HSTL I/O Sync SRAM Board Design Guidelines

System designers using high speed Synchronous Static RAM must consider how printed circuit board layout affects signal integrity when designing 100–300 MHz systems. Transmission line effects will affect signals even on short trace runs at these high speeds. Trace length and geometry are critical to maintaining signal integrity and ensuring error-free system operation.
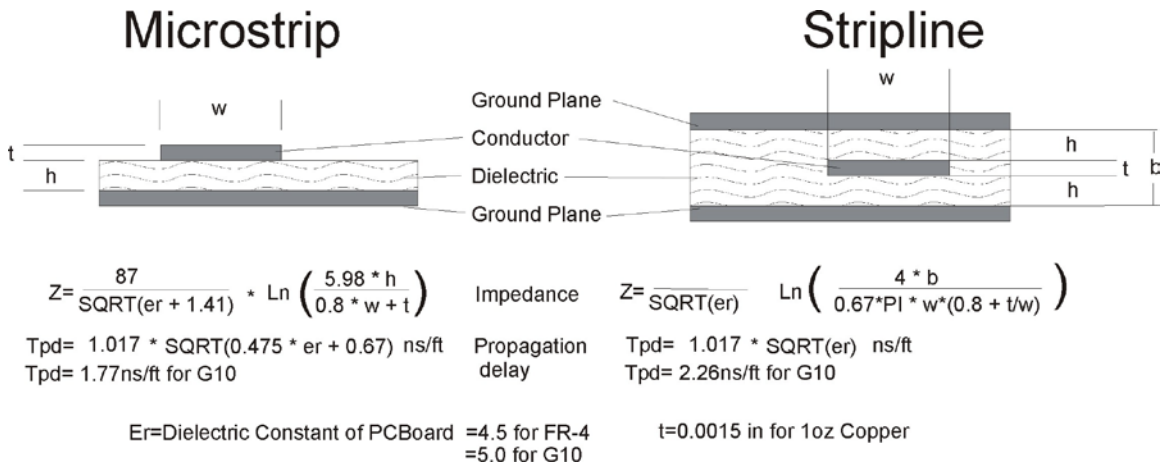
Transmission Lines

A transmission line is pair of conductors with constant parallel capacitance and constant series inductance per unit length. A signal that enters one end of a transmission line propagates down the line at a constant velocity and is not distorted or attenuated as it passes each point on the line. A transmission line may be constructed in a number of different geometries; the two that are important in printed circuit board design are Microstrip and Stripline. Microstrip transmission lines are those where one side of the trace is exposed to air, and the other side is separated from a ground plane by a printed circuit board dielectric material. Strip line transmission lines are sandwiched between two ground planes. The speed at which a signal travels down a transmission line is related to the dielectric constant of the material that surround the conductors. A typical printed circuit board substrate has a dielectric constant of 4.3 with respect to air, which has a dielectric constant of 1. The higher the dielectric constant, the slower the signal propagates, and the more capacitance per unit length.

A trace on the surface of a printed circuit board has one side exposed to air, so it's effective dielectric constant is lower than that of a trace buried inside the board. Signals propagate about 20% faster on surface traces and have about 20% less capacitance per unit length than buried traces.

 A trace on a printed circuit board with a ground plane on the next layer is treated as a transmission line when the signal driven at one end has a rise time that is less than 1/4 the

propagation delay of the trace. Typical CMOS output rise time are 1ns or less. Traces that are 1/4 of this length need to be considered transmission lines. The propagation delay of a 1700 mil long trace that is 8 mils wide and 5 mils above the ground plane is 250pS. This is a typical 50 Ohm Impedance printed circuit board trace in microstrip geometry. A stripline geometry trace that is sandwiched between two ground planes and adjusted to keep the impedance at 50 Ohms would have a width of 4 mils and a distance to the ground plane of 6 mils. However, the propagation delay is slower than the surface, stripline trace, so the length would be 1450 mils for a 250 pS delay. The approximate trace geometry for 50 Ohm impedance microstrip has a width that is twice the thickness of the printed circuit board material between the trace and the ground plane. For buried stripline traces, the width is a third of the distance between the ground planes with the trace placed halfway between the planes.
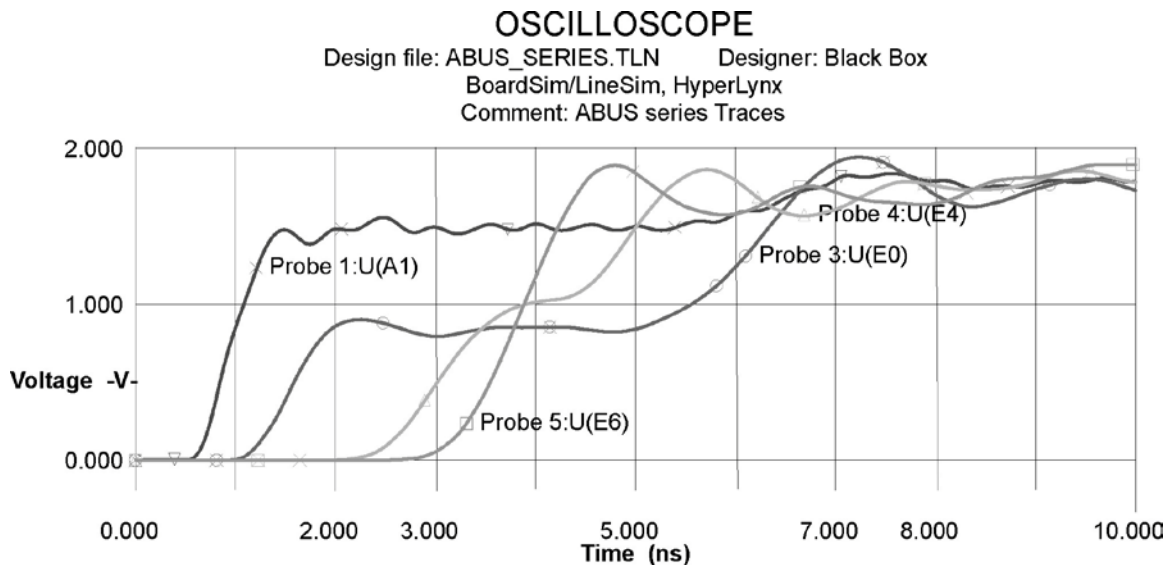
Microstrip

Stripline

Ground Plane
Conductor
Dielectric
Ground Plane

Impedance

$$Z = \frac{87}{\sqrt{e_r + 1.41}} * \ln\left(\frac{5.98 * h}{0.8 * w + t}\right)$$

$$Z = \frac{1}{\sqrt{e_r}} \ln\left(\frac{4 * b}{0.67 * PI * w * (0.8 + t/w)}\right)$$

Propagation delay

$Tpd = 1.017 * \sqrt{0.475 * e_r + 0.67}$ ns/ft
$Tpd = 1.77$ ns/ft for G10

$Tpd = 1.017 * \sqrt{e_r}$ ns/ft
$Tpd = 2.26$ ns/ft for G10

Er = Dielectric Constant of PCBoard = 4.5 for FR-4
= 5.0 for G10

t = 0.0015 in for 1oz Copper

Stripline geometry vs. Microstrip Geometry and equations for impedance, propagation delay
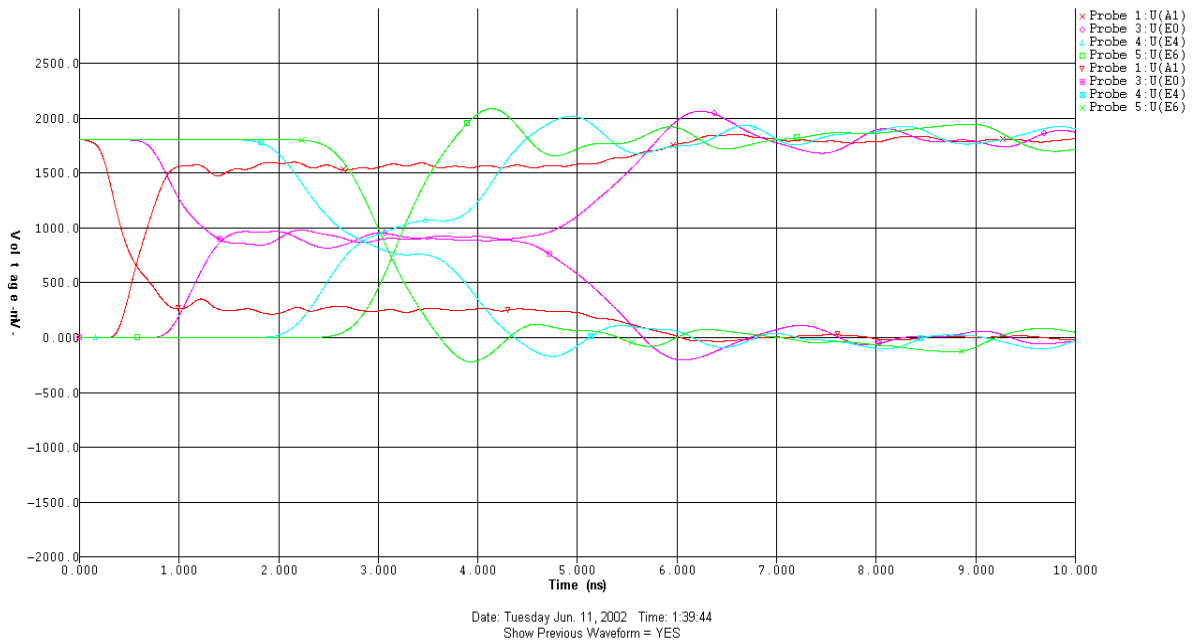
Termination Methods

Source Termination

An impedance matching resistor is placed near the single driver of the transmission line. The value of this resistor is selected such that the driver impedance plus the source termination resistance add up to the transmission line impedance. A 25-ohm driver will need a 25-Ohm Series resistor to properly match with 50-Ohm trace impedance. If there are multiple inputs along the transmission line, those in the middle will see only 1/2 of the full signal initially. An input at the end of the transmission line will see the full transition level at 1/2 of the round trip delay of the line, an input that is very close to the driver will see the full transition only after the initial transition has propagated from the driver, reflected from the open end of the line and returned to the driver. This form of termination is not recommended for clocks or other signals that drive several inputs or are edge triggered. Signals that are latched by a clock or other signal can be source terminated as long and the round trip delay is considered in the timing analysis for setup times.

OSCILLOSCOPE
Design file: ABUS_SERIES.TLN   Designer: Black Box
BoardSim/LineSim, HyperLynx
Comment: ABUS series Traces

Probe1 is connected to the source
Probe 3 is connected to the device closest to the source
Probe 4 is connected to the device closest to the end
Probe 5 is connected to the last device at the end

OSCILLOSCOPE
Design file: ABUS_PARALLEL.TLN   Designer: Black Box
BoardSim/LineSim, HyperLynx
Comment: ABUS series Traces

Probe 1:U(A1)
Probe 3:U(E0)
Probe 4:U(E4)
Probe 5:U(E6)
Probe 1:U(A1)
Probe 3:U(E0)
Probe 4:U(E4)
Probe 5:U(E6)

Voltage - mV -

Time (ns)

Date: Tuesday Jun. 11, 2002 Time: 1:39:44
Show Previous Waveform = YES

End Termination

A trace may be terminated at its end in several ways. An impedance matching resistor is placed at the end of the transmission line whose impedance is Z0. Its value R matches the impedance of the line. The other end of the resistor is connected to ground. This type of termination will require a driver capable of driving a current of V_{OH}/R, for logic at 3.3 V and a 50 Ohm system, the I_{OH} is 66 mA. Not many gates can drive this load. A more common approach is to use what is called a Split Termination where two resistors of value R1 + R2 = Z0 are placed at the end of the trace. On resistor is connected to V_{CC}, the other to ground. The values of the resistors are proportional to the I_{OH}/I_{OL} drive capability of the driving gate. This solution creates a constant current of $V_{CC}/(R1 + R2)$ through the terminations, in addition to the current flowing from the driver, which is $V_{OH}/R1$ or $V_{OL}/R2$. The resistors must be sized for the power dissipation. For CMOS 3.3 V logic with symmetrical outputs, the load current is 33 mA with outputs driven high or low. For 1.8 V logic, the drive current falls to 18 mA, which is marginal for most memory devices.

The HSTL (High Speed Transceiver Logic) specification uses a Thevenin equivalent termination to a termination voltage source. This is also known as active termination. The termination resistance is equal to the impedance of the trace, $R = Z0$. The termination voltage V_T is set to $V_{IO} / 2$. A single, end termination is used with HSTL Class I drivers. At an IO voltage VIO of 1.5 V, the load current is 15 mA to a 50 Ohm termination to $V_T = 0.75$ V. This type of termination works well with a net that has a single driver and multiple inputs along the net. It does not work well for nets that have bidirectional pins.
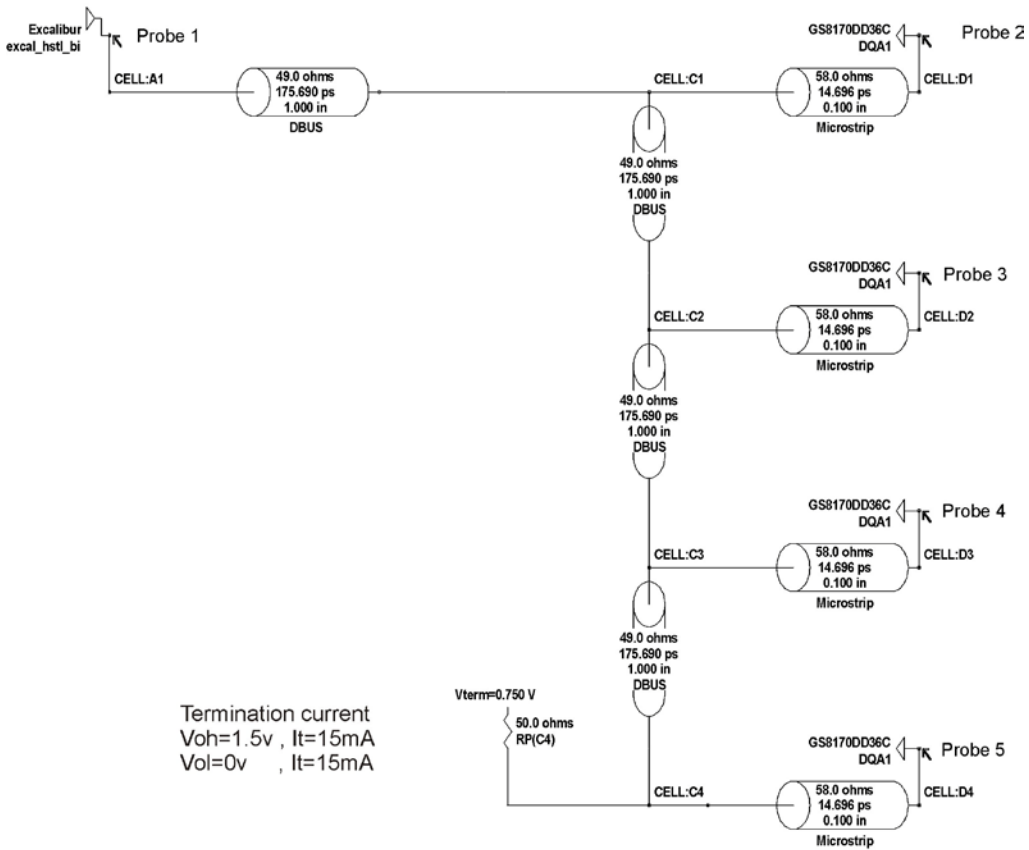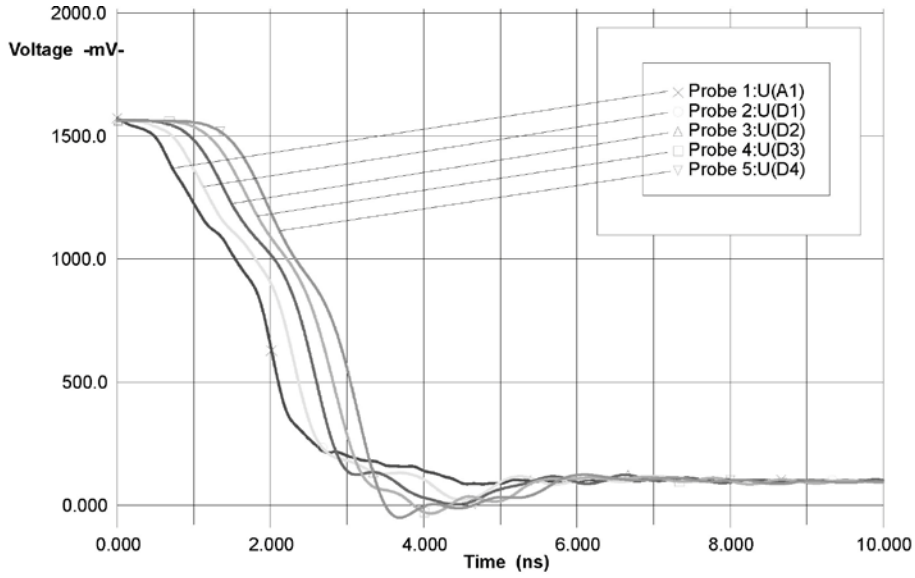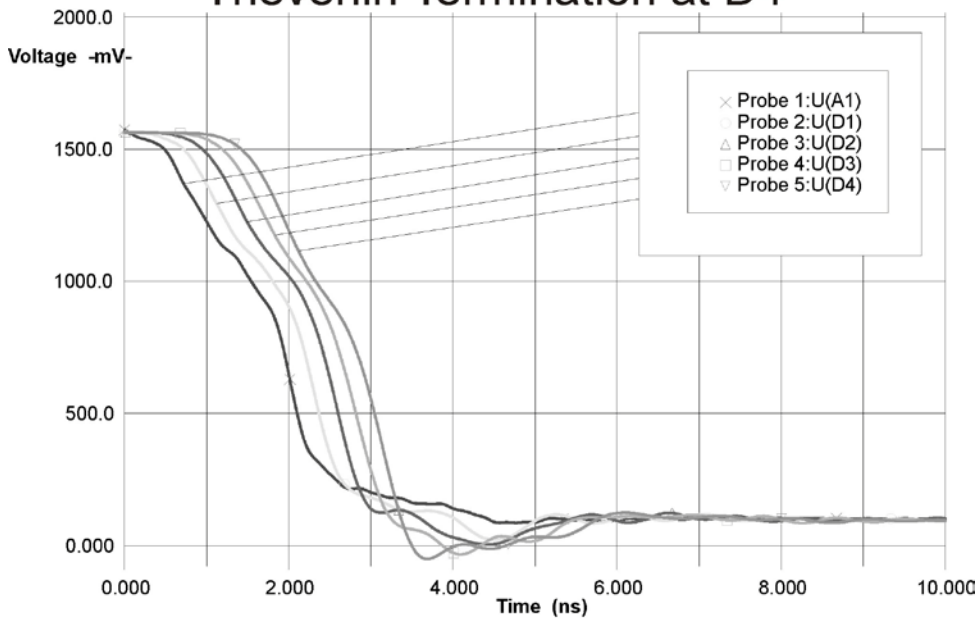
HSTL-I Termination

The inputs along the end terminated transmission line see the full transition as it passes each node until it reaches the termination and is damped. A properly terminated line will have no significant reflections. This type of termination works well with clock signals that are driven by strong, low impedance drivers. This type of termination will use more power than a source termination since the driver must supply current to the termination when the output voltage does not equal the termination voltage. The trade off is that it is faster for signals that go to more than one input since there is no reflection needed to build the signal to full amplitude. The driver is able to drive into a capacitive load directly without having a current limiting resistor in series with the load.

Thevenin Termination at D4

Series Termination at A1
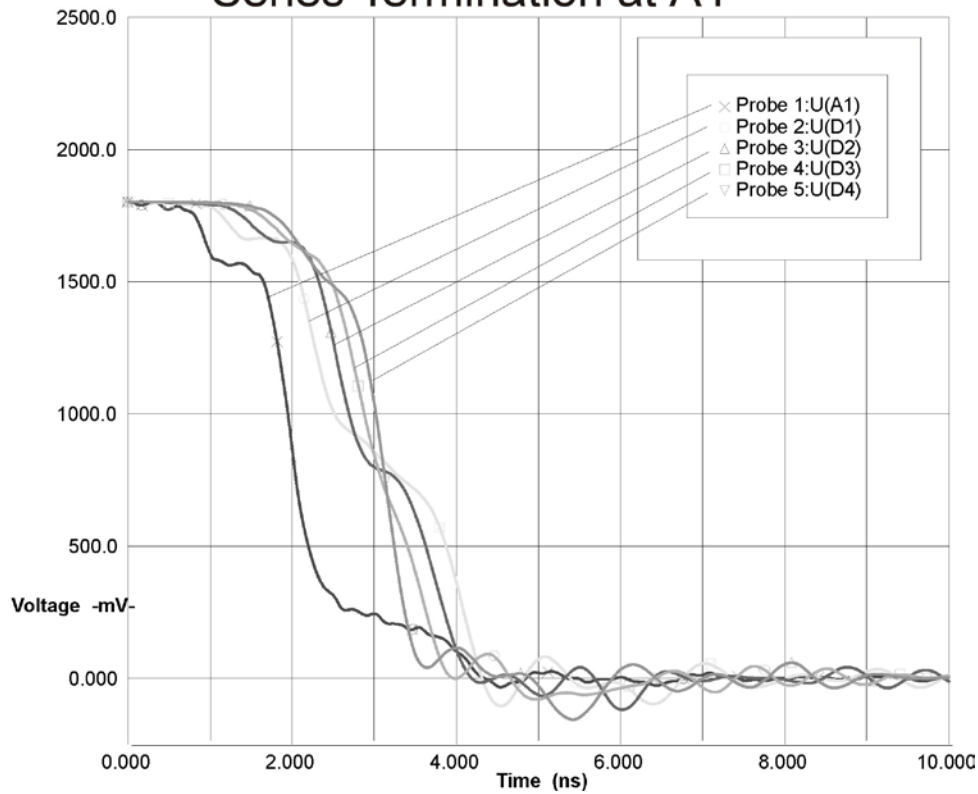
Signal Propagating Down Source Terminated Trace vs.
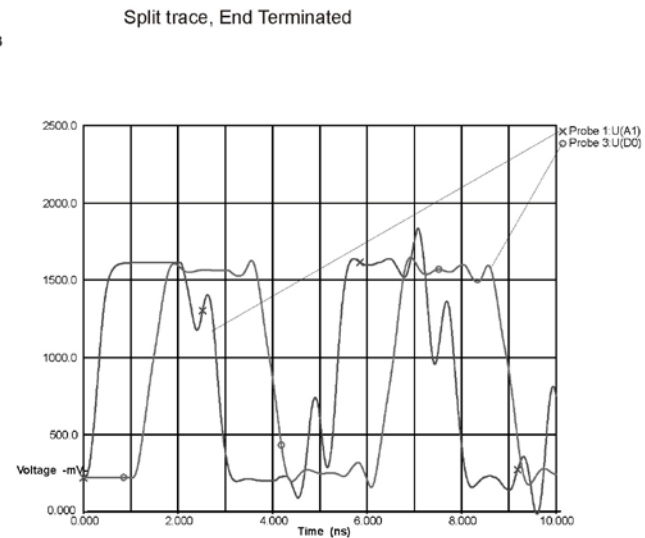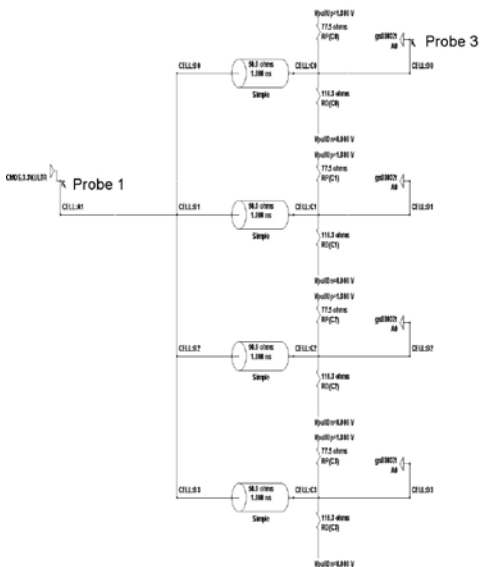
End Terminated Trace

Split Traces

If a signal is needed at multiple inputs, the trace can be split into multiple runs at the driver. Each run must be matched in length and loaded so that the reflected waves arrive at the termination at the same time and are damped out. This will work with a strong driver, such as a system clock driver. If source termination is used, each line is terminated by Rs, where Rs = Zline – (Rdriver * Nlines) . The matching is difficult to achieve in a system, board simulations should be done to verify correct operation.
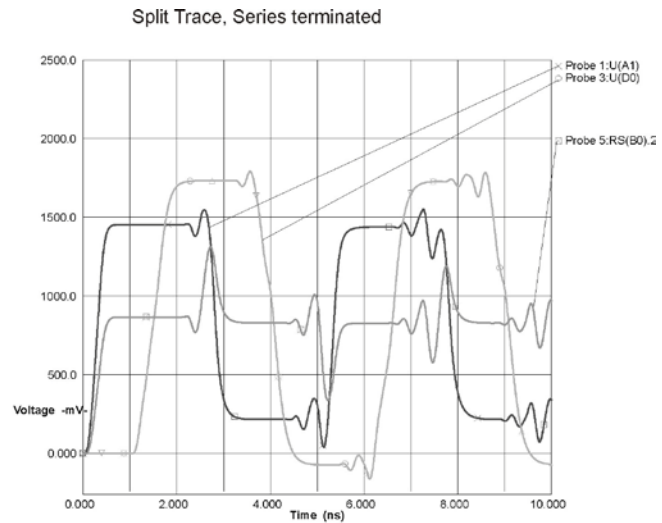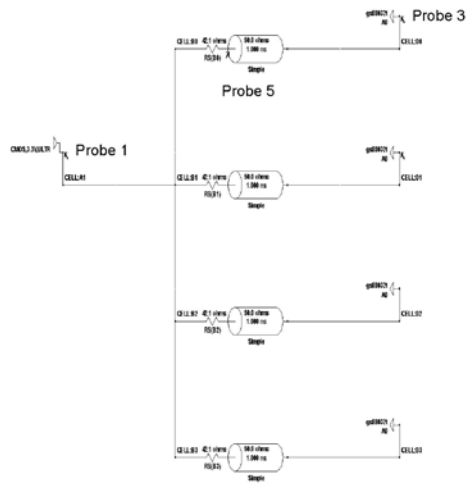
Note: Output impedance of CMOS driver is 3.8 Ohms

Source Terminated Split\End Terminated Split

If the traces are end terminated, the load on the driver will be Rterm/Ntraces, in the case of a 50 Ohm system; the driver will need to be able to drive into a 25 Ohm load. If a single driver is unable to drive the load, multiple drivers from the same device may be paralleled to increase the drive strength as long as the pin to pin skew is not excessive.

Short Traces

If the trace length is 1/4 of the transition time, the system can be considered lumped and treated as if all nodes change at the same time. The speed of a signal in FR4 is about 5 inches/nS. A rough estimate is that if a signal with a 1 nS transition time is driven into a 1 inch long or less trace, then it does not have to be treated as a transmission line. A memory subsystem may have several classes of signals. Each type of signal has a different topology and loading characteristic, and is treated differently. Termination techniques will differ, depending on the type of signal. A signal that is unidirectional and is latched by a separate clock signal may be source terminated. Signals that are sensitive to transitions and reflections should be end terminated.

The system controller drives the address bus to multiple devices. In a multiple device memory array there will be several high impedance gate loads distributed along each trace length. The system controller driver is typically a low impedance, strong output, capable of driving a large capacitive load. Its output impedance is lower than the impedance of the traces on the board. The loads on these traces can be the highest in the subsystem if there are multiple banks of memory.

Control signals such as chip selects, enables and read/write have similar characteristics to the address bus, and may be driven by the system controller to a single device or to multiple memories, depending on the system architecture. They are latched by the system controller clock at the memory device.

The system memory clock is a strong driver that drives several loads. There may be several clock driver outputs. This signal must have a clean transition when changing

states. Reflections in the clock line will cause false memory cycles to occur. Source termination is not advisable for clock signals.

The data bus is a bidirectional bus which can be driven by both the system controller and the SRAMs. The data driven by the system controller must be stable and meet the memory setup and hold time specifications during write operations. It must also meet the system setup and hold specifications during read operations when the memory device is driving the data bus. Timing margins can be critical to system operation.

SigmaRAM memory architecture has an echo clock, which is driven by the memory device to the system controller. This signal is used by the system controller to latch incoming data from the memory. Using this signal in memory arrays solves timing skew issues on memory reads. It allows for banks of memories to be connected to a single bus and increases timing margins.

During a system write, the address and control signals propagate from the controller. The system clock propagates along with them to the memory devices along the bus. Skew between the inputs and clock is maintained so that timing is consistent for all devices on the bus. When the controller drives data on the data bus, the clock and data timing is constant.

When a system read occurs, the controls and clock are driven by the controller and the timing skew is constant, however, the data driven by the memory back to the controller has a skew with respect to the controller-driven clock. This limits the speed and depth of systems unless an echo clock is provided by the memory. The echo clock maintains a minimum constant offset to the data bus for any memory in a bank. The length of the path does not affect the timing skew between the echo clock and the data bus as long as the trace lengths are kept matched. The limit on distance is related to the clock frequency. If the propagation delay is a significant fraction of the clock period, a read memory access from the furthest device may collide with a write access to the nearest device if back to

back cycles are attempted. A wait state may need to be inserted for bank to bank accesses.

Interfacing LVCMOS Memory to HSTL Controller

High Speed Transceiver Logic (HSTL) is a technology-independent standard for device IO pins. It specifies levels and drive strengths for single ended and differential IO pins. There are several Classes of HSTL outputs. HSTL Class I and Class II are the most compatible for interfacing with Low Voltage CMOS inputs.

HSTL Single Ended Input Specs

Spec	Parameter	Min	Typ	Max	Unit
VDDQ	Output Supply Voltage	1.4	1.5	1.6	V
Vref	Input Reference Level	0.6	0.95	1.3	V
VIHT	Logic Threshold High		1.00	Vref+50mV	
VILT	Logic Threshold Low	Vref-50mV	0.90		
VIH	Logic High Input			VDDQ+0.6	V
VIL	Logic Low Input	-0.3			V

HSTL Class Ib Outputs - Output Series Terminated

Spec	Parameter	Min	Typ	Max	Unit
VOH	Logic High Output	VDDQ			V
VOL	Logic Low Output			0	V
RS	Output Source Impedance		15		Ohms

HSTL Class II Outputs - Output and Input Series Terminated

Spec	Parameter	Min	Typ	Max	Unit
VOH	Logic High Output	VDDQ-0.2V			V
VOL	Logic Low Output			0.2	V
RS	Output Source Impedance		15		Ohms

Xilinx and other programmable logic manufacturers allow for 1.8 V logic levels with Their HSTL I/O pins. Verify the FPGA device datasheet you will be interfacing before using 1.8 V V_{IO} memory[1].

HSTL I, due to it's lower 1.5 V levels, has a slower edge rate than CMOS outputs that switch between 0 V and 2.5 V, this reduces the effects of reflections and allows for longer traces to be used before having to treat them as transmission lines.

SigmaRAM CMOS Input Specs

Spec	Parameter	Min	1.5 V	1.8 V	Max	Unit
VDDQ	Output Supply Voltage	1.4	1.5	1.8	VDD	V
VIH	Logic High Input	0.65*VDDQ	.975	1.17	VDD+0.3	V
VIL	Logic Low Input	-0.3	.525	.630	0.35*VDDQ	V

SigmaRAM Low Drive Output

| Spec	Parameter	Min	1.5 V	1.8 V	Max	Unit
VOH@-4mA	Logic High Output	VDDQ-0.4	1.10	1.40		V
VOL@4mA	Logic Low Output		0.40	0.40	0.40	V
RS	Output Source Impedance		50	50		Ohms

---

[1] See GSI App note "SigmaRAM Interface For Virtex-II"

SigmaRAM High Drive Output

| Spec	Parameter	Min	1.5 V	1.8 V	Max	Unit
VOH@-8mA	Logic High Output	VDDQ-0.4	1.10	1.40		V
VOL@8mA	Logic Low Output		0.40	0.40	0.40	V
RS	Output Source Impedance		25	25		Ohms

SigmaRAM with 1.5 V V_{DDQ} only has a 100 mV margin when driving HSTL with an 8mA load and HSTL Vref = 0.95 V. If the HSTL Vref is moved to 0.7 V, then there will be 300 mV of margin for V_{IH} and V_{IL} with the maximum load on the driver. This load is generally only observed with parallel termination. If series termination is used the DC load on a SigmaRAM output will be defined by the leakage currents of the inputs connected to that net.

HSTL output : CMOS input
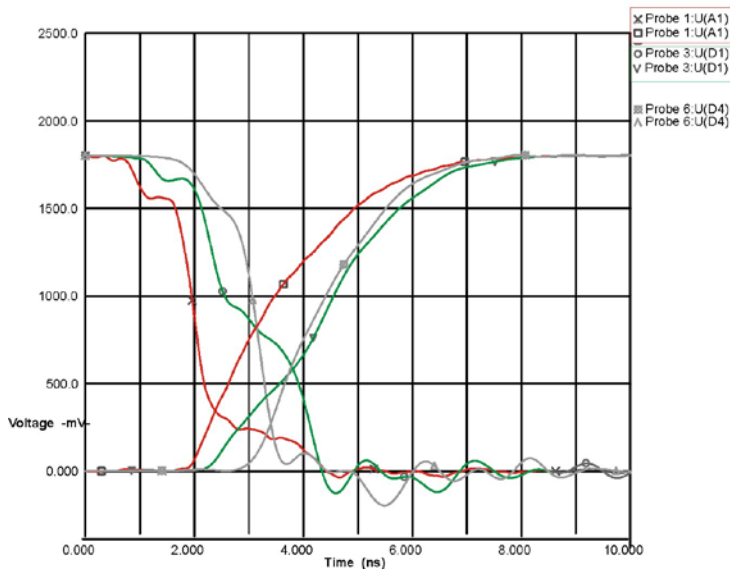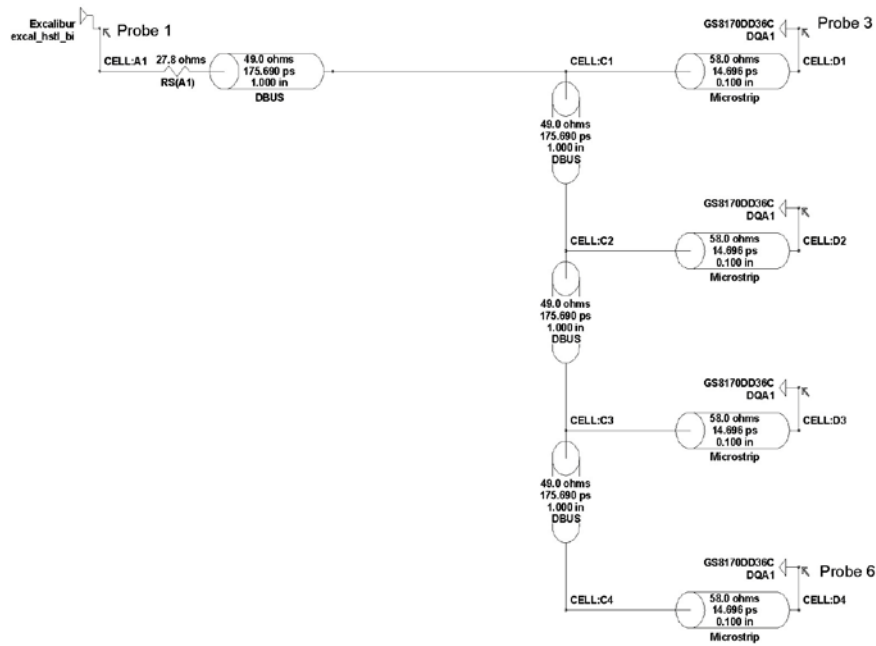
HSTL Output Driving CMOS Inputs

HSTL input : CMOS output
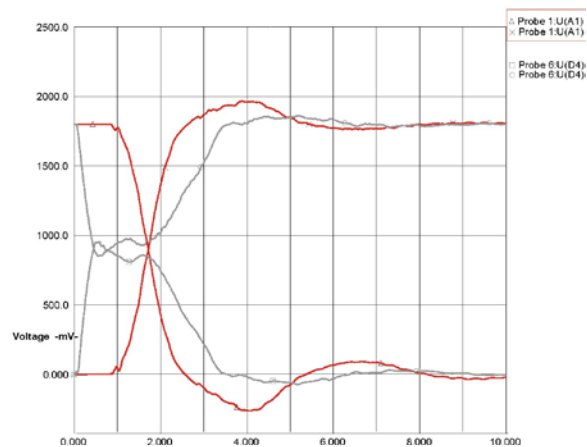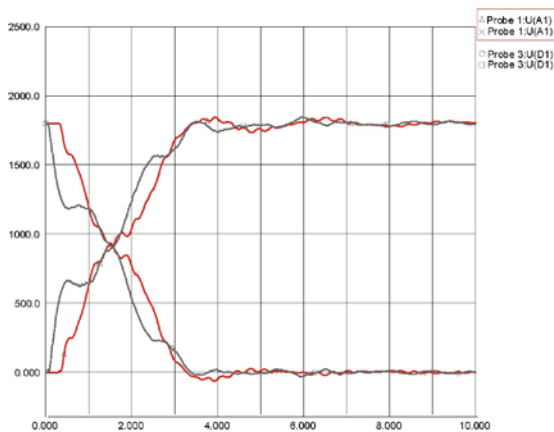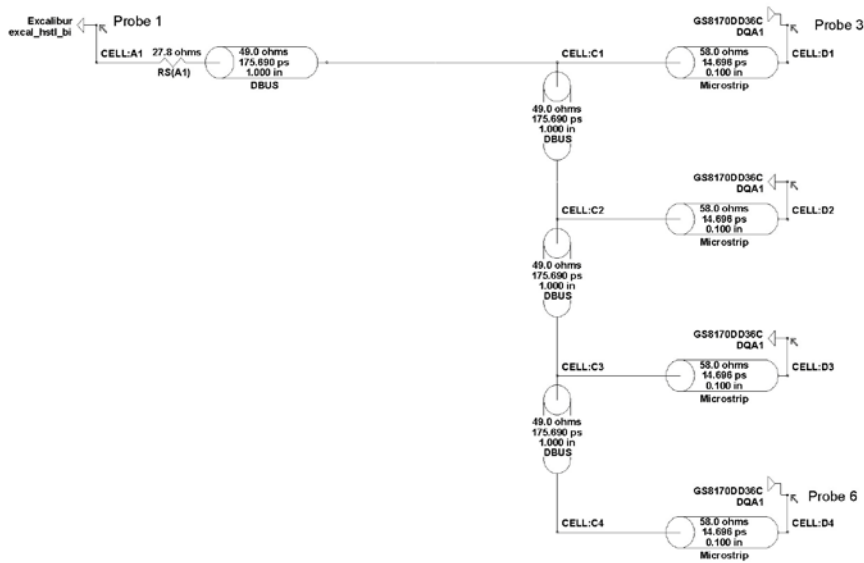
CMOS Output Driving HSTL Inputs

Layout

Placing memory devices on the top and bottom sides of the board is called 'clam shelling'. It allows the layout to share vias on address and data pins, reducing trace length and simplifying routing. The data bus pins on a memory device may be swapped if the memory is operating in a word wide mode since it does not matter which data bit is connected to a particular bit on the controller[2]. If the memory is to be run in a Burst

---

[2] See GSI App note "Memory Address Mismatch"

mode, the data bus may be grouped in bytes, but the bits may still be swapped within the byte.

The address lines may also be swapped to simplify layout. If the memory is running as a word access, and line may be swapped. If bursting is used, the least significant bits of the address bus must go to corresponding pins of the memory so the burst order is correct.

Stress keeping all trace lengths equal to improve timing margins. Serpentining the traces works well in high speed designs to add length to traces. All signals to the part that are latched by a clock should have the same length, not just address and data bus. If there are multiple devices on a bus, keep the order consistent, if a signal is routed to a particular device first, and then all signals must be routed to that device and then to the next device. If there are only two devices and the controller drivers are capable of driving a 25 ohm load, consider making separate, matched 50 Ohm traces to both parts.

Time Of Flight

In high speed systems, the propagation speed of signals on the board must be considered for a design to be successful. The system controller will drive address and control signals along with a clock to latch these signals in the memory devices. If these signals arrive at the part out of sync due to trace differences or loading differences, the memory device will latch incorrect data. Differences in terminations methods will affect timing. An end terminated signal is seen by each input as a full transition as it passes, until it reaches the termination and is damped. Source terminated signals are first seen at full amplitude by the last device on the trace. The devices that are closet to the driver will see the full signal last, in the case of an input that is physically next to the driver, it will take 2x the trace delay for the full signal to propagate to this input. Thus, time of flight becomes a limiting factor to system speed in the case where an output on a memory device is triggered or clocked by a signal from the controller and latched by the controller internally by the same clock. The time required for the clock to propagate to the memory and the time for the memory data to propagate back to the controller reduce the output setup time for the data bus.

EMC and High Speed Layout Guidelines

Place power and Ground planes together to take advantage of capacitive coupling. The makes the board itself acts as a distributed decoupling capacitor.

Add extra ground planes to isolate routing layers. This shields the signals, reducing emissions. Connect ground planes with scattered vias to create shorter paths for signal return currents. When the return path for a signal is broken, the magnetic field expands to reach the nearest path for the return current. This will create electromagnetic fields that radiate from the circuit. Pay attention to non ground vias through ground planes, if the clear out holes touch, it will create an obstruction for signal return currents.

SigmaRAM Layout Example

Using a memory system consisting of 8 SigmaRAMs with 36 data outputs each, the techniques discussed in this paper will be demonstrated. The part chosen for the sample design is the GS8170DW36B with 512K x 36 pinout in the 209-pin BGA package. This package has been designed for use in 'clamshell' layout. The pins are symmetrical when two parts are placed on both sides of a board in the same footprint.
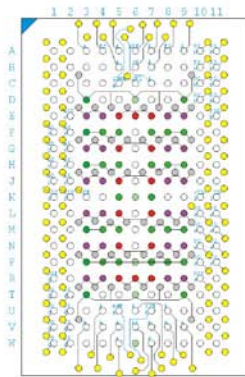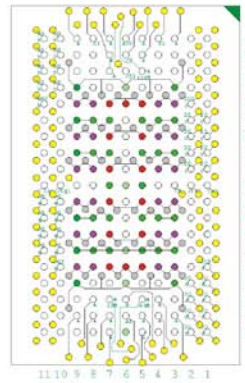
Fig. 1
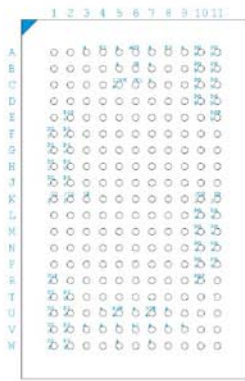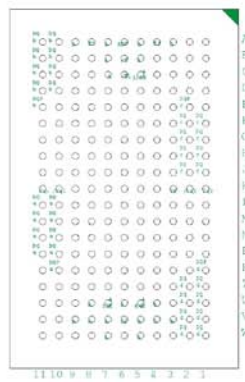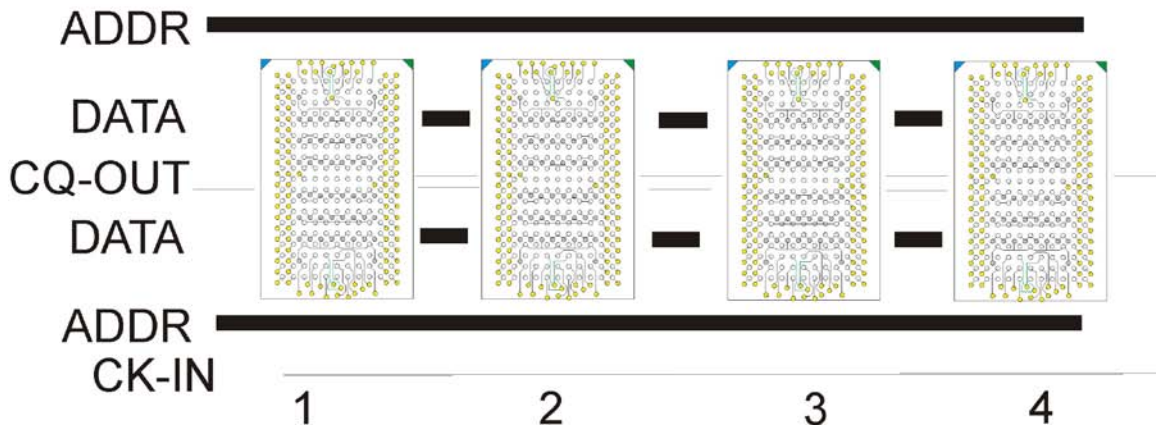

Fig. 2

Fig. 3

Fig. 4

Fig. 5

Fig. 6

In **Figure 1**, the top part is shown with a view looking through the part down at the pads on the top of the board. The BGA pads are the 11x19 array. BGA pads that are connected to V_{DD} core are red , pads connected to V_{DDQ} I/O supply are purple and those connected to ground are green. The vias carrying active signals are filled yellow, the vias carrying power are filled gray. **Figures 4** and **5** show pin function and pin number text in blue for

the device on the topside of the board and in green for the device on the bottom side of the board. Note that in **Figures 3** and **5** the column number ordering is reversed since this BGA is on the bottom of the board. The black traces connecting BGA pads to vias are shown for the top and bottom surfaces of the board only. Internal signal routing is not shown for clarity. The diagram in **Figure 2** shows only the thru-hole vias. Routing channels have been left in the layout for the data bus to be extended through the part to the next device for depth expansion. The clamshell layout of the SigmaRAM BGA facilitates 72-bit data bus width expansion. Comparing **Figures 1** and **3** or **Figures 4** and **5** it can be seen that none of the data pins overlap in the pad layout, this allows for a clean routing of data signals. In **Figure 3**, the four columns of vias are all data bits, each bit on the data bus makes one connection to a via per pair of BGAs. Power is brought into the part through the 6 rows of vias in the interior of the BGA footprint. The row orientation of power vias allow for easy routing of data lines through the footprint of the two devices. In **Figure 3** the center routing area also provides room for the CK clock input and CQ clock outputs. It is best that these signals are not routed next to other non-clock signals to avoid cross talk. The address bus and control signals are routed away from the component footprint and the address bus is split between the top and bottom of the part. The traces connecting the BGA pads to the vias that connect to the address bus need to be kept short to avoid signal reflections, they appear to the address bus drivers as small capacitive loads. SigmaRAM address signal A1 is connected to the system Least Significant Bit (LSB) of the address bus for proper burst operation. The other address bit can connect to any address line. The components share vias on address pins for all but the high order address lines reserved for memory size expansion these pins are not interchangeable on 32Mbit devices. On the top side of the BGA, some routing adjustments have to be made for the non-symmetry of the 128Mbit MSB and a regular address bit.

Figure 6 shows the basic layout of the address and data busses for an 8-chip memory system. The data bus is routed through the body on the component in the inner routing layers of the board, while the address and control bus go above and below the BGA footprints. If surface layers are used for signal routing, extra trace length must be added

to compensate for faster propagation of Microstrip transmission lines. The system clock input to the SigmaRAMs (CK-IN) is routed first to the device furthest from the controller and then to the other devices. It is parallel terminated at device number 1. The CK-IN utilizes incident wave switching because of the sensitivity of clock inputs to false switching. The address bus is series terminated at the system controller. Signals on the address bus use reflected wave switching. Inputs may be at indeterminate logic levels before the CK-IN signal arrives to latch the data. On the address bus, the first device that sees a full logic level is device number 4, the clock is routed to this device first. The trace length of CK-IN to device 4 must match that of the total address bus for proper timing. As the reflection of the address bus signals propagate back to the source, the full amplitude CK-IN signal also propagates with them, keeping the timing relationship constant until the CK-IN is terminated after device number 1. The address bus signals are damped at the source series terminations and do not reflect again. The SigmaRAM provides CQ-OUT signals that are used to synchronize the system controller input latches with the data bus outputs during memory reads. These signals propagate in parallel with the data bus back to the system controller. Since the memories used have bi-directional outputs, series termination is not practical at the BGAs. Source termination at the system controller will act to damp reflections from the SRAMs.