Maximizing ATA Disk Performance, Part 1

By: **Sean Daily** , **Windows IT Pro**

Should you use ATA (aka IDE) or SCSI hard disks in your Windows 2000 Professional systems? Whether to use ATA-based disks or their SCSI-based counterparts is one of the industry's oldest debates. In "SCSI and IDE: Defining the Differences," June 1997, I compared the technical underpinnings of the technologies' various flavors and discussed how to use either technology in Windows NT-based systems.
Since I wrote that article, the ATA and SCSI interfaces (not to mention NT) have evolved into newer, more powerful versions. The SCSI realm has expanded to include Ultra2 SCSI, fibre channel, Ultra160 SCSI, and the latest Ultra320 SCSI, providing maximum data transfer rates of 80MBps, 100MBps, 160MBps, and 320MBps, respectively. However, fast performance has always been a SCSI birthright, so perhaps a more significant development is how ATA specifications have bridged the price and performance gaps with their SCSI competitors.

ATA Bridges the Gap
The latest ATA/66 (aka Ultra DMA/66 or UDMA/66) and ATA/100 (aka Ultra DMA/100 or UDMA/100) standards offer fast burst and sustained data transfer rates (66MBps and 100MBps, respectively). The forthcoming SerialATA standard (scheduled to arrive in 2002 at initial speeds of 150MBps and later extend to 300MBps and 600MBps) will stretch ATA's performance boundaries even further. These developments make ATA a viable choice not only for consumer systems and low-end corporate desktops but also for systems that require excellent disk performance.

In real-world environments that use single-disk workstations, the aforementioned maximum data transfer rates aren't typically attainable. System constraints (e.g., ATA chipset limitations, the system bus architecture, physical disk limitations) often prevent such fast data delivery. Nevertheless, largely because of the price premium that SCSI commands (because of higher controller and disk costs), ATA is the primary storage for all but the highest-end desktop workstations. But maximizing the performance of ATA disks under Win2K isn't quite as simple as installing new disk drives and cables. In an effort to keep up with advancements in ATA storage technology, Microsoft has built some new capabilities into Win2K and has augmented Win2K's ATA support in various service packs and hotfixes. To get ATA devices working properly under Win2K, you need a basic understanding of the ATA interface, required hardware, and related software (e.g., Win2K service packs and hotfixes, built-in and third-party drivers).

Necessary Hardware
First, your system must contain a controller that supports your disks' advanced ATA modes. The most common standards that today's disks support are ATA/33 (aka Ultra DMA/33 or UDMA/33), ATA/66, and ATA/100. Virtually all ATA controllers are backward compatible with disks that use earlier standards. For example, an ATA/100 controller will typically support a disk that supports ATA/33 or even the older IDE or EIDE standards.

On most systems, this controller support is part of the motherboard's ATA chipset. (Intel commands most of the ATA chipset market, but chipset manufacturers such as VIAHardware.com also produce ATA controller chipsets.) However, in some cases, the controller might be an add-on PCI card, such as an ATA RAID controller.

The chipset—the most important component of your ATA system configuration—and its supporting driver control which features are available to disks and other devices attached to the controller. Therefore, an important first task is to thoroughly research the chipset on your motherboard or controller card and determine its capabilities. You can obtain this information from your system, motherboard, or controller-card vendor.

If your ATA chipset is embedded on your system motherboard, you need to verify that your system BIOS supports the ATA modes you intend to use. Although the chipset probably supports a particular ATA mode, the system's BIOS revision (which the system or chipset manufacturer typically updates regularly) might not support that mode. In such a case, you should be able to obtain a BIOS update from your system or motherboard manufacturer's support Web site.

Next, you need to ensure that your hardware offers the proper ATA support and is properly configured for optimal performance. First, make sure that all involved hard disks support the desired ATA modes of operation (e.g., ATA/66, ATA/100). Second, isolate the disks onto separate channels. This isolation is important because an ATA channel defaults to the speed of the slowest disk on the channel. If ATA/33- and ATA/100-capable disks reside on the same channel, the maximum attainable speed will be that of ATA/33. Therefore, place your slow ATA devices (e.g., CD-ROM, CD-R, CD-RW, Zip, older hard disks) on one channel and your fast ATA hard disks on another channel.

You also need to ensure that you're using the correct cables. The ATA/33, ATA/66, and ATA/100 specifications require the use of special 80-conductor ribbon cables, as opposed to the 40-conductor cables that earlier ATA disks use. The additional conductors function as grounding paths that improve the cable's signal-to-noise ratio. Finally, drive-placement requirements for the 80-conductor cables are different from earlier cables' requirements. You must place the master device (Drive 0) at the end of an 80-conductor cable, and you must place the slave device (Drive 1) at the middle connector. The blue connector at one end is for the host system or controller card, the gray connector in the middle is for the slave device, and the black connector at the other end is for the master device.

Maximizing ATA Disk Performance, Part 2

By: **Sean Daily** , **Windows IT Pro**

In "Maximizing ATA Disk Performance, Part 1," January 2002, InstantDoc ID 23297, I showed you how to obtain the best possible performance from ATA (aka IDE) disks in Windows 2000. I described the market's latest ATA standards and discussed some hardware pitfalls that you might experience while attempting to properly configure your ATA disks (e.g., ATA/33, ATA/66, ATA/100). This month, I conclude my two-part series with a discussion of software. I also attempt to fill in some of the information gaps in Microsoft's spotty Knowledge Base documentation about ATA. But before you implement this column's steps or advice, be sure that you've completed last month's hardware-specific steps.

A Little Perspective

To gain some perspective on the evolution of ATA disks under Win2K and Windows NT, you need to know a bit of history. Original IDE/ATA specifications used a data-transfer method called Programmed I/O (PIO), which relied heavily on the system CPU to handle data transfers from a disk's cache memory to system memory. The first PIO modes of operation (0, 1, and 2) fell under the umbrella of the original ATA standard. The later ATA-2 specification (aka EIDE or Fast ATA) included two new and faster PIO modes (3 and 4), as well as two modes that used direct memory access (DMA). By providing a less CPU-intensive data-transfer method, DMA addressed the burden that PIO imposed on the system CPU. To facilitate data-transfer operations, DMA also gave the IDE controller direct access to system memory. (Other controller types, including SCSI and Ethernet adapters, also use this methodology.) DMA eventually became the current ultra DMA (aka Ultra ATA, ATA/33, or DMA-33) standard, which supports a new mode called DMA Mode 3. This standard became so prevalent in the industry that the term UDMA—which refers to a data-transfer mode rather than a disk interface standard—is interchangeable with ATA in discussions about modern ATA disks (e.g., ATA/33 and UDMA/33, ATA/66 and UDMA/66). Also, for reasons of backward compatibility, modern ATA disks can support all data-transfer modes—including PIO and DMA—to their maximum capabilities.

Early UDMA disks and controllers were capable of operating at speeds of 33MHz, which meant that they could theoretically transport data at a maximum burst rate of 33MBps. Over time, enhanced and updated versions of UDMA supported operation at 66MHz/66MBps (ATA/66), using DMA Mode 4, and at 100MHz/100MBps (ATA/100), using DMA Mode 5. Virtually all the ATA disks available today are ATA/66 or ATA/100. These new capabilities have provided ATA with performance capabilities previously exclusive to SCSI disk subsystems and narrowed the gap between the ATA and SCSI standards. Also, because of ATA devices' lower price point, ATA has become preferable to SCSI for all but the most demanding applications (in which SCSI still holds some performance advantages).

Driver Developments
Despite these advancements in ATA technology and the ubiquitous presence of ATA disks in today's PCs, Win2K and NT are notoriously bad at optimizing the performance of ATA devices. Under NT, the root of this problem is the driver model that Microsoft implemented for ATA disks. Microsoft provided one pseudo-SCSI miniport driver—atapi.sys—which combines with scsiport.sys to form a pseudo-SCSI driver that supports all ATA devices on the system. This one-size-fits-all setup works fairly well and doesn't exact any significant performance penalties for older IDE controllers and devices. However, the atapi.sys driver often hamstrings newer controllers and devices because it doesn't take advantage of those devices' advanced PIO and DMA operation modes. Although several NT service packs provide atapi.sys new abilities to take advantage of the newer transfer modes, third-party drivers are still necessary to ensure optimum performance.

Under Win2K, Microsoft changed the ATA driver model from one that used atapi.sys as a monolithic, pseudo-SCSI miniport driver to one that implements a full-blown port driver (similar to scsiport.sys) for ATA. As a result, vendors can now implement specific enhancements (e.g., DMA, busmastering) in special MiniIDE drivers. Therefore, the OS can leverage each controller's special capabilities.

Although Microsoft has greatly improved the support for newer ATA controllers in Win2K, many ATA-support problems remain—particularly regarding support for the most recent DMA Mode 4 and 5 (ATA/66 and ATA/100, respectively) data-transfer methods. These problems are unfortunate because fully realizing the capabilities of ATA/66- and ATA/100-capable devices is impossible without enabling these DMA data-transfer modes.

ATA Drive Standards:

• ATA (IDE) - Supports PIO modes 0,1 & 2, SingleWord DMA modes 0,1 & 2, MultiWord DMA mode 0. Running at a maximum of 8.3MBps.
• ATA-2 (Enhanced IDE or Fast ATA) - Support for PIO modes 3 & 4, MultiWord DMA modes 1 & 2. Also supports LBA (Logical Block Addressing and block transfers). Running at a maximum of 16.6MBps.
• ATA-3 (Ultra-ATA, Ultra DMA, ATA33, DMA33) - Supports Ultra DMA mode 2 running at 33MBps. Adds improved reliability and drive security.
• ATA-4 (ATA66, Ultra DMA66, DMA66) - Supports DMA mode 4. Running at a maximum of 66MBps. (Requires 80 wire cable to acheive 66MBps [more information on cables below and comments on current drive speeds])
• ATA-5 (ATA100, Ultra DMA100, DMA100) - Supports DMA mode 5. Runs at a maximum of 100MBps.

• ATA-6 (ATA133, Ultra DMA133, DMA133) - Supports DMA mode 6. Runs at a maximum of 133MBps. Adds support for drives larger then 137GB.

All ATA standards are 16bit, however current BIOS's implement a feature that links two 16bit transfers together so half the 32bit PCI bus isn't wasted.

Although they get very confused and misinterpreted (as seen above) **ATA** is a drive interface standard. **IDE** (**I**ntegrated **D**rive **E**lectronics) and **EIDE** (**E**nhanced **IDE**) are drive design or configuration standards. **D**MA (**D**irect **M**emory **A**ccess) and **UDMA** (**U**ltra **DMA** are access methods for the drive(s), which allow transfer(s) to occur between the drive and RAM while bypassing the CPU (the data is 'controlled' by the hard drive controller and data is passed along a DMA Channel). **ATAPI** (**AT A**ttachment **P**acket **I**nterface) is an extension of ATA-2 (EIDE) that allows the interface to support CDROMs and tape drives.

Keep in mind that both ATA66 and ATA100 require 80 wire 40 pin EIDE cables, and that ATA33 only requires a 40 wire/pin IDE cable. The extra 40 strands in a ATA66/100 cable act as insulators between the 40 signalling strands to prevent/reduce crosstalk. You can use a ATA33 drive on a ATA66/100/133 channel and since all current controllers provide independent device timing, they will not affect each other's operation.

The current (Spring 2001) best/fastest drive on the market is the IBM Deskstar series (according to popular opinion and StorageReview.com). It only pushes 37.5MBps sustained, this is barely enough to justify the upgrade to ATA66 (for the 4.5MBps gain over ATA33). Some people may argue "Burst Transfer" rates being above this however the drive only achieves "Burst Transfer" rates when reading/writing it's own memory cache.

ATA133 is the current step, with speeds of 133MBps. Shortly we'll see SerialATA with speeds starting at 150MBps and progressing to 300MBps then 600MBps. The benefits of a 150MBps bus are quickly realised when you have multiple drives eating up 37.5MBps or more each since multiple drives can access the bus at the same time with SerialATA.

"INTEL DEVELOPER FORUM CONFERENCE, SAN JOSE, Calif. - August 24, 2000 - Seagate Technology, APT Technologies, Inc. and Vitesse Semiconductor Corporation today unveiled the first Serial ATA disc drive, giving a glimpse into the future of ATA disc drive technology. The drive is natively attached to an Intel® Pentium® 4 processor system through an APT Serial ATA PCI Host Bus Adapter, featuring a 1.5 Gbps transfer rate. The prototype demonstration combines technologies from Seagate, APT, Intel Corporation and Vitesse. It features a Seagate disc drive with its Serial ATA board, using APT's Serial ATA Link and Transport layers logic and Vitesse's 1.5 Gbs CMOS transceiver, attached via Serial ATA to APT's Serial ATA to PCI host bus adapter."

Benefits of SerialATA include: smaller cables (about the size of a CAT5 network cable (6 to 8 pins), hot plugging, multiple drives can access the channel at one (point-to-point), backward and future compatibility, compatibility with current OSs/software, reduced electrical requirements, and no jumpers.

SCSI (scuzzy) **S**mall **C**omputer **S**ystem **I**nterface, a parallel device interface standard (not limited to only storage devices, like IDE is) allowing for faster transfer rates then standard serial or parallel interface(s). It also allows multiple devices to access the same SCSI channel at the same time. (With IDE only one drive can be accessing the channel at a time).

SCSI Drive Standards:

• SCSI1 - An 8-bit bus, and supports data rates of 4 MBps
• SCSI2 (SCSI) - Same as SCSI1, but uses a 50-pin connector instead of a 25-pin connector, and supports multiple devices. This is what most people mean when they refer to plain SCSI.
• Wide SCSI - Uses a wider cable (168 cable lines to 68 pins) to support 16-bit transfers.

• Fast SCSI - Uses an 8-bit bus, with double the clock rate to support data rates of 10 MBps.
• Fast Wide SCSI - Uses a 16-bit bus, and supports data rates of 20 MBps.
• Ultra SCSI - Uses an 8-bit bus, and supports data rates of 20 MBps.
• SCSI3 - Uses a 16-bit bus and supports data rates of 40 MBps. Also called Ultra Wide SCSI.
• Ultra2 SCSI - Uses an 8-bit bus and supports data rates of 40 MBps.
• Wide Ultra2 SCSI - Uses a 16-bit bus and supports data rates of 80MBps.

RAID Redundant Array of Inexpensive (or Independent) Disks.

RAID Levels:
These Require 2 drives (or more):
• Level 0: Provides data striping which spreads the data blocks which make up your file(s) across multiple disks, resulting in improved performance but no fault tolerance.
• Level 1: Provides disk mirroring whereby data is written to multiple disks simultaneously. This results in almost complete fault tolerance (if one disk fails you simply read off the other exact copy).
These Require 3 drives (or more):
• Level 3: Same as Level 0, but also reserves one dedicated disk for error correction data. This results in good performance and partial fault tolerance.
• Level 5: Provides data striping at the byte level and also stripe error correction information. This results in excellent performance and good fault tolerance.
This requires 4 drives (or more):
• Level 0+1: Provides mirroring of striped arrays (4 or more 1 pair of striped disks mirrored onto another pair). This results in improved performance and almost complete fault tolerance.
RAID Info covers all levels including non-standard (1, 2, 3, 4, 5, 6, 7, 10, 53, & 0+1)

Note:
• The speeds listed above are all theoretical maximums.
• Standards which never made it to market or wide use are not covered

SATA Hard Drives

(Serial ATA) from Hitachi/IBM

- **SATA/150 Interface**
- **Up to 400 GB at 7200 rpm**
- **8MB buffer size**
- **Industry-leading 8.5ms average seek time**
- **Lowest power requirements**

These SATA hard drives extend Hitachi/IBM's long-standing tradition of performance leadership. The SATA Deskstar hard drives incorporate both new and proven technologies to boost the performance, capacity, and reliability to handle a wide range of advanced storage requirements.
Hitachi's standardized features in desktop solutions enable fast transfer rates, low power utilization and quiet acoustics. These features combine to make the Deskstar SATA hard drives ideal not only for standard PC systems but also for entry-servers

Making the Most of SATA Drives

By: **David Chernicoff** - Enterprise Storage **Web Exclusive**

Lately I've gotten a flurry of questions about the suitability of Serial ATA (SATA) drives for use in entry-level and midrange servers. Obviously, the easy answer is "yes, they work great and are an inexpensive alternative to SCSI." However, the more accurate answer isn't quite that simple.

The biggest problem I've seen with the use of SATA drives is that vendors or buyers of those devices tend to stick a couple of large drives in the computer and assume that capacity will fulfill disk requirements, especially in entry-level systems. Large-capacity (i.e., up to 400GB) SATA drives are so inexpensive that it seems to make sense to throw in the biggest drives you can afford. OEMs encourage this practice, offering their entry-level systems with a 120GB or 250GB drive as available configurations.

The problem is that, from a best-practices perspective, you shouldn't be sticking your OS and all your data on one drive. Too many things can go wrong with the OS--from simple crashes to nasty virus infections--to risk housing your data and OS on the same drive. Even when vendors try to address this concern, their solution isn't always optimal. For instance, one major server vendor recommends that its entry-level servers be equipped with at least two drives, which, on the surface, seems like a good idea. However, the vendor is actually recommending two drives so that they can be mirrored. Mirroring provides additional insurance against disk crashes but still leaves you open to virus infections and malicious software (malware) attacks.

If you have two SATA drives, I suggest that instead of mirroring them, you use the smaller one as a boot drive, configure the OS on it, and dedicate the second drive to data storage. If the server supports RAID (as do many servers, even entry-level machines), decide what type of RAID protection is suitable for the data drives. For the boot drive, which will change very little, choose a good backup and restore product that will ensure minimal downtime should you need to replace that drive. Or, if a virus infection or an OS failure occurs, you can use the backup to restore the OS to an operational state on the remaining functional drive.

SATA drives are a great boon to entry-level servers, and not because of their capacity or performance. (The new Seagate 400GB SATA drive is also available with an ATA 100 interface and effectively delivers the same performance over both interfaces.) The simplified cabling and increased cable lengths of SATA as compared with ATA mean that low-cost servers can have storage capacities once reserved for high-end boxes; it's now easy to deliver a server with more than a terabyte of SATA storage.

It's still important to remember that SATA isn't a replacement for SCSI in heavily utilized, high-performance environments. Although SATA offers improvements over ATA, SCSI's strengths still set it apart from SATA. If your primary need is to store large amounts of lightly used data, SATA is an excellent choice. For mail or database server applications, though, you're still better served--even at the entry level--by SCSI storage.