

DesignCon 2003

High-Performance Systems Design Conference

High-Speed Design Challenges for a 1.4GHz Network Processor

Barry Katz
Douglas Burns
Stephen Coe
Robert Moles
Daniel Nilsson

Signal Integrity Software, Inc.

www.sisoft.com

Abstract

To meet the current and future needs of the internet, faster network processors, capable of quickly analyzing large volumes of data, are required. These network processors require high-speed interfaces to memory, the IO fabric, and the inter-processor bus. While each interface is uniquely interesting, this paper will focus on the design issues involved with the implementation of the 500MHz data-rate QDR memory and a 1.2GHz SPI-4 port found on the Intel® IXP2800 network processor. The QDR introduces the concept of loop-back clocking that requires a sophisticated combination of synchronous and source-synchronous techniques to resolve timing margins. The SPI-4 port represents a high-speed multi-board challenge that requires extensive eye diagram analysis techniques. For each interface, design requirements, technical challenges, design tradeoffs, and the key methodology features are described. Waveform quality and timing results from both the pre-layout and post-route analysis are presented.

Biographies

Barry Katz has been at the forefront of high-speed design throughout his career, devoting his efforts to solving the problems faced by designers of leading edge high-speed systems. In 1995, Barry founded Signal Integrity Software, Inc. (SiSoft) where he has assembled a team of world-class experts committed to solving the most challenging high-speed design problems facing the industry today by delivering a comprehensive design methodology, software tools, and expert consulting. Barry has been a major influence to the signal integrity methodology utilized by numerous companies and has personally led multiple signal integrity design teams. He has expertise in all aspects of high-speed design including: timing analysis, interconnect analysis, crosstalk analysis, electromagnetic modeling of interconnect, packages and connectors, IO buffer analysis and selection, decoupling analysis, simultaneously switching output analysis, interconnect topology, termination selection, clock distribution and skew analysis, and high-speed bus design. He currently serves as chairman of the IBIS Quality committee. Barry holds a MSEE from Carnegie Mellon and a BSEE from University of Florida.

Douglas J. Burns graduated Magna cum Laude from the University of Massachusetts in 1981 with his BSEE and received his MSEE from Northeastern University in 1986. Between 1981 and 2000, Doug worked for Honeywell, Digital, and Compaq Computer. He has led VLSI design implementation teams, signal integrity teams, ASIC implementation teams, and influenced the system architecture of ALPHA systems. Doug consulted on SI and implementation issues across many groups within Digital/Compaq, and has a successful track record of bringing products to market. Current, Doug holds four patents in computer system design and has three additional patents pending. Doug's expertise spans a wide range of engineering disciplines including: system architecture, VLSI design, package design, timing analysis, interconnect analysis, cross-talk analysis, electromagnetic modeling, IO buffer analysis and selection, simultaneously switching output analysis, termination selection, clock distribution and skew analysis, and high-speed bus design.

Stephen R. Coe received his BS in Electrical Engineering from the University of Massachusetts in 1983. Stephen began his signal integrity career when he joined Digital Equipment Corporation in 1987. During his 13 years there, he solved the signal integrity issues for many of Digital's high performance computers. Stephen influenced the design of many high performance memory, PCI and AGP systems. He has been a major contributor to the signal integrity methodology utilized in the design of VAX and Alpha based workstations and servers. He holds patents on several Alpha server products. More recently, Stephen worked on the design of high-speed network appliances including routers, switches and hubs at Nortel Networks. Stephen is an expert in the analysis of interconnect, crosstalk and timing, as well as interconnect topology and termination selection, I/O selection, and electromagnetic modeling of interconnect, packages and connectors.

Robert Moles has worked with board designers in the computer industry on all aspects of signal integrity ranging from the chip IO cell at the transistor level through package, pc board, backplane to cable interfaces. This included board stackup design and layout rules, optimization of terminations, and modeling of packages, boards and connectors and their simulation in SPICE. At Honeywell Bull he provided technical leadership in SPICE modeling, interconnection technology, clock systems and timing analysis at VLSI chip, board, backplane and system interface design levels for midframes and minicomputers. He obtained compliance with "Corporation for Open Systems" 802.3 (Ethernet) LAN conformance tests, resulting in the company being the first to receive the COS Mark for 802.3 internetworking. Then, at Data General, he analyzed difficult signal integrity problems involving ECLinPS speed and other technologies for a high end minicomputer, long etch and cable expansion PCI buses for servers, also 500 Mbyte/s (4ns) SCI cable interfaces for NUMA using LVDS. More recently he has consulted with Compaq on DDR SRAM, 500 Mt/s system bus, package noise, clock distribution and PLL jitter sources, optimizing PCI driver characteristics. Robert has four patents in computers and data communications. He has a BSc (Eng) from London University, England, and taken selected graduate courses at Toronto, McMaster and Northeastern Universities.

Daniel Nilsson received his MS in Electrical Engineering from Luleå University Of Technology, Sweden in 1999. He started his career in 1998 at Ericsson Microwave Systems in Mölndal, Sweden when he joined their High Speed Board Design group. During the years at Ericsson, Daniel worked in defense electronics as well as telecommunication equipment projects. The tasks involved hardware design as well as software design. One of the major contributions as Ericsson was the system partitioning and signal integrity work on Ericsson's next generation AXE telecommunication switches based on the Alpha Microprocessor. Other projects involved a graphics interface for a Modular Airborne Computer System (MACS). His expertise lies within the areas of high speed interconnect design and analysis, system partitioning, and computer system administration. Daniel joined Signal Integrity Software Inc. in 2001 as a Signal Integrity Consultant.

The Intel® IXP2800 Network Microprocessor

The Intel IXP2800 is a state-of-the-art programmable network processor that operates at speeds up to 1.4 GHz. The IXP2800 is built around the XScale core with sixteen micro-engines that collectively provide more than 23.1 giga-operations per second. The external interface delivers OC-192 and 10 Gbps Ethernet performance rates and supports standard interfaces to media and fabric devices. Figure 1 shows a system level block diagram of a dual-processor configuration highlighting the major interfaces including Quad Data Rate (QDR) SRAM (1.6Gbytes/sec per channel), RAMBUS (1.6Gbytes/sec), PCI (64 Bit/66MHz), and SPI-4 (10 Gbps). The IXP2800 is packaged in a 1356 pin FCBGA (37.5 mm x 37.5 mm).

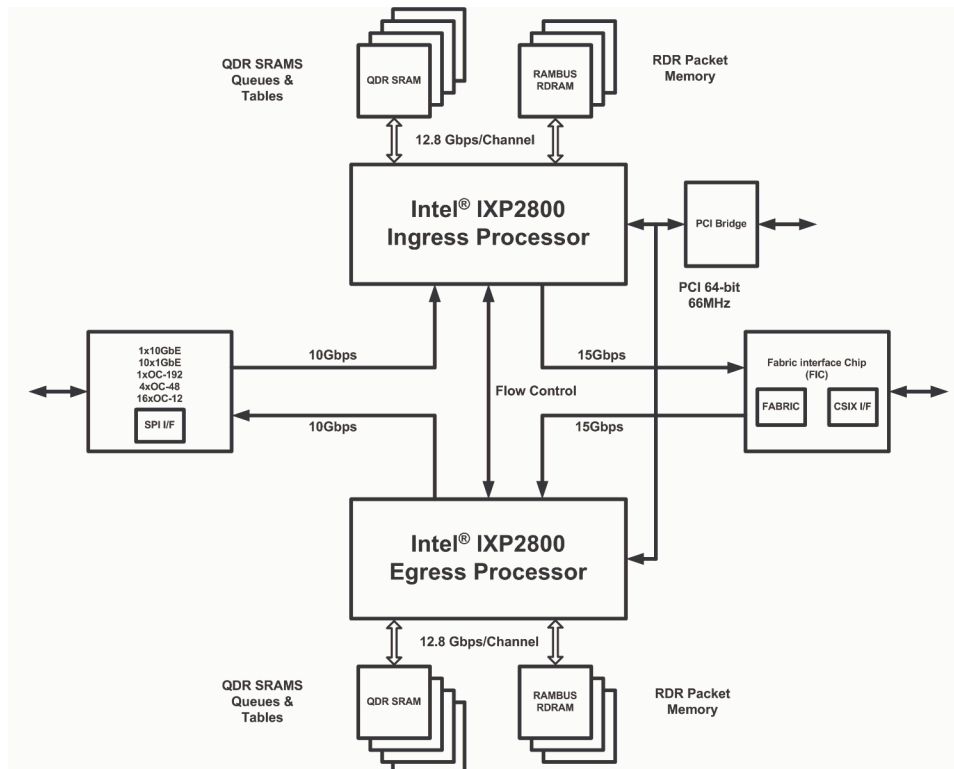


Figure 1 System-Level IXP2800 Block Diagram

The IXP2800 is an extremely high-performance component, both internally and externally, that requires careful consideration of signal integrity, timing, and power delivery from concept through implementation. This paper presents a high level discussion of the methodology that was utilized and high-speed design issues that were considered. The QDR and SPI-4 interfaces are discussed in greater detail to highlight the design challenges. The paper provides readers with key insights into the critical challenges that need to be considered when performing signal integrity and timing analysis for the IXP2800 QDR and SPI-4 interfaces. This paper also includes some general signal integrity tips, checklists and dependencies that should prove useful to other SI engineers as they encounter design analysis tasks of similar complexity.

High-Speed Design Methodology

Successful design of today's high-speed ICs, ASICs, packages, and systems require careful consideration of signal integrity, timing, and power delivery from design concept through physical implementation. Early in the design process, technology assessments are required. The first question to ask is: can it be done? This is a system level question, which spans all components of the implementation. Specifically, designers must evaluate silicon process, I/O cell libraries, package technology, connector technology, spatial constraints, and PCB technology to determine whether system performance targets can be met. Finally, technologies need to be selected and constraints generated for all components of the system. For example, timing budgets may be established and serve as constraints to the chip, package, and system design teams; or layout rules defining topology, termination, and length may be generated for the PCB layout team. This is an iterative process and continues on through the course of a project through implementation, often requiring tradeoffs between the different teams. Some of the high-level design tasks are:

- Signal integrity library development
- I/O buffer selection
- I/O buffer characterization
- IBIS model generation
- Package analysis and design
- Power delivery
- System decoupling requirements development
- Printed circuit board stack-up definition
- PCB layout rules generation
- Connector analysis and pinning
- Post-layout verification
- Integrated signal integrity and timing analysis

The execution of many of these tasks for the IXP2800 required massive amounts of simulation to cover the entire solution space of process, voltage, temperature, stimuli, topology, coupling, populations, and variants. Initial simulations were run with both HSpice and IBIS models to correlate IBIS model accuracy. Some end-to-end simulations required HSpice models to analyze timing from the input of the output buffer to the output of the input buffer (SiSoft's Core-to-Core™ methodology). SiSoft's SiAuditor™ product was used to perform the exhaustive solution space analysis, which required HSpice/IBIS co-simulation and integrated signal integrity and timing analysis.

Waveform Analysis

Switching buses at high data-rates results in signals that do not settle in one bit time. It is extremely important to consider inter-symbol interference (ISI) effects, coupling, and package variations. Furthermore, exciting the various energy states on a network requires the use of complicated stimulus patterns and requires waveform analysis on all edges of the output signal for quality and interconnect delay. Capturing all of the timing and waveform quality information requires a robust set of measurement criteria. Figure 2 shows the types of voltage levels required to ensure accurate capture of waveform information. Figure 3 shows the detailed eye diagram processing required to accurately capture critical eye diagram characteristics. The SiAuditor toolset utilizes all of these levels in

processing the signal waveforms and applies these levels to every simulation edge of every waveform. Using the method of Progressive Discovery, the tool presents multiple views of the data, in a clear and concise format, which eases the analysis effort.

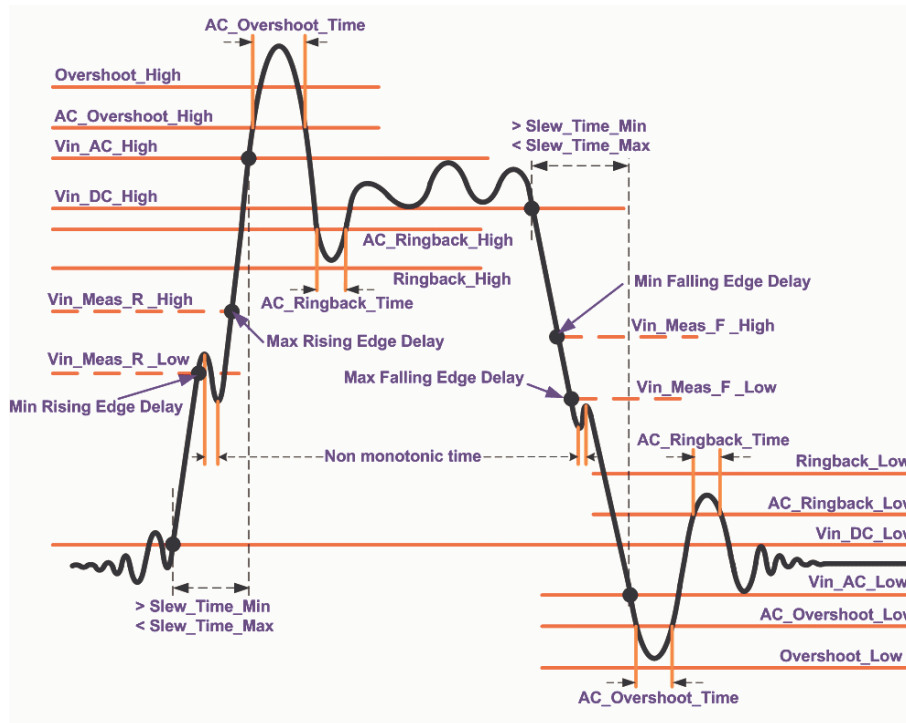


Figure 2 Waveform Processing Levels

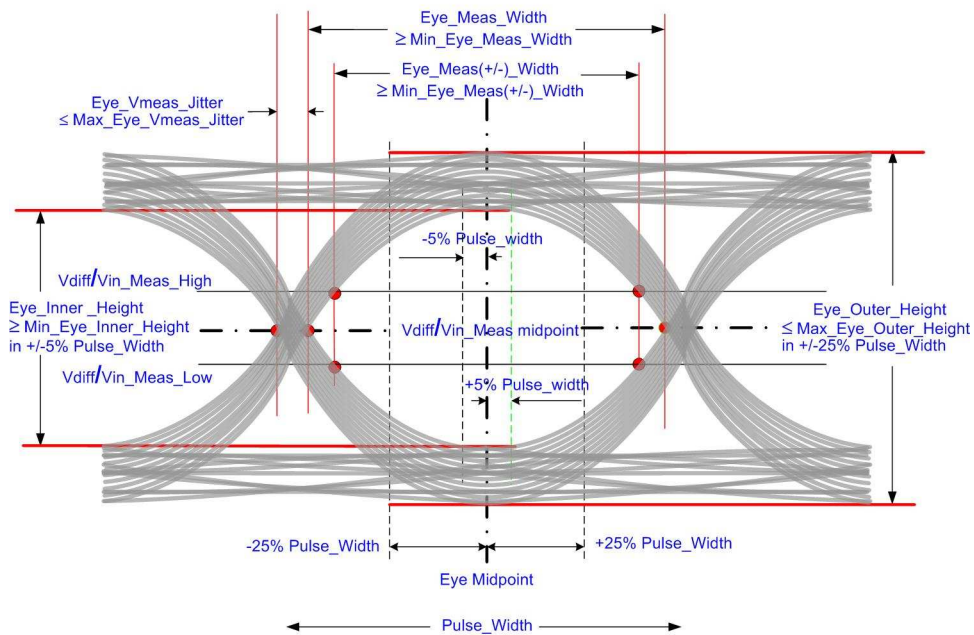


Figure 3 Eye Diagram Processing

QDR

The IXP2800 QDR interface is quite sophisticated. It can support up to 32 Mbytes with anywhere from one to four loads per interface at data rates up to 500MHz. The IXP2800 is frequency agile with programmable timing and drive strengths as well as programmable on-die termination for receiver cells. In addition, multiple clocking architectures are supported. This paper presents the challenges associated with a two-load 500MHz data-rate configuration with loop-back clocking. Figure 4 shows a block diagram of the interface.

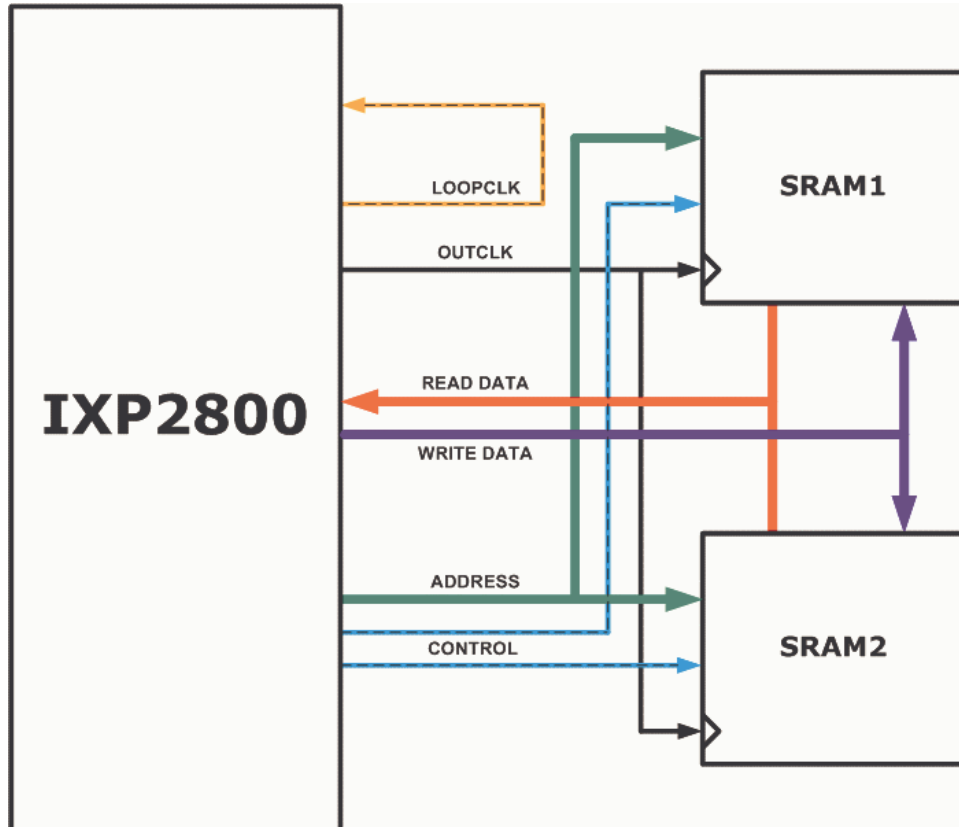


Figure 4 QDR Transfer Net Block Diagram

As implemented, this interface consists of six net-classes that are referred to as *Transfer Nets*. A Transfer Net is an implementation independent network description that defines high-level connectivity of components on a bus and valid transfers that can occur between these components. The transfer nets for this interface are as follows: Read_Data, Write_Data, Address, Control, Outclk, and Loopclk. As is shown in Figure 4, all paths are unidirectional. Outclk serves two purposes: it is the source-synchronous clock for Write_Data, Address, and Control; and it is also the synchronous clock that triggers the launching of the Read_Data from the QDR SRAM, which is then clocked in by Loopclk at the IXP2800 in a synchronous fashion. The major advantages of QDR are the independent read and write data paths, which enables simultaneous full-bandwidth utilization for read and write operation.

The critical design challenge associated with this interface was to define I/O drive strengths; IXP2800 programmable timing; PCB topology, termination, and interconnect lengths; and package tuning requirements to meet timing and signal quality constraints over a variety of SRAM configurations and performance targets.

QDR Timing

Overall timing margins on this interface are governed by the read and write operations. Figure 5 and Figure 6 show the write and read timing diagrams for the QDR interface. These diagrams are high-level representations of the uncertainties and timing constraints in the system. The uncertainties in the data valid windows are composed of uncertainty from AC specifications for the SRAM and IXP2800 components, package skew, and interconnect delay variations. For write operation, the IXP2800 bursts two words of data for every address and it clocks alternate data on the rising edges of Outclk_H and Outclk_L (these are complementary but not truly differential versions of Outclk). Alternate addresses are used for write and read operations so addresses switch as frequently as data and can be thought of as the same as data from a timing and simulation perspective. Control switches at one half the frequency of both Address and Write_Data and therefore does not dictate system performance.

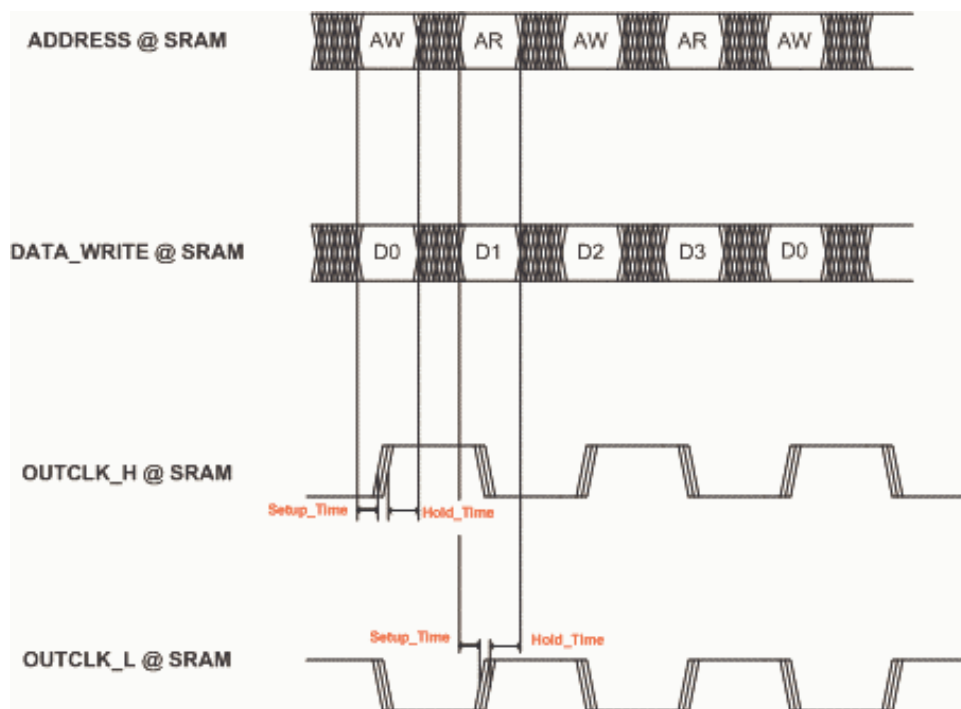


Figure 5 QDR Write Timing

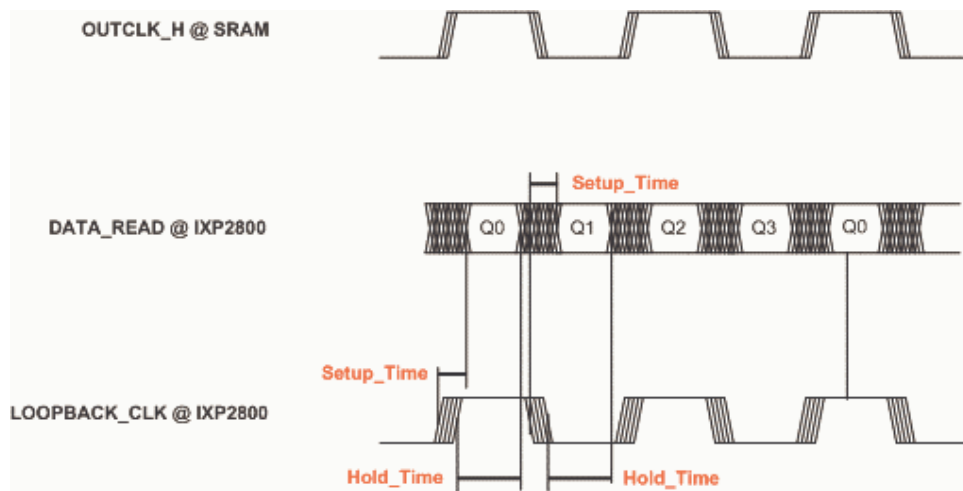


Figure 6 QDR Read Timing

Transfers from the IXP2800 to the SRAMs are source-synchronous with the timing margins dictated by Equation 1 and Equation 2. Conversely, transfers from the SRAM for read operations are synchronous between OUTCLK and LOOPCLK. Equation 3 and Equation 4 show the timing margin equations for the read operation. Table 1 shows the definitions for the parameters defined in these equations.

Equation 1 Write Setup Margin

$$\text{Write_Setup_Margin} = \text{Min_Outclk_Etch} - \text{Max_Data_Delay_Skew} - \text{Max_Data_Etch} - \text{SRAM_Setup_Time}$$

Equation 2 Write Hold Margin

$$\text{Write_Hold_Margin} = \text{Data_Rate} + \text{Min_Data_Etch} + \text{Min_Data_Delay_Skew} - \text{Max_Outclk_Etch} - \text{Cycle_to_Cycle_Jitter} - \text{SRAM_Hold_Time}$$

Equation 3 Read Setup Margin

$$\text{Read_Setup_Margin} = \text{Data_Rate} + \text{Min_Loopclk_etch} - \text{Out_to_Out_Skew} - \text{Cycle_to_Cycle_Jitter} - \text{Max_Outclk_Etch} - \text{Max_TCO} - \text{Max_Data_Etch} - \text{IXP2800_Setup_Time}$$

Equation 4 Read Hold Margin

$$\text{Read_Hold_Margin} = \text{Min_Outclk_Etch} + \text{Min_TCO} + \text{Min_Data_Etch} - \text{Max_Loopclk_Etch} - \text{Out_to_Out_Skew} - \text{IXP2800_Hold_Time}$$

Table 1 Definitions of Timing Margin Parameters

Data_Rate	Also commonly referred to as bit time. This is the time that corresponds to the maximum rate of switching.
Min_Outclk_Etch	Minimum interconnect delay relative to standard load for Outclk
Max_Outclk_Etch	Maximum interconnect delay relative to standard load for Outclk
Min_Data_Etch	Minimum interconnect delay relative to standard load for respective write or read operation
Max_Data_Etch	Maximum interconnect delay relative to standard load for respective write or read operation
Min_Loopclk_Etch	Minimum interconnect delay relative to standard load for Loopclk
Max_Loopclk_Etch	Maximum interconnect delay relative to standard load for Loopclk
Min_Data_Delay_Skew	This is the earliest Write_Data can occur at the IXP2800 relative to Outclk with both signals driving the standard load. This parameter is inclusive of package skew and programmable timing shifts. This typically contains the 90 degree phase shift of clock relative to data.
Max_Data_Delay_Skew	This is the latest Write_Data can occur at the IXP2800 relative to Outclk with both signals driving the standard load. This parameter is inclusive of package skew and programmable timing shifts. This typically contains the 90 degree phase shift of clock relative to data.
Min_TCO	Minimum clock-to-out time for the SRAM relative to standard load
Max_TCO	Maximum clock-to-out time for the SRAM relative to standard load
SRAM_Setup_Time	Amount of time data must be valid at SRAM prior to rising edge of OUTCLK_H or OUTCLK_L
SRAM_Hold_Time	Amount of time data must remain valid at SRAM after rising edge of OUTCLK_H or OUTCLK_L
IXP2800_Setup_Time	Amount of time data must be valid at IXP2800 prior to rising edge of Loopclk. This will contains a 90 degree programmable phase shift
IXP2800_Hold_Time	Amount of time data must be valid at IXP2800 after rising edge of Loopclk. This will contains a 90 degree programmable phase shift
Cycle_to_Cycle_Jitter	This represents the worst-case duty-cycle reduction that can occur between successive edges of a clock.
Out_to_Out_Skew	This is the amount of uncertainty of Outclk relative to Loopclk at the IXP2800 relative to standard load and inclusive of package skew.

Although the above equations correctly define timing margins in the system, careful consideration of correlative and un-correlative effects are required so as not to over-design the interface leaving critical performance on the table. A key point to note is that the above equations apply to simulations run at a given corner for a silicon process, etch process, voltage, and temperature. All corners must be run and analyzed independently. As an example, consider how Outclk and Loopclk drive affect read timing margins. Both of these networks are driving from the same chip and thus are in the same process corner (e.g. FF, TT, or SS). Combining the minimum and maximum interconnect delays across process for the read timing margin calculation would present an unrealistic and in some cases misleading uncertainty window. Due to compensation circuitry inside the I/O cells, it was not uncommon for the SS maximum delays to be less than the FF maximum delays.

Continuing with the Outclk/Loopclk example, etch properties are also correlative, provided networks are routed on same layers with comparable lengths. Etch is commonly known to have fast, typical, and slow characteristics (FE, TE, and SE). Typically, FE corresponds to higher characteristic impedance and faster propagation velocity and conversely SE corresponds to lower characteristic impedance and slower propagation velocity. The ranges of these characteristics are very much dependent on the material properties, temperature, stack-up construction, and manufacturing tolerances of the PCB. In many cases, etch correlation between traces used for clocks and data is not perfect. Often, networks must be partially routed on different layers or with different lengths. In these cases, uncertainty windows will increase as a function of the uncorrelated components of the network and must be accounted for. Finally, the etch process corners are independent for silicon process corners so all combinations silicon process and etch process must be considered over the entire solution space. Figure 7 below highlights the corner combinations run on this interface for post-layout analysis as shown in the SiAuditor GUI. Simulations should be run over full PVT, but independent analysis needs to be performed at each corner to quantify the setup and hold margins. Over all corners, the worst-case timing margins will dictate system performance.

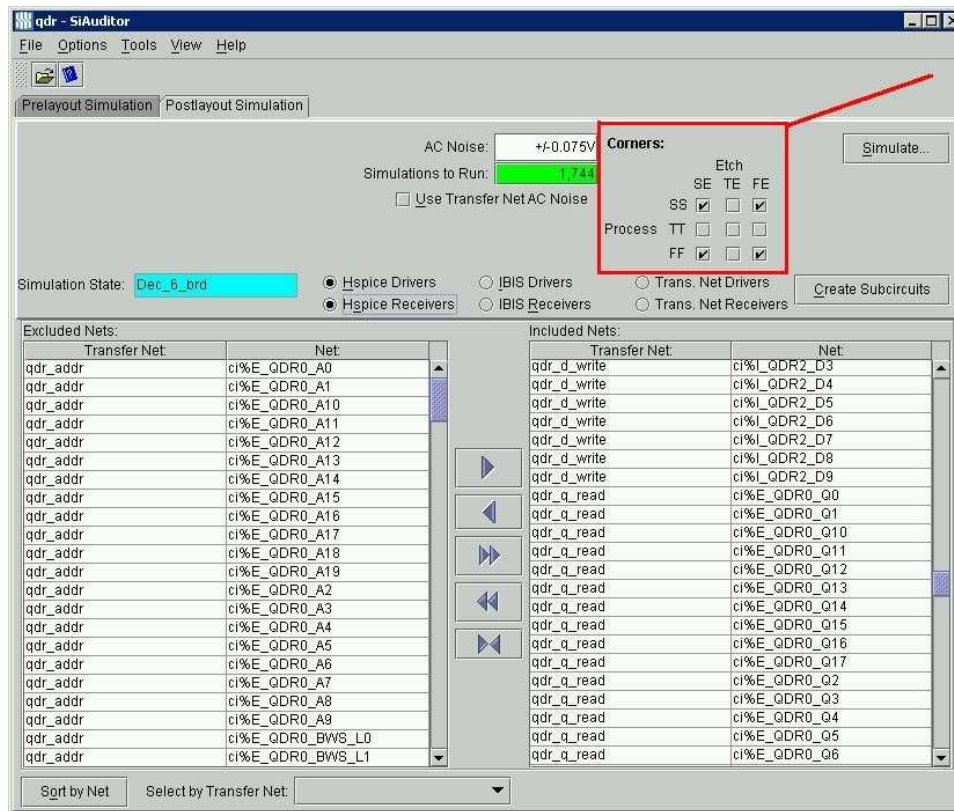


Figure 7 Corner combinations of Silicon and Etch Process

In determining timing margins, it was also required to run coupled simulations with complex stimulus patterns to excite inter-symbol interference (ISI) effects and include the effects of simultaneous switching output noise (SSO). Asynchronous noise of 75mV was incorporated to account for bus-to-bus asynchronous noise associated with simultaneous switching inputs (SSI), board crosstalk, and other system noise, which affects the threshold levels used in waveform analysis.

QDR Results

Solution space analysis was run over numerous topologies and length configurations for all transfer nets, configurations, and performance targets. The purpose of this was to prove operation, and drive layout constraints for both package and system groups as well as determine programmable timing and I/O strength selection. Post-layout signal integrity and timing verification was performed on actual system interconnect to verify system performance.

The QDR interface is an HSTL interface switching nominally between 0.0V and 1.65V. It is common to overdrive high-speed interfaces to get solid switching through the transition region to minimize interconnect skew resulting in slight overshoots at the receiver. This was problematic at the IXP2800 however, as this device could not tolerate significant overshoots due to device breakdown issues. Detailed specification of dynamic overshoot (see Figure 2) was required to properly constrain the system design. Another challenging aspect of this interface was getting incident wave switching on both Read_Data and Write_Data at 500MHz data rates for the four SRAM configuration. Figure 8 and Figure 9 show representative eye diagrams for write and read operation based on extracted post-layout simulations run using HSpice. As can be seen from these eye diagrams, acceptable signaling quality was achieved.

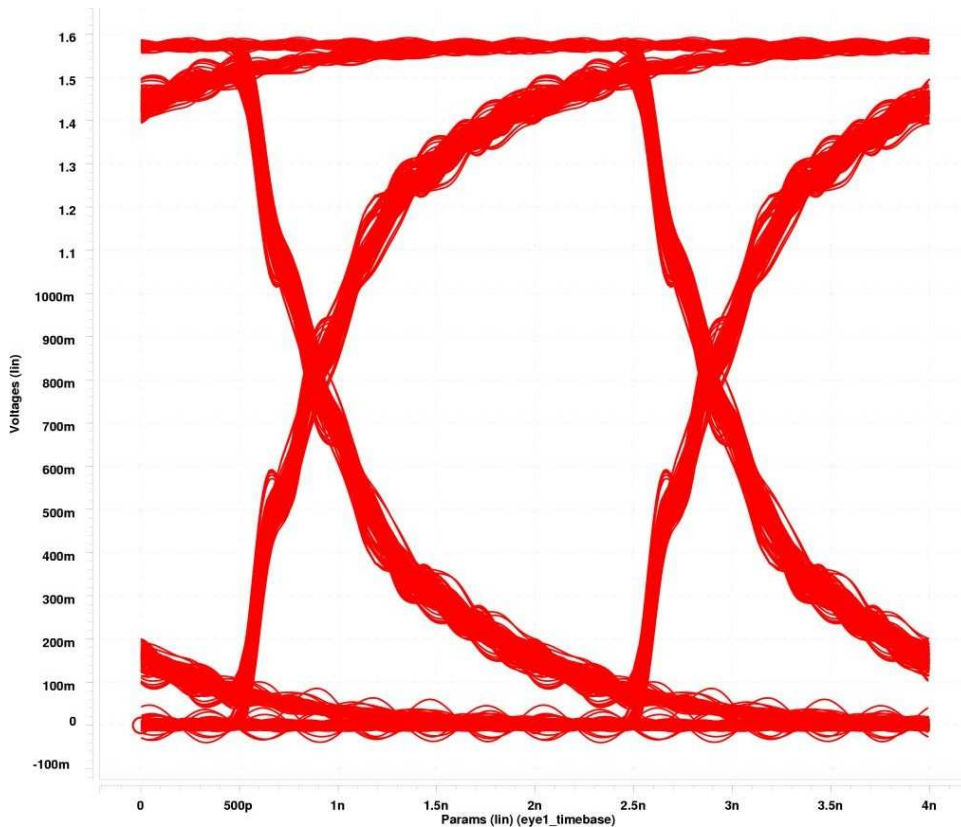


Figure 8 Data_Write Eye Diagram @ SRAM Pad

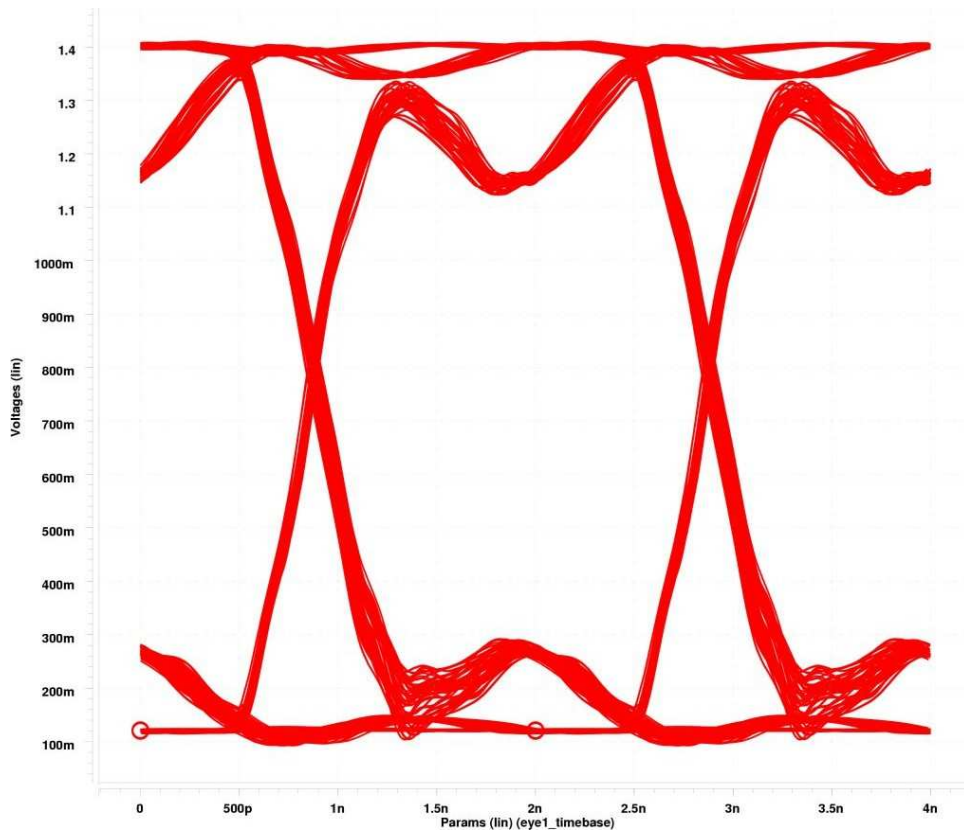


Figure 9 Data_Read Eye Diagram @ IXP2800 Pad

A key point to note is that for high-speed interfaces such the QDR with short edge rates as compared to the package routing delays, it is necessary to do all probing at either the Pad input or Core output of the receiver. This is due to the plateaus that will be seen at the pin, which are essentially a package round-trip in width.

Figure 10 shows post-layout timing results for an actual implementation of a two SRAM clam-shell configuration. The table shows the worst-case setup and hold margins for each transfer net taking into account the correlative and un-correlative affects discussed in previous sections of the paper.

	A	B	C	D	E
	Setup Margin (ns)	Hold Margin (ns)	Min Delay (ns)	Max Delay (ns)	Transfer Net
1	0.006	0.206	0.717	1.237	qdr_addr
2	0.37	2.132	0.555	0.881	qdr_ctl
3	0.008	0.23	0.698	1.215	qdr_d_write
4	0.018	0.068	0.596	0.898	qdr_q_read

Figure 10 QDR Timing Margins

An interesting observation that was made during the course of this analysis was that the IXP2800 setup time, hold time, and data output uncertainty windows were inversely proportional to frequency due to the use of the delay lock loops (DLLs) and other associated circuitry.

Another interesting observation was in the area of Read_Data timing. The AC timing specs of clock-to-out, setup, and hold for Read_Data are symmetric around the clock inputs to both the SRAM and IXP2800. Intuitively, a designer would expect to then match Loopclk and Outclk lengths to balance setup and hold timing margin. This was far from the case. Closer examination of the transfer nets showed the issue. Loopclk was a point to point net with the IXP2800 driving and IXP2800 receiving. Outclk was a three-drop net with the IXP2800 driving to two SRAM loads. Each SRAM load was equivalent to approximately two IXP2800 loads thus producing a loading imbalance between the two networks. Loopclk required additional interconnect to compensate for the loading imbalance.

SPI-4

The SPI-4 is a source-synchronous LVDS interface that operates up to a 1.25GHz data rate and serves as the primary interface to media and fabric devices. A particular challenging aspect of this interface was the requirement to achieve full-speed operation in a multi-board environment crossing two connectors. The IXP2800 has two modes of operation: static alignment and dynamic de-skew. Static alignment relies on precise alignment of PCB traces to allow capture of data, while dynamic de-skew uses per pin training cycles to eliminate pin-to-pin variations found in the package and PCB traces. One of the major benefits of dynamic de-skew is that it relieves etch matching requirement generally placed on source-synchronous interfaces. In the reference design, the SPI-4 interface path was up to 14 inches long. Simulations were run, apportioning this etch across all three boards. Figure 11 and Figure 12 shows the resulting eye diagram on the SPI-4 interface for an unmatched set of networks without and with dynamic de-skew.

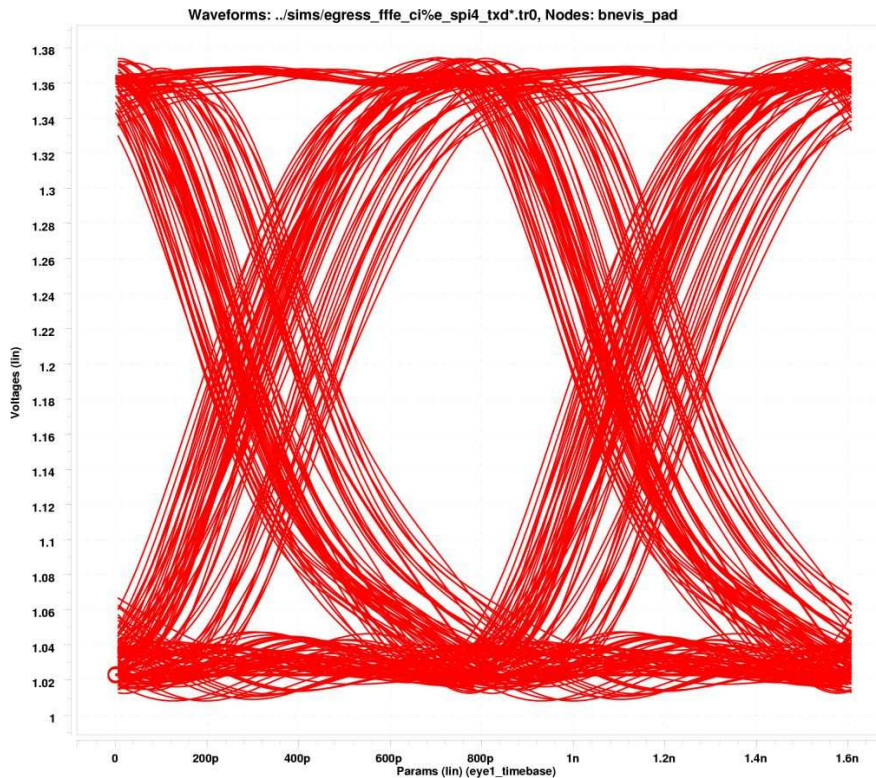


Figure 11 SPI-4 Unmatched Lengths Eye Diagram Before Dynamic De-Skew

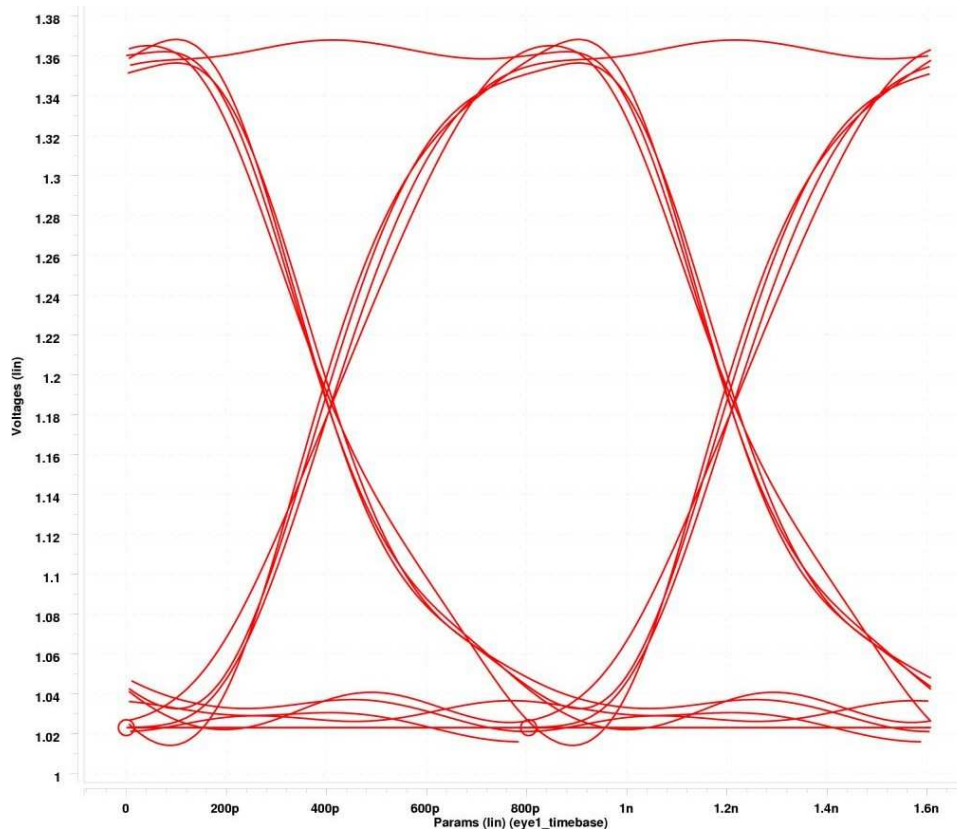


Figure 12 Unmatched Lengths After Dynamic De-Skew

As can be seen by the eye diagram, sufficient eye height and width were achieved in the final design. An interesting discovery was observed in a loop-back arrangement analysis; the length of etch on the loop-back card had a significant impact on the eye opening even when the total lengths across all boards remained constant. This effect was attributed to reflections due to impedance discontinuities and was resolved in actual implementation. In contrasting the analysis results between the static alignment and dynamic de-skew mode, it was found that etch tuning for the static alignment approach needed to be within 50 mils with matched via counts, via locations, and routing layers to achieve the full 1.25GHz data rate. In contrast, the dynamic de-skew approach allowed the design objectives to be met without the need to impose tight restrictions on the trace etch and vias.

Summary

High-speed design of multi-gigahertz systems requires careful consideration of signal integrity from design concept through implementation. This paper identified critical signal integrity and timing issues associated with high-speed interfaces such as those found on the IXP2800. A Core-to-Core design and analysis methodology was presented that addresses the high-speed design challenges associated with the QDR and SPI-4 interfaces in a complex multi-board configuration, including:

- Integrated signal integrity and timing analysis for both pre-layout and post-layout designs
- Use of mixed HSpice/IBIS models to achieve desired simulation accuracy and performance
- Rigorous waveform processing to extract interconnect delay, waveform quality, and eye diagram characteristics for every edge of every simulation
- Analysis of synchronous and source-synchronous timing in a loop-back clocking scheme