BladeCenter midplane and media interface card

J. E. Hughes
P. S. Patel
I. R. Zapata
T. D. Pahel, Jr.
J. P. Wong
D. M. Desai
B. D. Herrman

This paper describes the electrical architecture and design of the IBM eServer™ BladeCenter® midplane and media interface card. The midplane provides the redundant interconnects among processor blades, switch modules, media interface card, and management modules. It also serves as the redundant power distribution medium from the power modules to all blades and other devices. A major attribute of the BladeCenter electrical design is the redundant nature of the interconnects, which gives this product superior reliability and availability. The media interface card provides the interface between the CD-ROM and floppy disk drives and the blades that share these devices. The sharing of these devices was a key BladeCenter innovation. Also, to ensure that the architecture will be flexible enough to support multiple input/output fabric protocols, SerDes (serialized/deserialized) is used as the internal high-speed communication electrical interface. Since high-speed designs can easily result in higher implementation costs, a significant predesign simulation effort was undertaken to analyze and prioritize design guidelines in order to develop a high-speed midplane at a competitive cost. This paper highlights how we reduced board costs by finding solutions that overcame some of the challenges of 2.5-Gb/s data transmission over multiple printed circuit boards and connectors.

Introduction

The IBM eServer* BladeCenter* midplane provides the electrical interconnections among all of the chassis components, including the processor blades, switch modules, management modules, media devices, power modules, and blower modules. It also serves as the redundant power distribution medium from the power supplies to all of these components. The midplane provides up to 14 blade slots, four switch slots, one media bay, two management modules, and four power modules, all of which support hot-swap capability. Interconnections are redundant in order to provide high availability and maximize uptime for the customer.

To provide a multiblade system that supports multiple protocols, such as Gigabit Ethernet, Fibre Channel, and InfiniBand**, over a common copper-trace medium, a serialized/deserialized (SerDes) interface was used for this high-speed internal input/output (I/O) fabric. High-speed designs often require complex, high-speed printed circuit boards that can easily drive costs higher, usually through the use of overly stringent design methodologies or expensive printed circuit board materials that are not required for the system operating parameters. To achieve a cost-competitive midplane for the BladeCenter system, an extensive predesign modeling and simulation analysis was performed to determine the appropriate material and practices to be used in the design of the midplane. This paper discusses the implementation of solutions to overcome some of the challenges of 2.5-Gb/s data transmission over multiple printed circuit boards and connectors in order to reduce board costs.

A benefit of a multiserver chassis is that it provides the capability to share devices and hence lower the per-server cost. In the BladeCenter system, all 14 blades share a keyboard and mouse via the management module, and a Universal Serial Bus (USB) CD-ROM/DVD (compact

0018-8646/05/$5.00 © 2005 IBM

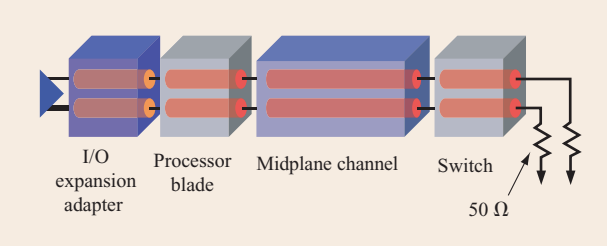


Figure 1

Typical BladeCenter link topology.

disk, read-only memory/digital video disk) drive and USB floppy disk drive (FDD) located in the front of the chassis. These devices are shared over redundant USB 1.1 buses [1] wired from the management module (keyboard and mouse) to the blades and from the media interface card (CD-ROM/DVD and FDD) to each of the blade Teradyne VHDM** (very high-density metric) connectors [2] via the midplane. A USB bus is typically wired as a point-to-point bus between a single host controller and a device (such as a CD-ROM); in the BladeCenter system, however, 14 host controllers share a single USB device. A bus design was created that would effectively create 14 point-to-point buses, one at a time via tristate buffers located on the midplane and under control of the management module and the baseboard management controller (BMC), which is the local systems management of the blade.

This paper describes the initial predesign analysis that was performed to establish the midplane design requirements for both material and design practices for the high-speed SerDes wiring. The methodology for sharing media, keyboard, and mouse is then presented, followed by a description of the midplane board, connectors, and signaling design. Finally, the media interface card, including media devices and the chassis front and rear panels, is covered.

Midplane predesign analysis

This section describes the electrical design challenges that were associated with the high-speed SerDes interfaces and were encountered during the definition, design, and verification of the BladeCenter system. SerDes is a full-duplex, high-speed serial bus comprising a single transmit and a single receive differential pair of signals. The BladeCenter design uses SerDes channels for interconnecting processor blades and high-speed switches that support Gigabit Ethernet, Fibre Channel, InfiniBand, and Myrinet** I/O fabrics. A comprehensive electrical design methodology, including accurate and detailed modeling and simulation of the complete design space, was required in order to achieve the speeds required by these standard interfaces while at the same time using low-cost printed circuit board material—mid-T_g FR4, hereafter known as FR4. Some of the obstacles and solutions to supporting 2.5-Gb/s data transmission over a copper midplane, multiple boards, and multiple connector technologies are highlighted.

The predesign analysis was performed to predict interconnect performance for relatively long trace lengths (across four boards and three connectors) in FR4 material in order to satisfy the high-speed design requirements and ensure the overall performance objective for this system. The predesign analysis was used to create board design guidelines that would prevent signal quality issues and enhance signal eye opening. The signal and power distribution integrity effects—including frequency-dependent losses, intersymbol interference (ISI), crosstalk, impedance discontinuities, and skew—were integrally analyzed across the complete design space. The signal loss was determined to be the prominent issue. ISI is a concern when the signal period is smaller than 2× the delay of the transmission line.

To develop design guidelines in a timely fashion before the first hardware prototype was built, the preroute simulation capability in Cadence SPECCTRAQuest** [3] was used to explore various topology configurations and to determine the limiting factors for total link loss. Given the system mechanical and electrical boundary conditions for the high-speed serial link, the topology shown in **Figure 1** depicts the BladeCenter end-to-end SerDes trace design, in which the differential pair traverses four boards and three connectors with data rates varying from 1 to 2.5 Gb/s. Simulations were performed to cover all permutations of design cases to determine the worst-case scenario. Behavioral device models supplied by external component suppliers were used in fine-tuning the solution space. A variety of tools and techniques were used to generate a channel model as part of the BladeCenter midplane predesign simulation and analysis. Cadence SPECCTRAQuest, Ansoft HFSS**, Synopsys Hspice**, TDA Systems IConnect**, and IBM CZ2D were the primary tools used for simulation. Channel models included the midplane VHDM connectors, I/O expansion adapter, Molex Plateau Technology High-Speed (HS) Mezz** connector [4], and board traces and vias for the I/O expansion adapter, blade, midplane, and switch module. The following sections briefly describe the methods and results associated with simulating the high-speed SerDes channels.

High-speed SerDes design challenges

Numerous challenges must be overcome to design a system that can reliably transmit and receive 2.5-Gb/s serial data over a backplane without errors. The crucial

first step in developing the set of design ground rules to be used in the development of the BladeCenter midplane was to understand high-speed design criteria and the effects design materials and methodologies would have. An initial investigation of high-speed design characteristics yielded four primary areas that had to be taken into consideration: printed circuit board losses, impedance discontinuities, crosstalk, and skew due to board routing. To develop design guidelines before the first hardware prototype was built, the preroute simulation capability in SPECCTRAQuest was used to quickly explore various topology configurations and to determine the limiting factors for total link loss. In addition, to validate some of the simulation results, a number of tests were performed using a standard VHDM backplane evaluation board to guide decisions for designing and building a midplane capable of supporting a bit rate of 2.5 Gb/s. Simulations that used a behavioral channel model were employed to process all of the permutations of design cases. The goal was to determine the worst-case scenario and then fine-tune the solution space. The mechanical and electrical constraints for the high-speed serial link, shown in the topology of Figure 1, represent the typical blade design, in which the signal traverses four boards and three connectors with data rates that vary from 1 to 2.5 Gb/s.

Printed circuit board loss

A transmitted signal must be of sufficient shape and size at the receiver input to be reliably recognized. The high-speed BladeCenter digital circuits operate at frequencies at which the risetime of a typical square wave interacts differently depending on the printed circuit board materials used (i.e., FR4, GETEK**, etc.). A consequence is that the printed circuit board signal path can directly introduce time-domain magnitude attenuation and indirectly introduce jitter, resulting in the degradation of a traveling signal beyond specified limits at the receiver input. The two electrical characteristics primarily responsible for frequency-dependent losses in printed circuit board traces are *skin effect* and *dielectric losses* [5].

Skin effect

The series transmission line elements of a trace, the resistance and inductance, can create a voltage drop. The resistive component can introduce a dc loss, which is independent of frequency. Normally the dc resistive drop is very small and negligible.

The inductance component can introduce an ac loss, which is dependent on frequency. The trace self-inductance, or negative feedback of the self-induced magnetic field, forces the current flow to the surface of the traces. This results in an increase of the ac resistance due to the reduced cross-sectional area. This skin effect causes the ac resistance to increase by the square root of the frequency. **Table 1** shows the skin-effect loss on stripline structure compared with frequency [6].

One approach to compensate for skin-effect losses is to increase the surface area by increasing the trace width. However, this simplistic solution is bounded by physical constraints. One is that wider traces require more board space to route them. As the pitch between component pins and pin-to-via spacing shrinks, the available wiring channels become narrower, reducing the number of traces through each channel. This results in additional wiring channels and probably additional board layers. Another limiting factor is that wider traces can require thicker cores to achieve greater separation between layers to maintain a controlled-impedance board design. Larger cores translate into thicker boards, larger via discontinuities, and higher cost. Wiring-density tradeoffs, simulation results, and cost targets were used to select an optimum trace width for the high-speed channels of 8 mils (0.008 in.). The simulation results are discussed briefly later in this paper.

Dielectric loss

The electromagnetic field associated with signal propagation down a printed circuit board trace resides external to the conductor and propagates through the surrounding medium—either the printed circuit board substrate or air, or both. The propagation speed of the signal is determined by the surrounding medium, which is characterized by the dielectric constant. With respect to practical signal propagation, air is considered equivalent to vacuum, with a dielectric constant $\epsilon = 1$ and a velocity of propagation of the speed of light. The velocity of propagation in all other media is slower than in air and can be gauged by the speed of light divided by the square root of the relative dielectric constant. For paramagnetic or diamagnetic materials, the dielectric constant is reduced to a measurement of the relative capacitance of a material to a vacuum.

When a dielectric medium is exposed to an electromagnetic field, as when a signal propagates down a printed circuit board trace, the polarization is experienced at the molecular level in the surrounding medium. The molecular dipole movement in the insulating substrate creates heat, which means that energy is dissipated. This kind of energy dissipation manifests itself as a loss of signal magnitude. A parameter relating to the energy dissipated is called *tangent loss*. The signal attenuation due to the surrounding insulating material is computed as the product of the tangent loss, the frequency, and the relative dielectric constant.

For FR4, the printed circuit board material used in the BladeCenter design, the dielectric constant varies little with frequency. However, FR4 has a significant tangent

Table 1 Loss effects compared with frequency for a 32-in. stripline.

Frequency	*Skin effect loss* (db)	*Dielectric loss* (db)	*Total loss* (db)
250 MHz	1.8	1.2	2.6
500 MHz	2.8	2.3	4.4
800 MHz	3.5	3.5	6.2
1 GHz	4	4.3	7.3
2	5.2	7.88	11.9
3	6.4	11.4	16.4
4	7.6	16.9	20.91
5	8.9	18.6	25.5
6	9.7	22.15	29.5
7	10.4	25.65	33.5
8	11.2	29.09	37.6
9	11.9	32.5	41.6
10	12.6	36	45.8

Table 2 Printed circuit board dielectric performance compared with cost.

Printed circuit board material	*Dielectric constraint*	*Loss tangent*	*Cost relative to FR4*
FR4	3.9–4.7	0.02–0.03	1
Isola FR406	4.2–4.35	0.013–0.018	1.15
GETEK	3.5–4.3	0.012	1.25
Nelco 4000-12	3.5–3.7	0.0116	2.1
Matsushita Megtron**	3.6	0.006	4.5
Rogers RO3003**	3	0.0013	5

loss value, which can result in large signal deterioration. Many materials that are used as printed circuit board laminates have better loss tangents, but even in volume production, these materials are still cost-prohibitive. **Table 2** shows the dielectric performance of various laminates and relative material cost.

Both dielectric loss and skin effect can contribute indirectly to ISI noise. ISI causes the ac attenuation of the signal high-frequency harmonics and thus dispersion or slowdown of the signal risetime, resulting in the delay of the high-speed time-domain waveform from reaching a full voltage up or down level within a periodic time. The reversal of the time-domain waveforms prior to reaching a steady-state level may produce greater undershoot or overshoot, resulting in a longer recovery time to reach a steady-state level for subsequent waveform changes. The negative consequence is that a longer bit time can occur for some edges in a data pattern and appear as jitter. One solution to the ac loss problem is to selectively increase the signal strength and transition time, which involves boosting only the high-frequency signal content when a waveform at the driver end changes from one voltage level to another. Unfortunately, that does not completely solve the problem of high-frequency rolloff. The pattern-dependent jitter may increase, and the signal may still not be able to achieve full strength within the blade slot. The overall power consumption of the transceiver will also increase as the buffer is required to drive more current.

Although dielectric loss was considered in the development of the design criteria for the BladeCenter midplane, industry data on high-speed interconnects [7] indicated that FR4 would provide an acceptable loss limit and would be the optimum choice for the midplane printed circuit board material.

Impedance discontinuities

Impedance discontinuities must be minimized as much as possible in the design of high-speed SerDes applications. The primary effect of an impedance discontinuity is the reflection of the signal energy back to the source, which adds to or subtracts from the incident signal. The magnitude of the reflection is directly proportional to the impedance mismatch; i.e., the larger the discontinuity, the larger the negative impact on the original signal.

Impedance discontinuities can be found at many points between the transmit and receive buffers along a signal path. Generally, an impedance discontinuity can be at any point where the signal encounters a geometric structure change in the media. A complex, high-density design such as the BladeCenter system is likely to have the potential for many discontinuities because of the number of printed circuit boards and connector systems through which the signals must transition. Connectors, plated-through-holes for connector pins, vias for transitioning layers, and differing board impedances all present impedance discontinuities that can have varying effects on high-speed signals.

The driving parameter for end-to-end impedance is typically a function of the SerDes I/O buffers requiring a constant differential trace impedance of 100 Ω ± 10%. This is also the case for the BladeCenter design, in which the SerDes channels must transition across four boards and three connectors: I/O expansion adapter, blade, midplane, switch module, one Plateau Technology HS Mezz connector, and two VHDM connectors. For all printed circuit boards and connectors that must support SerDes channels, a differential impedance of 100 Ω ± 10% is specified. Also, a criterion for connector use in the BladeCenter system was to have minimal impedance discontinuity into the complete SerDes topology. The

VHDM connector selected for the BladeCenter design ensured that the topology presented minimal and acceptable impedance discontinuities.

Impedance discontinuities are also created by plated-through-holes for the press-fit connectors used on the midplane. When a through-hole via is used to connect the connector signal pin to an internal signal layer, a portion of the via remains unused, creating a stub. The length of the stub depends on the distance from the midplane surface to the signal layer. These stubs can create an impedance discontinuity of as much as 40% to 50%, but this is generally an issue only at very high frequencies (above 5 Gb/s) and for very thick backplanes, where the size of the stub can be significant.

Two potential impedance discontinuities that must be considered for the required bit rates to be supported in the BladeCenter design are the standard signal vias to be used for transitioning layers and the resulting via stubs. **Figure 2** shows the simulation results for via stub effects with different stub lengths. In those cases where via stubs must be taken into account, the most common method of dealing with the stubs is to back-drill the via. Since this additional operation eliminates most of the stub but does incur additional cost, it should be used only if via stubs are determined to present a significant impedance discontinuity. The BladeCenter midplane bit-rate target is less than 5 Gb/s, so signal via stubs were not considered an issue. However, a goal of the BladeCenter midplane architecture was to maintain a clearly defined and tightly controlled design to ensure robust functionality and to provide more flexibility for I/O expansion adapter, blade, and switch module designs. Therefore, a design criterion for the midplane was to maintain consistent impedance from blade to switch module connector and completely eliminate layer transitions and therefore layer transition vias.

To pass 2.5-Gb/s signals through the midplane, reflections due to impedance mismatches must be minimized. This implies that a tight tolerance on the impedance specification is required for the SerDes signal traces. Differential impedance is normally restricted to 100 Ω ± 10%, which simulation results proved adequate for the BladeCenter midplane bit-rate targets. Tighter tolerances that lower reflections, such as ±5%, are achievable, but at a much higher cost. Typically, the printed circuit board manufacturer adjusts the design parameters to meet the specified differential trace impedance. They also account for etching tolerances, adjusting for the nonrectangular shape of the finished trace cross section, and make adjustments to accommodate the standard dielectric thicknesses in order to meet the specified impedance. As a result of these adjustments and manufacturing tolerances, the measured physical parameter values after processing will, with all likelihood, follow a Gaussian distribution around the mean values predicted by the manufacturer.

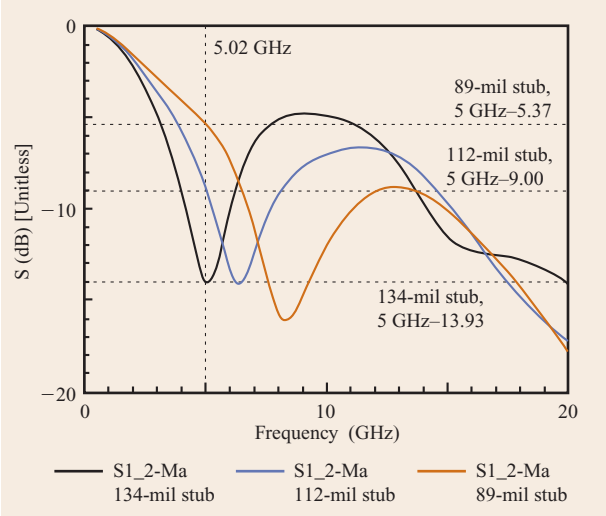


Figure 2

Simulation results for via stub effects with different stub lengths.

To ensure that BladeCenter boards fell within the specified impedance range, test coupons were added to the printed circuit board design before fabrication to provide the capability to measure the trace impedance and loss during the early debug and test phases. A time domain reflectometry (TDR) measurement then provided agreement with the simulated impedance calculation.

Crosstalk

Crosstalk occurs when a signal on a copper trace couples with a signal on another copper trace. This generally occurs when the two traces run in close proximity parallel to each other for an appreciable distance. Crosstalk can also occur between differential pairs of traces, such as the SerDes channels in a backplane, and is caused by inductive and capacitive coupling between the pairs. Crosstalk can be separated into two categories, near-end (NEXT) or reverse crosstalk and far-end (FEXT) or forward crosstalk [8]. The NEXT capacitance and inductance components have the same polarity and therefore do not cancel out (since only the stripline structure was used), while the FEXT capacitance and inductance components have opposite signs and thus tend to cancel out.

With careful placement of transmit and receive differential signal pairs, crosstalk can be reduced significantly without appreciably increasing printed circuit board cost. To eliminate crosstalk as a potential trouble area in the BladeCenter midplane, stringent

design guidelines were established for the high-speed SerDes traces. The separation between pairs of traces was maintained at a distance of $3d$ between traces within a differential pair, where d represents the distance between the signal layer and ground layer. Also, solid ground planes were used between all signal layers, thereby preventing any potential crosstalk from occurring within the midplane printed circuit board and providing a continuous return path to reduce any electromagnetic interference (EMI) effect. Other primary sources of crosstalk in any printed circuit board design are the connectors and module packages at the point where signal pins exit the package within very small distances. Crosstalk is reduced in the connector by surrounding each signal pair with a ground shield.

These shields are present both in the blade card section and in the mating interface to the backplane, providing continuous shielding across the connector. Breakout from module pins in close proximity kept crosstalk below specified levels. Differential signaling also helped in minimizing crosstalk by concentrating the electric field between the signal pairs.

Skew

Controlling the board skew within a differential pair is important at high data rates since the bit time decreases as data rates increase. Board skew contributes not only to loss of timing and voltage margins, but also to converting differential mode energy into common mode energy, leading to radiation and ac common-mode noise. Simulation shows that 50 ps of skew can have 0.4 V of the common-mode voltage. As the skew increases, the size of the common-mode voltage increases. Simulation showed significant harmonic content created on the common-mode currents from differential pair skew imbalances. As a result, the absolute skew had to be reduced so that common-mode noise would be minimized and electromagnetic compatibility performance optimized.

BladeCenter SerDes channels are matched within 20 mils for each differential pair. Note that matching SerDes trace lengths may not guarantee zero skew, as the bends may influence delay and the component package length may differ.

Connector technology selection

Selection of the right midplane connector technology for supporting high-speed differential signals at bit rates of 2.5 Gb/s and above is extremely important. Many factors were taken into consideration, including electrical performance, mechanical robustness, reliability, and routing impact to the printed circuit board to which the connector would be mounted. There were many connector vendors, and from each vendor, many connector types with varying characteristics and intended applications. Some connectors lend themselves to single-ended signals, some to differential signaling. High-speed midplane traces are typically differential, which suggests that the use of a connector technology that lends itself to differential signaling would be of the utmost importance. Also, both press-fit and surface-mount connectors are available. If press-fit connectors meet the data-rate requirements of the intended printed circuit board designs, they may be preferred over the surface-mount connectors because of their mechanical robustness.

For the BladeCenter midplane, only press-fit connectors were considered in the connector selection because of their mechanical strength. Although plated-through-hole (PTH) press-fit connectors do create a via stub, via stubs generally do not have an effect on channel performance until much higher speeds (above 5 Gb/s) are reached and board thickness grows beyond 0.200 in. Since the midplane thickness is in the range of 0.160 in., the effect of stub length was considered negligible.

Other major factors considered while selecting the connector were signal loss, impedance profile, crosstalk, skew within each differential pair, and header and receptacle footprints. When integrated into the SerDes topology, the footprints determine routability, pin density, and the intrinsic impedance of the PTH vias for a given printed circuit board cross section. From our analysis of these factors, it was determined that the VHDM connector technology met the electrical, mechanical, and cost requirements of the system. Even though the VHDM connectors were developed for high-speed backplane applications, trace-length adjustments were still needed on the blade and switch module printed circuit boards to compensate for deltas in the connector pin lengths. The differential pair is formed by two pins located on two different rows, resulting in different pin lengths within the differential pair. This skew was compensated for on the blade and switch cards to minimize the channel skew and to lower the conversion from differential to common mode.

Finalizing the high-speed channel topology and correlation

A behavioral channel model was used in fine-tuning the high-speed SerDes channel solution space for the midplane design. A complete end-to-end channel simulation analysis (time and frequency domain) was developed and carried out using a combination of SPECCTRAQuest and Hspice simulation tools that could be used to cross-check each other. Three approaches were taken in the routing topology investigation to determine the routing design rules and ensure that they would result in a fully functional first-pass design.

Approach 1: IBIS model simulation

One source of modeling data was the component supplier. Typically, component vendors supply I/O Buffer Information Specification (IBIS) [9] models to characterize their module I/O buffers. For IBIS models, one of the available simulators that can import IBIS code is SPECCTRAQuest, which was heavily used in the development of the BladeCenter electrical design. The component supplier provided an IBIS model that defined the receiver cell as a small capacitive load. This receiver model was configured to fit the component supplier's loose description. The IBIS code also included a representative package model of a module pin to the receiver die connection pad. The IBIS package code was in a single-line connector form, and the parasitic element values transferred directly into a resistive/inductive/capacitive matrix. Hspice simulation was done to verify a similar interface using an Hspice model. These device models were used for the Gigabit Ethernet SerDes interface. Not included in the IBIS file were the 100-Ω, on-chip termination and the package interconnection between the receiver and the termination, which meant that the differential impedance of the short package interconnect element could vary from 45 Ω to 70 Ω. Simulations were run in this range to substantiate that any Gigabit Ethernet controller would successfully operate under a wide range of receiver conditions.

An essential activity in any design that relies heavily on simulation, as was the case with the BladeCenter midplane design, is to correlate and verify simulation results with physical hardware. The basic approach to creating an eye diagram with physical hardware was done by triggering a single shot with every clock cycle, putting the scope in infinite-persistence mode, and letting it run repetitively. The simulation waveforms in **Figure 3(a)** were derived from the final BladeCenter SerDes channel layout and from the trace routing that was used to fabricate the first boards. Again, IBIS models were used to represent the Gigabit Ethernet SerDes device in the simulation.

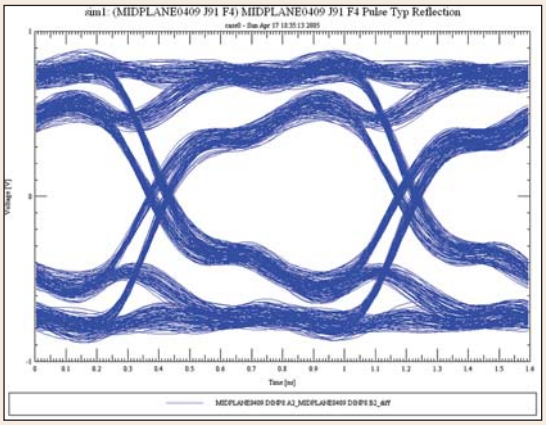
To correlate the eye diagrams captured from real hardware with the eye diagrams from simulation results, it was essential to use the simulation tools to generate eye diagrams using long random waveform patterns. SPECCTRAQuest provides a pseudo-random bit sequence (PRBS) pattern generator. With a click of the *random* button in stimulus edit form, a PRBS pattern can be created automatically and then used to produce pictures such as the eye diagrams generated with infinite-persistence mode. The simulation results correspond to the same module pins that were accessible in the hardware prototype. **Figure 3(b)** is a sample laboratory observation at a receiver pin taken at the printed circuit board via pad. The observed laboratory eye opening was adequate because of the low value of the receiver input threshold requirement. The simulated waveforms in Figure 3(a) and the observed waveforms in Figure 3(b) are very similar but not identical. For instance, the eye height measured in the laboratory was 724 mV compared with 750 mV in simulation. The eye width of 700 ps was observed in the laboratory and in simulation.

Despite some nominal peak-to-peak mismatch due to dc offset, the laboratory and simulation wave shapes shared the same characteristics. Both eye diagrams exhibited a symmetric shelf on the rising edge and the falling edge, which implied an impedance mismatch. SPECCTRAQuest simulation analysis suggested that the impedance mismatch was inside the chip. An investigation with the receiver module supplier confirmed the analysis. The manufacturer determined that the impedance of the interconnect element between the on-chip termination and the input buffer was manufactured in the lower range. This correlation of predicted simulation results with hardware measurements provided high confidence that the simulation methodology was correct and could be relied upon to develop the broader set of printed circuit board base design guidelines.

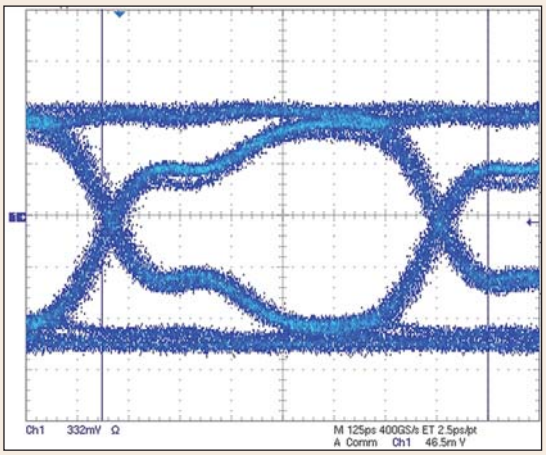
Approach 2: Piecewise-linear pulse simulation

To analyze circuits that operate at more than a 1-Gb/s data rate but for which no Hspice or IBIS models were available, the solution was to use a piecewise-linear pulse with 100 ps of risetime. The preroute simulation of a 1-V voltage source connected to a 100-Ω differential and a 50-Ω single-ended transmission line was performed. The transmission line model was created using a SPECCTRAQuest tool and also verified with other tool sets. The simulation using SPECCTRAQuest showed some loss but was limited to a 128-bit PRBS pattern to study the area of concern. Similar simulation was done using Hspice with a large PRBS pattern. Both results show that a 4-mil or 6-mil line width would produce significantly more loss than an 8-mil line width. After several iterations, the simulation of the 8-mil line width produced enough eye width at 2.5 Gb/s speed. The first prototype was developed using an 8-mil line width. The simulation predicted a worst-case eye opening of 450 mV with a 1-V output for 32 in. of line length. To close the loop and validate the simulation results, eye diagrams were obtained from initial hardware and correlated with the simulation results. The simulated waveforms in **Figure 3(c)** and observed waveforms in **Figure 3(d)** share very similar characteristics. For instance, both show a worst-case eye height of close to 450 mV except for some nominal peak-to-peak mismatch due to dc offset. Similarly, the eye width of 300 ps was observed both in the laboratory measurements and in simulation.

IBIS approach to tuning the high-speed SerDes channel

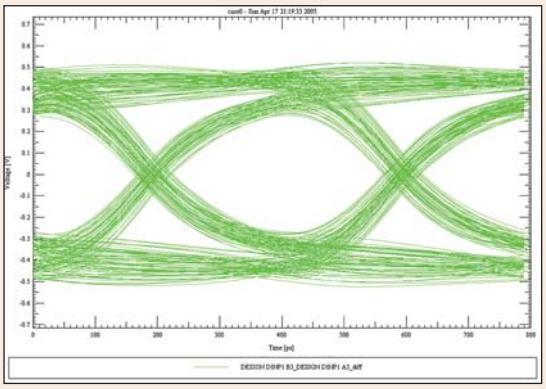


(a) Simulated eye diagram at data rate of 1 Gb/s.

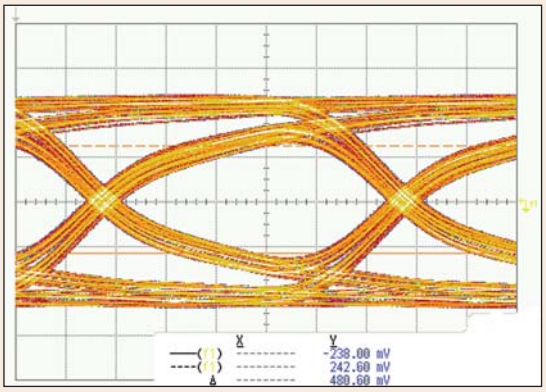


(b) Laboratory measurement at data rate of 1 Gb/s.

Piecewise-linear pulse approach to tuning the high-speed SerDes channel



(c) Simulated eye diagram at data rate of 2.5 Gb/s.



(d) Laboratory measurement at data rate of 2.5 Gb/s.

Figure 3

Correlation and verification of simulation results with physical hardware.

Approach 3: Frequency-domain simulation

Frequency-domain simulation was done to determine the effects of insertion loss and reflections. Signal loss is a key factor in high-performance interconnections. A trace-width analysis was conducted to determine the optimum trace width for the SerDes differential traces that would meet the overall system loss budget while maintaining effective wirability at an affordable cost. All simulation results indicated that an 8-mil trace width of 1-oz copper using a stripline configuration would satisfy the design requirements. The simulation predicted a worst-case eye opening of about 450 mV using a 1-V differential driver voltage and 100 ps of transition time for a 32-in. trace length for the SerDes channel at a data rate of 2.5 Gb/s.

In **Figure 4**, the simulated SDD21 plot shows a 7-dB loss at the fundamental frequency of 1.25 GHz (2.5 Gb/s), which indicates that the loss will be more than half of the output signal. The backplane was routed using this trace-width constraint, which helped yield a minimum layer count and printed circuit board cost.

USB subsystem

An objective of the BladeCenter architecture was to develop a method of sharing common I/O and media devices, such as a keyboard, mouse, CD-ROM/DVD, and FDD, thus avoiding duplication of hardware for each blade in the system. Deciding upon an appropriate bus was the first consideration in selecting a design for sharing the media devices.

Using interfaces for a PS/2 mouse, PS/2 keyboard, standard floppy, and integrated drive electronics (IDE) CD-ROM was determined to be undesirable. Each device uses many wires which, in sum, exceeded the number of pins allocated in the midplane connectors. A single common bus design was needed that could connect all of these devices; FireWire** and USB are natural candidates. The USB standard is fully supported in Microsoft Windows** and Linux** and is widely used. This implies greater maturity in the USB code stack in popular operating systems and more selections of USB hardware components. It was a more pervasive technology than FireWire at the time of the design decision to use USB.

After selection of the appropriate electrical interface to be used in sharing the devices, the method of sharing among 14 blades—either concurrently or one at a time—had to be determined. Initially, concurrent sharing of USB devices across all 14 blades was considered. Because USB protocol is designed to be point-to-point, traffic to USB devices from 14 separate USB host controllers would have had to be managed, along with accurate handling of 14 independent USB code stacks controlling all USB devices simultaneously. Therefore, sharing devices one at a time among 14 blades was selected over concurrent sharing.

Next, the selected method of sharing USB devices was investigated further. The preferred methodology was to connect and disconnect the USB bus to and from target blades in a manner identical to attaching and detaching a USB cable to and from a USB port. This was accomplished by the use of high-performance solid-state switches, switching one blade at a time on or off the bus. A prototype board was designed and tested to establish proof-of-concept for this sharing methodology and to determine the electrical characteristics required for the solid-state switches. Through this laboratory analysis it was determined that a switch of sufficient bandwidth, adequate to support 12 Mb/s and with a maximum capacitive load of 4 pF, was required to maintain the USB signals within the USB 1.1 specification. These solid-state switches are controlled by the management module in conjunction with the BMC. Commands from the management module direct the BMC to switch the blades on and off the bus in an orderly manner such that two USB host controllers are never on the USB bus at the same time.

Midplane physical design

The midplane dimensions are 426 mm × 254 mm (16.772 in. × 10 in.) with a thickness of 4.06 mm ± 0.04 mm (0.160 in. ± 0.01 in.). The board is constructed of FR4 and comprises 18 layers: two power, eight signal, and eight ground layers. The two middle layers are each 2-oz copper and are the main power distribution layers, providing redundant +12 V to all chassis devices. Four of the eight signal layers are dedicated to the wiring of the high-speed SerDes communication channels, while four are used for wiring the remainder of the buses. The solid ground layers are interspersed between the signal layers to eliminate any crosstalk issues. Additional information regarding the midplane can be found in [10].

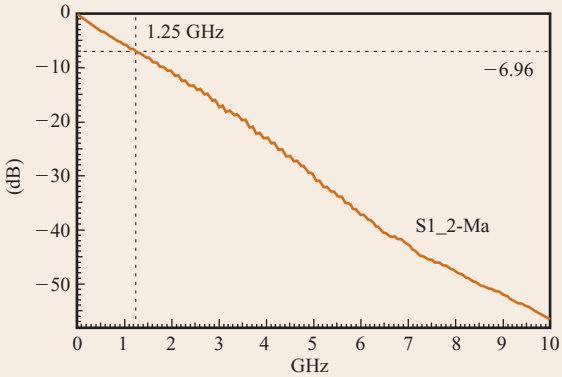


Figure 4

SDD21 plot for BladeCenter channel model.

Midplane connectors

To support the high-speed SerDes channels running at a bit rate of 2.5 Gb/s with sufficient margin to expand support to 3.125 Gb/s in the future, VHDM connectors were selected for the blade and switch module interfaces. The midplane connectors used for the 14 blade slots are six-row, 60-pin VHDM header connectors, two for each slot to accommodate the redundant control and communication signals. The four switch bays have two six-row, 60-pin VHDM connectors each, for a total of 120 pins per switch bay. To support the redundant power distribution, two Molex PowerPlus** [11] two-blade connectors are provided for each blade. Power to the switch bays is also redundant, so two PowerPlus connectors are used for each switch bay. The two management module bay connectors are 144-pin Teradyne HDM** connectors. Redundant power is supplied via three-blade PowerPlus power connectors.

Midplane interconnections

This section describes the signal interfaces for all of the blades and modules supported by the BladeCenter midplane. Refer to **Figure 5**.

Presence

All chassis blade and module slots have at least one presence pin so that the management module can detect

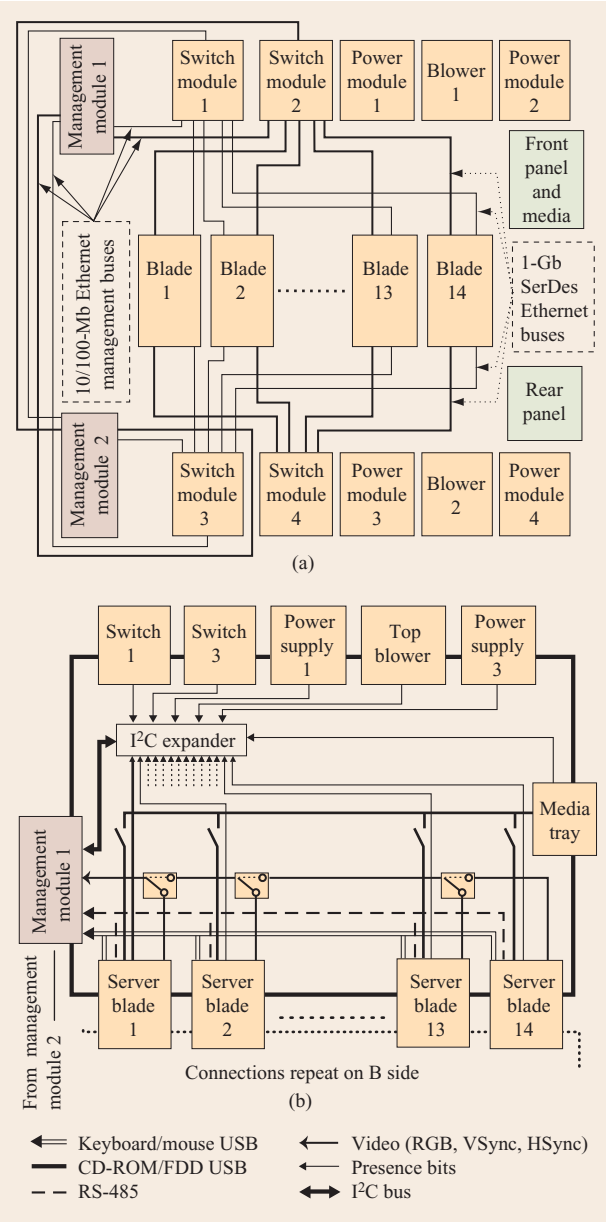


Figure 5

(a) Midplane SerDes and 10/100 Ethernet signaling. (b) Midplane control, USB, and video signaling.

the presence of blades, modules, power supplies, switches, and blowers in the chassis. Blades have redundant presence pins, one in each of the VHDM connectors, and modules and blowers each have a single presence pin. The presence signals are pulled up to +3.3 V on the midplane and, upon insertion into the chassis, the blades and modules pull their respective presence signals low to ground to indicate insertion into the chassis. The presence signals are then routed to redundant Inter-Integrated Circuit (I^2C) Serial Bus Interface registers on the midplane, which are readable by the management module. The management modules also have presence signals that indicate presence to other management modules.

Management module bus A or B selects

Many of the BladeCenter buses are redundant in order to ensure a highly available system. Redundant buses include the basic control and internal communication paths, such as the RS-485, USB, and I^2C buses. Because only one of the bus pairs can be active at a time, the management modules provide two signals, `MM_Sel_A` and `MM_Sel_B`, to indicate to the processor blades, switches, and media tray which redundant set of buses is active. The blades and modules decode the `MM_Sel_A/B` signals, selecting bus A if `MM_Sel_A` is high, or bus B if `MM_Sel_B` is high. If `MM_Sel_A/B` signals are both high or both low, it indicates an invalid condition, and the system components assume that no valid activity will occur on any of the internal control buses or that a functional management module is not present.

Slot addresses

All blade and module slots have multiple address pins that are hardwired to +3.3 V or ground on the midplane to set unique addresses for each slot and bay. The blade slots have four address bits (`A3-A0`) which are used by the blade for establishing its RS-485 bus address in communicating with the management module. The power modules and switch modules all use three address bits to set their base I^2C bus addresses. In addition, the switch modules use the address bits for setting their initial Internet Protocol address for communication with the management module over the internal 10/100-Mb Ethernet channels.

SerDes

The midplane provides the high-speed SerDes communication channels between the processor blades and the switch modules. The SerDes channels are full duplex differential pairs consisting of one transmit pair and one receive pair. The physical design requirements for the SerDes channels were established through modeling and simulation analysis. A trace width of 8 mils at 10-mil trace–trace spacing achieved the desired 100 Ω differential impedance. Additional design rules included the elimination of vias and matching trace lengths in a pair to within 0.1 in. Each blade slot has four SerDes channels; one channel per switch module is wired, which results in each blade having access to four switch modules.

USB 1.1

The USB buses are used to route signals for keyboard, mouse, CD-ROM, FDD, and chassis USB port functions. The keyboard and mouse USB signals are routed from the blades to the management module, while the CD-ROM, FDD, and USB port signals are routed from the blades to the media tray. The USB buses are wired redundantly, but only one pair of buses is active at one time. USB buses are normally point-to-point signal traces, meaning that only two devices are present on the bus. The BladeCenter USB implementation is unique in that the USB devices (keyboard, mouse, CD-ROM, and FDD) are shared among the 14 processor blades. This is accomplished by switching the blades on and off the bus one at a time, effectively creating a point-to-point bus. The inactive USB ports from the blades are electrically isolated from the USB bus by high-impedance switches that are directly controlled by the blade BMC.

I^2C bus

Redundant I^2C buses are routed from the management modules to each of the four BladeCenter switch bays. The I^2C bus provides initial configuration and status communication between the management and switch modules. Two I^2C buses are also routed to the four power modules for redundancy. These I^2C buses allow the management module to monitor the status of the power modules. Two I^2C buses are wired to the media tray and chassis rear panel to control the chassis light-emitting diodes (LEDs) and to read a thermal sensor mounted on the media interface card (MIC). In all cases, the active I^2C buses are indicated by the levels of the `MM_Sel_A` and `MM_Sel_B` signals.

RS-485 interface

The RS-485 bus provides the communication path between the management module and the processor blades. This interface consists of two data signals and one control signal. The two data signals form a two-wire bidirectional differential pair that runs at 57.6 Kb/s. The data signals (from each blade) are connected together on the midplane to form a multidrop interface that is routed to the management module. The slot identification of the blade (`Addr A3-A0`) is used to generate a unique RS-485 address to facilitate communication on this multidrop bus. A collision detection and avoidance algorithm is used for arbitration and recovery.

The third signal is used by the management module to force the blade off the RS-485 bus. The management module, via an I^2C I/O expander device on the midplane, can force any of the 14 blades off the interface. The above three signals are redundant and provide support for a second management module (an optional feature).

Management-module-to-switch-module 10/100 Ethernet

There are two 100-Mb/s Ethernet interfaces on a given Ethernet switch module, with one allocated to each of the chassis management modules. These interfaces can be used to initialize, configure, and monitor the switches, and to provide an Ethernet communication path to other system components. A management virtual local area network (VLAN) is defined to limit switch access to these ports only, but the ability exists to expand access to other ports by broadening the scope of the VLAN. Therefore, the management module is not initially visible from other ports until it is specifically configured in that manner.

Since the interfaces are routed on the internal midplane, the magnetics function typically required for 100-Mb/s links is not used for this application. This was done primarily to save space on the board.

Blower signals

Each blower in the system has three control signals that are routed to both management modules: `Tach_Blwr`, `Speed_Blwr`, and `Fault_Blwr`. The active management module monitors the speed of each blower using the tach signal and adjusts the blower speeds using the speed control lines. If the management module detects a blower failure, it sends the `fault_LED` signal to the failed blower, which then lights its fault LED.

Video: RGB, HSync, VSync

The midplane routes the redundant red, green, blue (RGB) analog video and horizontal sync (HSync) and vertical sync (VSync) signals from the processor blade VHDM connectors to both management module HDM connectors. Because the RGB signals are very sensitive to noise and changes in trace topology, the goal in routing the video signals on the midplane was to maintain point-to-point connections between the selected blade and the management module. To achieve the point-to-point trace requirement, Maxim MAX4259 Video Multiplexer–Amplifiers were used on the midplane. As blades are selected for active KVM, the multiplexers are used to switch sections of traces in and out of the video nets. The MAX4259s were selected because of their low input capacitance (2 pF), fast channel-to-channel switching time (20 ns), and wide bandwidth (130 MHz, 0.1 dB). Even though the HSync and VSync signals are not as susceptible to noise as the RGB signals, AHC125 buffers are used for these signals on the midplane to maintain the skew requirements between the RGB and HSync and VSync signals.

Power distribution

The midplane has two redundant power domains, A and B, which provide +12 V to all of the blade slots and

module bays [10]. Domain A is powered by redundant power modules 1 and 2, which power the chassis infrastructure, i.e., blowers, switch modules, management modules, and blades 1–6. Domain B is also powered by two redundant power modules, 3 and 4, and powers blades 7–14. All blade and module slots have redundant power inputs, one to each of the power modules. Power is distributed over two 2-oz copper planes in the midplane, which are each split in order to achieve the redundant distribution.

Media tray

The media tray assembly houses the media interface card, the front LED and control panel, FDD, CD-ROM drive, USB port, and an internal temperature sensor. The entire media tray is a hot-swappable component of the blade chassis. Figure 5(b) shows the interconnections between the media interface card and the BladeCenter midplane.

USB buses

The two USB connections to the media interface card are redundant and utilize Analog Devices dual-bus switches to select the active USB interface. The active USB signal is routed from the bus switch to a Cypress Semiconductor four-port USB 1.1 hub. Three of the four downstream ports are available to the blades. One of these ports is routed to an external USB type A receptacle, which allows a user to connect any USB 1.1-compatible device to a blade. This port is compliant with the USB 1.1 specification, supporting any self-powered or bus-powered device. The other two hub ports are connected to fixed devices on the media tray: a TEAC USB FDD and a Cypress Semiconductor USB/IDE bridge.

Diskette drive

The TEAC FDD is the same type of device found in external USB FDDs, commonly used for laptop or personal computers. Hence, this device provides full software compatibility with existing USB 1.1 FDD drivers for supported operating systems. Since this device is supported by the blade basic I/O system (BIOS), the FDD can be selected and used as a boot device.

CD-ROM drive

The USB/IDE bridge acts as an interface between the blade USB bus and a slim IDE CD-ROM drive, similar to those found in IBM ThinkPad* models. To the upstream USB host (blade server), the USB/IDE bridge looks like a mass storage device, providing full software driver compatibility with most popular operating systems. The bridge uses the AT Attachment Packet Interface (ATAPI) command set to initialize the IDE CD-ROM drive and, when it receives an ATAPI command on its USB interface, relays the command to the drive during normal operation. Utilizing a USB/IDE bridge for CD-ROM control allows the BladeCenter system to make use of a proven and inexpensive drive—used extensively in the mobile market—without the need for a custom interface. It should be noted that the connection between the bridge and the drive is a standard 40-pin ribbon-cable IDE interface; this brings up the interesting idea that in theory any IDE-compatible drive could be used instead of the CD-ROM: a DVD drive, CD-R drive, DVD-RW, etc. It should also be noted, however, that the USB/IDE bridge can function only at USB 1.1 speeds, so even a fast IDE drive could still operate at only a theoretical maximum that is roughly equivalent to a 6× CD-ROM drive. If supported by the blade BIOS, the CD-ROM drive can be used as a boot device.

Front LED panel

The front LED panel provides the following indicators: power on/off, chassis location, excessive chassis temperature conditions, chassis information (indicating to the customer to look in the management module log for additional information regarding alert conditions, configuration problems, etc.), and fault alarms. The LEDs are driven with a Philips 9551 I^2C bus-controlled LED driver. The fault LED is additionally hardwired to the management module in such a way that the fault LED will light if the management module is removed from the chassis. This is to alert the user of an unsupported configuration. The state of the front panel LEDs is mirrored by the rear LED panel.

Ambient temperature sensor

Mounted on the MIC circuit board is a National Semiconductor LM75 I^2C bus temperature sensor. The management module periodically polls this sensor to determine the ambient temperature for the chassis. This temperature reading is a key factor used in determining the chassis blower speed [10].

Hot-pluggable media tray

The media tray can be removed or inserted at any time during chassis operation, assuming that a user has determined there is no ongoing USB traffic to the FDD, CD-ROM, or any external device plugged into the USB port. This is possible because USB, by design, supports device hot-swapping, and I^2C bus signaling has no real dependence on device removals. Besides these two sets of signals, the only other connections to the midplane are for 12-V power and ground.

Summary

The electrical design challenges encountered before and during the design of this highly dense BladeCenter system were described. Since high-speed designs often generate

high implementation costs, the simulation prediction of the high-speed SerDes channel characteristics was used to establish a robust set of design guidelines and specifications that would result in an acceptable product cost. In addition to achieving the established printed circuit board cost target, implementation of these design guidelines resulted in a successful first-pass design. An acceptable signal loss of 7 dB was predicted and observed for a 32-inch trace length in FR4 material in the system. Also, simulations of the channel structure topology across four boards and three connectors resulted in the selection of 8-mil traces as the desired design point to support a bit rate in excess of 2.5 Gb/s. The wiring density yielded by the 8-mil-trace-width design requirement resulted in minimizing the midplane layer count. Minimizing the layer count and the use of low-cost FR4 printed circuit board material were essential to meeting the printed circuit board cost targets. Finally, the correlation between laboratory hardware and simulation predictions was possible with the use of both the custom stimulus function and system simulation model, which was created prior to the hardware prototype. The system was designed using trace constraint techniques that were successfully proven for impedance matching and minimizing signal loss, crosstalk, and skew.

Another feature of the BladeCenter design that required an initial prototype for proof-of-concept was the sharing of media, keyboard, and mouse devices via a USB 1.1 bus. USB is specified as a point-to-point bus, but the unique implementation of the USB subsystem allowed the sharing of USB devices via the use of high-bandwidth, low-capacitance solid-state switches. These switches are under the control of the management module in conjunction with the blade BMC. Detailed descriptions of the midplane redundant signal architecture, design, and connector technologies were provided. Also described was the design of the media tray, which included the media devices and the front LED panel.

*Trademark or registered trademark of International Business Machines Corporation.

**Trademarks or service marks of InfiniBand Trade Association, Teradyne Inc., Myricom, Inc., Cadence Design Systems, Inc., Ansoft Corporation, Synopsys, Inc., TDA Systems, Inc., Molex Incorporated, GE Electromaterials, Rogers Corporation, Apple Computer, Inc., Microsoft Corporation, Linus Torvalds, or Matsushita Electronic Materials, Inc. in the United States, other countries, or both.

References

1. Universal Serial Bus Specification, Revision 2, pp. 15–24 and pp. 119–170; see *http://www.usb.org/developers/docs/*.
2. VHDM Backplane Connector System; see *http://www.molex.com/cmc_upload/0/000/0–8/388/tab01vhdm.pdf*.
3. P. Patel, B. Herrman, J. Hughes, J. Wong, and M. Cobo, "Gigabit Ethernet Allegro PCB SI Simulation Matching Lab Measurements"; see *http://www.cadence.com/community/allegro/Resources/resources_pcbsi/si/TPIBM_gigabitsimulation.pdf*.
4. Plateau HS Mezz Connector System; see *http://www.molex.com/cgi-bin/bv/molex/index_login.jsp*.
5. A. Deutsch, G. V. Kopcsay, P. W. Coteus, C. W. Surovic, P. E. Dahlen, D. L. Heckmann, and D.-W. Duan, "Frequency-Dependent Losses on High-Performance Interconnections," *IEEE Trans. Electromagn. Compatibility* **43**, No. 4, 446–465 (November 2001).
6. "Using Pre-Emphasis and Equalization with Stratix GX," Altera Corporation, September 2003; see *http://www.altera.com/literature/wp/wp_pre-emphasis.pdf*.
7. R. Kollipara, G.-J. Yeh, B. Chia, and A. Agarwal, "Design, Modeling and Characterization of High Speed Backplane Interconnects," *Proceedings of DesignCon 2003—Design, Modeling and Characterization of High Speed Backplane Interconnects*; see *http://www.rambus.com/news/technical_docs/designcon2003_paper.pdf*.
8. A. Deutsch, "Electrical Characteristics of Interconnections for High-Performance Systems," *Proc. IEEE* **86**, No. 2, 315–357 (February 1998).
9. "I/O Buffer Information Specification (IBIS)," Version 3.2, September 1999; see *http://www.vhdl.org/pub/ibis/ver3.2/*.
10. M. J. Crippen, R. K. Alo, D. Champion, R. M. Clemo, C. M. Grosser, N. J. Gruendler, M. S. Mansuria, J. A. Matteson, M. S. Miller, and B. A. Trumbo, "BladeCenter Packaging, Power, and Cooling," *IBM J. Res. & Dev.* **49**, No. 6, 887–904 (2005, this issue).
11. PowerPlus (SSI) Connector System; see *http://www.molex.com/cgi-bin/bv/molex/index_login.jsp*.

Received December 16, 2004; accepted for publication February 22, 2005; Internet publication October 12, 2005

James E. Hughes *IBM Systems and Technology Group, 3039 Cornwallis Road, Research Triangle Park, North Carolina 27709 (jehughes@us.ibm.com)*. Mr. Hughes is a Senior Technical Staff Member working in BladeCenter Architecture and Hardware Development. He received a B.S. degree in electrical engineering from Pennsylvania State University in 1980 and joined IBM Endicott that same year. Between 1980 and 1995, he worked on several application-specific integrated circuit (ASIC), board, and system-level designs in the development of IBM System/370* and digital video products. In 1995 he joined the Personal Computer (PC) Server team in Research Triangle Park, where he was the lead engineer for several PC server and xSeries* server products. Mr. Hughes has been part of the BladeCenter architecture team since 2000 and is responsible for system and electrical architecture and design.

Pravin S. Patel *IBM Systems and Technology Group, 3039 Cornwallis Road, Research Triangle Park, North Carolina 27709 (pravinp@us.ibm.com)*. Mr. Patel is a Senior Engineer and technical leader working in IBM xSeries server development. He received a B.S. degree in electrical engineering from the New Jersey Institute of Technology in 1989. He is involved with time- and frequency-domain SI analysis/simulation for Intel-based server products. His other areas of responsibility include design and development of models for line cards, backplanes, traces, and vias; performing simulations for system-level voltage and timing budgets; and jitter characterization for high-speed serial link channels. Mr. Patel is currently working in the areas of architecture and analysis of SerDes interfaces for BladeCenter products.

Ivan R. Zapata *IBM Systems and Technology Group, 3039 Cornwallis Road, Research Triangle Park, North Carolina 27709 (izapata@us.ibm.com)*. Mr. Zapata received B.S. and M.S. degrees in electrical engineering from the University of Florida in 1998 and 2001, respectively, joining IBM shortly thereafter to work on xSeries hardware development. Mr. Zapata joined BladeCenter Hardware Development in 2002 and has since worked on various blade server components.

Thomas D. Pahel, Jr. *IBM Systems and Technology Group, 3039 Cornwallis Road, Research Triangle Park, North Carolina 27709 (pahel@us.ibm.com)*. Mr. Pahel received a B.S. degree in mathematics from the University of Carolina at Chapel Hill in 1989 and a B.S. degree in electrical engineering and an M.S. degree in computer engineering, both from North Carolina State University, in 1995 and 2002, respectively. Mr. Pahel joined IBM in 1995 and worked in the IBM Raleigh HelpCenter. He left IBM in 1997 and returned in 1999 to work on the RAS team for low-end two-way servers. In this capacity, he worked on the Software Rejuvenation Project under Dr. Richard Harper. He moved to hardware development for low-end two-way servers in 2000, and subsequently joined the BladeCenter development team in 2001. Mr. Pahel is currently working on future BladeCenter technologies. He is a member of Eta Kappa Nu.

Jack P. Wong *IBM Systems and Technology Group, 3039 Cornwallis Road, Research Triangle Park, North Carolina 27709 (jackwong@us.ibm.com)*. Mr. Wong received a B.S. degree in electrical engineering from Ryerson University. He joined IBM BladeCenter Hardware Development in 2001 and works on various blade server components and BladeCenter midplane.

Dhruv M. Desai *IBM Systems and Technology Group, 3039 Cornwallis Road, Research Triangle Park, North Carolina 27709 (ddesai@us.ibm.com)*. Mr. Desai is a Distinguished Engineer in IBM eServer xSeries Development, working as a BladeCenter system chief architect and strategist. He holds an M.S. degree in computer engineering from Nova Southwestern University and an M.S.E.E. degree from Texas A and M University. Mr. Desai has 24 years of experience in systems design and architecture of Microsoft Windows**/Intel-based systems. He holds 33 patents.

Bradley D. Herrman *IBM Systems and Technology Group, 3039 Cornwallis Road, Research Triangle Park, North Carolina 27709 (herrman@us.ibm.com)*. Mr. Herrman joined IBM after graduating from the University of Florida. He has been involved in printed circuit board design, ASIC logic design, integrated circuit design for memory modules, and circuit design for military applications. His current role is to consult with design groups on rapid implementation of high-speed networks, managing the electrical constraints for the layout design to achieve first-pass hardware prototype success, and making use of analytical applications to eliminate analog defects from printed circuit boards. Mr. Herrman has been awarded two patents.