Local Bus Compatibility Issues

Systems Engineering
Steve Reames
25-Jul-94

Introduction

The ATA Bus (AKA IDE bus) was a disk drive interface originally designed for the ISA Bus of the IBM PC/AT. With the advent of faster interfaces, the definition of the ATA Bus has been expanded to include new operating modes. Each of these modes, numbered zero through five, is faster than the one before it (higher numbers translate to faster transfer rates). Mode 3 has been implemented only recently by both disk drive and PC motherboard vendors.

Problem Statement

New PC motherboards and interface cards supporting local bus interfaces (both VLB and PCI) are running disk drives at the maximum transfer rate specified by ATA mode 3 (11.1 MB/s). The ATA Bus is currently specified as an unterminated bus. This lack of terminations causes ringing on the signals traveling between the host and the disk drive (see figure 1). Previous bus speeds were sufficiently slow that this ringing was not a problem. The new higher bus speeds available with ATA mode 3 – now being exercised by local bus interfaces – are causing system failures with Quantum disk drives.

Failures have been observed with AT&T/NCR, Acer, Daewoo/Leading Edge, IPC, Goldstar, and TMC motherboards. Problems have been traced to the Local Bus bridge chips used in these systems, particularly those manufactured by Appian, Adaptec, CMD, and PCTECH. The issue of supporting local bus operation with ATA mode 3 is a concern for the Thunderbolt, Lightning, Maverick, Roadrunner, and Daytona products. Common failure modes include system hangs during file transfers, data miscompares, and failure during cold boot.

Summary

The currently proposed minimum solution involves adding seven 82-ohm resistors to the drive in series with the control lines CS0-, CS1-, DA0, DA1, DA2, DIOR-, and DIOW-, and adding a delay circuit (two transistors, three resistors, one capacitor) to the IORDY line. The recommended solution includes adding eighteen additional series resistors to the following lines in addition to the above signals: DD0 through DD15, DMACK-, and DMARQ. The minimum solution will solve operation problems for all of the drives in the affected group for all systems tested. The full recommended solution is necessary to operate with systems using fast DMA. These solutions have not been tested on all systems in all configurations, but both theoretical analysis and lab testing show favorable results.

Table of Contents

Section 1 – Ringing on Unterminated Lines

Disk Drive Failure Modes

Most failures in the systems observed can be attributed to signal integrity problems on the control lines that go from the host to the drive. The problem appears most frequently on the DIOR- (read command) and DIOW- (write command) lines.

DIOR-

During a read cycle when DIOR- is asserted, it is possible for the ringing to create a short duration deassertion pulse (figure 1). This pulse occurs early in the read cycle. Inside the ATA interface portion of KONI or NEKO is a FIFO buffer that contains the data to be read. The extra pulse on the DIOR- line advances the FIFO pointer by one. This results in losing one word of data. The host system read operation receives one word too few, and the remaining bytes are shifted. A typical data sequence might look like . . .W7, W8, W9, W11, W12 . . . Notice that word 10 is missing from the returned data. This also means that the host will try to read one more word from the drive than the drive has remaining. Depending on the implementation of the BIOS, this may lock-up the system or simply return an extra byte of garbage at the end of the sector.

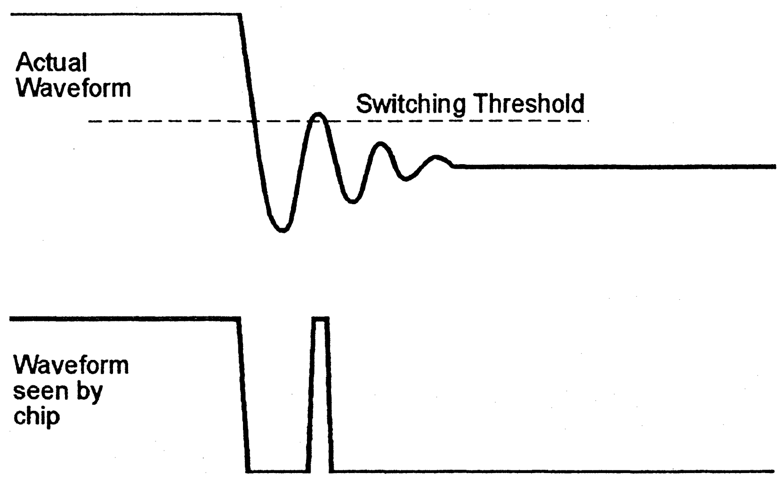


Figure 1 – Typical ringing on bus and its effect

DIOW-

Pulse slivers due to ringing on the DIOW- line cause a similar problem during writes. The pulse sliver advances the FIFO pointer by one unexpectedly, writing an extra word of garbage into the FIFO. Subsequent data bytes are shifted by one word. A typical stored data sequence on the drive might look like . . . W7, W8, W9, XX, W10, W11 . . . In this example an extra word was inserted during the write cycle for word 10. From the drive's point of view, the host is trying to write 514 bytes rather than the expected 512 bytes. The drive will throw away the final word and probably flag an error. A properly written BIOS will detect this error and indicate a problem to the user. The problems with the DIOR- and DIOW- lines have been seen primarily with PCI/ATA bridge chips manufactured by Appian and Adaptec (second source for Appian), though other bridge chips have shown similar symptoms.

Technical Discussion

The cause of the failures appears to be ringing on several signal lines that causes the drive to see false transitions. The ringing is due to the cable between the host and drive not being properly terminated. The cable acts as a transmission line, and as signals with fast edge speeds are applied, the system rings at its natural resonant frequency.

If the amplitude of the ringing is sufficient, then the voltage at the drive can cross the switching threshold and create a pulse sliver. Even ringing that itself does not cross the switching threshold can bring the voltage close enough to the switching point that the system becomes very susceptible to noise.

System measurements have shown the ringing problem is strongly associated with DIOR-, DIOW-, and CS0-. From the ATA specification, we can group all of the bus signals into seven basic structures (Figure 2). The ringing will be similar on signal lines with the same bus structure. Although problems have not yet been seen on some of the signals (such as DA0, DA1, or DA2), they are likely to have problems in the future because they are of the same bus structure as the problem lines. The effects of ringing will be different between signals depending on their function. The first group (DIOR-, DIOW-, CS0-, CS1-, DA0, DA1, DA2, DMACK-, RESET-) are control signals from the host to the drive. The signal DMARQ can be included in this group since it is a control signal from the drive to the host with a similar bus structure. DIOR- and DIOW- are symmetrical in their operation and will benefit from a similar fix. CS0-, CS1-, DA0, DA1, and DA2 are usually used in a combinatorial logic decode circuit in the drive. Any fix applied to these lines must either involve none of them or all of them. Although DMACK- and DMARQ have not shown any problems at this time, this is most likely because DMA at high speeds has not been used much at this time. This will change in the near future, so designers should consider implementing any fix to these lines also. Ringing on the RESET- line has no detrimental effect from a practical viewpoint, so this line can remain unterminated.

To study the effects of different solutions on these lines, a SPICE model has been constructed (figure 3). This model includes a host driver, 18 inches of ATA ribbon cable, and a disk drive. The resistors R100 and R200 are not currently in the system, and are part of the recommended solution. To model the system as it stands currently, both R100 and R200 are set to .001 ohms (SPICE does not permit zero-valued elements). Voltage monitoring is placed at the output of the host and at the input of the drive. The voltage at the input to the drive is the important signal for our purposes. The resulting waveforms are shown in figure 4.

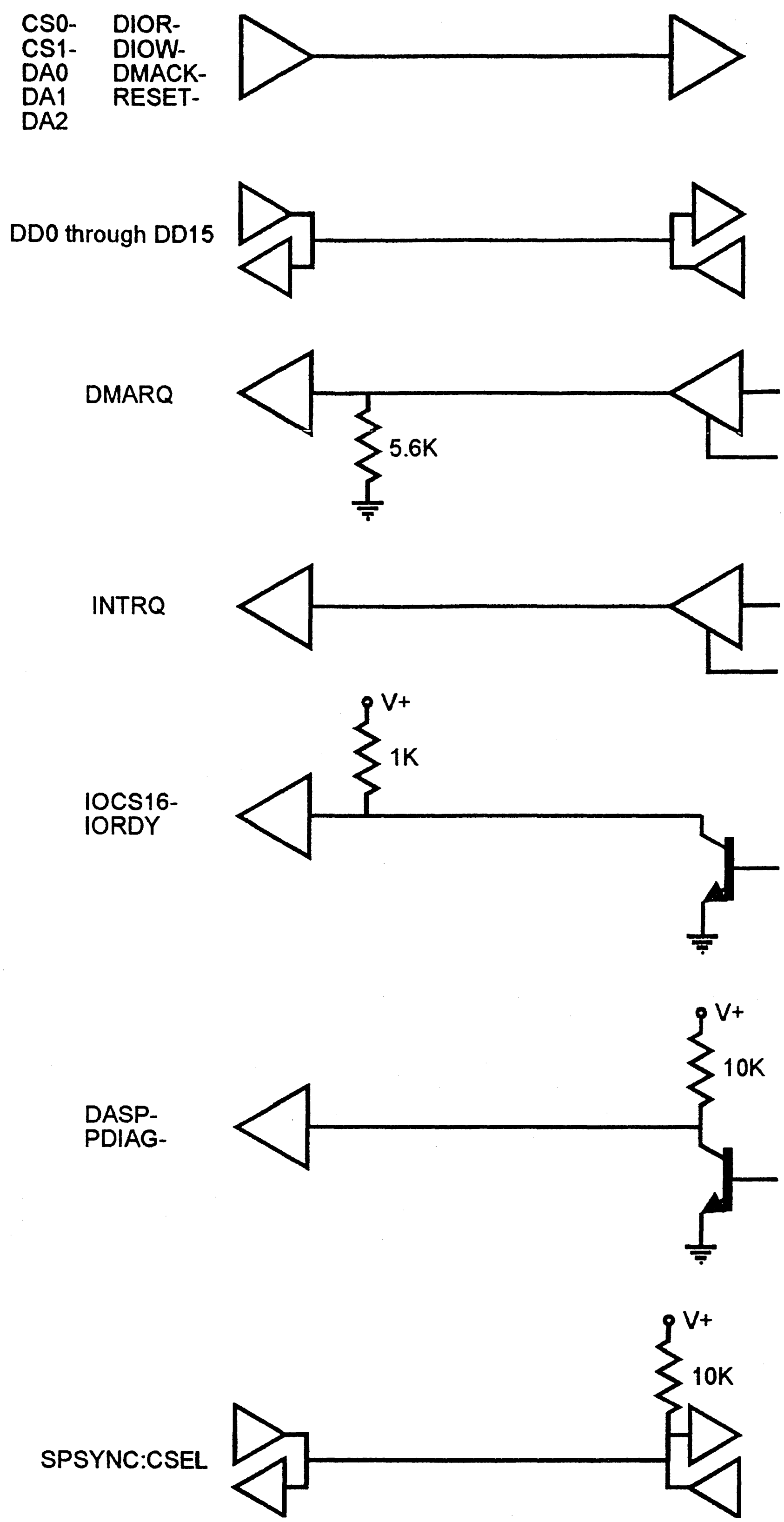


Figure 2 – The seven basic ATA bus structures

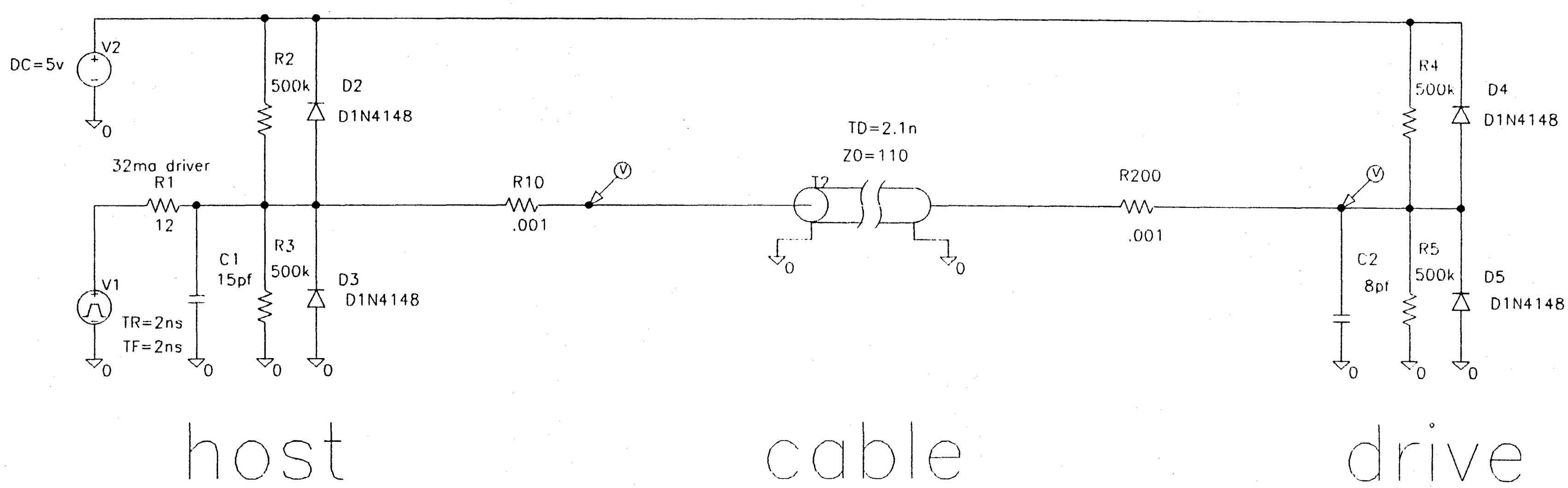


Figure 3

Quantum Corp.
500 McCarthy Blvd.
Milpitas, Ca. 95035

Page Size: A

ATA Interface Model

BY: Steve Reames

Revision: –

July 18, 1994 Page 1 of 1

* C:\STEVE\PSPICE\ATA1.SCH

Date/Time run: 07/18/94 16:35:23 Temperature: 27.0

(A) Figure 4

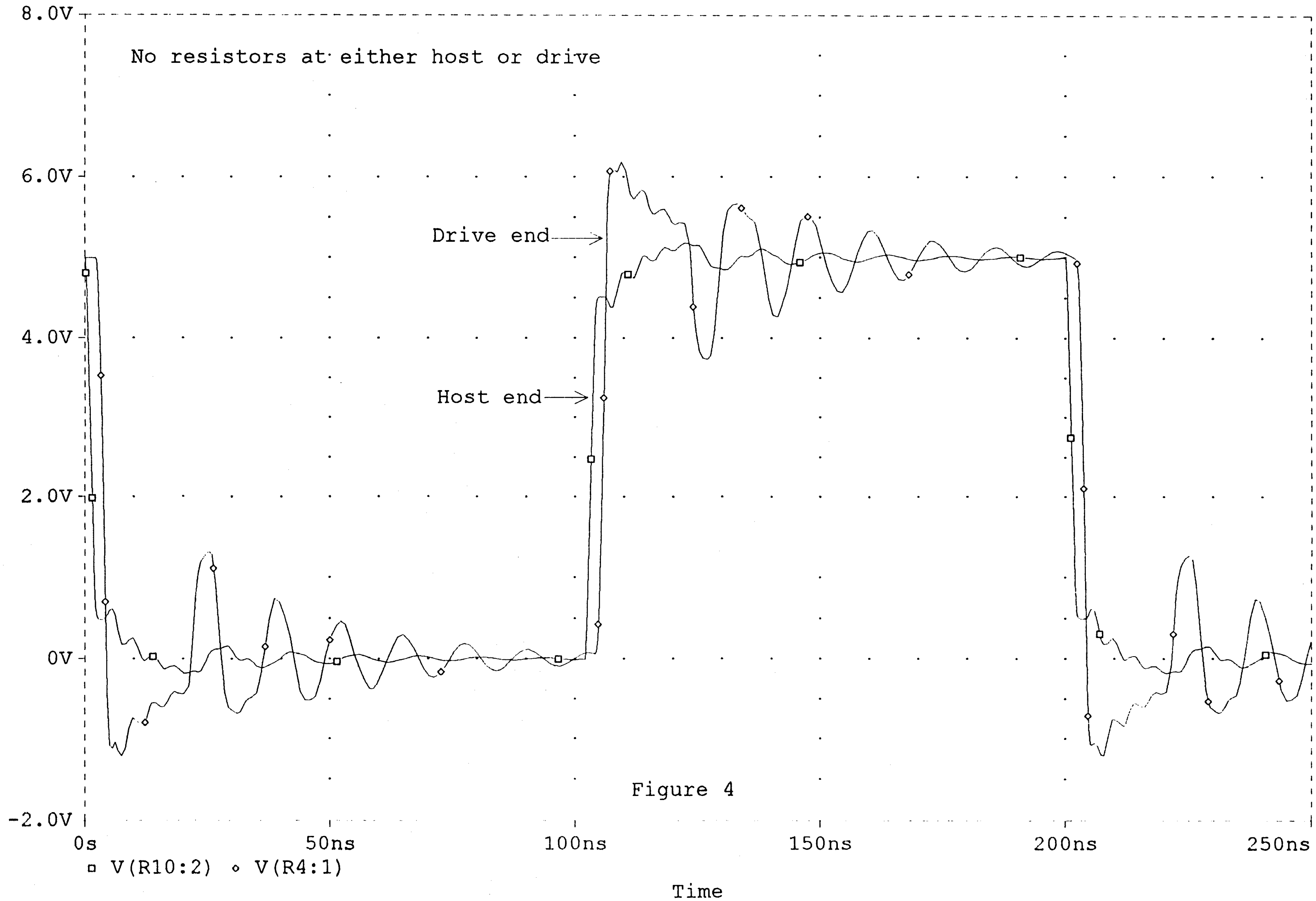


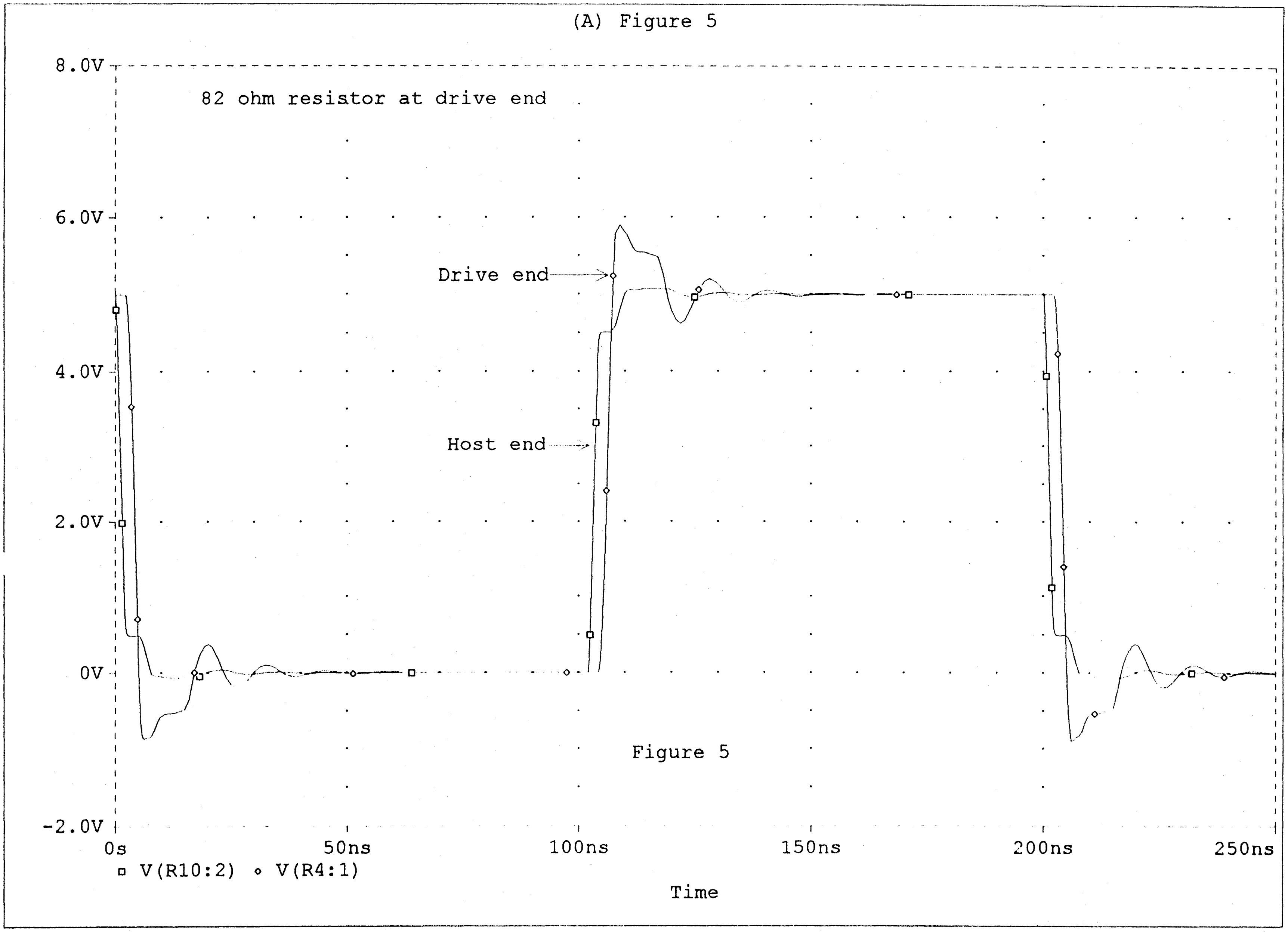
Figure 4

Date: July 18, 1994 Page 1 Time: 16:44:22

* C:\STEVE\PSPICE\ATA1.SCH

Date/Time run: 07/18/94 16:46:19

Temperature: 27.0



Date: July 18, 1994

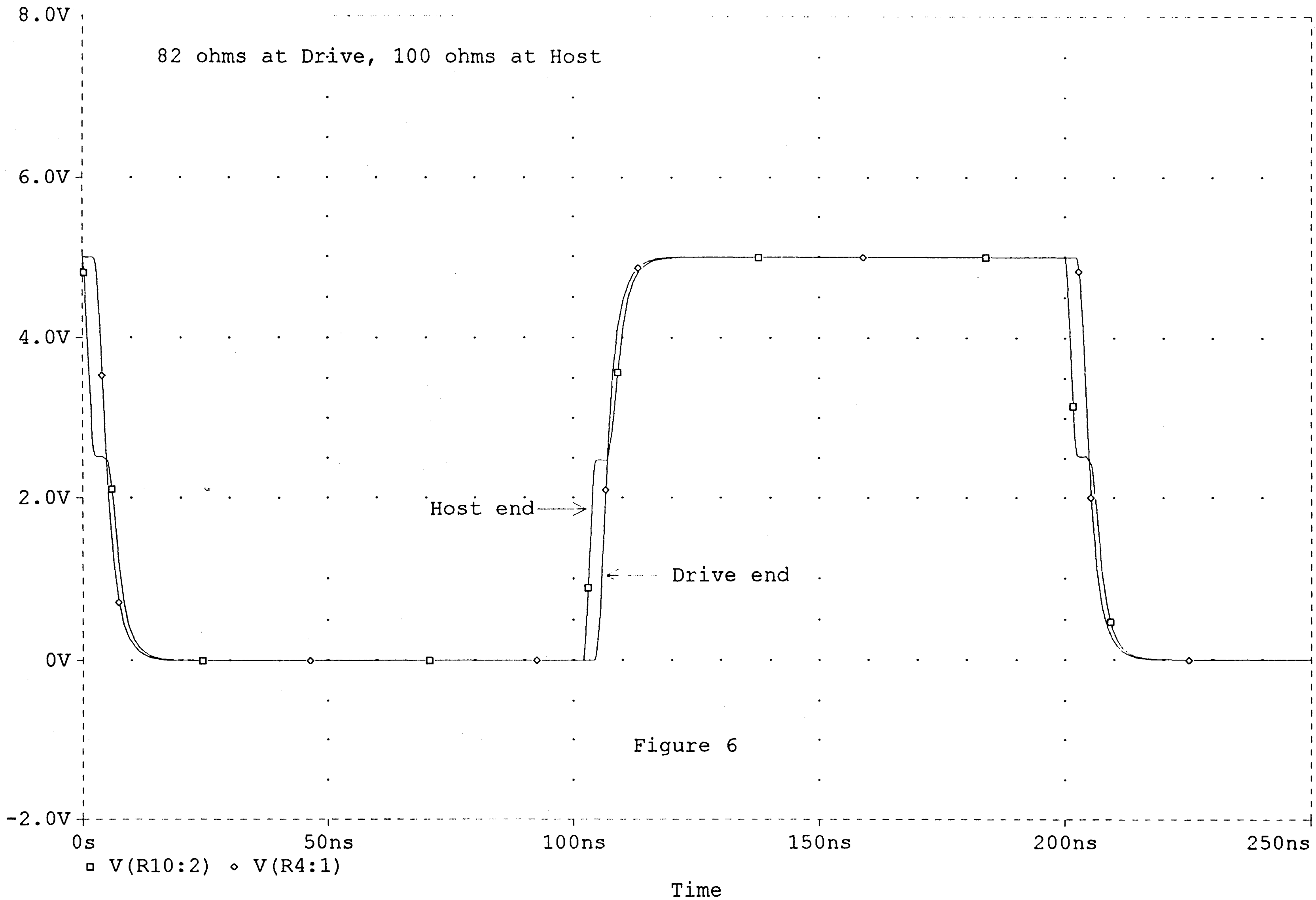
Time: 16:49:04

* C:\STEVE\PSPICE\ATA1.SCH

Date/Time run: 07/18/94 16:50:58

Temperature: 27.0

(A) Figure 6



Date: July 18, 1994

Time: 16:55:19

From a theoretical point of view the proper solution is to terminate the transmission line. Either a series termination at the source or a parallel termination at the drive would be acceptable. Unfortunately, we as a drive manufacturer do not have control of the source (host) end. Any solution we devise must accommodate various hosts which include source resistor values of zero to 110 ohms (33 ohms is a common value). To complicate matters, most manufactures have terminated only a few selected lines. A parallel resistor arrangement at the drive end (110 ohms) would be optimal for terminating the line but would cause excessive DC loading. Another solution would be to put a series RC network from the input to the drive to ground. If the C value is large enough, then the R would provide termination for AC signals, but would not cause loading at DC. This is an acceptable solution but is parts intensive.

The solution of terminating the transmission line is desirable because it is insensitive to cable length or various master/slave configurations. SPICE modeling shows that the system still operates properly even with competitor's drives in the master or slave location. The solution is also compatible with various host-end resistor values from zero to 110 ohms that may be used by motherboard manufacturers.

The preferred solution is to place a series resistor at the input to the drive. This solution depends on the fact that there is stray capacitance at the input of the drive. This stray capacitance tends to make the resistor look like it is terminated to ground at high frequencies. The value of this resistor should be near the impedance of the ribbon cable (approximately 110 ohms). From simulations it appears that any value from 70 to 100 ohms works satisfactorily. We have chosen 82 ohms as the recommended solution (figure 5).

Solutions that involve purely reactive elements (capacitors and inductors) are not recommended. Since the ringing is the result of a resonant system, adding reactive elements simply changes the frequency of oscillation. Although this may fix a given system problem, it has really just moved the interfering peaks to a different location, solving the problem for only that particular system. Proper solutions should include resistive elements to dissipate the energy stored in the transmission line.

Complicating Factors

There are several complicating factors that must be investigated before implementing any proposed solution. The first is that the solution must be considered with the various source terminations that might be implemented by the motherboard manufacturer. Although most systems currently do not have any termination resistors at the source, there are some systems using 33 ohms. Appian has recommended using 100 ohms with their ADI/2 PCI to ATA bridge chip, so we can expect to find future system using this value. Figure 6 shows the result of simulation with a 100 ohm host-end resistor and the proposed 82 ohm drive-end resistor.

Simulations have been performed with both one drive and two drives, mixed drives (e.g. one Quantum, one Conner), various source impedances, and with the host at the end of the cable or the host in the middle of the cable. Additional simulations have been done with different values of stray capacitance at the drive, with 1 ns, 2 ns, and 5 ns edge speeds, and with/without clamp diodes at the drive end. The waveforms were significantly improved in all cases. The 82 ohm series resistor solution was less effective if the total input capacitance of the drive was less than 8 pF. Slowing the transition time of the edges appeared to have minimal effect. Even edge speeds of 20 ns still had unacceptable levels of ringing. The clamp diodes are of interest because the VCC diodes have been removed in the next generation interface chip (LEO). When the bus is properly terminated, as in the recommended solution, the clamp diodes are not necessary.

DC Analysis

The current ATA specification proposal (April 27th, 1994) does not specify a DC input current requirement for the drive (or host). Most drives implement the ATA interface with CMOS process LSI chips, so the DC input current is negligible (modeled with 500K ohm resistors in the simulation). The worst case DC voltage drop occurs when the drive is driving the line back to the host (e.g. data lines during a read cycle). A likely worst-case load would be if someone placed a single F-series TTL buffer at the host. F-series logic specifies a high level input current of 20 microamps and a low level input current of 0.6 mA (maximum values). With an 82 ohm series resistor at the drive the voltage drop would be 0.05 volts. A similar resistor at the host would result in a total additional DC voltage drop of less than 0.1 volts. This is considered acceptable.

Other Signals

There are six other signal structures. The data lines DD0 through DD15 are basically similar to the control lines except that they are bi-directional. The issue of data line termination will be addressed in Section 2 of this document.

The DMA handshaking signals, DMACK- and DMARQ, have not demonstrated any problems at this time. These are time critical signals for DMA mode, and when DMA mode becomes widely used it is likely that we will see problems with this line. An 82 ohm termination resistor is recommended.

The signals DASP- and PDIAG- are not time critical signals, and do not need termination.

The INTRQ line does show ringing, but the way that interrupts are used in a system indicate that this is not a problem. No fix required.

The signals IOCS16- and IORDY are open collector signals with a 1K ohm pullup at the host. It is important to note that although the ATA specification calls out 1K for the pullup, many system integrators have been using 330 ohms to increase system speed. SPICE simulations show no ringing problems with these lines. No fix required.

The SPSYNC:CSEL line is either a vendor specific line (SPSYNC) or a DC status line (CSEL). In either case no termination resistor is called for.

Adverse Effects

Terminating the bus lines properly will decrease the edge speeds and therefore increase the delay of the signals. Simulations show that with an 82 ohm resistor at the drive and a 110 ohm resistor at the host, the incremental delay is 1.8 ns. This simulation was with a stray capacitance at the drive of 25 pF (max. allowed by ATA-2 proposed spec.).

A brief examination of the ATA-2 specification and the drive timing indicate that a delay of this magnitude is not a problem. To be completely sure that there is no timing violation, an engineer from the NEKO or KONI design teams would have to evaluate the timing change against the internal chip design limits. Until such time as an engineer becomes available, we will assume that delays of 2 ns or less are acceptable.

Section 2 – Problems with IORDY Signal

Disk Drive Failure Modes

The problems seen with the PCTECH bridge chips are a bit more complex to describe. The problem only appears when a read cycle occurs and the drive finds it necessary to delay the host read using the IORDY line. Eventually a byte becomes available, and the drive asserts IORDY, telling the host that data is present on the data lines. The current Quantum drives (using KONI or NEKO) have a zero nanosecond specified setup time from data to assertion of IORDY. This means that the data is placed on the bus at exactly the same time as IORDY is asserted (figure 7). The current versions of KONI and NEKO violate their data setup times and occasionally deliver data as late as 15 ns after IORDY goes high.

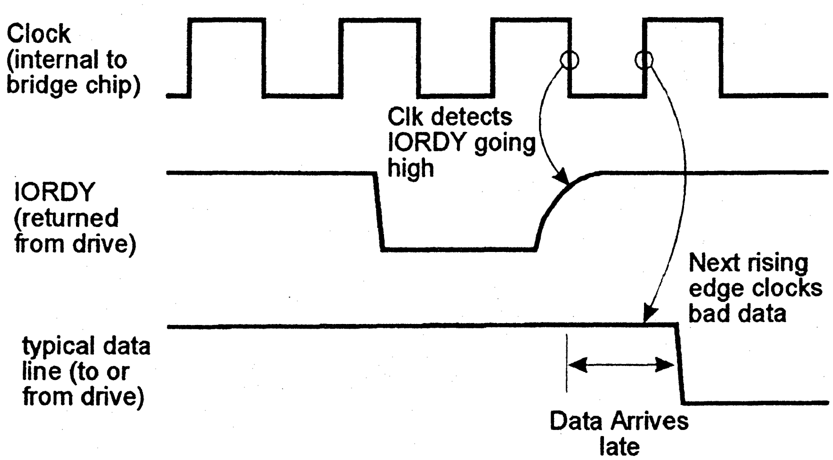


Figure 7 – Late data problem with PCTECH bridge chip

The PCTECH chip samples IORDY on both rising and falling edges of its internal clock. The clock speed is usually 33 MHz. Data is always read into the chip on the rising edge of the clock. If the assertion of IORDY is timed such that the chip detects it with a falling edge, then the next rising edge – only 15 ns later – will capture the data. Since the KONI and NEKO chips can deliver data as much as 15 ns late, then the data is not ready when the PCTECH chip latches in the data. Ringing on the data lines, as described in Section 1 above, aggravates the problem. The data lines are often not stable until 30 ns or more after IORDY is asserted. Other chips from different manufacturers only sample IORDY on rising edges of the clock so there is at least 30 ns before the next rising edge that is used to capture the data. This problem scenario has been confirmed with engineers at PCTECH.

Technical Discussion

The solution is to delay the assertion of IORDY. If the assertion of the IORDY signal is delayed relative to the data then the KONI/NEKO chip will have enough time to drive the data lines and the data lines will have more time to settle. This delay can be accomplished several different ways, but it is important not to significantly increase the deassertion of IORDY (falling edge). The down side of this solution is that it will increase the cycle time for every access that requires IORDY. This may decrease system performance by a small amount. The solution currently being implemented is a two-transistor circuit as shown in figure 8. Although the late setup of data has shown to be a problem only with the PCTECH chip, it likely that most other systems and chips are very close to their timing margin and may demonstrate occasional failures in the field.

The timing margin on IORDY can be improved with both a delay circuit and adding series resistors to eliminate ringing on the data lines. The series resistors on the data lines will also help during fast DMA transfers, where the data setup time is critical but the IORDY signal is not used. Although we are not currently seeing any use of fast DMA, Compaq and others have stated that they intend to use it in the future.

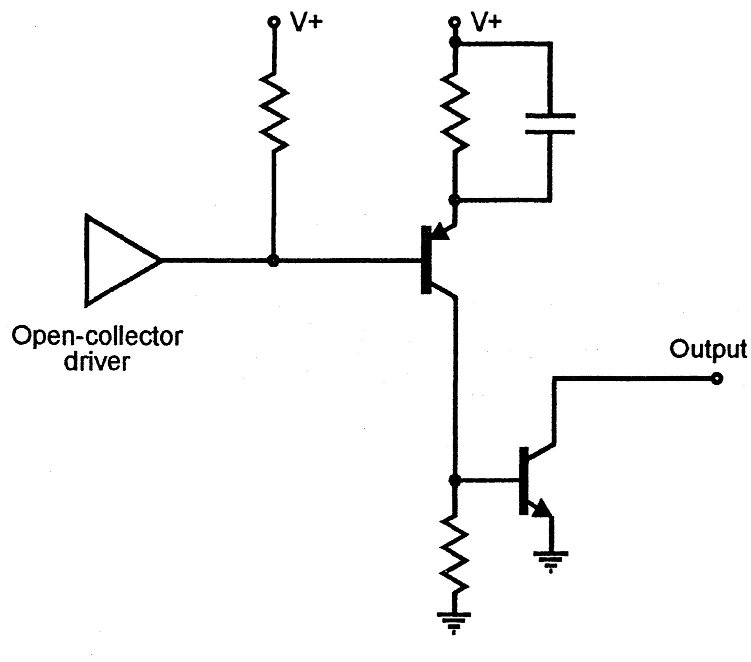


Figure 8 – IORDY delay circuit for PCTECH bridge chip

There is one more non-problem that is worth mentioning for the sake of completeness. There are certain conditions on the drive which will cause severe ringing on the IORDY line. If the host performs a read cycle to the drive, and the

FIFO on the drive is empty, then the drive will deassert IORDY in an effort to hold-off the host. Data coming off the heads is an asynchronous process when compared to host accesses, so it possible that a byte can be put in the FIFO just a couple nanoseconds after IORDY was deasserted (figure 9). When this happens, the drive sees that data is now available, so it reasserts IORDY. Since the IORDY signal from the drive is open collector, and the pulse width is less than the round-trip delay of the ATA cable, large amplitude ringing is set up on the cable. Although the signal looks terrible, it is not a problem in reality. IORDY is not sampled by the host until much later in the read cycle, and by that time the ringing has subsided. Therefore the distorted waveform is never seen by the host. As a point of reference, this issue has been addressed in the next generation data path controller (LEO) and will not occur with the new design.

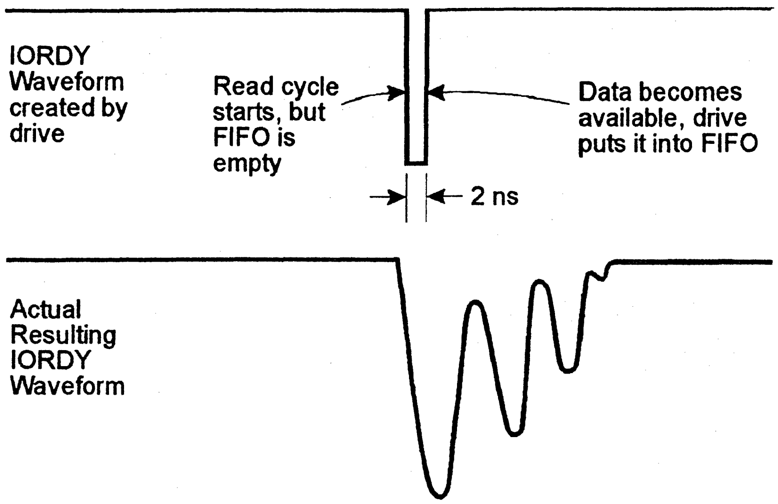


Figure 9 – Pulse sliver on IORDY causing ringing

Adverse Effects

The adverse effects of adding resistors to the data lines is the decrease in timing margin during read and write cycles. The timing delay is the same as discussed in Section 1 above: 1.8 ns. During writes this is no problem since the drives have adequate data setup time. During reads it is possible, but not likely, that we fail to meet the timing requirements, but a detailed timing simulation of KONI/NEKO would be required to know for certain.

The adverse effect of adding delay to the IORDY signal is a slight decrease in performance. This decrease is less than might be expected. First of all, IORDY cycles only occur on approximately 5% of all cycles. The execution time of these cycles will increase. But as the execution time increases, then more bytes get loaded into the FIFO by the data path controller. This means that the drive will be able to stream data for longer before having to use an IORDY cycle again. This tends to offset much of the performance loss. Quick estimates put the total performance decrease at around 1 or 2 percent.

Section 3 – Narrow Pulses on Interrupt Line

Disk Drive Failure Modes

Quantum drives have been shown to fail when run on motherboards using Appian bridge chips. Failures occur during cold boot. The usual symptom is a system hang. The problem has been traced to the presence of narrow pulses on the interrupt line (less than 5 microseconds). This was traced to a problem with the BIOS drivers supplied by Appian and has since been fixed.

Technical Discussion

The system hang problem occurs when the system is waiting for an interrupt that never comes. The drive generates an interrupt, but the pulse width is too narrow to be recognized by the host. The host continues to wait for an interrupt, and the drive waits to be serviced by the host. This results in a system deadlock.

The problem was found in a coding error in the Appian BIOS drivers. There are two registers defined in the ATA interface that return the status of the drive: the Status Register and the Alternate Status Register. The contents of the two registers is identical, but reading the Status Register will clear any pending interrupts. If an application wishes to poll the status of the drive it should read the Alternate Status Register to avoid inadvertently clearing a pending interrupt.

The Appian drivers were waiting for an interrupt and polling the Status Register while they were waiting. The drive and the host operate asynchronously. Eventually there would come a time when the drive set the interrupt line true at about the same time that the software would read the Status Register. This causes the interrupt line to be set and almost immediately cleared. The resulting narrow pulse is not seen by the processor.

Although this problem was fixed by Appian with a new release of their driver, it is interesting to note the Conner drives did not show any problems. It seems that Conner has done something to guarantee a minimum pulse width of approximately 15 microseconds on the interrupt line. This of course confused the debugging process – "if a Conner drive works and a Quantum doesn't, the problem certainly can't be in the drivers." It may be worth looking into what Conner has done to understand why they enforce a minimum pulse width.

Section 4 – Deassertion of IOCS16-

Disk Drive Failure Modes

Quantum drives have been observed to deassert the IOCS16- line after DIOR- or DIOW- are deasserted. The ATA specification states that the IOCS16- line should be decoded based on the address lines and chip select, and not on DIOR- and DIOW-. The address lines remain asserted well after the end (deassertion) of DIOR- and DIOW-, so IOCS16- is being deasserted by the drive too soon. This problem has been found by HP Grenoble and Siemens-Nixdorf

Inc. (SNI). Neither customer has requested that Quantum take any action. No system failures due to this problem have been observed in the field.

Technical Discussion

Although this problem has been found by two customers it probably won't adversely effect any known system. When an address is first driven from the host, the drive has to decide whether the accessed register is 16-bits or 8-bits wide. The IOCS16- line is returned by the drive to tell the host that this cycle will be 16-bits wide. Obviously, this eight or sixteen bit decision must be made at the beginning of the read or write cycle. It makes no sense for the drive (or the host) to change its mind halfway through. Likewise, when the read or write has been completed (DIOR- and DIOW- deasserted) the state of the IOCS16- signal should be irrelevant.

For these reasons it is unlikely that any local bus bridge chips will sample IOCS16- after DIOR- and DIOW- have been asserted. Therefore we should not expect to see any problems in the field. But the ATA specification does state that the IOCS16- line should be held longer than we are driving it, so future designs should correct this problem.

Quantum

Date: 27-Jul-94

From: Steve Reames (x6443), Systems Engineering

To: All ASIC designers

Subject: Local Bus issues and future ASICs

Systems Engineering has been following the issues with Quantum drives in PCs with local busses. The problems are all related to fast ATA. No problems are known with local busses and SCSI. The problems occur when systems attempt to run the ATA interface in mode 3 (transfer rates up to 11.1 MB/s). The problem is not actually with the local bus itself; the problem is with ringing on the ATA bus signals. The same difficulties appear in both PCI and VLB systems. The attached chart classifies the problems seen in the field and organizes them by root cause.

The attached document titled "Local Bus Compatibility Issues" covers the known local bus problems in detail. Solutions to the problems identified have already been incorporated into LEO-AT with the exception of the bus termination resistors (which must be added to the PCB).

Conclusions

1) No changes are recommended for the LEO-AT chip.
2) PCBs in future products should include the recommended termination resistors.

Some Ideas to Consider

The ATA bus ringing problem currently requires a number of external components, but could be aided by clever ASIC design in the future. Anything that can be done to reduce bus ringing or increase ringing immunity would improve the reliability of Quantum drives. Controlled rise and fall times can help reduce the ringing seen at the host on signals driven by the disk drive. Active diode clamps, perhaps biased to 0.6 V from supply and ground, could clamp the ringing on host signals seen by the drive. These and other techniques should be investigated for future designs.

Steve

Steve Reames

Apparent Problem in Field	QuIT Record Number	Customer Name	Adapter Chip Vendor	Root Cause	Maverick & Lightning solution	Roadrunner & Thunderbolt solution	LEO-AT Current Solution	LEO-AT Recommended Solution
Data corruption	1422	Research Machines	Adaptec	Ringing	Series resistors on PCB	Series resistors on termination board	Increase rise/fall time to 5ns min.	same
System hang	1430	Austin Computer	unknown	Ringing				
Boot failure	1627	IPC	Appian	Ringing				
Rejects super-mode cmds	6327	Quantum AE	DTC	Ringing				
Boot failure	6735	Acer	Appian	Ringing				
Boot failure	6801	AT&T/NCR	Appian	Ringing				
Data corruption	5705	AST	PC Tech	IORDY setup time	Delay circuit on PCB	Delay circuit on termination board	Insure positive data/IORDY setup time	same
Data miscompare	6191	Intel	PC Tech	IORDY setup time				
none	none	HP, Siemens-Nixdorf		IOCS16 hold time	No action taken	No action taken	Derive IOCS16 from address lines as opposed to DIOR and DIOW	same
date: 26-Jul-94								