Motherboard Design Rules for the IBM 6x86 P200+ Microprocessor *...and Beyond!*

Application Note

Author: Duane K. Becker

Revision Summary: This is the initial release version of this application note.

Introduction

This application note is intended to help motherboard designers properly design a motherboard to accomodate the IBM 6x86 P200+ Microprocessor.

As x86 motherboard bus speeds get faster, more attention must be paid to board characteristics that adversely affect high frequency operation. Characteristics such as trace to trace capacitance, transmission line effects, power supply decoupling, thermal constraints, and clock distribution are all more critical at higher frequencies. These characteristics and the methods of reducing their effects are discussed here.

Trace to Trace Capacitance

The impedance of a capacitor gets smaller at higher frequencies (neglecting effective series inductance and resistance). Specifically, the impedance from one board trace to another gets smaller as board frequencies go up. For example, at 20 Mhz, a pair of traces with 5 picoFarads capacitance between them, has an interconnective impedance of 1/jwc = 1/(j * 2 * pi * freq * cap) = 1591 ohms impedance magnitude connecting the AC between the two lines. If the characteristic impedance of the lines is around 150 ohms or so, this is a fairly negligible amount of cross coupling, and the effects will be swamped out by the transmission line impedance and the source impedance driving it. However, at 75Mhz, the intertrace impedance drops to 1/(j * pi * 2 * 75M * 5pf) = 424 ohms magnitude. At these frequencies and higher, the trace to trace impedance is getting closer to the characteristic impedance of the source and transmission line and will cause intertrace crosstalk to become more pronounced. At 100Mhz, the impedance is 318 ohms, only twice the characteristic impedance of the line. In a high speed design, this kind of cross connection should be accounted for, minimized, and validated in signal integrity analysis. This amount of intertrace capacitance will manifest itself as what appears to be noise on the line, as neighboring lines switch at different times. This can aggravate setup/hold times and cause unstable operation.

One way to reduce the problem is to limit the trace to trace capacitance. It is tempting, from a layout point of view, to run busses next to each other from a source to a destination, but this is the worst thing you can do electrically since it closely couples one line to its neighbor. The two keys to reducing inter trace capacitance is to keep the traces as short as possible, and to never run two traces next to each other for an extended distance. If the two lines must run a fairly long distance, choose two different paths for the traces to run, or if layout prevents that, then periodically shuffle the bus using vias and jump-overs so that no two lines follow each other for more that a centimeter or two at a time. This will have the effect of reducing the capacitance from any one line to any other. Although there will be many more cross couplings to a given line due to line shuffling, no one given line will have the bulk of all the capacitance, and many small switching contributions will have minimal effect on the given line. The key to reducing trace to trace capacitance is short runs, and bus shuffling along the route chosen.

Other methods of reducing the intertrace capacitance are choosing a board material with a lower dielectric coefficient, increasing the spacing between conductors, and increasing the interplane board material thickness.

Transmission Line Effects

The basics of transmission line effects, and what you can do to reduce them are discussed here.

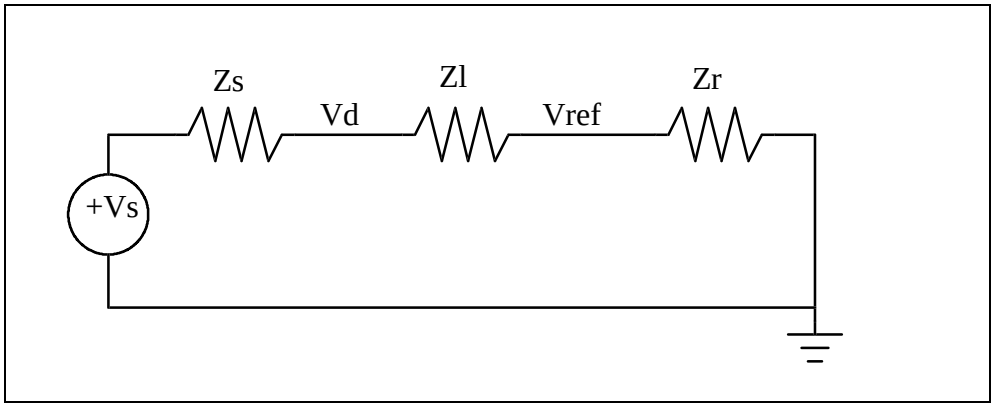

Figure 1: Electrical Model of a Transmission Line

When a wavefront of source voltage Vs is launched into a wire, it takes on the voltage:

$$Vs * \frac{Zs}{Zs + Zl} = Vd \qquad \text{(Equation 1)}$$

where Zs is the source impedance and Zl is the characteristic impedance of the transmission line, and Vd is the drive voltage at the pin of the driver (that is after Zs).

.

This smaller wave then propagates down the line to the end of the transmission line where it encounters Zr, the receiver impedance. However, when a wavefront hits an impedance, it divides down in amplitude, then reflects, doubling the voltage difference after being divided by its impedance mismatch. The reflection amplitude is given by:

$$Vref = Vi + (Vw - Vi) * 2 * \frac{Zt}{Zt + Zo} \qquad \text{(Equation 2)}$$

where Vi is the initial voltage at the reflection point, Vw is the incident wavefront voltage, Zt is the impedance the wavefront hits, and Zo is the characteristic impedance of the line the wave was following. For example, if a 1V wave on a line impedance of 100 ohms hits a 200 ohm impedance, and the 200 ohm impedance already has ½ volt on it, the apparent voltage at the 200 ohm impedance, and the amplitude of the reflected wave will be:

1/2V + (1V - 1/2V) * 2 * 200/(200 + 100) =1.166 volts.

This reflected wave then propagates back and the procedure repeats on the transmission side, reflecting back and forth until the steady state line voltage is reached. Depending upon how mismatched the source impedance, line impedance, and receiver impedance are, and the propagation time on the line, the reflections can take a long time to settle out.
When the source voltage has a rise time associated with it, the reflections that bounce back and forth are not distinctly identifiable, since the source wave may still be rising during the reflections. This will cause the reflected waves to take on a more rounded shape, which gives rise to the phenomenon commonly known as ringing, overshoot, undershoot, and ringback. These characteristics should not be considered as noise, since they are an inherent part of the signal propagation, however, the ringing can cause erroneous operation in a motherboard. Additionally, the ringing can couple to other circuits through capacitive paths between traces. The key to preventing these kinds of problems is to provide proper termination for the transmission lines. A properly terminated transmission line will not reflect any signal, and no overshoot, undershoot, or ringback will occur.

Other line characteristics that affect clean signal transmission are turning radiuses of wires, and extraneous vias in a line. 90 degree turns and every via in a wire create an impedance mismatch or discontinuity in a transmission line. Each such discontinuity provides a certain amount of reflection in the signal. To minimize this effect, use as few vias as possible, and use 45 degree turns instead of 90 degree turns.

Transmission Lines

There are a number of ways to reduce reflections on a transmission line. Diode clamping, series dampening resistors, balanced and unbalanced termination resistors, and resistive-capacitive termination are among the methods commonly used. Ideally, if the driver source impedance, the transmission line impedance, and receiver impedance were all the same, the signal would propagate to the end cleanly with no reflections. This is not often the case, and a proper termination technique should be employed. Note that on a PCI bus, reflections are a required part of the bus specification, so the following techniques should only be exercised on signals where reflections and ringback need to be kept to a minimum.

Diode Clamping is a commonly used technique where a pair of normally reverse biased diodes are wired from the end of a transmission line to the Vcc and Ground planes. When a wave hits the end of the line, it starts to reflect off the high impedance and tries to double in voltage, however, the fast diode starts to conduct and holds the overshoot or undershoot to 1 diode voltage drop above or below the associated power rail. Thought of another way, the impedance of the diode, being variable according to its forward bias, will fairly closely match itself to the transmission line impedance, thereby limiting the reflection to a diode drop. This method does not prevent the reflection, but limits the overshoot on the first reflection. Successive reflections back will not be suppressed since they won't fall above the diode's bias requirement. The Diode Clamping Technique is shown in Figure 2.

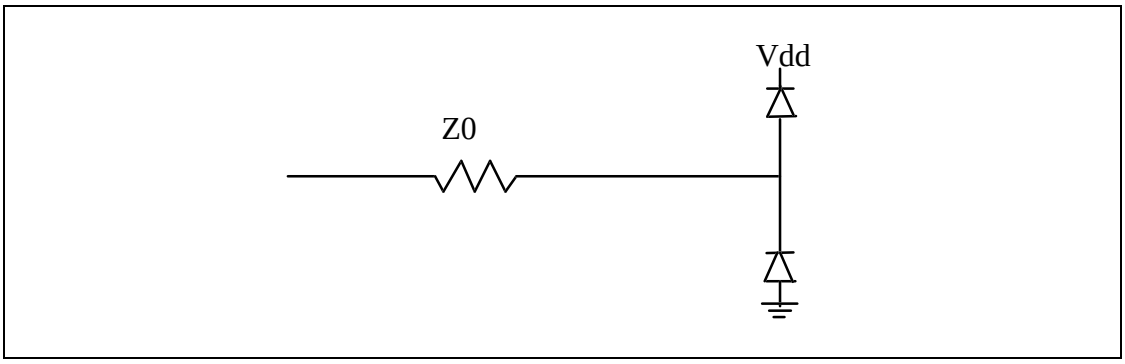

Figure 2: Diode Clamp Termination

Series resistive dampening is another technique used to limit ringback and reflections. By putting a small valued resistor (10 to 30 ohms) in series with the drivers, the rise time of the signal at the transmission line is increased due to the time required to to charge the line capacitance. This allows the reflection to be reduced in amplitude. This method also allows the source impedance of the driver to be more closely matched with the characteristic impedance of the transmission line. This does not address the reflection at the high impedance end of the line, though.

Additionally, series resistors can limit the drive levels on a heavily loaded net to the point where a logical high level may not have the noise immunity headroom for the circuit being driven. This method is shown in Figure 3.

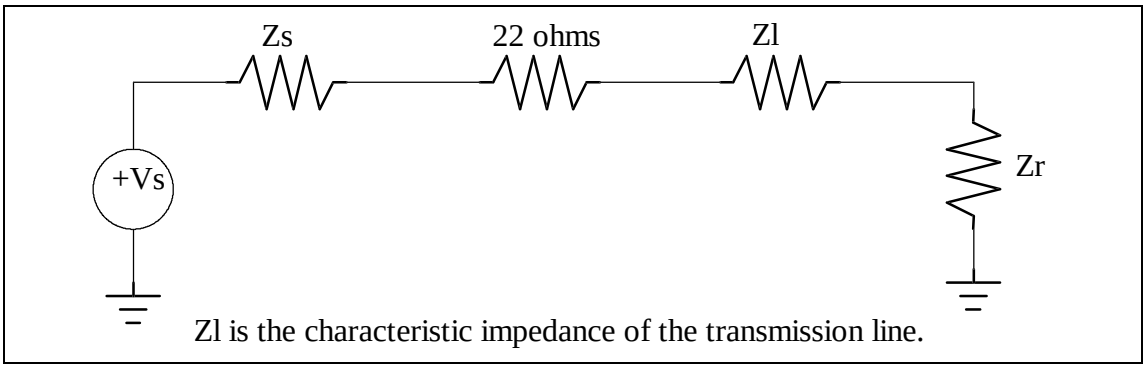

Figure 3: Series Resistor Termination

Another method of terminating a transmission line is by the use of unbalanced termination resistors. This technique involves placing resistors at the end of the transmission line to Vcc or more preferably the ground bus. By selecting a resistor that comes close to the characteristic impedance of the line, the reflected wavefront voltage will equal out with the incident wavefront voltage per equation 2 and no apparent reflection will occur. The problem with this method is that a termination resistor of 100 ohms or so will heavily load down the line driver, possibly holding the steady state level of the signal below the required high level voltage. It also has the effect of forcing a normally tri-stated bus into a logical one or zero state, and increasing the current requirements during steady state signaling. The unbalanced resistor termination technique

is shown in Figure 4. Zt should be chosen so that Zr in parallel with Zt equals Zl, the transmission line impedance.

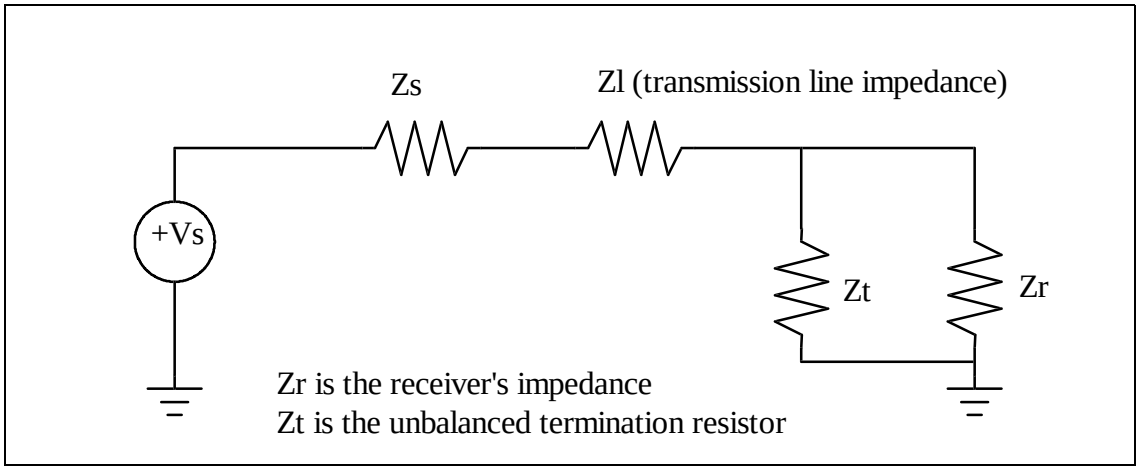

Figure 4: Unbalanced Resistor Termination

Balanced termination resistors

This technique is like the single resistive technique, except that by using two resistors each of twice the value of the line impedance, and wiring the bus to both Vcc and to ground in a "T" network, the drive low current and the drive high current required to pull the bus is reduced by a factor of 2. It also allows the bus to "float" to the midpoint between the voltage rails. It does share the disadvantage of making the drivers work harder in steady state, but to a lesser extent.

RC series termination

This technique involves attaching an RC series network to the end of the transmission line. The RC series network is designed to provide an impedance that is matched to the characteristic impedance of the transmission line, but at a frequency determined by the rise and fall times of the incident wave. This will suppress any reflections while not wasting the DC drive capability of the source drivers. This method has the disadvantage that the line being terminated should be well categorized as to it's characteristic impedance. The RC method is shown in Figure 5. The value of C should be (3*Tr)/Zo where Tr is the rise time of the incident signal, and Zo is the characteristic impedance of the transmission line. This is derived from the following:

At the primary frequency of 1/(2*Tr), you want a very low impedance path provided by C, and the rest of the characteristic termination impedance to be taken up by R. This is to allow the AC component to be absorbed by the termination, while allowing the DC portion of the wave to remain unloaded. If Zo is the transmission line impedance, and a factor of 10 lower AC impedance is chosen for the termination capacitor, then

|Zc| = 1 / (2*pi*f*C) = Zo/10 (equation 3)

Substituting 1/2*Tr for f, and solving for C we get:

10 Tr / (pi * Zo) = C or approximately **3*Tr/Zo = C.**

Then to choose an appropriate Series Resistor, we need to compute the total impedance of the RC network, and solve for the resistance value.

|Zo| = Sqrt(Zr*Zr + |Zc|*|Zc|) (magnitude = root of sum of squares) **(equation 4)**

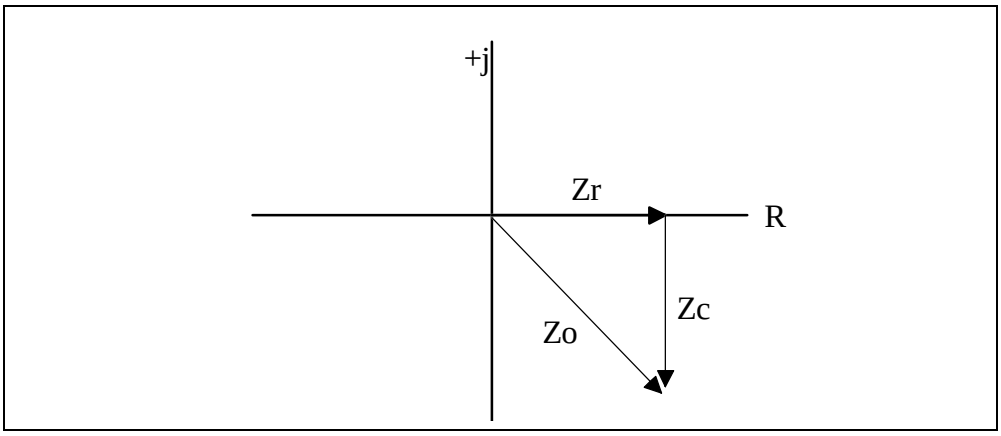

Figure 5: Phasor Diagram for Series RC Termination

For a 100 ohm characteristic line impedance and 4 ns rise time, C = 120 pF, and Zc is found by:

Zc = 1/jwC = -j/wC. At 1/ (2*4ns) = 125Mhz, Zc= -j/(2*pi*125M*120pF) = -10.6j

Solving equation 4 for R, we get

R=sqrt(Zo*Zo - |Zc|*|Zc|) = sqrt(100*100 - 10.6*10.6) = 99 ohms.

So the complete RC termination circuit for a characteristic line impedance of 100 ohms with a signal rise time of 4 ns is shown in Figure 6.

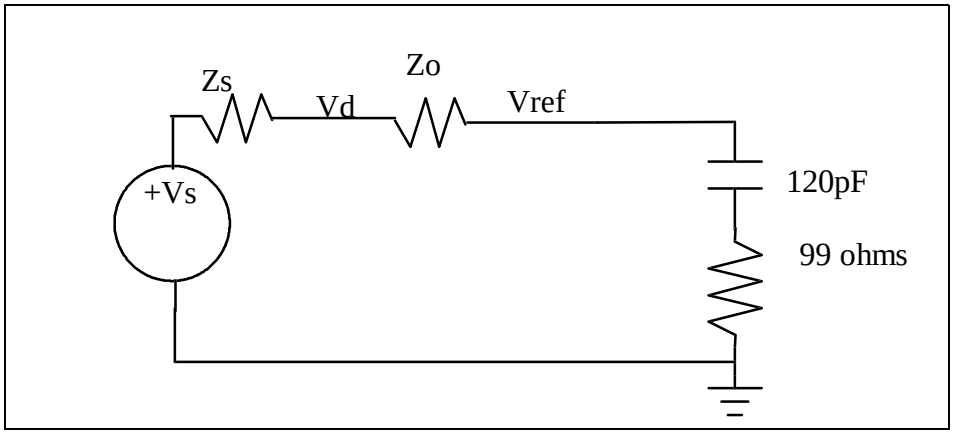

Figure 6: Series RC Termination

Of course, if the receiving element has a termination impedance that is within a factor of ten of the characteristic line impedance, it's impedance should be considered in parallel with the termination network and the network should be recalculated appropriately.

This termination network will generally work well over a 60% to 150% range of the desired characteristic impedance. In other words, if the network is designed for 100 ohms, it will work well from 60 ohms to 150 ohms line impedance.

Power Supply Decoupling

At higher bus frequencies, internal circuits and bus lines switch more often. Additionally, faster parts of a given technology often have faster signal rise and fall times. The more frequent and faster transitions dump more electrical energy through board capacitance and internal CMOS gate capacitance. Since all real life capacitors are lossy, this creates the need for a fast response from the power supply to keep the voltage supplying the circuits at a steady state.

At 75 Mhz and beyond, almost all of the current being used by a CMOS circuit is due to switching the internal capacitances on and off very quickly, and the value of the current becomes more proportional to the operating frequency. To supply a steady power supply voltage to the quickly switching circuits, the power supply must provide very low impedance to the high speed frequencies involved. This is accomplished using an adequate quantity and value of low ESR (effective series resistance) capacitors.

There are usually three values of decoupling capacitors used on motherboards, each addressing a different frequency spectrum required of the power supply. The large, bulk decoupling capacitors, which usually have a high ESR are placed on a board to compensate for voltage regulator recovery delays and long leads from the power supply. They provide the ability to detune and decouple the low frequency changes in current. One such example is the entry and exit from suspend mode where motherboard current can swing from near zero to 5 Amperes or more. These large capacitors are selected to hold the power bus reasonably stable during the surge until the regulator can stabilize at the new current draw. The next smaller valued capacitors, usually in the 100 nF range of value, is used to help decouple some of the higher frequency components that the self-inductance of the larger, bulk capacitors cannot deliver in microsecond times. These are usually chip capacitors with about 1 nH of self inductance. These type capacitors usually have a self resonant frequency given by:

1 / (2*pi*sqr(LC)) (equation 5) or
1 / (2*pi*sqr(1e-9 * 100e-9) = 16Mhz

*(Note: 100e-9 means 100 * 10**-9 or 100 nF)*

Above these frequencies, the 100 nF capacitors are not as effective at shunting out the higher frequencies. Additionally, the inductance of the copper lines or planes from the capacitors to the actual device pins adds a high frequency impedance to the capacitor and further reduces the capacitor's ability to decouple higher frequencies. At higher bus speeds, however, even these chip capacitors have too much self inductance in series with their capacitance to meet the demands of

high speed, high current switching circuitry, and a third, smaller value of capacitor, specifically selected for high frequency decoupling is required. These are usually 10 nF capacitors with very low ESR (about 500 picohenries) whose self resonance is given by
equation 5 or:

1 / (2*pi*sqr(500e-12 * 10e-9)) = 71Mhz

In the following example, the high frequency portion of the power grid of a motherboard is shown.

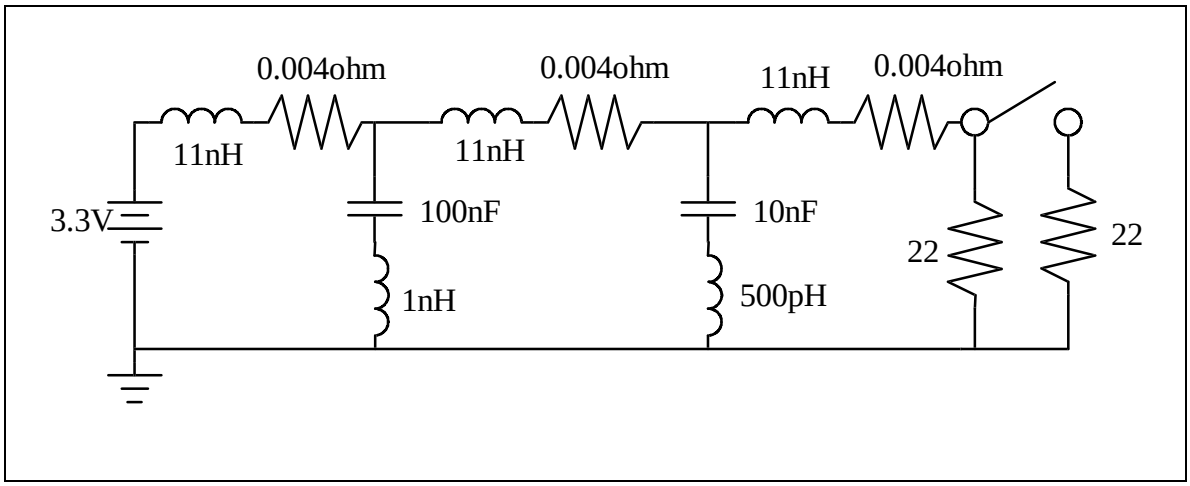

Figure 7: High Frequency Model of Power Grid

In this circuit example, the switch selects between a 150mA and a 300 mA load, emulating the nominal current switching of 3 power pin pairs of an IBM 6x86 processor during normal operating conditions. The 11nH inductances in series with the 4 milliohm resistors approximates the voltage and ground plane interconnects between decoupling capacitors and the IBM 6x86 processor power pins. The inductances in series with the capacitors approximate the decoupling capacitors, complete with their series ESR. The 3.3V battery emulates the voltage regulator and it's bulk filtering capacitors. This circuit, when simulated in SPICE, results in voltage spikes, as measured at the switch simulated pin, of 1.5 volts, and approximately 1 nanosecond in duration. Fifteen capacitor pairs provides sufficient decoupling of the high frequency noise induced by the switching logic in IBM 6x86 processor. In general, use one pair per 300mA of load current unless the wiring to the capacitors is excessively long.

See Applications Note 40215 for more detail on selecting the bulk, voltage regulator filter capacitors.

Thermal concerns

The higher in frequency a board runs, the more power it will need to dissipate. In order to design a board to run at higher frequencies, the intersignal capacitance must be reduced by making signal lines as short as possible. This brings up a thermal issue. As the lines get shorter the packing density of the high speed circuits increases, and it gets more difficult to provide adequate airflow over the components to keep them within operating temperature. There are many parameters that

affect the power dissipation capability of the motherboard in a system. Air flow, case volume, ambient temperature, plug in cards, and the number and thicknesses of signal and power planes are among the variables. Assuming natural convection cooling in an operating environment temperature of 35 C, an average junction temperature of devices not exceeding 100 C, a minimum of two signal layers of 1 oz/square foot thickness, 2 power planes of 3 oz/square foot thickness, the ideal maximum power density to put on a motherboard is 0.04 watts per square centimeter, well distributed on the board. If adequate airflow is included in a design, this dissipation recommendation can be made higher.

See Applications Notes 40209 and 40216 for more detail on thermal solutions and thermal management designing considerations

Clock Distribution

At high bus frequencies, the time allowed per cycle drops dramatically. Since the chipsets, CPU, and peripheral components all use the same clock, it is critical that the clock signal be correctly distributed in a way that minimizes chip to chip skew, clock reflections and noise, and that keeps the clock waveform shape as close to ideal as possible.

Since all wires have a finite and predictable speed of conduction, any mismatch in the clock distribution flight times will result in skew, or difference in arrival time, of the edges of the clock signal. This will have the result of reducing data setup or hold times. Alternately, if a design requires additional setup or hold time, purposely introducing delay in the appropriate clock lines can alleviate a timing problem. Ideally, a good design should start with equal clock flight times to every chip using the clock signal.

There are three distinct distribution methods for a clock line: random spawning, daisy chaining, and radial distribution. Each has their merits and deficiencies and are discussed below.

Random spawning distribution is essentially randomly picking off a piece of a clock signal wherever it is most convenient. This method is shown below in Figure 8.

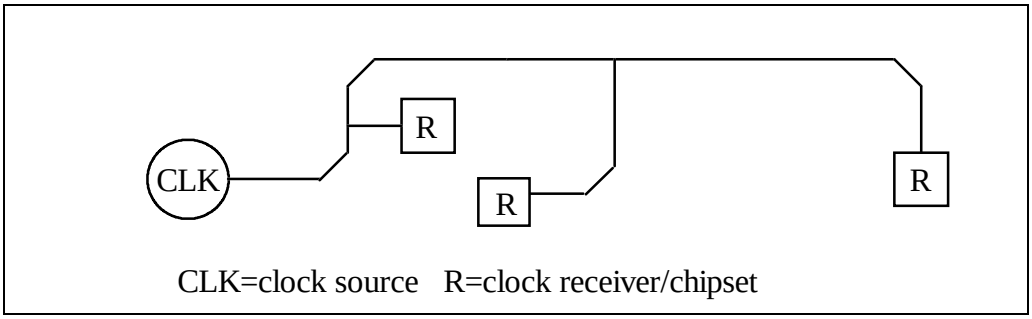

Figure 8: Random Spawning Clock Distribution

In this example, the clock is distributed from one point, and receiving components pick off the clock signal from the net as needed. This method allows the most convenience for routing a clock signal, however it results in so many impedance mismatches along the route, unterminated ends,

and mismatched flight times, that it will cause timing skew errors, reflections, and noise pickup from nearby lines. This distribution technique is only valid for circuits with high noise margins, and clock rise times in excess of the longest flight time. This distribution technique is deadly to high speed motherboard design and should be avoided.

Daisy Chaining, shown in Figure 9, is a cleaned up version of the random spawning method of wiring a clock signal.

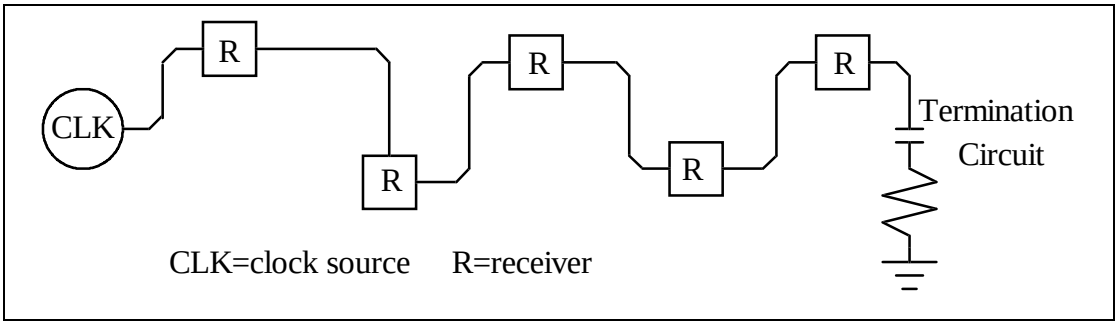

Figure 9: Daisy Chaining Clock Distribution

In this example, a single transmission line goes from source to receiver to receiver, without creating any unnecessary stubs along the way. When the clock launches a wavefront, it will propagate all the way to the termination circuit, which will prevent reflections and maintain the original clock waveform along the entire circuit. Often, a small valued resistor of 10 to 22 ohms is placed immediately after the clock source. This has several benefits. First, it helps to match the impedance of the clock source to the clock transmission line, thereby reducing signal launch ringing and reflections. Secondly, the series resistance, coupled with the distributed capacitance of the transmission line will slow down rise and fall times of the wave by filtering out the high frequency harmonics of the clock line. This will have the effect of “Rounding the edges” of the clock signal. This will reduce electromagnetic radiated noise from the motherboard, and also reduce the amount of noise the clock line introduces into nearby PCB conductors since higher frequency harmonics couple to other signal lines much easier than the primary frequency of the clock. Thirdly, if reflections are produced off the end of the transmission line due to an inexact match between the characteristic line impedance and the termination circuit impedance, the series resistor will help in dissipating the resulting ringback. Using the daisy chain method of clock distribution will help minimize impedance mismatches along the transmission line by eliminating stub-outs. If a stub out is required in a daisy chain distribution, it should be kept to no longer than 2 centimeters in length.

One flaw in the daisy chain distribution technique is that each receiver gets clocked at a slightly different time since the wave takes a finite time to propagate to the end of the transmission line. If each chip’s setup and hold times have sufficient slack in them relative to all the other drivers and receivers in the other chips using the clock, then the daisy chain solution will work. If cycle timings are so close that the total flight time approaches the slack between chipsets, then setup or hold time violations are imminent.

Radial Clock Distribution is shown in Figure 10. It addresses most of the deficiencies of the above distribution techniques.

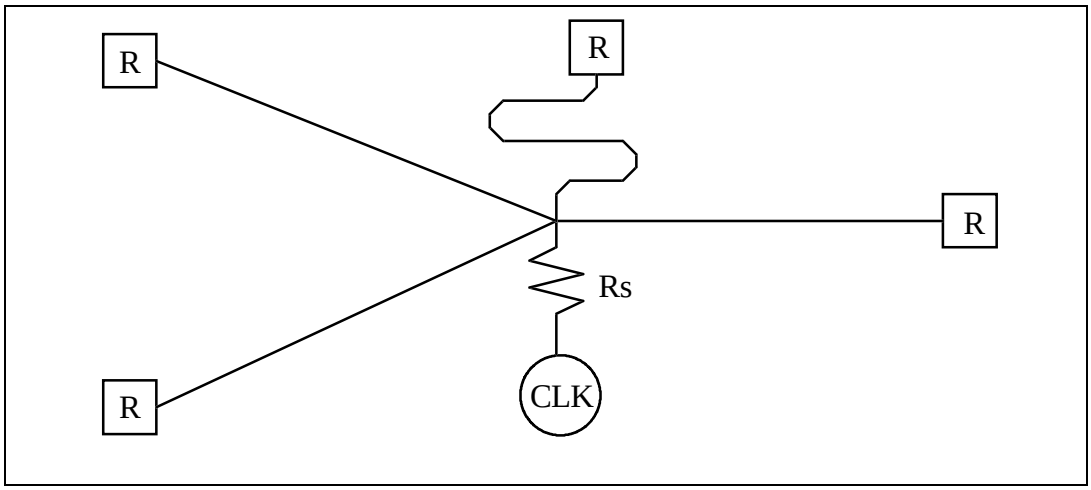

Figure 10: Radial Clock Distribution

In this distribution scheme, the clock generator is centralized, and distribution from one common point is done. Additionally, if any physical distance from the clock to the receivers are significantly shorter, a path must be chosen to that receiver to compensate for the shorter flight times. The goal in the radial distribution technique is to provide the same wave flight time to each receiving element in the circuit. That way, clock skew is minimized and setup and hold times are equalized, resulting in a more robust design. A small source resistor is also sometimes included in the layout for the same reasons given in the daisy chain method. Additionally, an RC termination circuit should be added to each receiving element reflections and ringback needs minimization. An RC termination is preferred since it does not induce any DC loading to the clock source. One disadvantage to this method is that the clock layout needs to be carefully considered and the layout will be more difficult.

Regardless of which distribution is used, it is important that reflections and ringback be minimized in any design. Reflections will appear as periodic noise on the clock line and will cause noisy coupling to nearby circuits, setup/hold time violations, data corruption, and radiated electromagnetic field emissions. The safest way to eliminate this is by properly terminating all clock transmission lines.

Another good design practice in clock distribution is to surround the clock line by two ground traces along its flight. See Figure 11 below for an example.

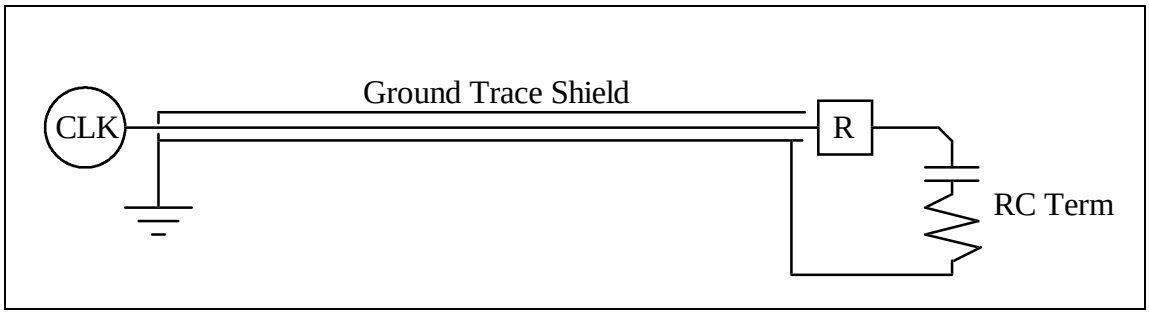

Figure 11: Clock Shielding

In this example, a pair of ground lines are run adjacent to the clock line, and the termination circuit completes the transmission line. It is important to ground the shield at the source and not at the receiving end of the line. If this looks like a coaxial solution, you're on the right track. At high frequencies, a shield around the clock line will help prevent transmission of clock harmonics to surrounding circuitry, as well as prevent surrounding circuitry from coupling their signals into the clock line. Ideally, the clock and shield should be located next to the ground plane, and there should be no signals on adjacent planes in the vicinity of the clock line.

There are a number of different ways of distributing the clock line in a board design. The most important thing to remember are to reduce clock skew to different receivers, minimize reflections and ringback by keeping the impedances of the line and termination consistent, retain only those harmonics of the clock wave that are needed to meet rise and fall time specifications and duty cycles, and to minimize coupling of signals and clocks.

General Design

There are a number of other design considerations to account for in a high speed design. Any turns taken by a high speed signal should be made smoothly, or at 45 degree angles and never at 90 degree angles, except for via (plane to plane connection) holes. Additionally, every via hole for a signal introduces two 90 degree turns which will appear as slight impedance discontinuities. Excessive use of vias or turns on a signal will increase reflective ringback on the line.

When designing a motherboard for use with the IBM 6x86 processor and Intel** Pentium** processor, it is very important to compare the timing specifications of the two parts, megahertz for megahertz, and design for the worst case of each AC signal timing. Most of the signal valid, setup, and hold times are compatible, but a few differences exist, and can vary from revision to revision of the parts. For example, if a certain signal has a setup requirement of 3 ns on a Pentium processor, and 2 ns on an IBM 6x86 processor , then the worst case of 3 ns should be used in design and simulation. If a clock to valid time for a signal is 5 ns for an IBM 6x86 processor and 3 ns for a Pentium processor, then the 5 ns time should be used for design. Generally, use the longest setup times, the longest output launch valid times, and the shortest hold times. Be sure to consider all required frequency variations of parts when choosing the correct design timings. Since design timing errors are difficult to prevent in most high speed designs, it is very useful to include debugging aids in the first passes of a design. For example, putting a zero ohm jumper in

line with the control signals of the CPU allows easier probing of the signals and the ability to insert a delay or resistive element if needed. Putting test points in a first pass design also allows control signals and data busses to be probed easier without introducing excessive bus capacitance. Once a design is firmed up, small test points and debugging aids can be extracted from the circuit to reduce fabrication costs.

Summary

At high bus frequencies, motherboard design becomes geometrically more difficult since more and more design factors interact to defeat the integrity of the design. Thermal capabilities, signal integrity, timing issues, and power supply issues all require a higher level of sophistication and simulation. By following the recommendations given here, high speed design can be made more reliable and easier to bring up in the lab environment.
