



Serial ATA

A Comparison with Ultra ATA Technology

In past years, increasing hard disk transfer rates have forced the ATA interface specification to be continuously updated to avoid becoming the limiting factor in disk I/O performance. As consumers embrace new usage models such as digital video creation and editing, digital audio storage and playback, file sharing over high-speed networks, and other data intensive applications, demands on hard drive throughput are expected to increase even further. To keep pace, the storage interconnect must be developed beyond existing Ultra ATA technology. The new approach is Serial ATA, a serial implementation of the parallel Ultra ATA interface. With this paradigm shift in I/O design, the roadmap of ATA will be extended beyond the theoretical limits of the Ultra ATA bus.

The purpose of this document is to educate the reader on the technical differences between Ultra ATA and Serial ATA technology, and to provide explanation for the transition from a parallel to serial bus architecture. The key design points of each technology will be described and compared, followed by an overview of the system level and end-user advantages of Serial ATA technology. The ATA protocol itself will not be discussed, as in this sense there is no difference between the technologies. Serial ATA is software compatible with the ATA interface and thus will appear to the OS as a standard ATA device. Note that it is assumed the reader has an understanding of electrical engineering design principles; the paper is intended primarily for OEMs, system designers, and product manufacturers who are considering adding Serial ATA capability to their designs.

Technology Introduction

Ultra ATA is the primary internal storage interconnect for the desktop, connecting the host system to peripherals such as hard drives, optical drives, and removable magnetic media devices. Ultra ATA is an extension of the original parallel ATA interface introduced in the mid 1980's and maintains backward compatibility with all previous versions of this technology. The latest revision of the Ultra ATA specification accepted by the ANSI supported INCITS T13 committee, the governing body for ATA specifications, is ATA/ATAPI-6, which supports up to 100Mbyte/sec data transfers. Development of the ATA/ATAPI-7 specification, an update of the parallel bus architecture that provides up to 133Mbytes/sec, is currently being finalized. (www.t13.org)

Serial ATA is the next-generation internal storage interconnect designed to replace Ultra ATA technology. Serial ATA is the proactive evolution of the ATA interface from a parallel bus to a serial bus architecture. This architecture overcomes the electrical constraints that are increasing the difficulty of continued speed enhancements for the classic parallel ATA bus. Serial ATA will be introduced at 150Mbytes/sec, with a roadmap already planned to 600Mbytes/sec, supporting up to 10 years of storage evolution based on historical trends. Though Serial ATA will not be able to directly interface with legacy Ultra ATA hardware, it is fully compliant with the ATA protocol and thus is software compatible. (www.SerialATA.org)

Parallel vs. Serial Bus Architecture Overview

Ultra ATA Bus Architecture

Bus Design – The latest revision of the ATA specification, ATA/ATAPI-6 where Ultra ATA 100 is defined, maintains backward compatibility with all previous ATA revisions, using the standard 16-bit wide parallel data bus and 16 control signals across a 40-pin connector.

Bandwidth – To understand the 100Mbytes/sec throughput, several factors must be considered. With a 16-bit data bus, two bytes are transmitted per bus transaction. Thus to achieve a throughput of 100

Mbytes/sec, the data bus must be clocked at 50MHz. To minimize strobe design complexity, Ultra ATA uses a “double data rate” or double-edge clocking mechanism for all Ultra DMA transfers. Using this technology, data is registered both on the rising and falling edges of the data strobe, halving the required strobe frequency. Thus the bandwidth is:

$$\begin{array}{rcl}
 & 25\text{MHz strobe} & \\
 \times & 2 \text{ for double data rate clocking} & \\
 \times & 16 \text{ bits per edge} & \\
 \hline
 & 8 \text{ bits per byte} & \\
 = & 100 \text{ Mbytes/sec} &
 \end{array}$$

Timing – As mentioned above, data must be clocked at 50MHz, or every 20ns. Note that because of data setup and hold times, all data lines must in fact switch and settle within approximately 10ns (see *Figure 1*). It is this worst case switching time that designers must meet.

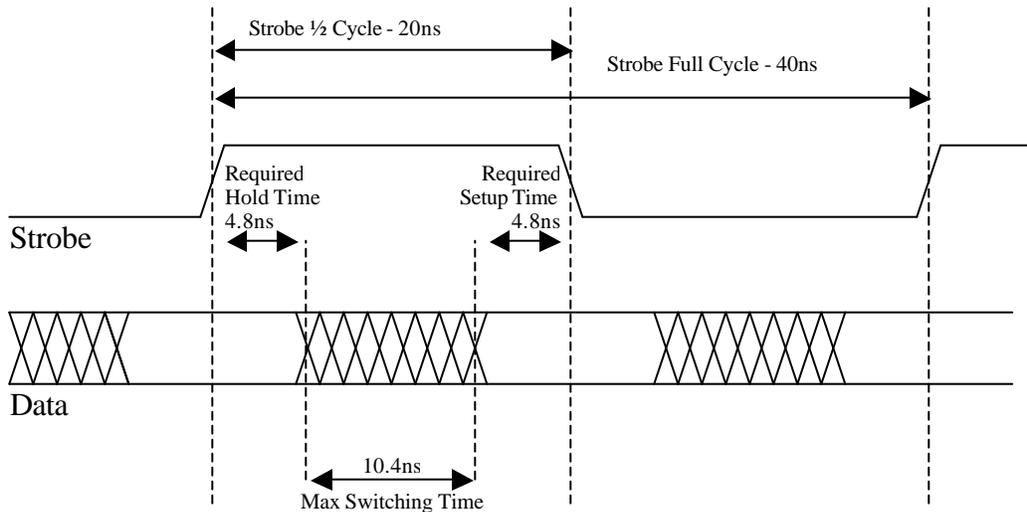


Figure 1 – Ultra DMA Burst Transfer Timing

Serial ATA Bus Architecture

Bus Design – In contrast to Ultra ATA’s parallel bus design, Serial ATA uses a single signal path to transmit data serially, or bit by bit, and a second serial path to return receipt acknowledgements to the sender. Because each of these signal paths is a 2-wire differential pair, the Serial ATA bus consists of 4 signal lines per channel. Control information is transmitted either as short predefined bit sequences that are distinguishable from data, in packet format, or using out-of-band signaling (control signals sent using on/off signal pulses, similar to Morse code), and thus does not require separate transmission lines.

Bandwidth – The 16-bit wide parallel Ultra ATA bus is capable of transmitting two bytes of data per clock. Though Serial ATA transmits only a single bit per clock, the serial bus may be run at a much higher speed to compensate for the loss of parallelism. Serial ATA will be introduced with a bandwidth of 1500Mbits/sec, or 1.5Gbits/sec. Because data is encoded using 8b/10b encoding (an 80% efficient encoding used with digital differential signaling to maintain a constant average “DC” bias point), the effective maximum throughput is 150Mbytes/sec.

$$\begin{array}{rcl}
 & 1500\text{MHz embedded clock} & \\
 \times & 1 \text{ bit per clock} & \\
 \times & 80\% \text{ for 8b10b encoding} & \\
 \hline
 & 8 \text{ bits per byte} & \\
 = & 150 \text{ Mbytes/sec} &
 \end{array}$$

Timing – The 1.5Gbits/sec transfer rate requires bit transitions and reception to occur within 0.667ns. The maximum allowed switching time is 0.273ns, much faster than the 10ns transition time allowable for Ultra ATA. However, as seen in the following sections the electrical design parameters of the serial bus are more tightly controlled. Serial ATA can thus meet and surpass the timing required to deliver throughput comparable to that of Ultra ATA.

Electrical Design Constraints – Overcoming Parallel Design Complexities with a Serial Bus

Optimization of any high-speed digital bus design in fact requires careful consideration of analog design issues. Undesired analog effects associated with parallel data busses such as crosstalk, ground bounce, ringing, and clock skew have become major design constraints for the Ultra ATA interface, which is forced to maintain compatibility with legacy parallel technology. These same issues are expected to become critical roadblocks to further Ultra ATA speed increases. Serial ATA alleviates many of these problems by transitioning to a serial data bus.

The intent of the following section is to first present the design methods required to achieve the current Ultra ATA data rate and illustrate the complexities of further speed enhancements to the parallel technology. In each case, the paper will describe how the Serial ATA bus architecture overcomes these complexities to extend the performance roadmap of ATA. Though full explanation of the analog design issues discussed is beyond the scope of this document, a brief overview of these noise and error sources follows to provide the reader with the background necessary to compare the electrical properties of each bus design:

- Crosstalk results from magnetic fields generated from transitioning currents being coupled into neighboring current loops, similar to the functionality of a transformer. The magnitude of the crosstalk is proportional to the rate of the change in the current and the amount of coupling between the current loops. Thus it is most apparent in parallel busses where multiple adjacent lines may be switching in the same direction at the same time and inject a noise voltage onto a victim signal.
- Ground bounce is most problematic when several signals switch at the same time or when using high-speed drivers, both common with parallel data busses. The instantaneous power draw is such that the decoupling capacitors for the device cannot supply the necessary current and the supply voltage sags. If the voltage decreases enough, the change can be mistaken as a bit transition.
- Ringing results from impedance changes in a signal path in systems in which signal rise time is close to or faster than the path propagation delay. When this condition holds true, the signal path must be seen as a distributed system, implying that all points on the signal path may not be at the same voltage at the same time. As the signal propagates down a path, the voltage magnitude is related to the “effective” impedance up to the point in the signal’s flight. If this impedance suddenly changes, the voltage temporarily increases to maintain the current flow. This voltage “reflects” back along the transmission line to the source, where if not fully damped may reflect again to the receiver, and may repeat until effectively damped. This causes oscillations in the voltage, or ringing, on the transmission line.
- Clock skew results from discrepancies in transmission path delays between clock and data signals or signal degradation of the clock signal. If the clock trace is shorter than the data lines, for example, the clock signal may arrive at the receiver before the data lines have stabilized, thus registering incorrect data. Alternatively, ringing or noise on the clock line can delay the clock transition relative to the data switching, possibly violating data hold time.

Electrical Design Overview

Ultra ATA – Maintaining Compatibility with a Legacy Parallel Design

The PC industry adopted the ATA-1 standard as the primary storage I/O interface in the mid ‘80s, and has maintained backward compatibility to this original standard ever since. The protocol has scaled well, from 3.3Mbytes/sec with ATA-1 to 100Mbytes/sec with ATA/ATAPI-6. However, this radical increase in speed has come with added design complexity. In the following sections, many of the key design advances to reach 100Mbyte/sec transfer rates will be discussed and their limitations for further speed increases explained.

Serial ATA Electrical Design Considerations

Due to the challenges in increasing the speed of the Ultra ATA specification further, a shift in design strategy is required. Serial ATA addresses this need by making the transition to a high-speed serial bus. To mitigate many of the design problems associated with high-speed single-ended and/or parallel busses, Serial ATA uses low voltage differential signaling. With this approach, each data “signal” is in fact transmitted over two lines which carry equal and opposite versions of the signal. The receiver then decodes the signal based on the differential voltage between these lines. The “common-mode” voltage, or the voltage the lines use as a DC reference plus any noise injected equally into both lines, is rejected at the receiver. This common-mode voltage may change over time, though the variations above a certain frequency may be injected into the receiver as noise. These excellent electrical properties provide many of the key design advantages that enable Serial ATA to extend beyond Ultra ATA speeds.

Clocking Strategy

Ultra ATA – Non-interlocked clocking

Because of the high data rates and relatively high board and cable propagation delays, Ultra ATA uses non-interlocked clocking, also known as source-synchronous clocking. In typical synchronous clocking designs, data is transmitted from the source and clocked at the receiver using a local clock signal. With non-interlocked clocking, the clock or data strobe signal is generated at the source and sent with the data. Assuming identical trace or cable lengths and characteristics, both data and strobe arrive at the receiver at the same time.

This technique allows more flexibility in total propagation delay along the bus, but introduces additional complexities. Because the strobe is sent with the data, it is subject to ringing and reflections. If this noise is great enough, data may be “double clocked” if the strobe signal crosses the switching threshold. Data settling time is often more critical, as the strobe transition is typically more aggressively aligned with the data transition.

Clock skew becomes more complicated as well, as both data and clock propagation times may vary from the predicted delay. To reduce problems related to skew, 100MHz transfers (Ultra DMA Mode 5) must use 3.3V signaling (vs. 5V with previous ATA specs) so that signal transitions are more symmetric about the 1.5V switching threshold. Termination impedances are also more tightly constrained to reduce ringing in the signals that could cause plateaus or bumps in the signal edge, resulting in delayed threshold crossings.

Serial ATA – Embedded Clocking

Unlike the parallel ATA bus, Serial ATA does not have a separate signal dedicated as a strobe or clock. Instead the clock is “embedded” in the data stream itself. When no data is being sent across the bus, a “101010...” pattern is transmitted so that both devices may synchronize their internal receivers with the incoming bit transition timing. This synchronization is maintained during data transmissions. The 8b10b encoding enforces several bit transitions per 10 bits even during data transmission; clock drift is minimized by continuously tracking these transitions. Embedded clocking provides the timing benefits of source-synchronous clocking without introducing problems associated with clock skew.

Cabling

Ultra ATA – 80-wire ribbon cable

Until ATA/ATAPI-3, or Ultra ATA 33, the ATA interface used a 40-wire cable to transmit data, of which only 7 signals were ground. Because crosstalk is proportional to the size of the mutual current loops between signal lines, the large separation between each signal and its respective return ground line results in significant crosstalk over this cable. For transfer rates greater than 33MHz (Ultra DMA Mode 3), the original 40-wire cable has been replaced by an 80-wire version with alternating ground and signal lines. This greatly minimizes crosstalk among signals and helps to balance the effective impedance of each line at high frequencies. However, the cable is specified to be at most 18” to minimize signal integrity issues.

Serial ATA – 4-wire cable with support for optional shield / drain wires

Serial ATA uses a minimum 4-wire cable that includes differential pairs for transmitting and receiving data. To minimize impedance and crosstalk, many cables incorporate additional shielding ground drain lines which function similarly to the 40 interspersed ground lines in the 80-pin Ultra ATA cable. The Serial ATA connector supports 3 independent ground return paths. Serial ATA cables are specified to be at max 1m in length.

Connectors

Ultra ATA – 40-pin dual row header

Though the cable has been updated for use in high-speed data transfers, the ATA connector has remained the standard 40-pin dual row header to maintain backward compatibility. The 40 extra ground wires in the cable are tied to the 7 ground pins in the connector. Because additional ground lines have not been added, inductive coupling in the connector introduces a significant amount of crosstalk during switching. The effect of crosstalk is greatest on a signal that remains constant while all neighboring signals transition in the same direction, up to 1V worst case. This generated noise is most problematic when transmitting to the device in the middle of the cable, or when receiving from this device. As crosstalk is proportional to the change in current over time, it can be reduced by limiting the rise and fall times, or slew rate, of the bus drivers. However, this strategy forces lower clock speeds and thus is not conducive to bus speed increases.

Serial ATA – 7-pin custom connector

The 0.5” wide cable connector directly connects the 4 signal wires and 3 ground lines to the receiving terminal in a single row. Because the connector includes the shielding ground pins, very little crosstalk is introduced. Note that the receiving terminal uses extended connectors for the 3 ground signals so that the ground reference between the device and host can be shared prior to signals being applied at the input. A similar mating sequence is enforced with the new 7/8” wide 15-pin single row power connector. This feature is necessary to accommodate hot-plugging.

Termination Strategy

Ultra ATA – Source Termination

Ultra ATA configurations are source terminated to minimize ringing. Using this termination scheme, a series resistor is placed at the driver’s transmit output. The magnitude of this resistance is chosen so that the resistance plus the output impedance of the transmitter match the controlled trace and/or cable impedance. This produces a voltage divider at the output of the driver that effectively halves the strength of the emitted signal. When this signal reaches the receiver, the signal reflects at the impedance mismatch formed between the controlled impedance transmission line or cable and the very high input resistance of the receiver. This reflection doubles the strength of the signal at the receiver, and thus the signal returns to the original amplitude. When the reflected signal propagates back to the source, it sees a continuous impedance path to ground and thus is completely damped assuming perfect termination. If the termination resistor is poorly chosen, some of this reflected signal will be again reflected and will cause ringing in the signal. If the ringing is of great enough amplitude, settling time of the signal may be affected.

This termination scheme is very effective when using a single driver and receiver at opposite ends of the connecting cable. The standard ATA cable, however, allows a second device to be attached in the middle of the signal path. At this point in the cable, the signal experiences a “plateau” as the transmitted signal will arrive at half strength and the reflected signal must propagate backwards from the receiver before full voltage swing is achieved. If overshoot of the initial signal is great enough, the half-amplitude signal may cross the switching threshold on or more times before the reflected signal arrives. This overshoot can also be controlled by limiting output slew rate, but as mentioned above this solution will be problematic if greater bus speed is desired.

Serial ATA – Tighter Impedance Specifications / Support for Automatic Impedance Matching

Since Serial ATA uses only 4 signal lines per channel, proper termination of all lines is less costly, both in design complexity and dollars. All devices are required to provide precision termination impedances. Support is also provided for active impedance matching circuits that ensure exact matching to any cable or device. Though Serial ATA utilizes the same source termination scheme as parallel ATA, many of the problems are reduced by this near-perfect termination and the point-to-point connection topology, which ensures that the only receiver is at the endpoint of the transmission line.

PCB Routing

Ultra ATA – Parallel Data Bus and Clock Routing Constraints

Because the ATA interface uses 32 signal lines for each channel, with a typical configuration of 2 Ultra ATA channels per board 64 signals must be routed from the I/O controller to the ATA connector. To minimize crosstalk and ringing, trace widths and separations must be carefully controlled. Typically a maximum routing length of 8” is specified. All trace lengths are typically required to be within 0.5” from that of the strobe line to minimize clock skew.

Serial ATA – Differential Pair Routing Constraints

Each Serial ATA channel consists of two differential pairs, a total of 4 transmission lines. Each pair should be routed with the differential lines at constant separation and with equal length. Because of the high signaling speed and limitations

of the common FR4 PCB routing material, trace lengths should typically be less than 6". Though proper routing is more important with Serial ATA than it is with Ultra ATA, there are far fewer signals to route. Each Serial ATA channel requires 4 signal traces, thus support for 4 drives would require a total of 16 routed signals vs. 74 for Ultra ATA.

Signaling

Ultra ATA – Legacy 5V Tolerance

As mentioned above, devices that use Ultra DMA Mode 5, or 100 Mbytes/sec, transmissions must use 3.3V signaling to balance high-to-low and low-to-high transition times. To be fully backward compatible, Ultra ATA devices and hosts must be 5V tolerant to avoid being damaged when connected to ATA/ATAPI-5 or earlier devices. This 5V tolerance must be considered during IC design, as it is expected to become more difficult to support with newer CMOS design processes.

Serial ATA – Low Voltage Differential Signaling

The noise rejection capabilities of differential pairs allow for low voltage signaling. With Serial ATA, the voltage swing is 0.125V about the common-mode voltage, with a minimum common-mode voltage of 0.25V. Because Serial ATA does not maintain hardware compatibility with previous ATA specifications, the 5V tolerance constraint is removed.

Serial ATA – Beyond Better Signaling Characteristics...

Connection Methodology

Ultra ATA – Master/Slave Shared Bus

Ultra ATA technology supports up to 2 drives per channel via a shared bus. Though the two attached devices are referred to as "master" and "slave", there is no difference in operation between or priority given to the connected devices. The host bus adapter is the true bus master and uses the master/slave status of the drives to route requests to the correct device and to determine the boot device. Though Ultra ATA supports a command queuing algorithm, it is rarely incorporated into devices, and thus the data bus is locked if a command to either drive is outstanding. Because of this, the bus bandwidth is shared between the master and slave devices when both are actively interacting with the host.

Serial ATA – Point to Point Connections for Dedicated Bandwidth

Serial ATA uses a point-to-point connection topology, meaning that each source is connected to one destination. Each channel has the capability to work independently so that there is no contention between drives and thus no sharing of interface bandwidth. This connection strategy also negates the need for master/slave jumper settings on devices.

Cable and Connector Specification

Ultra ATA – Legacy Wide Ribbon Cable

Ultra ATA uses a 2" wide 80-wire cable with a maximum specified length of 18". To maintain backward compatibility with prior ATA revisions, Ultra ATA devices use a 40-pin dual-row header connector. This also implies that a 40-pin connector can be plugged in to incompatible Ultra ATA systems. Power is typically supplied by a separate 4-pin power plug. A smaller hardware-incompatible 46-pin version of the cable/connector assembly is available to deliver power and control through a single connector for use with small form factor drives.

Serial ATA – Thin, Flexible Cable with Small-Footprint Connector

Typical Serial ATA cables are at most 0.25" wide and can be up to 1m in length, allowing for proper routing to minimize restriction to airflow and reduce clutter. The 0.5" wide 7-pin data connector requires little area on board or device, important for upcoming 2.5" hard disk drives. The Serial ATA signal connector and additional power connector are thin enough to be used unmodified on all hard-drive sizes available today, negating the need for multiple connector types. To illustrate the reduced board space requirement, note that in a four drive system, Serial ATA connectors use 25% of the board space required by Ultra ATA.

Added Features

Ultra ATA – Basic Reliability Support

The ATA/ATAPI-6 revision includes support for a CRC error check on data transmissions to ensure that the data sent was received correctly. Control signal transmissions are not protected.

Serial ATA – Enhanced Reliability, Hot-Swap Support, First Party DMA

Serial ATA adds 32-bit CRC error correction for all bits transmitted, as opposed to only data packets in Ultra ATA. Also, Serial ATA supports hot-swapping via hardware support and by design of the connector (variable length pads, minimal insertion force design, and a specified location on device backing plates). With the proper software drivers Serial ATA devices may be hot-plugged internal to the box or blind-mated to a backplane or device bay. Serial ATA also provides built in support for first-party DMA, removing bottlenecks associated with on-board DMA controllers.

Conclusions

The intent of this document was to explain the transition to a new internal storage I/O bus design, and to illustrate the benefits of this transition. The paper described several electrical design challenges that complicate further speed increases to the Ultra ATA interface and provided explanation of how Serial ATA overcomes these issues. Next discussed were benefits of Serial ATA technology from the serial bus architecture, added features, and improved design.

Serial ATA Achieves and Surpasses Bandwidth of Parallel Ultra ATA through Design Enhancements

- Differential signaling allows rejection of crosstalk and ground-bounce as common-mode noise
- Embedded clocking provides advantages of source-synchronous clocking without problems related to clock-skew from poor strobe signal integrity and mismatched trace lengths
- More controlled cable design, allowing longer cables
- Serial ATA connector properly designed to reduce crosstalk, a major issue with the legacy Ultra ATA connector
- Automatic impedance matching greatly reduces ringing and thus settling time
- Less complexity and less board space required for trace routing
- Low voltage signaling (as low as 0.5V max voltage) allows for compatibility with future design processes
- Serial ATA bus architecture enhancements allow for an initial speed of 150Mbytes/sec, with a planned roadmap up to 600Mbytes/sec

Serial ATA Offers Advantages over Ultra ATA Technology

- Point to point connection topology ensures dedicated 150Mbytes/sec to each device
- Thinner, longer cables for easier routing
- Fewer interface signals require less board space and allow for simpler routing
- Better connector design for easier installation and better device reliability
- CRC error checking on all data and control information
- Hot-swap capability
- First-party DMA support

References

1. “Serial ATA: High Speed Serialized AT Attachment Revision 1.0”, Serial ATA Working Group
2. “Information Technology: AT Attachment with Packet Interface – 6 Rev. 3b”, T13 Committee
3. Johnson, Howard - “High-Speed Digital Design: A Handbook of Black Magic” – Prentice Hall, 1993