added to the shared bus. So, in a typical ATA or SCSI RAID system, adding a hard drive will increase the total system throughput by some amount less than the throughput of the disk. With Serial ATA, on the other hand, each added hard drive can deliver its maximum throughput. Point-to-point connectivity offers the added benefit of simpler configuration. Dedicated links make a Serial ATA RAID system easy, fast and relatively inexpensive to set up.

Point-to-point connectivity introduces an important issue relative to scalability. With Serial ATA, the capacity supported by the configuration is a function of the number of available point-to-point connections. Traditionally, the architecture would determine the number of devices supported, and the number of connectors on the cable attached would dictate the number of drives that may be connected. With SATA, scaling is achieved simply by adding more point-to-point links in the system at the host level, with each connector having one cable and one drive connected to it. The number of links included is dictated by the number a vendor has included on the system board. Additional links can be added via a controller card or a RAID card.

Cyclical Redundancy Checking (CRC) Error Detection

Perhaps the most beneficial and most significant improvement of Serial ATA over Parallel ATA is Cyclical Redundancy Checking (CRC) on commands. While both Serial and Parallel ATA have CRC, the CRC is different in SATA due to the fact that SATA has CRC for the entire command (FIS - command code, data packets, etc.), where Parallel ATA only has CRC for the data transfer. CRC has been available in SCSI since Ultra160 technology was introduced, and significantly improves data integrity by checking all transferred data and verifying that it is received correctly.

CRC greatly increases error detection reliability during the execution of high-speed transfers and hot-plugging or hot-swapping (insertion and removal of drives without interfering with the data transfers taking place; see “Hot-Plug Support” on page 3). CRC is calculated on a per-burst basis by both the host and the hard drive, and is stored in their respective CRC registers. At the end of each burst, the host sends the contents of its CRC register to the hard drive, which then compares it against its own register’s contents. If the hard drive reports errors to the host, then the host retries the command containing the CRC error.

With Serial ATA, each protocol layer has the capability to identify errors and can perform recovery and control actions as well as forward information to the next higher layer in the stack. Each layer has a means to be aware of—and recover from—errors in the layer below it. A detailed explanation of this feature is covered in “Protocol and Interconnect Level Reliability” on page 4.

Hot-Plug Support

SATA supports hot-plugging (also known as “hot-swapping”), the ability to swap out a failed hard drive without having to power down the system or reboot. This capability contributes to both data availability and service-ability without any associated downtime, making it a critical feature for extending SATA into enterprise applications.

The Serial ATA 1.0 specification requires staggered pins for both the hard drive and drive receptacles. Staggered pins mate the power signals in the appropriate sequences required for powering up the hot plugged device. These pins are also specified to handle in excess of the maximum allowed inrush current that occurs during drive insertion. SATA-compliant devices thus need no further modification to be hot-pluggable and provide the necessary building blocks for a robust hot plug solution, which typically includes:

- Device detection even with power downed receptacles (typical of server applications)
- Pre-charging resistors to passively limit inrush current during drive insertion
- Hot-plug controllers to actively limit inrush current during drive insertion
- Receptacle and plug guides for alignment during drive insertion

Improved Cabling

Cabling sits near the top of the list of hardware-related service calls. It also is a significant consideration in air flow design and hardware assembly. Serial ATA features an enhanced cable design that offers important benefits that add up to reduced service calls and efficient manufacturing.

Though it appears simple, the basic Serial ATA connector design is a remarkably efficient and practical design offering a number of notable features/benefits:

- The "L" shaped Serial ATA data and power connector make plug orientation very obvious to the end-user, thus preventing incorrect mating.
- The extrusion has "ears" which guide and align the plug during the mating process.
- The conductors are engineered for hot-plugging; they connect in three stages—first pre-charge, then ground, then power.
- The connector locations on the back of both 2.5" and 3.5" devices are identical, allowing design of backplanes that accommodate either size device.

The Serial ATA connector represents a substantial improvement over Parallel ATA, which has a long history of problems with bent pins. And it represents a significant improvement over SCSI's daisy-chain topology, where, if one cable disconnects or fails or if a terminator is missing, the entire group of drives will not perform properly.

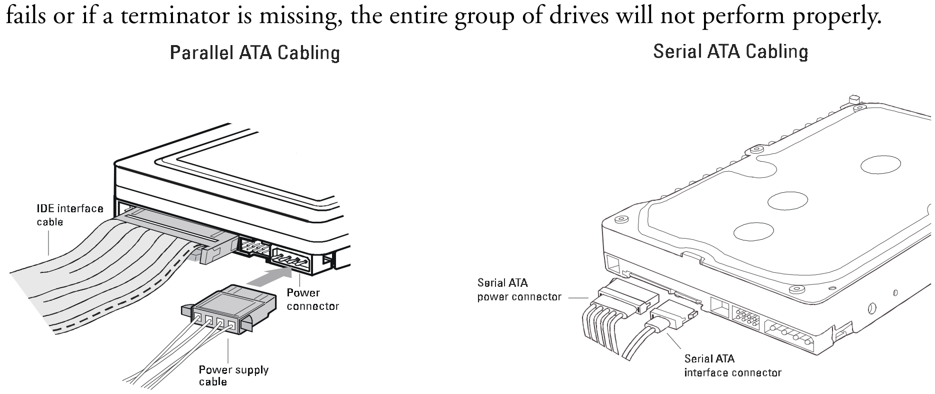

Figure 2. Parallel ATA vs. Serial ATA Cabling

➤ SERIAL ATA RELIABILITY

Serial ATA features greatly enhance storage system reliability compared to ATA. Unlike its predecessor, SATA provides important protection and recovery features at four levels: the *protocol and interconnect* level, the *device* level, the *sub-system,* and the *system* level.

Protocol and Interconnect Level Reliability

Thanks to its CRC feature (see "Cyclical Redundancy Checking (CRC) Error Detection" on page 2), SATA provides solid error detection at every layer:

- PHY layer detects and handles raw signal and 10-bit stream problem conditions such as no device, OOB signaling failures, and PHY internal errors. Information pertaining to these errors is also visible to the link system.
- Link layer detects and handles 10-bit stream and frame problem conditions such as invalid state and data integrity errors. Information pertaining to these errors is also visible to the transport layer.
- Transport and software layers detect and handle command and status problem conditions such as internal, frame, protocol, and state errors, including error handling for command block status registers and command failed/timeout. Information pertaining to the transport layer is visible to the software layer.

As explained previously, SATA features hot-plug ability plus an enhanced cable design that helps eliminate connection errors and bent pins. Together, these features contribute to improved reliability and reduced downtime resulting in fewer service calls.

Device Level Reliability

Hard drive reliability is expressed in number of hours as Mean Time Between Failure (MTBF). It is important to note that Western Digital's first SATA drive will have the same MTBF (1.2 million hours) as a SCSI drive. Keeping this in mind, also consider that the statistical mean represented by MTBF provides information about a population, but not particular drives. The failure of a particular drive cannot be predicted—hard drive failure is an ever-present reality. Therefore, it is reasonable to expect the reliability of SATA hard drives to approximate that of SCSI drives.

To address the reality of potential hard drive failure, storage vendors utilize RAID (Redundant Array of Inexpensive Disks) to provide disk mirroring and parity data protection. In the 20 years since RAID was created, the "inexpensive" in the acronym has evolved to "independent," reflecting the evolution toward more expensive drives to achieve higher RAID performance.

Serial ATA offers price and performance that puts the "inexpensive" back in RAID. In addition to its low cost, Serial ATA has characteristics that make it particularly good for RAID including dedicated point-to-point channels with a manageable cabling topology and staggered spin-up. The result is an affordable storage solution that delivers more than satisfactory uptime for a range of applications *not* considered "mission-critical."

Sub-System Level Reliability

The physical hardware of a storage system also impacts its availability. The enterprise storage market is typified by storage arrays that provide dense, rack-optimized solutions with redundant components such as power supplies and fans. Systems manufacturers are striving for ever-denser solutions, requiring smaller form factors, to meet customer demand for more storage in limited space. This means the number of drives requiring cooling is increasing. This growing challenge for storage enclosure manufacturers is exaggerated when utilizing ultra high-density SATA drives.

Temperature is a critical factor impacting the reliability of drives and other components—the cooler the drives, the greater the reliability. As enclosure manufacturers develop solutions based on these ultra high-density drives, it is critical to ensure proper cooling and fan redundancy to preserve system availability. SATA is well-positioned to handle these environments since its thinner cabling allows better airflow through the system.

System Level Reliability

To achieve high system availability (as well as scalability), many solution providers employ cluster configurations. There are two basic cluster configuration models: the shared disk model and the shared nothing model. Serial ATA can be employed effectively in either of these two models.

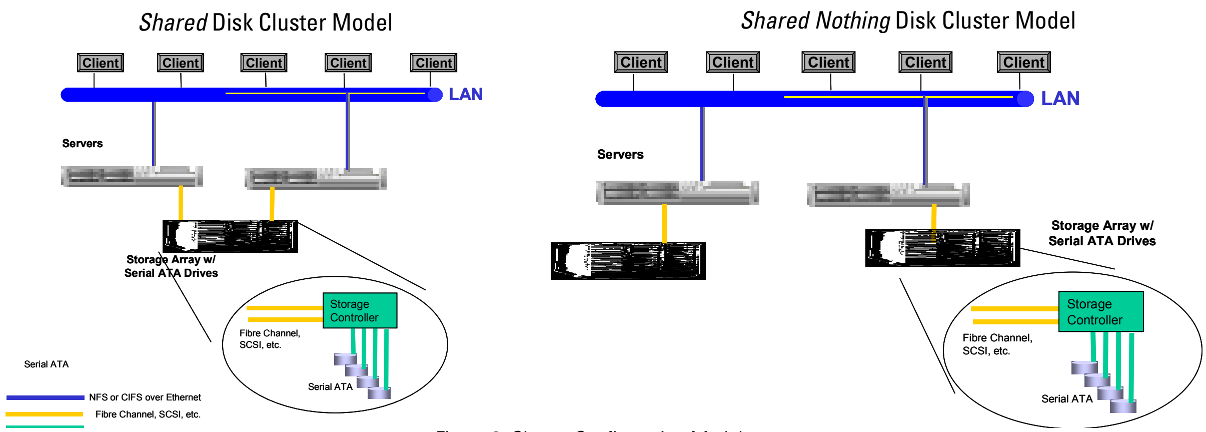

Figure 3. Cluster Configuration Models

Both clustering models eliminate the server (or appliance) as single points of failure. And, in both cases, Serial ATA RAID arrays can eliminate the hard drive as single point of failure.

Summary—Serial ATA Reliability

Based strictly on MTBF specifications, SCSI drives clearly deliver a higher level of uptime than SATA drives. However, it is system uptime that really matters. Since no hard drive is infallible, the only way to meet the need for 24/7 availability is to create a system that can tolerate a failure of any component, without loss of connectivity. The only way to achieve this is to create a system that has redundant *everything*—from multiple I/O controllers and servers to multi-path cabling.

Dual port SCSI drives can add value to such a redundant solution; if it is in a RAID volume, no data or connectivity is lost if the drive itself fails. However, a single port drive can achieve the same objective through RAID mirroring, such as RAID 5+1 (this is striping with parity + mirroring). Although mirroring requires twice as many drives, many IT professionals intentionally implement this method for enhanced performance, offering the ability to access the same data in parallel from both drives (~2x performance).

Additional technologies are also used to aid IT professionals in the ability to predict failures and to quickly swap out faulty components. System solutions include such implementations as rigid backup policies, hot spares, hot-plug, enclosure management/error prediction (SMART), or any combination of these system solutions and others as determined by the needs of each business.

Thus, if one can achieve the system goal of guaranteed data reliability and data availability with either SCSI or SATA technology in a RAID array, then the deciding factor will likely be cost. In such a scenario, SATA offers an economic advantage that is hard to ignore.

➤ SERIAL ATA PERFORMANCE

Since devices based on SATA are still in development, the performance of these devices has yet to be fully tested. However, it is clear that SATA will usher in an entirely new level of ATA performance, utilizing a data transfer rate of 1.5 Gb/second (150 MB/s). This surpasses today's mainstream ATA performance by 50% or more, with a clear roadmap for future performance increases. Recall that in a system combining multiple devices, such as a RAID configuration, SATA offers the advantage of additive device performance, devoting full bandwidth to each drive in the system.

A typical misconception about SATA is that its serial nature makes it inherently slower than parallel bus designs. However, witness the trend toward serial technology for all desktop/server data transport mechanisms including PCI Express, Fibre Channel SANs, Infiniband Architecture, and others. Indeed, even SCSI is moving from parallel to serial technology, with the future Serial Attached SCSI (SAS) draft standard. Given the high speeds required by today's technology, coupled with the synchronization constraints of parallel data buses, high-speed serial links provide a practical basis for future technologies.

How SATA Improves Performance

The key to SATA's higher performance (at least compared to ATA) is its point-to-point topology. SATA does not have to share the ATA bus as in the traditional ATA master/slave topology. Add to that SATA's dedicated 1.5 Gb/second (150 MB/s) maximum performance per device, and it can be seen that the bus already has room to spare when today's best-of-breed drives are hard pressed to deliver 100 MB per second. Note that this is just the starting point for Serial ATA; the specification developers are planning increased speed transitions for SATA over the next several years.

Performance and Price

While it is likely that first-generation SATA drives may not match the throughput of the best high-end SCSI drives, it is important to place this in the context of real-world storage solutions. Few storage solution vendors actually use high cost, best-of-breed SCSI devices in their systems. The SATA cost advantage enables vendors to use best-of-breed SATA products with adequate performance at a cost lower than if they used "mainstream" SCSI devices. This may in reality deliver performance comparable to or less than the best SATA units. The important point about SATA is that it narrows the gap between ATA and SCSI performance, while retaining the traditional ATA price advantage. And, in today's IT marketplace, adequate performance at the best price point is the name of the game for a growing number of storage applications.

➤ CONCLUSION

Clearly, SATA represents an important extension in proven ATA interface technology—one that targets SATA squarely at networked storage. The question is not *if* SATA will penetrate the enterprise, but *how far* can it go? While its attractive cost has already driven ATA technology into the enterprise space, SATA delivers performance and reliability that promise to solidify ATA's presence in the entry level/blade server category and extend to other business critical systems. SCSI will likely continue to be used for mission-critical storage applications. SATA, however, is ideally positioned to satisfy the requirements for a large number of enterprise storage applications, including:

- Web hosting
- Firewalls
- E-mail servers
- Small business and remote office business applications
- File sharing
- Media streaming
- Tape backup replacement
- Performance workstations

From a system point of view, SCSI and Fibre Channel will likely continue in high-end Storage Area Network (SAN) segment. However, despite the growth of SANs, Direct Attach Storage (DAS) still dominates the storage market. These DAS configurations are attractive to organizations seeking moderate performance at a reasonable cost. This is a market space where Serial ATA-based storage solutions can provide a superior value proposition, appealing strongly to what is still the largest segment of the external storage market.

DAS customers can take advantage of new and attractive options that Serial ATA offers. Now they can choose enterprise-class storage subsystems with enhanced power, packaging, acoustic, and cooling characteristics within a fully redundant/ fault-tolerant configuration; a feature set superior to existing SCSI JBOD arrays. In addition, the use of low-cost Serial

ATA external RAID controllers within such subsystems in combination with Serial ATA drives will provide a new entry point in terms of cost per gigabyte in the RAID storage market space.

Storage solutions with higher performance and higher price tags than ATA technology will continue to receive serious consideration in the storage industry. The question for storage solution providers and customers is this: *Is the information of such critical importance that it warrants paying the significant premium SCSI- and Fibre Channel-based systems will command over a SATA based solution?*

Given today's emphasis on IT costs, the exploding growth of storage capacity and the attractive performance roadmap and reliability features offered by Serial ATA, it is likely that many IT managers, especially in the Small and Medium Enterprise (SME) space, will say "Yes" to SATA for a growing number of enterprise storage applications.

➤ FREQUENTLY ASKED QUESTIONS (FAQS)

Is Serial ATA backward compatible with Parallel ATA?

Serial ATA is specified for software compatibility with Parallel ATA. In other words, no changes are required to operating systems and PATA drivers to use SATA hard drives. Computer vendors who would like to incorporate the advantages of Serial ATA will provide adapters that enable backward compatibility for SATA hard drives in today's computer systems.

Why is Serial ATA an acceptable alternative to SCSI for enterprise storage?

Serial ATA meets the demanding requirements of server and networked storage systems; needs that go beyond those of desktop systems. These critical requirements include:

- **Performance**–SATA provides high throughput (high data transfer speeds) to serve up large files quickly, maximizing storage system utilization to enhance workgroup productivity.
- **Reliability**–Lost data or lost access to data is a critical concern for any enterprise system, including those handling applications and data not technically considered "mission critical."
- **Availability**–SATA delivers high uptime and effective failover protection to meet the expectations of storage consumers who expect that their data will always be present and available.
- **Data Integrity**–SATA keeps data accurate and free from corruption.
- **Scalability**–SATA supports the ability to build onto an existing system, in terms of both performance and capacity, while preserving storage investments.
- **Interoperability**–SATA provides an enterprise storage solution that interoperates readily with other storage network elements, systems, and applications, leveraging industry standards.

The other critical enterprise priority today is *total cost of ownership* (TCO). Reducing the cost of satisfying the requirements listed above has become an issue of strategic importance both for storage system vendors and for their customers. The reduced cost of SATA implementation compared to SCSI addresses this enterprise concern.

Western Digital
20511 Lake Forest Drive
Lake Forest, California 92630

2579-001066-A01 02/04

For service and literature:
800.ASK.4WDC USA
+31.20.4467651 Europe
support.wdc.com
www.westerndigital.com