Roger A. Dame

The AlphaServer 4100 Low-cost Clock Distribution System

High-performance server systems generally require expensive custom clock distribution systems to meet tight timing constraints. These clock systems typically have expensive, application-specific integrated circuits for the bus interface and require controlled etch impedance for the clock distribution on each module in the server system. The DIGITAL AlphaServer 4100 system utilizes phase-locked loop circuits, clock treeing, and termination techniques to provide a cost-effective, low-skew clock distribution system. This system provides multiple copies of the clock, which allows off-the-shelf components to be used for the bus interface, which in turn results in lower costs and a quicker system power-up. Component placement and network compensation eliminated the need for controlled-impedance circuit boards. The clock system design makes it possible to upgrade servers with faster processor options and bus speeds without changing components.

Every digital computer system needs a clock distribution system to synchronize electronic communication. The primary metric used to quantify the performance of a clock distribution system is clock skew. Synchronous systems require multiple copies (outputs) of the same clock, and clock skew is the unwanted delay between any two of the copies. In general, the lower the skew, the better the clock system. Clock skew is one of several parameters that affect bus speed. Bus length, bus loading, driver and receiver technology, and bus signal voltage swing also affect bus speed. If problems arise that jeopardize meeting timing goals, though, these additional parameters are difficult to change because of physical and architectural constraints.

The DIGITAL AlphaServer 4100 clock distribution system is a compact, low-cost solution for a high-performance midrange server. The clock system provides more copies of the clock than machines in the same class typically need. The distribution system allows expansion on those module designs where more copies of the clock are needed with minimal skew. The system is based on a low-cost, off-the-shelf phase-locked loop (PLL) as the basic building block. The simple application of the PLL alone would not provide low clock skew, though. Signal integrity techniques and trade-offs were needed to manage skew throughout the system. The technical challenges were to design a low-cost system that would (1) require only a small area on the printed wiring boards (PWBs), (2) be adaptable to various speed grades (options) of CPUs, and (3) have good performance, i.e., low skew. This paper discusses the techniques used to optimize the performance of an off-the-shelf PLL-based clock distribution system.

Design Goals

Based on its experience with previous platform designs, the design team considered a clock skew under 10 percent of the bus cycle time a reasonable target for a midrange server system. The cycle time design target of the AlphaServer 4100 system was 15 nanoseconds (ns); consequently, the skew goal was 1.5 ns or less. This goal would allow a total of 13.5 ns for clock to output of the transmitting module (Tco) (the time the

transmitting module needs to drive data to a stable state from a clock edge); setup and hold time requirements for the receiving module (the minimum time that data needs to be stable at the receiver [flop] before and after the local clock edge); and bus settling time. The following is a breakdown of the timing based on the selection of components for the bus interface:

Bus cycle	15.0 ns
Transmitting module (Tco)	5.1 ns
Setup and hold time for the receiving module	1.5 ns
Clock skew	1.5 ns
Time allocated for bus settling	6.9 ns

The selection of components was based on availability, speed, cost, and size. The goal was to eliminate the need for costly application-specific integrated circuits (ASICs) and still meet the critical timing performance.

The AlphaServer 4100 bus is a simple distributed bus, 305 millimeters (mm) long, with 10 loads (modules) and parallel termination at both ends. The first-order estimate of bus settling time assumed one full reflection or twice the loaded velocity of propagation delay end to end. The estimate took into account bus timing optimization, which is discussed later in this paper. It was also estimated that 25 copies of the clock would be required for the processor modules, and 46 copies of the clock would be required for certain memory modules (synchronous dynamic random-access memory [SDRAM]–based designs). Only the rising edge of the clock could be used for critical timing. If the falling edge were used for latches, then clock skew would dramatically increase because of the duty cycle distortion associated with PLLs. The memory module design allowed very little space for clock circuitry and needed more copies of the clock than any other module design in the system. Further, the physical size of the memory module determined the actual size of the server box. Trade-offs had to be made in the design and timing to make the off-the-shelf solution work. The key goal was to optimize the solution to get the worst-case skew as close as possible to the 1.5 ns estimated goal and to find system trade-offs to allow higher module-to-module skew for a 15 ns bus.

A survey of custom clock circuits available within DIGITAL and off-the-shelf, commercially available PLLs suggested that a custom circuit was required. Unfortunately, the circuits that would be available within our project schedule were costly, consumed far too much circuit board area, required emitter-coupled logic (ECL) or positive emitter-coupled logic (PECL) inputs, and dissipated substantial power. The best off-the-shelf solution was cost-effective, required less space than custom circuits, and provided adequate fan-out. The skew performance, however, ranged from 2 ns to 4 ns, which exceeded the design goal. Given the project time constraints and the design benefits of the off-the-shelf solution, it was paramount that we make the off-the-shelf solution work.

Bus Trade-offs

The design philosophy of using stock components for the bus interface allowed some latitude in the bus design. Typical bus interfaces use large ASICs, each handling up to 50 percent of the data bits. Such a design results in a relatively long dispersion etch from the connector to the ASIC. These devices can range in size from 200 to 400 pins and can require up to 38 mm of etch from the ASIC to the connector. SPICE simulations demonstrated that the length of each module's dispersion etch or bus "stubbing" had a profound effect on bus settling time.[1] Figure 1 shows bus settling time (worst-case driver-receiver combination) as a function of module dispersion etch. The bus trunk length was fixed at 305 mm.

The designers used an 18-bit-wide transceiver in a low-profile surface mount package with a pin pitch of 0.5 mm. The location of the I/O pins for the bus connections on the interface transceiver (located on the same side of the package, which allows the device to be placed very close to the bus connector) and the connector pitch facilitated short dispersion etch (less than 13 mm). This design decreased by 1 ns the settling time typically found on ASIC-based interfaces with comparable trunk lengths and loading.

Bus termination is another parameter that designers can manipulate to further improve settling time. We used parallel terminators at both ends of the bus on the AlphaServer 4100 system. The bus protocol has two features that allow aggressive termination, approaching the unloaded impedance of the trunk. We placed an anticontention cycle between the module that relinquishes the bus and the module that begins to drive the bus. This arrangement reduces the possibility for driver contention (stress) as well as the possibility of generating ringing on the bus caused by large changes in current after contention. The bus "parking" feature forces the last driving module to continue driving the bus to

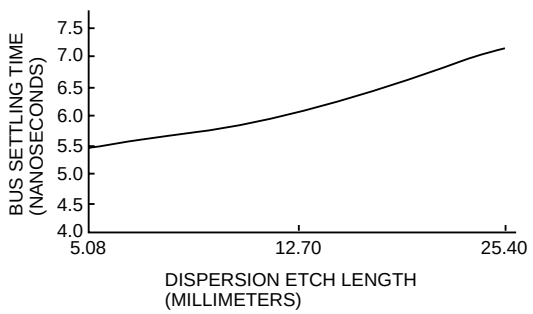
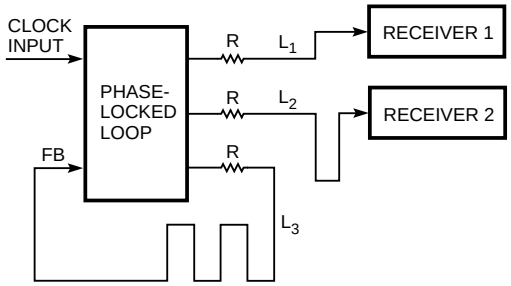


Figure 1
Bus Settling Time As a Function of Dispersion Etch Length

a logic state during long idle times until another module wants to use the bus. Without this feature, the bus would settle at the terminator Thevenin voltage if no modules were driving the bus. Both protocols allow for Thevenin voltage to be close to the thresholds of the receivers. Normally this is avoided if the bus is left idle, because the receivers can go metastable, i.e., arrive at the unstable condition where its input voltage is between its specified logic 0 and logic 1 voltage levels, resulting in uncontrolled oscillation. Centering the Thevenin voltage in the normal full voltage swing had two advantages: (1) it balanced the settling time for both transitions, and (2) it reduced the driver current. The reduced driver current allowed for a lower Thevenin resistance, which brought the terminators closer to the unloaded (no modules) impedance of the bus, thus ensuring that the bus would settle within 6 ns.

The Basic Building Block

Texas Instruments' CDC586 clock distribution circuit was chosen as the basic building block for the system because of its low cost and functionality. The device has a fan-out of 12 outputs with a single compensation loop and a frequency range of 25 megahertz (MHz) to 100 MHz, and is a 3.3-volt (V) bipolar complementary metal-oxide semiconductor (BiCMOS) part. Process skew is 1 ns between any two parts with the same reference input clock, and root mean square (RMS) jitter is 25 picoseconds (ps).[2] The CDC586 has a built-in loop filter, which reduces the number of support components. Unlike custom clock circuits with multiple, independent compensation loops, the simple, single loop design required critical attention to the layout of each module design to ensure the best possible skew performance. The circuit board layout designer had to determine the maximum etch length from the PLL to the receiver. All copies of the clock had to be precisely matched in length to the maximum length found, and routed on the same etch layer with 0.51 mm (20 mil) spacing to other etches and minimum etch crossovers from other etch layers on dual strip-line lay-ups. Typical strip-line etch in multilayer PWBs is a signal layer that has reference planes, usually assigned to power or ground, in the layer above and the layer below. This design allows better impedance control and eliminates cross talk from other signal layers. PWB thickness and cost constraints often result in modified forms on the inner layers, however. Dual strip-line etch is often used in these cases. This design consists of two signal layers sandwiched between reference planes in the layers above and below. Generally the dielectric thickness between the two signal layers is greater than the dielectric thickness between either signal layer and its related (nearest) reference plane to minimize cross talk between the two signal layers. Figure 2 illustrates a typical application.



KEY:

FB	FEEDBACK LOOP INPUT FOR THE PHASE-LOCKED LOOP
R	SERIES TERMINATING RESISTOR
L_1, L_2, AND L_3	ETCH LENGTHS

Figure 2
Typical Phase-locked Loop Connection

Etch Layout

The PWB lay-ups used on various modules in the AlphaServer 4100 system contain microstrip etch (surface etch) and dual strip-line etch. Ideally, single strip-line etch would be optimum for clock etch; however, it requires more layers at higher cost for PWB material. One drawback to dual strip-line lay-ups is etch crossover. A crossover is a point along an etch trace where another etch, one on a different layer not separated by a reference plane, crosses. The crossover forms small capacitance patches, which can load the clock etch and affect its impedance and velocity of propagation. The result is additional skew from clock etch to clock etch. Designers avoided crossovers on all clock etch, and the design does not permit parallel etch on the other layer within the dual strip-line, which could induce cross talk.

Figure 2 shows matched etch lengths L_1, L_2, and L_3. On some module designs, this etch can be fairly long. The layout designers would generally "serpentine" or "trombone" these long etch runs to comply with the aforementioned layout rules. Spacing between the loops on the same etch run in the serpentine or trombone is critical. If the spacing is too close, then coupling will occur, thus changing the velocity of propagation as well as signal quality. Designers used simulation to determine a minimum etch-to-etch spacing for each PWB lay-up. The maximum allowable cross-talk noise level for any minimum spacing was 400 millivolts (mV). This level is within the maximum transistor-transistor logic (TTL) low-state level of 800 mV. Larger spacings were used where no other layout rules would be affected.

The Use of External Series Terminating Resistors

External series terminating resistors (also called terminators), denoted by R, are used at the source (see Figure 2). Although Texas Instruments offers another version of the PLL, namely CDC2586, which has

built-in series terminators, the AlphaServer 4100 designers did not use this variation for the following reasons:

- Some forms of clock treeing (a method of connecting multiple receivers to the same clock output) require multiple source terminators.
- The nominal value for the internal series terminator was not optimum for the target impedance of the PWBs.
- The tolerance of the internal series terminators over the process range of the part could be as high as 20 percent compared to 1 percent for external resistors.

Local Power Decoupling

PLLs are analog components and are susceptible to power supply noise. One major point source for noise is the PLL itself. Most applications require all 12 outputs to drive substantial loads, which generates local noise. A substantial number of local decoupling capacitors (one for every four output pins) and short, wide dispersion etch on the power and ground pins of the PLL were required to help counter the noise. Designers also used tangential vias to minimize parasitic inductance, which can severely reduce the effectiveness of the decoupling capacitors. Typical surface mount components have dispersion etch, which connects the surface pad to a via. Tangential vias attach directly to the pad and eliminate any surface etch that can act like inductance at high frequency. The PLLs were also located away from other potential noise sources such as the Alpha microprocessor chip.

Analog Power Supply Filter

The most important external circuit to the PLL is the low-pass filter on the analog power pins. Typically, PLL designs have separate analog and digital power and ground pins. This allows the use of a low-pass filter to prevent local switching noise from entering the analog core of the PLL (primarily the voltage-controlled oscillator [VCO]). If a filter is not used, then large edge-to-edge jitter will develop and will greatly increase clock skew. Most PLL vendors suggest filter designs and PWB layout patterns to help reduce the noise entering the analog core. The CDC586 PLL was introduced at the beginning of the AlphaServer 4100 design, and the vendor had not yet specified a filter for the analog power input. It is important to note that if any new PLL is considered and preliminary vendor specifications do not include details about the analog power, the designer should contact the vendor for details.

Two forms of low-pass filters were considered: L-C and R-C. The L-C filter consists of a series inductor L from the power source to the analog power pins of the PLL and a capacitor C from the same power pins to ground. The R-C filter consists of a series resistor R from the power source to the analog power pins of the PLL and a capacitor C from the same power pins to ground.

The L-C filter can be implemented in two ways: (1) by using a surface mount inductor and (2) by using a length of etch for the inductor. In either case, the Q of the circuit has to be kept low to prevent oscillation. Q is a dimensionless number referred to as the quality factor and is computed from the inductance L and resistance R (in this case the inductor's resistance) of a resonant circuit using the formula $Q = \omega L/R$, where ω equals $2\pi f$, and f is the frequency. A low-value resistor in series with the inductor can help. Extreme care should be taken if the length-of-etch (used to generate inductance) implementation is considered. The etch must be strip-line-etch isolated from any other adjacent etch or etch on other layers not separated by power or ground planes. A two-dimensional (2-D) modeling tool should be used to calculate the length of etch needed to get the proper inductance value for the filter. Simple rules of thumb for inductance will not work with reference planes (i.e., power and ground planes).

The R-C filter is limited to PLLs with moderately low current draw on the analog power pins. The current generates an IR drop (the voltage drop caused by the current through the resistor) across the resistor R. Typical PLL analog power inputs require less than 1 milliamp (mA), which would allow a reasonable value resistor R. Two capacitors should be used in the R-C type filter: a bulk capacitor for basic filter response and a radio frequency (RF) capacitor to filter higher frequencies. Bulk capacitors are any electrolytic-style capacitor 1 microfarad (μF) or greater. These capacitors have intrinsic parasitics that keep them from responding to high-frequency noise. The benefit of the L-C filter is that, although a single capacitor can be used (two are still suggested with this style filter), the reactance of the inductor increases with frequency and helps block noise. Both filter styles were used in the AlphaServer 4100 system.

System Distribution Description

The AlphaServer motherboard has four CPU slots, eight memory slots, and an I/O bridge module slot. Each module in the system, including the motherboard, has at least one PLL. The starting point of the system is the CPU that plugs into CPU slot 0. Each CPU module has an oscillator and a buffer to drive the main system distribution, but the CPU that plugs into slot 0 actually drives the system distribution. A PLL on the motherboard receives the clock source generated by the CPU in slot 0 and distributes low skew copies of the clock to each module slot in the system. Each module in the system has one and in some cases two PLLs to supply the required copies of the clock locally. Figure 3 shows the basic system flow of clocks.

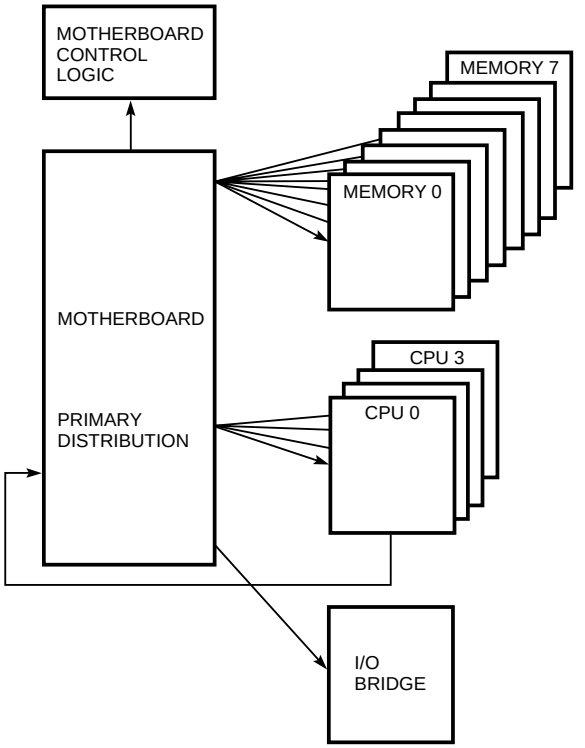


Figure 3
System Clock Flow Diagram

The Alpha microprocessor used on all CPU options for the AlphaServer 4100 system has its own local clock circuitry. The microprocessor uses a built-in digital PLL that allows it to lock to an external reference clock at a multiple of its internal clock.[3] In the context of the AlphaServer 4100 system, the reference clock is generated by the local clock distribution system. The AlphaServer 4100 is fully synchronous.

Each CPU in the system has two clock sources: one for the bus distribution (system cycle time) and one for the microprocessor. This design may appear to be a costly one, but this approach is extremely cost-effective when field upgrades are considered. When new, faster versions of the Alpha microprocessor become available, new CPU options will be introduced. To remain synchronous, the Alpha microprocessor internal clocks need to run at a multiple of the system cycle time. Although the system cycle time goal is 15 ns, the cycle time needs to be adjusted to the speed of the CPU option used. Placing the bus oscillator, which drives the primary PLL for the clock system (cycle time), on the CPU module and designing the clock distribution system to function over a wide frequency range makes field upgrades as simple as replacing the CPU modules. The motherboard does not need to be changed.

Skew Management Techniques

The AlphaServer 4100 system had four design teams. Each team was assigned a portion of the system. Signal integrity techniques had to be developed to keep the skew across the system as low as possible. These techniques were structured into a set of design rules that each team had to apply to their portion of the design. To develop these rules, designers explored several areas, including impedance range, termination, treeing, PLL placement, and compensation.

Impedance Range

Controlled impedance (+/−10 percent from a target impedance) raises the PWB cost by 10 percent to 20 percent, depending on board size. Each raw PWB has to be tested and documented by the PWB suppliers, which results in a fixed charge for each PWB, regardless of size. Therefore, smaller PWBs have the highest cost burden. The AlphaServer 4100 uses relatively small daughter cards. Since low system cost was a primary goal, noncontrolled impedance PWBs had to be considered. Unfortunately, allowing the PWB impedance range (over process) to spread to greater than +/−10 percent makes the task of keeping clock skew low more difficult. Specification of mechanical dimensions with tolerances was the only way to provide some control of the impedance range with no additional costs.

Table 1 contains the results of simulations performed using SIMPEST, a 2-D modeling tool developed by DIGITAL, for a six-layer PWB used on one of the AlphaServer 4100 modules. The PWB dimensions and tolerances specified to the vendors were used in the simulations. The dielectric constant, the only parameter not specified to the vendor, ranged from 3.8 to 5.2, which overlaps the typical industry-published range of 4.0 to 5.0 for FR4-type material (epoxy-glass PWB).[4] Since our PWB material acceptance with the vendor is based on meeting dimension tolerances, we used the 6σ impedance range on all SPICE simulations, thus ensuring that all acceptable PWB material would work electrically.

Table 2 shows the impedance range for a controlled impedance PWB for the target impedance reported in

Table 1
Vendor Impedance Ranges Specifying Dimensions Only

	4σ Yield	6σ Yield
Mean target impedance	71 ohms	71 ohms
Impedance range	62 ohms to 83 ohms	57 ohms to 89 ohms

Table 2
Vendor Impedance Range for an Impedance Tolerance of +/−10 Percent

	+/−10 Specification Range
Mean target Impedance	71 ohms
Impedance range	64 ohms to 78 ohms

Table 1. The difference in impedance range between specifying dimensions and impedance is −7 ohms to 11 ohms. The simulations suggested that the range differences have a minor impact on signal behavior.

The target impedance was based on nominal dimensions and dielectric constant. The target of 71 ohms was chosen to optimize routing density and to keep the layer count down for most designs. Another advantage was that keeping the minimum impedance above 50 ohms would minimize loading. The impedance range covers the full mechanical dimensions and dielectric constant ranges. Properly implemented, the PLLs would effectively eliminate local etch delay module to module over the full process range of the PWBs. The main challenge was to adequately terminate without sacrificing skew performance at the extreme process range (6σ) of the PWB material.

Termination

The designers used series termination on all clocks in the system. Parallel terminators would have exceeded the drive capability of the CDC586. Diode clamping was not practical when so many copies of the clock were required because of PWB surface area constraints. Normally, the optimal termination value is one that provides critical damping for the case where the driver's impedance is the lowest and the etch impedance is the highest. Designers can then make adjustments at the other extreme corner (high driver impedance and low etch impedance) to avoid nonmonotonic behavior such as plateaus. This generally introduces slope delay uncertainty at the slow corner (high driver impedance and low etch impedance), which can be substantial. To minimize this effect, designers selected terminator values that allow overshoot and some bounce-away from the threshold region at the extreme process corner. Overshoot can reach the maximum specified alternating current (AC) input of the receivers over the worst-case process range. Some receivers have built-in diode clamping to their power supply rails as a result of ESD circuits in their input structures (ESD circuits are used for static discharge protection). In these cases, the clock signal is clamped, which in turn dampens bounce. The injection currents caused by clamping would be tested in SPICE simulations to be sure that the parts were not stressed. If the tests indicated stressed parts, designers would adjust the terminator value accordingly.

Treeing

Treeing is a method of distributing clocks from a single source driver to many receivers. This practice, which is well known to memory designers, was used on the AlphaServer 4100 memory modules, bus interface logic, and primary distribution clocks on the motherboard. The designers used two basic forms of treeing: the balanced H tree and the shared output tree. The balanced H tree is best suited for fixed loads (receivers) of the same type (i.e., memories, transceivers, etc.). A single, series-terminated clock output feeds a trunk line to a via and then branches to each load. Each branch is equal in length. The total compensated path includes the pre-terminator stub, the main trunk, and the branch extending to the load. Figure 4 illustrates the clock treeing topology. The shared output tree was used where various module configurations could alter clock loading. Specifically, the distribution on the motherboard is restricted to one PLL to keep the clock skew low. Consequently, some outputs needed to drive more than one slot. A single output driver drove two terminators—one for each load. The low driver impedance isolated reflections from perturbing a module when a module slot was left open.

PLL Placement

Placement of the PLL on each module is critical. Figure 5 is a simplified view of the primary distribution up to and including the PLL on a module. The AlphaServer

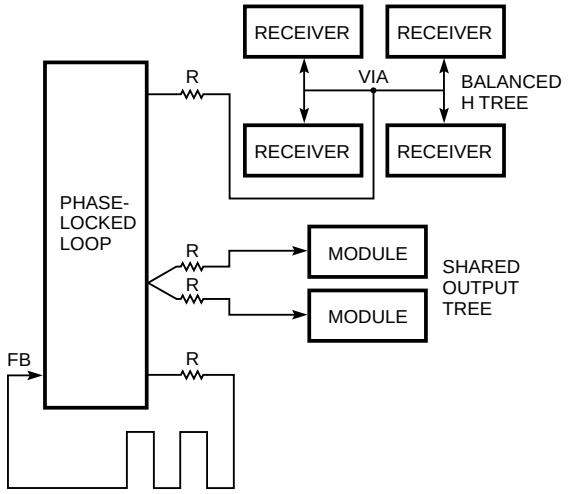


Figure 4
Clock Treeing

4100 system has two types of module connectors: a Metral connector (Futurebus+-style connector) is used on the CPU modules and the I/O bridge module, and an Extended Industry Standard Architecture (EISA) connector is used on the memory modules. Intrinsic delay on these connectors could differ substantially depending on pinning and the signal-to-return ratio in the application. The Metral connector is a right-angle, two-piece connector with four rows of pins: rows A, B, C, and D. The row A pins are the shortest, and the row D pins are the longest. The EISA connector is an edge connector with two rows of pins with minor length differences pin to pin on either side of the connector. Designers had to balance the pinning of these connectors for the clock circuits in such a way that the module-to-module skew would not be affected. The Metral connector was pinned to replicate the loop inductance of the EISA connector.

Dispersion etch is required on each module to connect the PLL to the connector. This etch can have different impedance and velocity of propagation from module to module as a result of PWB process range, which translates into additional module-to-module clock skew. Designers can deal with this problem in two ways.

First, adding the same dispersion length L_3 (see Figure 5) to the compensation loop L_2 nulls this error. This becomes obvious if you look at the PLL's basic function. The insertion delay T_{id} from the clock-in pin of the PLL to the input pin of the receiver is approximately 0 ns if $L_1 = L_2$, or

$$T_{id} = (T_{L_1} + T_{L_3}) - T_{L_2}.$$

For $T_{L_1} = T_{L_2}$ (equal etch lengths), $T_{id} = T_{L_3}$.

Adding T_{L_3} to the compensation path yields

$$T_{id} = (T_{L_1} + T_{L_3}) - (T_{L_2} + T_{L_3}).$$

For $T_{L_1} = T_{L_2}$ (etch equal lengths), $T_{id} = 0$ ns,

where

T_{id} = the insertion delay from the connector pin to the receiver input
T_{L_1} = the etch delay from the PLL output to the receiver input
T_{L_2} = the etch delay of the PLL compensation loop
T_{L_3} = the dispersion etch delay connector to the clock-in of the PLL.

One drawback to this method is that the etch lengths could get fairly large, which would result in edge rate degradation. AlphaServer 4100 designers did not use this placement method on the current set of modules; however, they will consider using it on new designs that require a different location for the PLL.

The second way of dealing with the dispersion etch from the module connector to the clock-in pin of the PLL is to make the dispersion etch very short and to take a skew penalty over the PWB process. Placement studies on the various module designs suggest that a 25-mm dispersion etch would allow reasonable placement of PLLs. The additional skew is just under 50 ps, based on a velocity of propagation range of 5.59 ps/mm to 7.36 ps/mm.

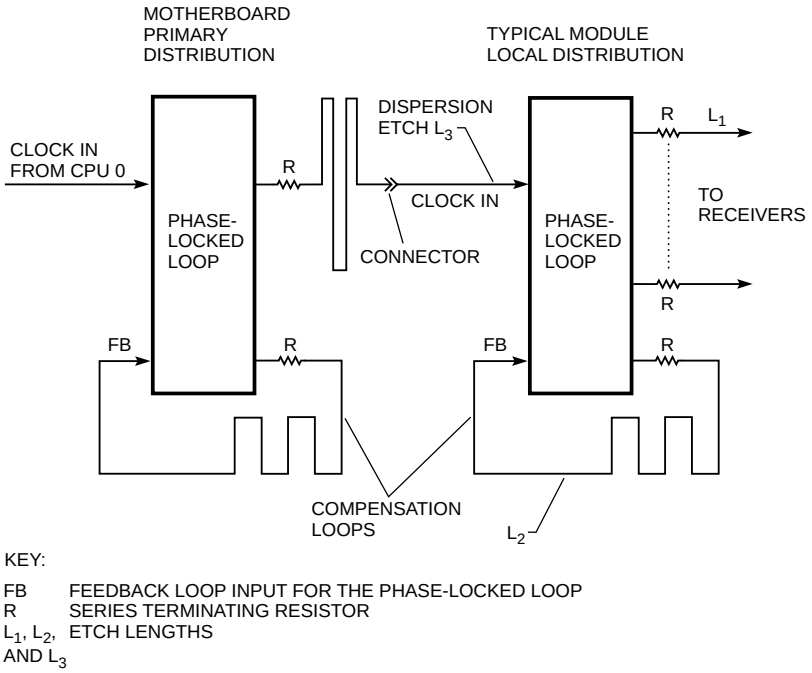


Figure 5
Primary Distribution

Compensation

Some modules have a wide variety of circuits receiving clocks that, because of input loading, do not balance well with the various treeing methods. Designers used dummy capacitor loading to help balance the treeing. This approach was particularly useful on memory modules, which could be depopulated to provide different options using the same etch. Surface-mount pads were added to the etch such that if the depopulated version were built, a capacitor could be added to replicate the missing load on the tree, thus keeping it in balance. The CPU modules have a wide variety of clock needs, which results in two forms of skew: (1) load-to-load skew at the module and (2) control logic–to–CPU skew, for control logic located on the motherboard. The local load-to-load skew is acceptable because only one PLL is used and the output-to-output skew is only 500 ps. Motherboard-to-CPU control logic skew, though, is critical because of timing constraints.

Dummy capacitor loading at each lightly loaded receiver would have reduced the skew, but to compensate for just one heavily loaded receiver would have required many capacitors. PWB surface area and the requirement of simplicity dictated the need for an alternative. The solution was to keep the clock edges as fast as possible (by adjusting the series terminators) and to add a compensation capacitor at the input (the feedback [FB]) of the PLL's compensation loop. This effectively reduced the skew from the slowest load on the CPU to the control logic on the motherboard. Figure 6 shows the disparity between light and heavy loading from T_1 to T_2. Without feedback compensation, the PLL self-adjusts to the lightly loaded receiver. This adjustment results in skew T_1 to T_2 from the heavy load to the control logic on the motherboard. A capacitor on the FB input of the PLL split the difference between T_3 to T_2 and T_3 to T_1 and minimized the perceived skew.

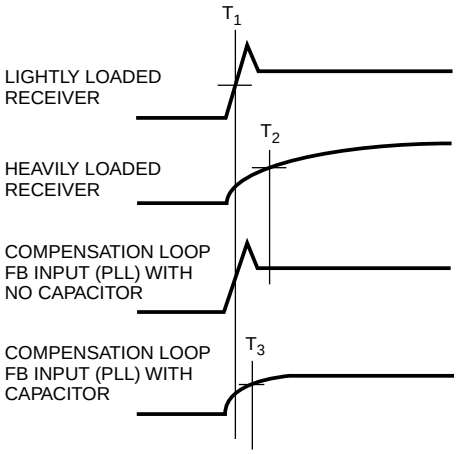


Figure 6
Feedback Loop Compensation

Skew Target

Designers generated the worst-case module-to-module clock skew specification for the AlphaServer 4100 from vendor specifications, SPICE simulations, and bench tests using the techniques discussed in this paper. The worst-case skew goal is 2.2 ns and is summarized in Table 3.

The reader will note that eight times the vendor's specification may appear to be a rather conservative specification. The use of this value was based on two concerns: (1) the PLL was new at the time, and experience suggested that the vendor's specification was aggressive; and (2) some level of padding was required if the exception to the rules was needed. Actual system testing bore out these concerns. The vendor had to relax the jitter specification from 25 ps to 70 ps RMS, and there were some difficulties getting good load balance. The specification did not change, however. Reassessing the allocated bus settling time yields the following:

Bus cycle	15.0 ns
Transmitting module (Tco)	5.1 ns
Setup and hold time for the receiving module	1.5 ns
Clock skew	2.2 ns
Time allocated for bus settling	6.2 ns

SPICE simulations for a fully loaded bus with the worst possible driver receiver position yielded a bus settling time of 5.7 ns. The relaxed skew of 2.2 ns maximum was acceptable for the design.

Comparative Analysis

A comparison of clock distribution systems between two other platforms best summarizes the AlphaServer 4100 system. The AlphaServer 4100 has a price and performance target between those of the AlphaServer 2100 and the AlphaServer 8400 systems. Table 4 compares the basic differences among these systems relating to clock distribution for a CPU module from each platform.

Both the AlphaServer 2100 and the AlphaServer 8400 systems have large custom ASICs for their module's bus interface. The AlphaServer 4100 and the AlphaServer 8400 systems have bus termination; the AlphaServer 2100 system does not. Allowing a bus to

Table 3
Worst-case Clock Skew

Stage	Source	Skew Component
Motherboard	Out-to-out skew	500 ps (vendor specification)[2]
Inputs to modules	Load mismatch	100 ps (simulation/bench test)
Module to module	PLL process	1,000 ps (vendor specification)[2]
Inputs to receivers	Load mismatch	200 ps (simulation/bench test)
Inputs to receivers	PLL jitter	400 ps (eight times the vendor specification)[2]
Total clock skew		2,200 ps = 2.2 ns

Table 4
Clock Distribution Comparison of Three Platforms

	AlphaServer 2100 System	AlphaServer 4100 System	AlphaServer 8400 System
Bus width	128 + ECC	128 + ECC	256 + ECC
Bus speed	24 ns	15 ns	10 ns
Clock skew	1.5 ns	2.2 ns (max.)	1.1 ns (max.)
Inputs requiring clocks	10	25	14
Clock drivers used	12	13	11
Number of clock phases	4	1	1

settle naturally (with no termination), as in the case of the AlphaServer 2100 system, requires a tighter skew budget from the clock system. The trade-off is higher cost, power, and PWB area for lower bus speed. Higher performance systems, such as the AlphaServer 8400 and AlphaServer 4100 systems, generally require faster bus speeds with terminators. The AlphaServer 4100 has shorter bus stubbing (module transceiver to connector dispersion etch) and slower bus speed than the AlphaServer 8400, which allows larger skew (as a percentage of the bus speed).

Table 5 is a comparison of board area needed and cost for the clock system. Designers analyzed an entry-level system consisting of one CPU module, one memory module, and one I/O bridge or interface module. The board area shows the space required by the active components only (the digital phase-locked loops, PLLs, drivers, etc.).

Both Tables 4 and 5 show that the clock system design for the AlphaServer 4100 system requires only one-third the space of either the AlphaServer 2100 system or the AlphaServer 8400 system at a fraction of the cost and distributes more copies of the clock.

Conclusions

An effective, low-cost, high-performance clock distribution system can be designed using an off-the-shelf component as the basic building block. DIGITAL AlphaServer 4100 system designers accomplished this by optimizing the bus and developing simple techniques structured in the form of design rules. These rules are

- Use positive edges for critical clocking.
- Match delay through different connectors using appropriate pinning.
- Use a fixed dispersion etch length from the connector to the PLL.
- Route and balance all clock nets on the same PWB layer.
- Minimize adjacent-layer crossovers and maximize spacings.
- Use minimum value terminators.
- Use tree and loop compensation where needed.
- Use conservative local decoupling and a low-pass filter on the PLL (analog power).

Table 5
Board Utilization and Cost Comparison

	AlphaServer 2100 System	AlphaServer 4100 System	AlphaServer 8400 System
Board area used*	352.8 square centimeters	111.4 square centimeters	371.3 square centimeters
Normalized cost	1.00	0.46	4.40

*Note that these measurements do not include decoupling capacitors and terminators.

The worst-case lab measurement of clock skew between any two modules in a fully configured system was 1.1 ns, which is well within the 2.2 ns calculated maximum skew.

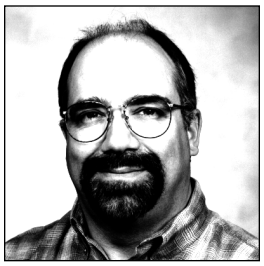
Acknowledgments

Terry Skrypek and Bruce Alford assisted with the prototyping and measurements. Cheryl Preston, Andy Koning, Steve Coe, George Harris, and Larry Derenne worked with the designers to ensure compliance with the signal integrity rules. Darrel Donaldson, Don Smelser, Glenn Herdeg, and Dan Wissell provided invaluable technical guidance.

Note and References

1. SPICE is a general-purpose circuit simulator program developed by Lawrence Nagel and Ellis Cohen of the Department of Electrical Engineering and Computer Sciences, University of California at Berkeley.
2. CDC—*Clock Distribution Circuits,* Data Book (Dallas, Tex.: Texas Instruments Incorporated, 1994).
3. *Alpha 21164 Microprocessor Hardware Reference Manual* (Maynard, Mass.: Digital Equipment Corporation, September 1994).
4. C. Guiles, *Everything You Ever Wanted to Know About Laminates. . . But Were Afraid to Ask,* 4th ed. (Maitland, Fla.: Arlon, Inc., January 1989).

Biography



Roger A. Dame
A principal signal integrity engineer in the Midrange Servers group, Roger Dame is currently working on the AlphaServer 4100 project. During the 10 years he has been with this group, he has also contributed to the VAX 6000, VAX 5800, VAX 7000, DEC 7000, and DEC 10000 projects. In earlier work at DIGITAL, in the Industrial Products group, he developed analog-to-digital process control system interfaces. Roger joined DIGITAL in 1971. He holds an A.S.E.E.T. degree from Springfield Technical Community College and a B.S.E.E.T. (summa cum laude) from Central New England College. Roger is coinventor of the laser bus used in the DEC 7000 and DEC 10000 systems.