BladeServer

Base Specification and Design Guide for SERDES High-Speed Electrical Signaling

Document Version: 2.45 Document Date: May 12, 2010

Classified Document

The information in this document is IBM/Intel Confidential and is protected by a Confidential Disclosure Agreement when disclosed to other individuals of other corporations or entities.

Preface	2	4
1.1	Introduction	4
1.2	Document Control	4
1.3	Version Levels	4
1.4	Document Change History	4
1.5	Change Frequency	4
2 P	reface	5
3 D	Device (SERDES) Selection	6
3.1	1.25Gb/s Gigabit Ethernet Agents	7
3.2	2.125 Gb/s Fibre Channel Agents	8
3.3	2.5Gb/s InfiniBand TM Agents	9
3.	.3.1 InfiniBand TM Receiver Pin Spec Addendum.	9
3.4	Multi-Gigabit Repeaters	10
3.5	4.25 Gb/s Fibre Channel Tx/Rx Agents	12
3.6	4.25Gb/s Fibre Channel Repeater Agents on Switch	16
3.7	4.25Gb/s Fibre Channel Equalizer Agent on Daughter card	18
3.8	4x3.125Gb/s Ethernet Tx/Rx Agents	20
3.9	8.5Gb/s Fiber Channel Tx/Rx Agents	22
3.10	8.5Gb/s Fibre Channel Repeater Agents on Switch	25
3.11	8.5Gb/s Fibre Channel Equalizer Agent on Daughter card	27
3.12	2 4x10 Gb/s Infinaband Tx/Rx Agents	29
3.13	10 Gb/s KR Ethernet Tx/Rx Agents	31
3.14	Test Setups	32
3.15	Plug to Plug Channel Loss (SDD21) Guide for 1Gb/s to 8.5Gb/s	34
3.16	Plug to Plug Channel Loss (SDD21) Guide for 4x3.125Gb/s to 10Gb/s KR	35
Board ((Interconnect) Specification	
3.17	V System Compliance	37
3.18	S Specification Method	
3.	.18.1 Pass/Fail Threshold	
3.	.18.2 Trend Curve	
3.19	Interconnect Specification Data	40
4 B	Board Design Guidelines	42
4.1	Imperative Layout Rules	43
4.	.1.1 Route all high-speed switch fabric signals first!	43
4.	.1.2 Routing Length Guidelines and PWB Characteristics (except for 4.25Gb/s, 8.5Gb/s and	4x3.125
G	sb/s) 43	
4.	.1.3 Routing Length Guidelines and PWB Characteristics (4.25 Gb/s)	45
4.	.1.4 Routing Length Guidelines and PWB Characteristics (8.5 Gb/s)	46
4.	.1.5 Routing Length Guidelines and PWB Characteristics (3.125 Gb/s)	48
4.	.1.6 4 X 10 Gb/s Infinaband Routing Length Guidelines and PWB Characteristics	50
4.	.1.7 Routing Length Guidelines and PWB Characteristics (10Gb/s KR)	51
4.	.1.8 Clocks and Internal PLL's	54
4.	.1.9 Line Width vs. Impedance	54
4.	.1.10 Spacing	54
4.	.1.11 Guard Traces	54
4.	.1.12 Additional design guidelines for all 4 X 3.125 Gb/s interfaces	55
4.	.1.13 Additional Design Guideline for 8.5Gb/s Fiber Interface	57
4.	.1.14 Split Planes, Plane Referencing, Voids	60
4.	.1.15 Metal Fastener Spacing	63
4.	.1.16 Corners	64
4.	.1.17 Pin Grid Array Routing	64
4.	.1.18 Thermal Relief	66
4.	.1.19 Do not route TX or RX pairs parallel to a near edge of the board	66

BladeServer Bas	se Specification SERDES Design	12 May 2010
4.1.20	Vias, Anti-Vias	67
4.1.21	DC Blocking Caps	69
4.1.22	Capacitor Placement	69
4.1.23	Stubs and Test Pads	69
4.1.24	Power Delivery.	70
4.1.25	Termination schemes	70
4.1.26	Connectors	70
4.1.27	Net Naming Convention	
4.2 Margi	inal Layout Rules (Layout Suggestions)	
4.2.1 Sys	mmetrical Via "Route-By"	70
4.2.2 Bre	eakout (Length Matching)	71
4.2.3 Sys	mmetrical Route Around a Non-Signal Via	72
5 Appendix		73
5.1 Blade	e Configuration Diagrams	73
5.1.1 Bla	ade w/ Daughter Card and Repeater	73
5.1.2 Bla	ade w/ Repeater, but no Daughter Card	73
5.1.3 Bla	ade w/ no Daughter Card and no Repeater	74
5.1.4 Do	buble Wide Blade w/ Daughter Card and Repeater	74
5.2 Backp	plane (Mid-board) System Concept	75
5.3 Diele	ctric Weave-Effect Reduction	76
5.4 Extern	nal Termination Schemes	77
5.4.1 Ter	rminator Resistor to Ground - One simple example	77
5.4.2 Par	rallel Termination - One simple example	78

Preface

1.1 Introduction

This document describes the high speed signals that go through the BladeServer midplane.

1.2 Document Control

All approved levels are 1.x and higher. The document is only available in PDF format.

1.3 Version Levels

Version	Date	Reason
2.42	09/28/09	Update for QDR and KR info.
2.43	02/18/2010	Version level raised by one for synchronization with other architecture documents.

1.4 Document Change History

Document change history will be maintained for versions 1.x and greater.

Version	Date	Reason
1.03	08/19/2004	Approved version
1.05	09/01/2005	Added 4.25 Gb/s information
1.06	10/10/2005	Table 3.14 – RLTX-DIFF changed 18 to 15, and 2.125 to 2.5 GHz
		Table 3.15 – VTX-DIFF changed A to B, 400 to 50, removed 1000
		5.1.2 – changed 10 mils to 5, added statement about parasitic
		4.3 and 5.1 - Changed connector part numbers to ROHS parts
2.00	11/10/2005	Added 4x3.125 Gb/s information
2.21	09/01/2007	Added 8.5Gb/s design requirements
2.41	03/20/2009	Added 8.5Gb/s repeater design requirements
2.42	09/28/2009	Added QDR Infinaband and 10 Gb/s KR Ethernet
2.43	02/18/2010	Version level raised by one for synchronization with other architecture
		documents.
2.44	05/12/2010	Version skipped.
2.45	05/12/2010	Updated BOSSC Web site link.

1.5 Change Frequency

This document will be updated to reflect changes and updates that are approved by the joint Intel/IBM Collaboration Architecture Review Board.

2 Preface

This document is intended to support the design and specification of BladeServer Daughter Cards, Blades, and Switches. The goal is to provide guidelines and electrical specifications which can be used to design these plug-in modules without the need to perform simulation.

The following SERDES electrical switch fabrics are supported by this document:

- Gigabit Ethernet (GbE) 1.25 Gb/s
- Fibre Channel (FC) 2.125 Gb/s
- InfiniBand (IB) 2.5 Gb/s
- Fibre Channel (FC) 4.25 Gb/s (see Note 1 below)
- 10 Gigabit Ethernet 4x3.125 Gb/s
- Fiber Channel (FC) 8.5 Gb/
- QDR InfiniBand –4x10 Gb/s
- 10 Gigabit Ethernet (KR) --- 10Gb/s

It is understood that some designs will challenge/exceed the limits that are specified within these guidelines. In these cases, vendors must provide signal integrity simulation results for review. Both Enterprise and Telco midplane topologies must be accounted for if simulation is required. Simulations must encompass all pessimistic worst case scenarios and include the following:

- Vias, connectors, dielectric weave effects, 3-dimensional coupling.
- Trace impedance and loss for each signal trace, including X-talk to adjacent signals and ground path interruptions from pin-field routing
- All package variations including mounting pads
- Any additional auxiliary part and mounting variations (Ex/ DC Blocking Capacitor)

Data used for the design guide section is derived from HSPICE simulations, ANSOFT HFSS, and lab measurements. It is intended to be sufficient to guarantee operation. Further rule refinements are possible but must be validated with simulation and compliance measurements.

This document is divided into three sections: Device Selection, Interconnect Specification, and a Board Design Guidelines.

The intent of the Device Selection section is to enable the acquisition of SERDES parts that will operate in all cases when their carriers are employed in a BladeServer system. As a standalone document, the Device Selection enables device usage without requiring a BladeServer designer to validate the device's electrical performance. The Device Selection applies to all Blades, Switches, and Daughter Cards. The device needs to do auto-negotiation of the speed.

The Interconnect Specification is also a standalone document. It enumerates a very specific type of loss budget that the board (interconnect) must comply to. Each board type: Switch, Blade, and Daughter Card have their own specification.

The Design Guidelines provide recommendations and details of board design that will enable robust electrical operation of the BladeServer. Most of the guidelines are general and apply to all boards. However, there are sections that pertain to each board individually regarding line length and PWB stackup fabrication.

Note 1: The IBM HS-40 (machine type 8839) blade does not support 4.25Gb/s Fibre channel I/O adapters in I/O expansion slot-1.

3 Device (SERDES) Selection

The following are the electrical characteristics that must be used for fabric device selection and design.

Observe the following figures for reference. The figures display the Transmit and Receive pins voltage and timing characteristics as they are seen on an eye diagram.

Figure 3.1 - TX Eye Voltage and Timing



Figure 3.2 - RX Eye Voltage and Timing



3.1 1.25Gb/s Gigabit Ethernet Agents

Table 3.1 Differential TX	Coutput Coutput	Recommendations
---------------------------	-----------------	-----------------

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		800		ps	Each UI is 800ps +/- 100ppm	2
VTX-DIFFp-p	Differential Peak to Peak Output Voltage	800		1600	mV	VTX-DIFFp-p = 2* VTX-D+ - VTX-D-	1,5
TTx_RJ	Random Jitter			0.125	UI	Observed at BER 1E-12	1,3
TTx DJ	Deterministic Jitter			0.125	UI		1,3
VTX-CM-ACP	AC Peak Common Mode Output Voltage			25	mV	VTX-CM-ACp = RMS {[(VTX- D+) + (VTX-D-)] / 2 - VTXdcavg}	1
RLTX-DIFF	Differential Return Loss	12			dB	50MHz-625MHz	4
RLTX-COMM	Common Mode Return Loss	6			dB	50MHz-625MHz	4

Notes:

1 Measured with test load show in Figure 3.5

2 Measured over 3000 UI

3 BER 1E-12 is at +/-7 sigma

4 Measured at the device pins on each channel with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation

5 Normal operating voltage is 1200 mV. If the device had pre-emphasis setting, then set to medium level.

Table 3.2 Differential RX Input Recommendation	
--	--

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		800		ps	Each UI is 800ps +/- 100ppm	2
VRX-DIFFp-p	Differential Peak to Peak Input Voltage	200		1600	mV	VRX-DIFFp-p = 2* VRX-D+ - VRX-D-	1
TRX_RJ	Receiver Random Jitter			0.175	UI	Observed at BER 1E-12	1,3
TRX_DJ	Receiver Deterministic Jitter			0.175	UI		1,3
VRX-CM-Acp	AC Peak Common Mode Input Voltage			150	mV	VRX-CM-ACp = RMS {[(VRX- D+) + (VRX-D-)] / 2 - VRXdcavg}	
RLRX-DIFF	Differential Return Loss	12			dB	50MHz-625MHz	4
RLRX-COMM	Common Mode Return Loss	6			dB	50MHz-625MHz	4
Vsql	Squelch			75	mV	signals below this level are not received	1

Notes:

1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system voltage pin specification can be determined by multiplying the Rx spec by (1-2 ρ), where ρ is the reflection coefficient looking into the package and die combination. Return loss is 20*log($|\rho|$).

- 2 Measured over 3000 UI
- 3 BER 1E-12 is at +/-7 sigma
- 4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

3.2 2.125 Gb/s Fibre Channel Agents

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		470.6		ps	Each UI is 470.6ps +/- 100ppm	2
VTX-DIFFp-p	Differential Peak to Peak Output Voltage	800		1600	mV	VTX-DIFFp-p = 2* VTX-D+ - VTX-D-	1,6
VTX-DE- RATIO	De-Emphasized Differential Output Voltage (Ratio)	3		6	dB	This is the ratio of the VTX- DIFFp-p of the second and following bits after a transition divided by the VTX-DIFFp-p of the first bit after a transition.	1,5
TTx_RJ	Random Jitter			0.125	UI	Observed at BER 1E-12	1,3
TTx_DJ	Deterministic Jitter			0.125	UI		1,3
VTX-CM-ACP	AC Peak Common Mode Output Voltage			25	mV	VTX-CM-ACp = RMS {[(VTX- D+) + (VTX-D-)] / 2 - VTXdcavg}	
RLTX-DIFF	Differential Return Loss	12			dB	50MHz-1.0625GHz	4
RLTX-COMM	Common Mode Return Loss	6			dB	50MHz-1.0625GHz	4

Table 3.3 Differential TX Output Recommendations

Notes:

¹ Measured with test load show in Figure 3.5.

² Measured over 3000 UI

³ BER 1E-12 is at +/-7 sigma

4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA) (per channel). Meeting differential return loss does not guarantee operation

⁵ Do not use unique equalization based off of pre-determined blade ordering

6 Normal operating voltage is 1200 mV. If the device had pre-emphasis setting, then set to medium level.

Table 3.4 Differential	RX Inp	out Recommendation
-------------------------------	--------	--------------------

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		470.6		ps	Each UI is 470.6ps +/- 100ppm	2
VRX-DIFFp-p	Differential Peak to Peak Input Voltage	200		1600	mV	VRX-DIFFp-p = 2* VRX-D+ - VRX-D-	1
TRX_RJ	Receiver Random Jitter			0.175	UI	Observed at BER 1E-12	1,3
TRX_DJ	Receiver Deterministic Jitter			0.175	UI		1,3
VRX-CM-Acp	AC Peak Common Mode Input Voltage			150	mV	VRX-CM-ACp = RMS {[(VRX- D+) + (VRX-D-)] / 2 - VRXdcavg}	1
RLRX-DIFF	Differential Return Loss	12			dB	50MHz-1.0625GHz	4
RLRX-COMM	Common Mode Return Loss	6			dB	50MHz-1.0625GHz	4
Vsql	Squelch			75	mV	signals below this level are not received	1

Notes:

1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by (1-2 ρ), where ρ is the reflection coefficient looking into the package and die combination. Return loss is 20*log($|\rho|$)

² Measured over 3000 UI

³ BER 1E-12 is at +/-7 sigma

4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA) (per channel). Meeting differential return loss does not guarantee operation.

3.3 2.5Gb/s InfiniBand[™] Agents

The InfiniBand High-Speed Electrical Signaling Specification states that:

"Backplane connections **shall** use either passive equalization or driver preemphasis to compensate for the maximum loss as specified in (IB-Spec) Table 18. This equalization or pre-emphasis **shall** be provided on the InfiniBand module which contains the transmitter. Backplane connections which exceed the specified loss **shall** provide additional equalization on the backplane."

The unique architecture of a BladeServer environment requires Transmitter de-emphasis because there is no other form of equalization on the backplane (see below). All other electrical characteristics can be found in the InfiniBandTM Architecture Specification, Chapter 6.

Table 3.5 Differential TX Output Recommendation

Symbol	Parameter	Min	Nom	Мах	Units	Comments	Notes
UI	Unit Interval		400		ps	Each UI is 400ps +/- 100ppm	2
VTX-DE- RATIO	De-Emphasized Differential Output Voltage (Ratio)	3		6	dB	This is the ratio of the VTX- DIFFp-p of the second and following bits after a transition divided by the VTX-DIFFp-p of the first bit after a transition.	1,3

Notes:

- 1 Measured with test load show in Figure 3.5.
- ² Measured over 3000 UI
- ³ Do not use unique equalization based off of pre-determined blade ordering

3.3.1 InfiniBand[™] Receiver Pin Spec Addendum.

The in situ receiver eye pin specification can be determined by multiplying the Rx eye spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20*\log(|\rho|)$

9

3.4 Multi-Gigabit Repeaters

Table 3.6 Differential TX C	Dutput Recommendation
-----------------------------	------------------------------

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval	400		800	ps	UI must be 100ppm. Below 400ps is acceptable, but not required.	2
VTX-DIFFp-p	Differential Peak to Peak Output Voltage	800		1600	mV	VTX-DIFFp-p = 2* VTX-D+ - VTX-D-	1,6
VTX-DE- RATIO	De-Emphasized Differential Output Voltage (Ratio)	3		5	dB	This is the ratio of the VTX- DIFFp-p of the second and following bits after a transition divided by the VTX-DIFFp-p of the first bit after a transition.	1
TTx_RJ	Random Jitter			0.125	UI	Observed at BER 1E-12	1,3
TTx_DJ	Deterministic Jitter			0.125	UI		1,3
VTX-CM-ACP	AC Peak Common Mode Output Voltage			25	mV	VTX-CM-ACp = RMS {[(VTX- D+) + (VTX-D-)] / 2 - VTXdcavg}	
RLTX-DIFF	Differential Return Loss	12			dB	50MHz-625MHz	4
RLTX-COMM	Common Mode Return Loss	6			dB	50MHz-625MHz	4
PSRR	Power supply rejection ration			-20	dB	This is the dB ratio of respective differential noise component on the TX pair output to a AC signal on the power supply between 50MHZ and 1.25GHz	5

Notes:

1 Measured with test load show in Figure 3.5. Jitter specification assumes 90% deterministic jitter of incoming signal is removed by receiver of repeater

² Measured over 3000 UI

³ BER 1E-12 is at +/-7 sigma

4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

5 See Figure 3. to measure the chip's ability to block power supply noise over the voltage range of the repeater device.

6 Normal operating voltage is 1200 mV. If the device had pre-emphasis setting, then set to medium level.

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
•						Repeater must be at least	
UI	Unit Interval			470.6	ps	2.5Gb/s at 100ppm"	1
	Differential Peak to					VRX-DIFFp-p = 2* VRX-D+ -	
VRX-DIFFp-p	Peak Input Voltage	200		1600	mV	VRX-D-	1
						The maximum Interconnect and	
						Transmitter jitter that can be	
						tolerated by the Receiver can be	
	Minimum RX					derived as TRXMAX- JITTER =	
TRX-EYE	Eye Width	0.35			UI	1 - TRX-EYE= .65 UI	1
	Amount of channel						
	jitter eliminated						
TJTT_BLOCK	from Rx to Tx	90%			%	This blocks jitter amplification.	6
	AC Peak Common					VRX-CM-ACp = RMS {[(VRX-	
VRX-CM-Acp	Mode Input Voltage			150	mV	D+) + (VRX-D-)] / 2 - VRXdcavg}	1
	Differential Return						
RLRX-DIFF	Loss	12			dB	50MHz-1.25GHz	3
	Common Mode						
RLRX-COMM	Return Loss	6			dB	50MHz-1.25GHz	3
						signals below this level are not	
Vsql	Squelch			75	mV	received	1
						Output disabled if VRX-DIFFp-p	
LOS	Loss of Signal			100	mV	is in LOS range specified	1,5
						This is the dB ratio of respective	
						differential noise component on	
						the TX pair output to a AC signal	
	Power supply					on the power supply between	
PSRR	rejection ration			-20	dB	50MHZ an 1.25GHz	7

Table 3.7 Differential RX Input Recommendation

Notes:

1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20^{*}\log(|\rho|)$.

² Measured over 3000 UI

3 Measured with at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

⁴ This is the differential voltage at which the output will be disabled

⁵ If any in band signal goes below a settable range, the device must provide the capability to disable outputs. This can be implemented either internally (desirable) or externally.

- 6 Repeater jitter blocking is validated with setup as shown in Figure 3. It is expect that retiming or receiver FIR equalization may be employed. However, repeaters must work the entire range of data rates.
- ⁷ See Figure 3.8 to measure the chip's ability to block power supply noise.

3.5 4.25 Gb/s Fibre Channel Tx/Rx Agents

All 4.25Gb/s Daughter Card and Switches require 850pf series DC blocking capacitors on transmit and receive lines. DC blocking requires a broad range of low loss, resonant free, frequency coverage. The upper limit of bandwidth over which insertion loss meets specification is determined by the location of parallel resonances. For maximum resonant-free bandwidth, custom broadband blocks such as the DEL C06BLBB2X5UX family components are recommended. Avoid placement in the middle of the trace route.

Table 3.8 Differential TX Output Recommendations

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		235.3		ps	Each UI is 235.3ps +/- 100ppm	2
VTX-DIFFp-p	Differential Peak to Peak Output Voltage	850		1600	mV	VTX-DIFFp-p = 2* VTX-D+ - VTX-D-	1
TTx_RJ	Random Jitter			0.125	UI	Observed at BER 1E-12	1,3
TTx_DJ	Deterministic Jitter			0.125	UI		1,3
VTX-CM-ACP	AC Peak Common Mode Output Voltage			25	mV	VTX-CM-ACp = RMS {[(VTX- D+) + (VTX-D-)] / 2 - VTXdcavg}	1
RLTX-DIFF	Differential Return Loss	12			dB	50MHz-2.125GHz	4
RLTX-COMM	Common Mode Return Loss	6			dB	50MHz-2.125GHz	4
De-emphasis	Post-cursor tap	5		6		Pre-cursor tap coefficient	5

Notes:

1 Measured with test load show in Figure 3.5.

2 Measured over 3000 UI

3 BER 1E-12 is at +/-7 sigma

4 Measured at the device pins on each channel with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation but it is necessary not to exceed this specification.

⁵ 2 Tap FIR filter only on Daughter card device. **DO NOT turn on PRE-EMPHASIS or DE-EMPHASIS setting for switch output for 4.25Gb/s interface**

BladeServer Base Specification

Table 3.9 Transmitter Jitter Specifications

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		235.3		ps	Repeater must be at least 4.25Gb/s at 100ppm"	1
DJ	Deterministic Jitter			0.125	UI		1
TJ	Total Jitter			0.33	UI	The maximum Interconnect and Transmitter jitter that can be tolerated by the Receiver can be derived as TRXMAX- JITTER = 1 - TRX-EYE= .57 UI	1
Eye Mask	X1					1/2 the total Jitter.	
	X2			.4	UI	X1+0.19	
	A	425			mV		
	В			800	mV		

Notes:

1

Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20^{10}(|\rho|)$.

Figure 3.3 - Jitter compliance eye mask at 4.25Gb/s for transmitter



BladeServer Base Specification

SERDES Design

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		235.3		ps	Each UI is 235.3ps +/- 100ppm	2
VRX-DIFFp-p	Differential Peak to Peak Input Voltage	200		1600	mV	VRX-DIFFp-p = 2* VRX-D+ - VRX-D-	1
TRX_RJ	Receiver Random Jitter			0.175	UI	Observed at BER 1E-12	1,3
TRX_DJ	Receiver Deterministic Jitter			0.175	UI		1,3
VRX-CM-Acp	AC Peak Common Mode Input Voltage			40	mV	VRX-CM-ACp = RMS {[(VRX- D+) + (VRX-D-)] / 2 - VRXdcavg}	
RLRX-DIFF	Differential Return Loss	12			dB	50MHz-2.125GHz	4
RLRX-COMM	Common Mode Return Loss	6			dB	50MHz-2.125GHz	4
Vsql	Squelch			75	mV	signals below this level are not received	1

Table 3.10 Differential RX Input Recommendation

Notes:

1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system voltage pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20*\log(|\rho|)$. Conversely the design target at the die would be $(1+2\rho)$ time the Rx spec.

2 Measured over 3000 UI

3 BER 1E-12 is at +/-7 sigma

4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

BladeServer Base Specification

SERDES Design

12 May 2010

The Fibre Channel Jitter specification at reference Beta t (receiver pin) is shown in Table 3.11 and the compliance mask is shown in Figure 3.4. Total Jitter (TJ) consists of Random jitter (RJ), Duty Cycle Distortion(DCD), Period jitter(PJ), and Inter symbol Interference(ISI)

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval	235.3	470.6	941.2	ps	Repeater must be at least 2.5Gb/s at 100ppm"	1
DJ	Deterministic Jitter			0.33	UI		1
TJ	Total Jitter			0.52	UI	The maximum Interconnect and Transmitter jitter that can be tolerated by the Receiver can be derived as TRXMAX- JITTER = 1 - TRX-EYE= .48 UI	1
Eye Mask	X1					1/2 the total Jitter.	
	X2			.4	UI	X1+0.19	
	A	100			mV		
	В			800	mV		

Table 3.11 Receiver input Jitter Specifications

Notes:

1

Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20^{*}\log(|\rho|)$.

Figure 3.4 - Jitter compliance eye mask at 4.25Gb/s for Receiver



3.6 4.25Gb/s Fibre Channel Repeater Agents on Switch

Repeaters will be required if the Tx or Rx agents cannot meet the prior Tx/Rx specifications. Here is the topology which shows the placement of Repeater when Daughter card is driving and Switch is receiving.



Table 3.12 Repeater: Differential TX Output Recommendation

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		235.3		ps	UI must be 100ppm. Below 400ps is acceptable, but not required.	2
VTX-DIFFp-p	Differential Peak to Peak Output Voltage	800		1250	mV	VTX-DIFFp-p = 2* VTX-D+ - VTX-D-	1
Gb/s	Rate Supported		1.0625 2.1250 4.2500		Gb/s		
TTx_RJ	Random Jitter			0.125	UI	Observed at BER 1E-12	1,3
TTx_DJ	Deterministic Jitter			0.125	UI		1,3
VTX-CM-ACP	AC Peak Common Mode Output Voltage			25	mV	VTX-CM-ACp = RMS {[(VTX- D+) + (VTX-D-)] / 2 - VTXdcavg}	
RLTX-DIFF	Differential Return Loss	12			dB	50MHz-2.125GHz	4
RLTX-COMM	Common Mode Return Loss	6			dB	50MHz-2.125GHz	4
5055	Power supply					This is the dB ratio of respective differential noise component on the TX pair output to a AC signal on the power supply between	_
PSRR	rejection ration			-20	dB	50MHZ an 2.125GHz	5

Notes:

1 Measured with test load show in Figure 3.5. Jitter specification assumes 90% deterministic jitter of incoming signal is removed by receiver of repeater

² Measured over 3000 UI

³ BER 1E-12 is at +/-7 sigma

4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

5 See Figure 3.8 to measure the chip's ability to block power supply noise over the voltage range of the repeater device.

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
						Repeater must be at least	
UI	Unit Interval		235.3		ps	4.25Gb/s at 100ppm"	1
	Differential Peak to					VRX-DIFFp-p = 2* VRX-D+ -	
VRX-DIFFp-p	Peak Input Voltage	100		1700	mV	VRX-D-	1
						The maximum Interconnect and	
						Transmitter jitter that can be	
						tolerated by the Receiver can be	
	Minimum RX					derived as TRXMAX- JITTER =	
TRX-EYE	Eye Width	0.35			UI	1 - TRX-EYE= .65 UI	1
	Amount of channel						
	jitter eliminated						
TJTT_BLOCK	from Rx to Tx	90%			%	This blocks jitter amplification.	6
	AC Peak Common					VRX-CM-ACp = RMS {[(VRX-	
VRX-CM-Acp	Mode Input Voltage			150	mV	D+) + (VRX-D-)] / 2 - VRXdcavg}	1
	Differential Return						
RLRX-DIFF	Loss	12			dB	50MHz-2.125GHz	3
	Common Mode						
RLRX-COMM	Return Loss	6			dB	50MHz-2.125GHz	3
						signals below this level are not	
Vsql	Squelch			75	mV	received	1
						Output disabled if VRX-DIFFp-p	
LOS	Loss of Signal			100	mV	is in LOS range specified	1,5
						This is the dB ratio of respective	
						differential noise component on	
						the TX pair output to a AC signal	
	Power supply					on the power supply between	
PSRR	rejection ration			-20	dB	50MHZ an 2.125GHz	7

Table 3.13 Repeater: Differential RX Input Recommendation

Notes:

1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20^{*}\log(|\rho|)$.

² Measured over 3000 UI

3 Measured with at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

⁴ This is the differential voltage at which the output will be disabled

⁵ If any in band signal goes below a settable range, the device must provide the capability to disable outputs. This can be implemented either internally (desirable) or externally.

- 6 Repeater jitter blocking is validated with setup as shown in Figure 3.9. It is expect that retiming or receiver FIR equalization may be employed. However, repeaters must work for the entire range of data rates.
- ⁷ See Figure 3.8 to measure the chip's ability to block power supply noise.

3.7 4.25Gb/s Fibre Channel Equalizer Agent on Daughter card

Repeaters will be required if the Tx or Rx agents cannot meet the prior Tx/Rx specifications. Here is the topology which shows the placement of Equalizer when Switch is driving and Daughter card is receiving.



30 in. Tline with 3 connectors



Table 3.14 Equalizer: Differential TX Output Recommendation.

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		235.3		ps	UI must be 100ppm. Below 400ps is acceptable, but not required.	2
VTX-DIFFp-p	Differential Peak to Peak Output Voltage at C	550		750	mV	VTX-DIFFp-p = 2* VTX-D+ - VTX-D-	1
Gb/s	Rate supported		1.0625 2.1250 4.2500		Gb/s		
TTx_RJ	Random Jitter			0.125	UI	Observed at BER 1E-12	1,3
TTx_DJ	Deterministic Jitter			0.125	UI		1,3
VTX-CM-ACP	AC Peak Common Mode Output Voltage			25	mV	VTX-CM-ACp = RMS {[(VTX- D+) + (VTX-D-)] / 2 - VTXdcavg}	
RLTX-DIFF	Differential Return Loss	15			dB	50MHz-2.5GHz	4
RLTX-COMM	Common Mode Return Loss	9			dB	50MHz-2.125GHz	4
	Power supply					This is the dB ratio of respective differential noise component on the TX pair output to a AC signal on the power supply between	
PSRR	rejection ration			-20	dB	50MHZ an 2.125GHz	5

Notes:

1 Measured with test load show in Figure 3.5. Jitter specification assumes 90% deterministic jitter of incoming signal is removed by receiver of repeater

² Measured over 3000 UI

³ BER 1E-12 is at +/-7 sigma

4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

5 See Figure 3.8 to measure the chip's ability to block power supply noise over the voltage range of the repeater device.

SERDES Design

Table 3.15 Equalizer	: Differential RX In	put Recommendation
----------------------	----------------------	--------------------

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		235.3		ps	Repeater must be at least 2.5Gb/s at 100ppm"	1
VTX-DIFFp-p	Differential Peak to Peak Input Voltage at B	50			mV	VRX-DIFFp-p = 2* VRX-D+ - VRX-D-	1
TRX-EYE	Minimum RX Eye Width	0.35			UI	The maximum Interconnect and Transmitter jitter that can be tolerated by the Receiver can be derived as TRXMAX- JITTER = 1 - TRX-EYE= .65 UI	1
TJTT_BLOCK	Amount of channel jitter eliminated from Rx to Tx	90%			%	This blocks jitter amplification.	4
RLRX-DIFF	Differential Return Loss	15			dB	50MHz-2.125GHz	3
RLRX-COMM	Common Mode Return Loss	7			dB	50MHz-2.125GHz	3

Notes:

1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20^{*}\log(|\rho|)$.

² Measured over 3000 UI

3 Measured with at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

4 Repeater jitter blocking is validated with setup as shown in Figure 3.9. It is expect that retiming or receiver FIR equalization may be employed. However, repeaters must work the entire range of data rates.

3.8 4x3.125Gb/s Ethernet Tx/Rx Agents

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		320		ps	Each UI is 320ps +/- 100ppm	2
VTX-DIFFp-p	Differential Peak to Peak Output Voltage	800		1600	mV	VTX-DIFFp-p = 2* VTX-D+ - VTX-D-	1
TTx_RJ	Random Jitter			0.27	UI	Observed at BER 1E-12	1,3
TTx_DJ	Deterministic Jitter			0.17	UI		1,3
VTX-CM-ACP	AC Peak Common Mode Output Voltage			25	mV	VTX-CM-ACp = RMS {[(VTX- D+) + (VTX-D-)] / 2 - VTXdcavg}	1
RLTX-DIFF	Differential Return Loss	10			dB	50MHz-1500MHz	4
RLTX-COMM	Common Mode Return Loss	6			dB	50MHz-1500MHz	4
De-emphasis	Post-cursor tap	5		6		Pre-cursor tap coefficient	5
Mataa							

Table 3.16 Differential TX Output Recommendations

Notes:

1 Measured with test load show in Figure 3.5

2 Measured over 3000 UI

3 BER 1E-12 is at +/-7 sigma

4 Measured at the device pins on each channel with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation but it is necessary not to exceed this specification.

⁵ 2 Tap FIR filter

Table 3.17 Transmitter Jitter Specifications

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		320		ps	Repeater must be at least 3.125Gb/s at 100ppm"	1
DJ	BladeCenter Deterministic Jitter			0.17	UI		1
TJ	BladeCenter Total Jitter			0.35	UI	The maximum Interconnect and Transmitter jitter that can be tolerated by the Receiver can be derived as TRXMAX- JITTER = 1 - TRX-EYE= .65 UI	1
Eye Mask	X1					1/2 the total Jitter.	2
	X2			.4	UI	X1+0.19	2
	А	400			mV		2
	В			800	mV		2

Notes:

1 Measured with test load as shown in Figure 3.5.

2 See figure 3.3

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		320		ps	Each UI is 320ps +/- 100ppm	2
VRX-DIFFp-p	Differential Peak to Peak Input Voltage	200		1600	mV	VRX-DIFFp-p = 2* VRX-D+ - VRX-D-	1
TRX_RJ	Receiver Random Jitter			0.125	UI	Observed at BER 1E-12	1,3
TRX_DJ	Receiver Deterministic Jitter			0.125	UI		1,3
VRX-CM-Acp	AC Peak Common Mode Input Voltage			40	mV	VRX-CM-ACp = RMS {[(VRX- D+) + (VRX-D-)] / 2 - VRXdcavg}	
RLRX-DIFF	Differential Return Loss	12			dB	50MHz-1500MHz	4
RLRX-COMM	Common Mode Return Loss	6			dB	50MHz-1500MHz	4
Vsql	Squelch			75	mV	signals below this level are not received	1

Table 3.18 Differential Rx Input Recommendations

Notes:

- 1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system voltage pin specification can be determined by multiplying the Rx spec by (1- 2ρ), where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20^{10}(|\rho|)$. Conversely the design target at the die would be $(1+2\rho)$ time the Rx spec.
- 2 Measured over 3000 UI
- 3 BER 1E-12 is at +/-7 sigma
- 4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

Table 3.19 Receiver Input Jitter Specifications

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		320		ps	Repeater must be at least 3.125Gb/s at 100ppm"	1
DJ	BladeCenter Deterministic Jitter			0.17	UI		1
TJ	BladeCenter Total Jitter			0.50	UI	The maximum Interconnect and Transmitter jitter that can be tolerated by the Receiver can be derived as TRXMAX- JITTER = 1 - TRX-EYE= .50 UI	1
Eye Mask	X1					1/2 the total Jitter.	2
	X2			.4	UI	X1+0.19	2
	A			100	mV		2
	В			800	mV		2

Notes:

Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20^{10}(|\rho|)$.

2 See figure 3.4

1

3.9 8.5Gb/s Fiber Channel Tx/Rx Agents

The Fiber channel device should meet following criteria to have a successful operation in BladeCenter environment.

- 1- FFE: 10 taps with T/2 spacing
- 2- DFE: 5 taps with T spacing.
- 3- Supports 8 FC auto speed negotiation requirements.

4- Adaptive convergence algorithm that optimizes the link performance to be lower than 1E-15 bit error rate.

If the device meet above criteria, then look at section 5.1.4 to implement the OSIBA network on the switch design.

If the device does not meet above criteria, then the HBA or Switch will be required to have an external equalizer.

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
						Each UI is 118ps +/-	
UI	Unit Interval		118		ps	100ppm	2
	Differential						
VTX-	Peak to Peak					VTX-DIFFp-p = 2* VTX-	
DIFFp-p	Output Voltage	800		1200	mV	D+ - VTX-D-	1
	Random Jitter			0.13	UI	Observed at BER 1E-15	1,3
	Deterministic						
	Jitter			0.17	UI		1,3
	AC Peak					VTX-CM-ACp = RMS	
VTX-CM-	Common Mode					{[(VTX-D+) + (VTX-D-)] /	
ACP	Output Voltage			15	mV	2 - VTXdcavg}	1
	Differential						
RLTX-DIFF	Return Loss	10			dB	50MHz-4250MHz	4
RLTX-	Common Mode						
COMM	Return Loss	6			dB	50MHz-4250MHz	4
De-							
emphasis	Post-cursor tap	5				Pre-cursor tap coefficient	5

Table 3.20 Differential TX Output Recommendations

Notes:

1 Measured with test load show in Figure 3.5

2 Measured over 3000 UI

3 BER 1E-12 is at +/-7 sigma

4 Measured at the device pins on each channel with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation but it is necessary not to exceed this specification.

⁵ 3 Tap FIR filter

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		118		ps	Repeater must be at least 8.5Gb/s at 100ppm"	1
Rise / Fall time 20-80		40ps		N/A	Ps		1

Table 3.21 Transmitter Jitter Specifications

BladeServer	Base Specification	l	SEI	RDES I	Design	12
%						
DJ	BladeCenter Deterministic Jitter		0.17	UI		1
TJ	BladeCenter Total Jitter		0.30	UI	The maximum Interconnect and Transmitter jitter that can be tolerated by the Receiver can be derived as TRXMAX- JITTER = 1 - TRX-EYE= .65 UI	1
Eye Mask	X1				¹ / ₂ the total Jitter.	2
	X2		.4	UI	X1+0.19	2
	Α	400		mV		2
	В		600	mV		2
Notes:						

1 Measured with test load as shown in Figure 3.5.

2 See figure 3.3

12 May 2010

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		118		ps	Each UI is 118ps +/- 100ppm	2
VMA	Differential Peak to Peak Input Voltage	300	470	540	mV		1
TRX_RJ	Receiver Random Jitter			0.25	UI	Observed at BER 1E-15	1,3
TRX_DJ	Receiver Deterministic Jitter			0.47	UI		1,3
VRX-CM- Acp	AC Peak Common Mode Input Voltage			40	mV	VRX-CM-ACp = RMS {[(VRX-D+) + (VRX-D-)] / 2 - VRXdcavg}	
RLRX- DIFF	Differential Return Loss	12			dB	50MHz-4250MHz	4
RLRX- COMM	Common Mode Return Loss	6			dB	50MHz-4250MHz	4
	Bit Error ratio		10Ē- 12				
Notes:							

Table 3.22 Differential Rx Input Recommendations

1 Voltage Modulation Amplitude is measured at the input to the receiver device under test using the procedure defined in Annex A.8 FC-PI-4. Receiver Jitter tolerance testing using minimum VMA and maximum jitter should be used. At BER 10E-12

2 Measured over 3000 UI

3 BER 1E-12 is at +/-7 sigma

4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

Table 3.23	Receiver	Input Jitter	Specifications
-------------------	----------	---------------------	----------------

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		116		ps	Repeater must be at least 8.5Gb/s at 100ppm"	1
DJ	BladeCenter Deterministic Jitter			0.17	UI		1
TJ	BladeCenter Total Jitter			0.50	UI	The maximum Interconnect and Transmitter jitter that can be tolerated by the Receiver can be derived as TRXMAX- JITTER = 1 - TRX-EYE= .50 UI	1
Eye Mask	X1					1/2 the total Jitter.	2
	X2			.5	UI	X1+0.19	2
	A			175	mV		2
	В			425	mV		2

Notes:

1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by (1-2ρ), where ρ is the reflection coefficient looking into the package and die combination. Return loss is 20*log(|ρ|).

2 See figure 3.4

3.10 8.5Gb/s Fibre Channel Repeater Agents on Switch

Repeaters will be required if the Tx or Rx agents cannot meet the prior Tx/Rx specifications. Here is the topology which shows the placement of Repeater when Daughter card is driving and Switch is receiving.



The repeater or equalizer should meet following criteria:

- 1- FFE: 10 taps with T/2 spacing
- 2- DFE: 4 taps with T spacing.
- 3- Bypass mode for 4G and 2G operation.
- 4- Supports 8 FC auto speed negotiation requirements.
- 5- Adaptive convergence algorithm that optimizes the link performance to be lower than 1E-15 bit error rate.

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		118		ps	UI must be 100ppm. Below 200ps is acceptable, but not required.	2
VTX-DIFFp-p	Differential Peak to Peak Output Voltage	350	650	850	mV	VTX-DIFFp-p = 2* VTX-D+ - VTX-D-	1
Gb/s	Rate Supported		2.125 4.250 8.500		Gb/s		
TTx_RJ	Random Jitter			0.35	UI	Observed at BER 1E-15	1,3
TTx_DJ	Deterministic Jitter			0.20	UI		1,3
VTX-CM-ACP	AC Peak Common Mode Output Voltage			15	mV	VTX-CM-ACp = RMS {[(VTX- D+) + (VTX-D-)] / 2 - VTXdcavg}	
RLTX-DIFF	Differential Return Loss			-12	dB	50MHz-4.25GHz	4
RLTX-COMM	Common Mode Return Loss		-8		dB	50MHz-4.25GHz	4
PSRR	Power supply rejection ration					Measured PSRR per XFP standards	6

Table 3.24 Repeater: Differential TX Output "C" Recommendation

Notes:

1 Measured with test load show in Figure 3.5. Jitter specification assumes 90% deterministic jitter of incoming signal is removed by receiver of repeater

- ² Measured over 3000 UI
- ³ BER 1E-15 is at +/-7 sigma
- 4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.
- 5 See Figure 3.8 to measure the chip's ability to block power supply noise over the voltage range of the repeater device.
- 6 PSRR is measured for noise levels that vary from 10 Hz to 10 MHz range. The spec is 2% from 10Hz to 1MHz and 3% between 1MHz to 10MHz.

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		118		ps	Repeater must be at least 8.5Gb/s at 100ppm"	1
VRX-DIFFp-p	Differential Peak to Peak Input Voltage	30		700	mV	VRX-DIFFp-p = 2* VRX-D+ - VRX-D-	1
TRX-EYE	Minimum RX Eye Width			0.65	UI	The maximum Interconnect and Transmitter jitter that can be tolerated by the Receiver can be derived as TRXMAX- JITTER = 1 - TRX-EYE= .65 UI	1
TJTT BLOCK	Amount of channel jitter eliminated from Rx to Tx	90%			%	This blocks jitter amplification.	6
VRX-CM-Acp	AC Peak Common Mode Input Voltage	7			mV	VRX-CM-ACp = RMS {[(VRX- D+) + (VRX-D-)] / 2 - VRXdcavg}	1
RLRX-DIFF	Differential Return Loss			-12	dB	50MHz-4.25GHz	3
RLRX-COMM	Common Mode Return Loss		-8		dB	50MHz-4.25GHz	3
LOS	Loss of Signal					Firmware Implementation	
PSRR	Power supply rejection ration					Measured using PSRR per XFP standards	7

Table 3.25 Repeater: Differential RX Input "B" Recommendation

Notes:

1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20*\log(|\rho|)$.

² Measured over 3000 UI

3 Measured with at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

⁴ This is the differential voltage at which the output will be disabled

5 If any in band signal goes below a settable range, the device must provide the capability to disable outputs. This can be implemented either internally (desirable) or externally.

Repeater jitter blocking is validated with setup as shown in Figure 3.9. It is expect that retiming or receiver FIR equalization may be employed. However, repeaters must work for the entire range of data rates.

7 PSRR is measured for noise levels that vary from 10 Hz to 10 MHz range. The spec is 2% from 10Hz to 1MHz and 3% between 1MHz to 10MHz.

3.11 8.5Gb/s Fibre Channel Equalizer Agent on Daughter card

Repeaters will be required if the Tx or Rx agents cannot meet the prior Tx/Rx specifications. Here is the topology which shows the placement of Equalizer when Switch is driving and Daughter card is receiving.



The repeater or equalizer should meet following criteria:

1-FFE: 10 taps with T/2 spacing

2-DFE: 4 taps with T spacing.

3-Bypass mode for 4G and 2G operation.

4-Supports 8 FC auto speed negotiation requirements.

5-Adaptive convergence algorithm that optimizes the link performance to be lower than 1E-15 bit error rate.

Table 3.26 Repeater:	Differential TX	Output "C"	Recommendation

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		118		ps	UI must be 100ppm. Below 200ps is acceptable, but not required.	2
VTX-DIFFp-p	Differential Peak to Peak Output Voltage	350	650	850	mV	VTX-DIFFp-p = 2* VTX-D+ - VTX-D-	1
Gb/s	Rate Supported		2.125 4.250 8.500		Gb/s		
TTx_RJ	Random Jitter			0.35	UI	Observed at BER 1E-15	1,3
TTx_DJ	Deterministic Jitter			0.20	UI		1,3
VTX-CM-ACP	AC Peak Common Mode Output Voltage			15	mV	VTX-CM-ACp = RMS {[(VTX- D+) + (VTX-D-)] / 2 - VTXdcavg}	
RLTX-DIFF	Differential Return Loss			-12	dB	50MHz-4.25GHz	4
RLTX-COMM	Common Mode Return Loss		???		dB	50MHz-4.25GHz	4
PSRR	Power supply rejection ration					Measured PSRR per XFP standards	6

Notes:

1 Measured with test load show in Figure 3.5. Jitter specification assumes 90% deterministic jitter of incoming signal is removed by receiver of repeater

- ² Measured over 3000 UI
- ³ BER 1E-15 is at +/-7 sigma
- 4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

5 See Figure 3.8 to measure the chip's ability to block power supply noise over the voltage range of the repeater device.

6 PSRR is measured for noise levels that vary from 10 Hz to 10 MHz range. The spec is 2% from 10Hz to 1MHz and 3% between 1MHz to 10MHz.

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		118		ps	Repeater must be at least 8.5Gb/s at 100ppm"	1
VRX-DIFFp-p	Differential Peak to Peak Input Voltage	30		700	mV	VRX-DIFFp-p = 2* VRX-D+ - VRX-D-	1
TRX-EYE	Minimum RX Eye Width			0.65	UI	The maximum Interconnect and Transmitter jitter that can be tolerated by the Receiver can be derived as TRXMAX- JITTER = 1 - TRX-EYE= .65 UI	1
TJTT BLOCK	Amount of channel jitter eliminated from Rx to Tx	90%			%	This blocks jitter amplification.	6
VRX-CM-Acp	AC Peak Common Mode Input Voltage	7			mV	VRX-CM-ACp = RMS {[(VRX- D+) + (VRX-D-)] / 2 - VRXdcavg}	1
RLRX-DIFF	Differential Return Loss			-12	dB	50MHz-4.25GHz	3
RLRX-COMM	Common Mode Return Loss	??			dB	50MHz-4.25GHz	3
LOS	Loss of Signal					Firmware Implementation	
PSRR	Power supply rejection ration					Measured using PSRR per XFP standards	7

Table 3.27 Repeater: Differential RX Input "B" Recommendation

Notes:

1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20*\log(|\rho|)$.

² Measured over 3000 UI

3 Measured with at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.

⁴ This is the differential voltage at which the output will be disabled

5 If any in band signal goes below a settable range, the device must provide the capability to disable outputs. This can be implemented either internally (desirable) or externally.

Repeater jitter blocking is validated with setup as shown in Figure 3.9. It is expect that retiming or receiver FIR equalization may be employed. However, repeaters must work for the entire range of data rates.

7 PSRR is measured for noise levels that vary from 10 Hz to 10 MHz range. The spec is 2% from 10Hz to 1MHz and 3% between 1MHz to 10MHz.

3.12 4x10 Gb/s Infinaband Tx/Rx Agents

This section describes the signaling that allows for InfiniBand**TM** link operation at 10 Gbits/s (QDR) on HSSM connection in BladeCenter System.

The frequency dependent attenuation of the interconnection media degrades the signal and thus produces Inter-Symbol Interference or Data Dependent Jitter which is a component of the Deterministic Jitter. The effects of high frequency attenuation can be reduced by techniques such

as: • Pre-distortion or Pre-emphasis of the signal produced at the driver. Adaptive equalization techniques such as partial response or DFE will require at the receiver.

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		100		ps	Each UI is 100ps +/- 100ppm	2
VTX-DIFFp-p	Differential Peak to Peak Output Voltage	800		1600	mV	VTX-DIFFp-p = 2* VTX-D+ - VTX-D-	1
TTx_RJ	Random Jitter			0.27	UI	Observed at BER 1E-15	1,3
TTx_DJ	Deterministic Jitter			0.15	UI	Without pre-emphasis	1,3
TTx_TJ	Total Jitter			0.30	UI	Without pre-emphasis	1,3
VTX-CM-ACP	AC Peak Common Mode Output Voltage			25	mV	VTX-CM-ACp = RMS {[(VTX- D+) + (VTX-D-)] / 2 - VTXdcavg}	1
RLTX-DIFF	Differential Return Loss			-10	dB	100MHz-6.25GHz	4
RLTX-COMM	Common Mode Return Loss			-10	dB	100MHz-6.25MHz	4
RLTX-DC	Differentail to Common mode			-20	dB		
RLTX-COMM	Common Mode Return Loss	10			dB	100MHz-6.25MHz	4
De-emphasis	Post-cursor tap					Pre-cursor tap coefficient	5
Sdbtb	Skew ot any two physical Lanes			500	ps		

Table 3.28 Differential TX Output Recommenda	itions
--	--------

Notes:

1 Measured with test load show in Figure 3.5

2 Measured over 3000 UI

3 BER 1E-15 is at +/-7 sigma

4 Measured at the device pins on each channel with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation but it is necessary not to exceed this specification.

⁵ 3 Tap FIR filter

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		100		ps	Each UI is 100ps +/- 100ppm	2
BER	Bit Error Ratio		10E-15			With Minimum Input	
VRX-DIFFp-p	Differential Peak to Peak Input Voltage	30			mV	VRX-DIFFp-p = 2* VRX-D+ - VRX-D-	1,5
TRX_RJ	Receiver Random Jitter			0.125	UI	Observed at BER 1E-15	1,3
TRX_DJ	Receiver Deterministic Jitter			0.125	UI		1,3
VRX-CM-Acp	AC Peak Common Mode Input Voltage			40	mV	VRX-CM-ACp = RMS {[(VRX- D+) + (VRX-D-)] / 2 - VRXdcavg}	
RLRX-DIFF	Differential Return Loss			-8	dB	100MHz-6.25MHz	4
RLRX-COMM	Common Mode Return Loss			-8	dB	100MHz-6.25MHz	4
RLRX-COMD	Differential to Common Mode Return Loss			-20	dB	100MHz-6.25MHz	4

Table 3.29 Differential Rx Input Recommendations

Notes:

- 1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system voltage pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20*\log(|\rho|)$. Conversely the design target at the die would be $(1+2\rho)$ time the Rx spec.
- 2 Measured over 3000 UI
- 3 BER 1E-12 is at +/-7 sigma
- 4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.
- 5 3 TAP DFE equalizer

1 4010 010 0 10	Table 5.56 Receiver input sitter specifications									
Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes			
UI	Unit Interval		100		ps	Repeater must be at least 3.125Gb/s at 100ppm"	1			
DJ	BladeCenter Deterministic Jitter			0.17	UI		1			
TJ	BladeCenter Total Jitter			0.50	UI	The maximum Interconnect and Transmitter jitter that can be tolerated by the Receiver can be derived as TRXMAX- JITTER = 1 - TRX-EYE= .50 UI	1			

Table 3.30 Receiver Input Jitter Specifications

Notes:

1

Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20^{10}(|\rho|)$.

3.13 10 Gb/s KR Ethernet Tx/Rx Agents

The 10 Gb/s –KR transmitter includes programmable equalization to compensate for frequency dependant loss in the BladeCenter environment. This equalization may be accomplished with a three-tap FIR structure. The 10Gb/s KR channel should meet bit error rate of "10E-15" in BladeCenter environment. The HSSM channel has been used for KR operation.

For TX Output and RX input recommendation please refer to IEEE 802.3ap KR specification.

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		100		ps	Each UI is 100ps +/- 100ppm	2
BER	Bit Error Ratio		10E-15			With Minimum Input	
VRX-DIFFp-p	Differential Peak to Peak Input Voltage	20			mV	VRX-DIFFp-p = 2* VRX-D+ - VRX-D-	1,5
TRX_RJ	Receiver Random Jitter			0.125	UI	Observed at BER 1E-15	1,3
TRX_DJ	Receiver Deterministic Jitter			0.125	UI		1,3
VRX-CM-Acp	AC Peak Common Mode Input Voltage			40	mV	VRX-CM-ACp = RMS {[(VRX- D+) + (VRX-D-)] / 2 - VRXdcavg}	
RLRX-DIFF	Differential Return Loss			-8	dB	100MHz-6.25MHz	4
RLRX-COMM	Common Mode Return Loss			-8	dB	100MHz-6.25MHz	4
RLRX-COMD	Differential to Common Mode Return Loss			-20	dB	100MHz-6.25MHz	4

Table 3.32 Differential Rx Input Recommendations

Notes:

1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system voltage pin specification can be determined by multiplying the Rx spec by (1- 2ρ), where ρ is the reflection coefficient looking into the package and die combination. Return loss is 20*log($|\rho|$). Conversely the design target at the die would be (1+ 2ρ) time the Rx spec.

- 2 Measured over 3000 UI
- 3 BER 1E-15 is at +/-7 sigma
- 4 Measured at the device pins with a 50 ohm referenced Vector Network Analyzer (VNA). Meeting differential return loss does not guarantee operation.
- 5 3 TAP DFE equalizer

Table 3.33 Receiver Input Jitter Specifications

Symbol	Parameter	Min	Nom	Max	Units	Comments	Notes
UI	Unit Interval		100		ps	Repeater must be at least 3.125Gb/s at 100ppm"	1
DJ	BladeCenter Deterministic Jitter			0.17	UI		1
TJ	BladeCenter Total Jitter			0.50	UI	The maximum Interconnect and Transmitter jitter that can be tolerated by the Receiver can be derived as TRXMAX- JITTER = 1 - TRX-EYE= .50 UI	1

Notes:

1 Measured with test load as shown in Figure 3.6. Device validated with test in Figure 3.7. The in situ system pin specification can be determined by multiplying the Rx spec by $(1-2\rho)$, where ρ is the reflection coefficient looking into the package and die combination. Return loss is $20^{*}\log(|\rho|)$.

3.14 Test Setups

Figure 3.5 Transmitter Test Setup



The setup in Figure 3.5 is for testing the transmitter specification only. The test load shall be 50 ohms to ground, DC blocked by a 10nF capacitor. The data pattern shall be a valid 8b10b sequence of repeating K28.5. Devices shall provide this out-of-system test. If cabling is required to make this measurement between the indicated nodes, results must be de-embedded.

Figure 3.6 Receiver Eye Test Load Calibration Setup



The setup in Figure 3.6 is for calibrating a channel to test the receiver specification. The goal of this calibration step is to define and determine a channel to be used for receiver testing. A test channel is DC-blocked by 10nF capacitors. The differential transmitters have nominal transmitter specifications (as found earlier in this chapter). The calibration transmission line is 50 ohm +/-10%. When creating a model, the Spec user shall create enough loss through the transmission lines to create a minimum receiver eye opening at the RX pins (RX voltage levels are seen earlier in this chapter). This is achieved by adjusting the length of the transmission line. Typical differential reference channels have about 12dB of Insertion Loss. The data pattern shall be a valid 8b10b PSRB pattern.

Note: The reason an independent low-amplitude differential source is not used to test the RX sensitivity of a device is because a low-amplitude differential source does not exhibit the common mode signal qualities of a long lossy channel.

Figure 3.7 Receiver Eye Compliance Set up



This calibration test channel supplies valid DC balanced 8b10b PSRB data to the receiver device using the same "tuned" transmission line model as was used in the Receiver Eye Test Load Calibration Setup. The receiver device must receive the sent data, with no data errors. Figure 3.7 introduces channel jitter that is similar to the jitter characteristics observed in a system. (These electrical characteristics of channel jitter are significantly different in nature than UI jitter produced at the output of any transmitter.)

Figure 3.8 Power Supply Rejection Measurement Setup



Figure 3.8 (or an equivalent circuit) shall be used to determine the power supply rejection ratio. The device must be provided a DC offset level equal to the device's specified voltages with a superimposed AC sine wave that is swept in frequency from 50MHz to 2.125GHz (to represent noise). The device is held in a quiescent high or low state and the amplitude of the AC output voltage is measured. The result is the power supply rejection ratio (dB) of the measured ac output voltage to the AC input voltage



Figure 3.9 Deterministic Jitter Block Setup for Multi-Gigabit Repeaters

The setup in Figure 3.9 shall be used determine a channel jitter blocking coefficient (TJTT_BLOCK). The channel used for this test is defined in Figure 3.6.



3.15 Plug to Plug Channel Loss (SDD21) Guide for 1Gb/s to 8.5Gb/s

Figure 3.10 Daughter Card to Switch SDD21 channel specific



Figure 3.11 SDD21 Plug to Plug Channel Test Point Locations

Figure 3.10 is to be used as a reference guide for loss between the test locations shown in Figure 3.11. This is not intended to replace the aforementioned device specifications. The test channel used in device testing is adjusted to produce the minimum Rx eye when terminated into the test load.

3.16 Plug to Plug Channel Loss (SDD21) Guide for 4x3.125Gb/s to 10Gb/s KR



Figure 3.12 Daughter Card to Switch SDD21 channel specification



Figure 3.13 SDD21 Plug to Plug Channel Test Point Locations

Figure 3.12 is to be used as reference guide for loss between test locations annotated in Figure 3.13. This is not intended to replace the aforementioned device specifications. The test channel used in device testing is adjusted to produce the minimum Rx eye when terminated into the test load.

IBM/Intel Confidential

Board (Interconnect) Specification

The Board Specification is described by equivalent electrical parametric terms rather than PWB layout parameters. PWB parameters are suggested in the Board Design Guide section. The following voltage transfer function defines how much voltage is delivered from transmit to receive:

$$Voltage_transfer = \frac{\frac{s21}{2} \cdot (1 - \Gamma_s) \cdot (1 - \Gamma_L)}{1 - s11\Gamma_s - s22 \cdot \Gamma_L - s21 \cdot s12 \cdot \Gamma_L \cdot \Gamma_s + s11 \cdot s22 \cdot \Gamma_L \cdot \Gamma_s} - \sum_n s2n$$

Equation 1 - Voltage Transfer

- s11, s12, s21, and s22 are differential s-parameters of the channel as shown in Figure 4.1
- s2n is the differential far-end crosstalk from the differential nth port when n is 6,8,10...
- Γ_s and Γ_L are the complex differential reflection coefficient of the source and load and are determined from the allowable interface to the channel



Figure 4.1 - Generalized Board showing (n) Differential Trace Pairs

The purpose of the Board (Interconnect) Specification is to specify a bounding transfer function which comprehends loss effects due to mismatched interface, crosstalk, and insertion loss. This gives board developers the most flexibility of design choices. Note that board connectors count toward the loss budget.

It turns out the simplest way to create this bounding transfer function is to specify the channel specification in terms of a transmission line (see Figure) with loss specified for two given frequencies. The simplest practical implementation involves specification of a length of coax or reference board transmission line. The precedent for this is the cable fibre test plan for optical communications. The compliance channels are 500hm single ended impedance. Being uncoupled, they have 100 ohm differential impedance. Line length may be used to dial in loss to achieve the eye spec at the receiver. A similar method is use to determine system level compliance.





IBM/Intel Confidential
3.17 System Compliance

To determine spec TX compliance, follow these steps (see diagrams below):

- 1. For Through Blade TX test only, select driver with appropriate spec minimum TX, otherwise use drive strength of part being used in system.
- 2. Terminate end of each port of the channel with 50 ohms
- 3. Use worst-case bit pattern on all TX lines. For simulation, this is defined by lone pulse analysis in. Measurement is any combination PSRB patterns. Blade-to-Switch communication must be successful to a BER of 1E-12.
- 4. Measure eye against the appropriate spec for reflection coefficient = 0 (50 ohm termination)



Figure 4.3 - Blade or Switch TX Compliance



Figure 4.4 Daughter Card TX Compliance



Figure 4.5 Through Blade Interconnect TX Compliance

BladeServer Base Specification

SERDES Design

To determine spec Rx compliance, follow these steps (see diagrams below):

- 1. Select drivers with appropriate spec minimum TX.
- 2. For the Through Blade test
 - a. Terminate the end of each port of the channel with 50 ohms
 - b. Use worst-case bit pattern on all TX lines. For simulation, this is defined by lone pulse analysis in. Measurement is any combination PSRB patterns. Blade to switch communication must be successful to a BER of 1E-12.
 - c. Measure eye against the appropriate spec for the reflection coefficient = 0.
- 3. For other Rx board tests
 - a. Terminate the end of the channel with the board under test
 - b. Use worst-case bit pattern on all TX lines. For simulation, this is defined by lone pulse analysis. For measurement, this is any combination PSRB patterns. Blade to switch communication must be successful to a BER of 1E-12.
 - c. Measure eye at the input to the SERDES chip against the appropriate spec adjusted by the reflection coefficient for the SERDES part.



Figure 4.6 Blade or Switch RX Compliance



Figure 4.7 Daughter Card RX Compliance



Figure 4.8 Through Blade Interconnect RX Compliance

3.18 Specification Method

Observe the below methods to understand how the board specification functions.

3.18.1 Pass/Fail Threshold

Each fabric (Ex/FC, GbE, IBA) can be specified to pass by meeting a frequency-based loss equivalent value. To accurately define the frequency response curve, two data points are given in **Point A** and **Point B** for every data rate. **Point B** is defined at the maximum data frequency allowed. **Point A** is defined at the maximum data frequency divided by 5 (ex/ five bits in a row are a logic "1" or "0").





3.18.2 Trend Curve

Typical interconnect loss curves exhibit resonance bumps that can make the pass/fail comparison challenging. By fitting a line between the frequency of point A and point B with an RMS trending method, additional margin can be gained by curves that would pass, if not for their resonances. To pass compliance, the trend curve must be above the Pass/Fail Threshold. In other words, the Actual Loss Curve can dip below the Pass/Fail Threshold, and the system can still achieve compliance as long as the Trend Loss curve is passing. The limit of deviation to this line must be < 0.3 dB.



Figure 4.10 - Trend Loss Curve Promotes Additional Margin

3.19 Interconnect Specification Data

There are 4 major channel specification categories. Using either one or a combination of the following specs will allow a designer to meet the loss specifications of any Blade, Daughter Card, or Switch configuration. The VHDM connectors imply connection to a mid-plane. This channel spec requires that the Device Electrical Specifications from the preceding chapter must be met.

The loss specification for each board does not include interface connectors or chip packages. Those losses are included in the backplane loss, or the blade model loss in the case of the Daughter Card. Landing zones such as vias and pads are included, however.

No Spec Compliance Channel Loss is listed for the backplane because it is not being externally designed.

	GbE (1.25	Gb/s)		FC (2.125 Gb/s)					IBA (2.5Gb/s)		
Poin	t A	Poin	t B	Point A	١	Point B		Point	Α	Point	В
125 MHz	-1.8dB	625 MHz	-2.4dB	212.5 MHz	-0.5dB	1.0625 GHz	-1.6dB	250 MHz	-0.6dB	1.25 GHz	-1.9dB



Info: This board has the SERDES down on the board. (Ex/ GbE)

Table 4.2 Switch Interconnect for On-Board Device

	GbE (1.2	25Gb/s)			FC (2.1	25 Gb/s)	IBA (2.5Gb/s)				
Point	А	Point	В	Point A	A	Point B		Point	А	Point H	3
125 MHz	IHz -1.8dB 625 MHz -2.4dB		212.5 MHz	-0.5dB	1.0625 GHz	-1.6dB	250 MHz	-0.6dB	1.25 GHz	-1.9dB	



Info: This switch has the SERDES down on the board. (Ex/ GbE, FC, IBA)

Table 4.3 Daughter Card Interconnect

(GbE (1.25	Gb/s)			FC (2.125 Gb/s) IBA (2.5Gb/s)					.5Gb/s)	
Point	A	Poin	t B	Point A		Point B		Point	Α	Point	В
125 MHz	-0.5XdB	625 MHz	-1.0dB	212.5 MHz	-0.13dB	1.0625 GHz	-0.4dB	250 MHz	-0.15dB	1.25 GHz	-0.5dB



Mezz Conn. Molex #75005-2100 Molex #75003-2100 Info: The Daughter Card has a SERDES device driving to the mezzanine connector. (Ex/ GbE, FC, IBA)

(GbE (1.2	25Gb/s)			FC (2.1	25 Gb/s)	IBA (2.5Gb/s)				
Point	A	Point	В	Point A		Point B		Point	А	Point E	3
125 MHz	-0.4dB	625 MHz	-1.2dB	212.5 MHz	4dB	1.0625 GHz	-1.2dB	250 MHz	5dB	1.25 GHz	-1.5dB

 Table 4.4 Blade Through Interconnect for Daughter Card



Info: After the Daughter Card has a SERDES device driving down through its mezzanine connector out to the VHDM connector. (Ex/ GbE, FC, IBA)

41

4 Board Design Guidelines

All rules are not equal. This document is sorted into two major sections to assist layout risk assessment. The first section lists all of the absolutely *Imperative Layout Rules* that don't offer exceptions. Following these rules will lead to clean routing and excellent signal quality. However, in the real world, there is always going to be layout quandaries. Therefore, other design recommendations that have a marginal effect on signal quality are in a later section called *Marginal Layout Rules*.

The Severity Rating is as follows:



Also, observe these notation marks:



For Example:

This A Means: Must do this! This A Means: Must NOT do this!

4.1 Imperative Layout Rules

4.1.1 Route all high-speed switch fabric signals first!

This is not absolutely required, but given the opportunity should be done every time. As a strong recommendation, it helps alleviate many risks in the board layout process.

4.1.2 Routing Length Guidelines and PWB Characteristics (except for 4.25Gb/s, 8.5Gb/s and 4x3.125 Gb/s)

The following tables represent maximum segment lengths that will guarantee operation without the need to verify with simulation. These tables also assume that the PWB characteristics and other subsequent guidelines are followed.

In the below tables, "x" refers to the dielectric thickness of a layer between a signal and its nearest ground plane. The basic stack-up configurations require all signals to be ground referenced. Asymmetric striplines can be loosely referenced to a power plane as long as the distance from the power plane to the trace is greater than the distance from the ground plane to that trace. Length matching between TX and RX pairs is not required; however, signals within a pair need to be kept under 5 mils per board. Also, parasitics (of vias, padstacks) along the lines in the pairs must be matched.

Note: Simulation tolerances are different than impedance coupon test tolerances. Simulation tolerance should be used to qualify the robustness of a design, test tolerances are used as an acceptance criteria for a manufactured board (measured on the impedance coupon).



Table 5.1 Blade Interconnect for On-Board Device Trace Spec

Blade	Interco	onnect	for (On-B		GbE (1.25Gb/s	FC (2.125 Gb/s)	IBA (2.5Gb/s)					
pwb	er	tand		W	Т	S	D	Z0D	sim tol	test tol	Max length	Max length	Max length
	1GHz	1GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches	inches	inches
fr4	4.1	<0.02	mS	7-9	1.5	>1x	>5x	100	20	10	9	7	7
fr4	4.1	<0.02	SL	7-9	1	>1x	>3x	100	15	10	10	8	8

Info: This board has the SERDES down on the board. (Ex/ GbE)





BladeServer Base Specification SERDES Design

 Table 5.2 Switch Interconnect for On-Board Device Trace Spec

Swite	witch Interconnect for On-Board Device Trace Spec b er tand W T S D ZOD sim tol test tr											FC (2.125 Gb/s)	IBA (2.5Gb/s
pwb	er	tand		W	Т	S	D	Z0D	sim tol	test tol	Max length	Max length	Max length
	1GHz	1GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches	inches	inches
fr4	4.1	<0.02	mS	7-9	1.5	>1x	>5x	100	20	10	9	7	
fr4	4.1	<0.02	SL	7-9	1	>1x	>3x	100	15	10	10	8	8

VHDM Molex #74074-9993 Molex #74030-9973





Info: This switch has the SERDES down on the board. (Ex/ GbE, FC, IBA)

T 11 C 2	D 14	0 1	T / /	T	a
I anie 5 4	Liguionter	1 ard	Interconnect	Irace	Nnec.
1 auto 5.5	Dauginer	Caru	mucrounicer	Trace	Spuce
	4 /				

Daug	aughter Card Interconnect for Trace Spec											FC (2.125 Gb/s)	IBA (2.5Gb/s)
pwb	er	tand		W	Т	S	D	Z0D	sim tol	test tol	Max length	Max length	Max length
	1GHz	1GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches	inches	inches
fr4	4.1	<0.02	mS	7-9	1.5	>1x	>5x	100	20	10	4	2	2
fr4	4.1	<0.02	SL	7-9	1	>1x	>3x	100	15	10	4	2	2



Info: The Daughter Card has a SERDES device driving to the mezzanine connector. (Ex/ GbE, FC, IBA)

Table 5.4 Blade In	erconnect for Daughter	Card Trace Spec

Blade	lade Interconnect for Daughter Card Trace Spec											FC (2.125 Gb/s)	IBA (2.5Gb/s)
pwb	er	tand		W	Т	S	D	Z0D	sim tol	test tol	Max length	Max length	Max length
	1GHz	1GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches	inches	inches
fr4	4.1	<0.02	mS	7-9	1.5	>1x	>5x	100	20	10	8	6	6
fr4	4.1	<0.02	SL	7-9	1	>1x	>3x	100	15	10	8	6	6



Info: After the Daughter Card has a SERDES device driving down through its mezzanine connector out to the VHDM connector. (Ex/ GbE, FC, IBA)

4.1.3 Routing Length Guidelines and PWB Characteristics (4.25 Gb/s)

The following tables represent maximum segment lengths that will guarantee operation without the need to verify with simulation. These tables also assume that the PWB characteristics and other subsequent guidelines are followed.

In the below tables, "x" refers to the dielectric thickness of a layer between a signal and its nearest ground plane. The basic stack-up configurations require all signals to be ground referenced. Asymmetric striplines can be loosely referenced to a power plane as long as the distance from the power plane to the trace is greater than the distance from the ground plane to that trace. Length matching between TX and RX pairs is not required; however, signals within a pair need to be kept under 5 mils per board to minimize the common mode effect on over all system. Connector skew matching need to be implemented near the connector.



Note: Simulation tolerances are different than impedance coupon test tolerances. Simulation tolerance should be used to qualify the robustness of a design; test tolerances are used as acceptance criteria for a manufactured board (measured on the impedance coupon).

	Swi	tch Inter	rconne	ect for	On-Bo	oard D	evice	Trace	Spec		FC (8.5 Gb/s)		
pwb	pwb er tand W T S D Z0D Sim tol Test tol												
	1GHz	2GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches		
Low Loss	4.1	<0.014	mS	7-8	1.5	>1x	>5x	100	20	10	6		
Low Loss	Loss 4.1 <0.014 SL 7-8 1 >1x >3x 100 15 10												

Table 5.5 Switch	Interconnect	for On-Board	Device Trace S	Spec
			201100 11000	



Info: This switch has the SERDES down on the board. (Ex/ GbE, FC, IBA)

BladeServer Base Specification

	FC (4.25 Gb/s)											
pwb	Max length											
	1GHz 1GHz mils oz x hj x h ohm +/-% +/-%										inches	
fr4	fr4 4.1 <0.02 mS 8-9 1.5 >1x >5x 100 20 10											
fr4	fr4 4.1 <0.02 SL 8-9 1 >1x >3x 100 15 10											

Table 5.6 Daughter Card Interconnect Trace Spec



Info: The Daughter Card has a SERDES device driving to the mezzanine connector. (Ex/ GbE, FC, IBA)

	FC (4.25 Gb/s)												
pwb	pwb er tand W T S D Z0D Sim tol Test tol												
	1GHz	1GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches		
fr4	4.1	<0.02	mS	7-9	1.5	>1x	>5x	100	20	10	5		
fr4	fr4 4.1 <0.02 SL 7-9 1 >1x >3x 100 15 10												

 Table 5.7 Blade Interconnect for Daughter Card Trace Spec



Info: After the Daughter Card has a SERDES device driving down through its mezzanine connector out to the VHDM connector. (Ex/ GbE, FC, IBA)

4.1.4 Routing Length Guidelines and PWB Characteristics (8.5 Gb/s)

The following tables represent maximum segment lengths that will guarantee operation without the need to verify with simulation. These tables also assume that the PWB characteristics and other subsequent guidelines are followed.

In the below tables, "x" refers to the dielectric thickness of a layer between a signal and its nearest ground plane. The basic stack-up configurations require all signals to be ground referenced. Asymmetric striplines can be loosely referenced to a power plane as long as the distance from the power plane to the trace is greater than the distance from the ground plane to that trace. Length matching between TX and RX pairs is not



Note: Simulation tolerances are different than impedance coupon test tolerances. Simulation tolerance should be used to qualify the robustness of a design; test tolerances are used as acceptance criteria for a manufactured board (measured on the impedance coupon).

	FC (8.5 Gb/s)											
pwb	Max length											
	1GHz	2GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches	
Low Loss	Low Loss 4.1 <0.014 mS 7-8 1.5 >1x >5x 100 20 10											
Low Loss	Loss 4.1 <0.014 SL 7-8 1 >1x >3x 100 15 10											



Info: This switch has the SERDES down on the board. (Ex/ GbE, FC, IBA)

Note: If the Switch ASIC meets the compliance section's 3.9 TX / RX specification, then OSIBA network need to be implemented on the Switch design.



BladeServer Base Specification SERDES Design

	FC (8.5 Gb/s)											
pwb	Min length											
	1GHz	2GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches	
Low Loss	4.1	<0.014	mS	7-8	1.5	>1x	>5x	100	20	10	4.5	
Low Loss	Loss 4.1 <0.014 SL 7-8 1 >1x >3x 100 15 10											





Info: The Daughter Card has a SERDES device driving to the mezzanine connector. (Ex/ GbE, FC, IBA)

	Blade Interconnect for Daughter Card Trace Spec											
pwb	er	Tand		W	Т	S	D	Z0D	Sim tol	Test tol	Max length	
	1GHz	2GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches	
Low Loss	4.1	<0.014	mS	7-8	1.5	>1x	>5x	100	20	10	4.5	
Low Loss	4.1	<0.014	SL	7-8	1	>1x	>3x	100	15	10	3.5	

 Table 5.10 Blade Interconnect for Daughter Card Trace Spec



Info: After the Daughter Card has a SERDES device driving down through its mezzanine connector out to the VHDM connector. (Ex/ GbE, FC, IBA)

4.1.5 Routing Length Guidelines and PWB Characteristics (3.125 Gb/s)

The following tables represent maximum segment lengths that will guarantee operation without the need to verify with simulation. These tables also assume that the PWB characteristics and other subsequent guidelines are followed.

BladeServer Base Specification

SERDES Design

12 May 2010

In the below tables, "x" refers to the dielectric thickness of a layer between a signal and its nearest ground plane. The basic stack-up configurations require all signals to be ground referenced. Asymmetric striplines can be loosely referenced to a power plane as long as the distance from the power plane to the trace is greater than the distance from the ground plane to that trace. Length matching between TX and RX pairs is not required; however, signals within a pair need to be kept under 5 mils per board. Also, the 4 pairs of TX and the 4 pairs of RX must match within 1000 mil 1500 mil. Compensate connector pin mismatch on the Switch and Daughter card.

Note: Simulation tolerances are different than impedance coupon test tolerances. Simulation tolerance should be used to qualify the robustness of a design, test tolerances are used as acceptance criteria for a manufactured board (measured on the impedance coupon).



Table 5.11 Switch Interconnect for On-Board Device Trace Spec

	(4x3.125 Gb/s)											
pwb	Max length											
	1GHz	1GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches	
fr4	4.1	<0.02	mS	7-9	1.5	>1x	>5x	100	20	10	6	
fr4	fr4 4.1 <0.02 SL 7-9 1 >1x >3x 100 15 10											





Info: This switch has the SERDES down on the board. (Ex/ GbE, FC, IBA)

Table 5.12 Daughter Card Interconnect Trace Spec

	(4x3.125 Gb/s)										
pwb	pwb er tand W T S D Z0D Sim tol Test tol										
	1GHz	1GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches
fr4	4.1	<0.02	mS	7-9	1.5	>1x	>5x	100	20	10	4
fr4	fr4 4.1 <0.02 SL 7-9 1 >1x >3x 100 15 10										



Info: The Daughter Card has a SERDES device driving to the mezzanine connector. (Ex/ GbE, FC, IBA)

4.1.6 4 X 10 Gb/s Infinaband Routing Length Guidelines and PWB Characteristics

The following tables represent maximum segment lengths that will guarantee operation without the need to verify with simulation. These tables also assume that the PWB characteristics and other subsequent guidelines are followed.

In the below tables, "x" refers to the dielectric thickness of a layer between a signal and its nearest ground plane. The basic stack-up configurations require all signals to be ground referenced. Asymmetric striplines can be loosely referenced to a power plane as long as the distance from the power plane to the trace is greater than the distance from the ground plane to that trace. Length matching between TX and RX pairs is not required; however, signals within a pair need to be kept under 5 mils per board. Also, the 4 pairs of TX and the 4 pairs of RX must match within 1000 mil 1500 mil. Compensate connector pin mismatch on the Switch and Daughter card.

Note: Simulation tolerances are different than impedance coupon test tolerances. Simulation tolerance should be used to qualify the robustness of a design, test tolerances are used as acceptance criteria for a manufactured board (measured on the impedance coupon).



 Table 5.13 Switch Interconnect for On-Board Device Trace Spec

		(4x10 Gb/s)										
pwb	Max length											
	1GHz	1GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches	
IS415	IS415 3.7 <0.015 mS 7-9 1.5 >1x >5x 100 20 10											
IS415	IS415 3.7 <0.015 SL 7-9 1 >1x >3x 100 15 10											



Info: This switch has the SERDES down on the board. (Ex/ GbE, FC, IBA)

IBM/Intel Confidential

	(4x10Gb/s)											
pwb	pwb er tand W T S D Z0D Sim tol Test tol											
	1GHz	1GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches	
IS415	3.7	<0.015	mS	7-9	1.5	>1x	>5x	100	20	10	4	
IS415	IS415 3.7 <0.015 SL 7-9 1 >1x >3x 100 15 10											



Info: The Daughter Card has a SERDES device driving to the mezzanine connector. (Ex/ GbE, FC, IBA)

4.1.7 Routing Length Guidelines and PWB Characteristics (10Gb/s KR)

The following tables represent maximum segment lengths that will guarantee operation without the need to verify with simulation. These tables also assume that the PWB characteristics and other subsequent guidelines are followed.

In the below tables, "x" refers to the dielectric thickness of a layer between a signal and its nearest ground plane. The basic stack-up configurations require all signals to be ground referenced. Asymmetric striplines can be loosely referenced to a power plane as long as the distance from the power plane to the trace is greater than the distance from the ground plane to that trace. Length matching between TX and RX pairs is not required; however, signals within a pair need to be kept under 5 mils per board. Compensate connector pin mismatch on the Switch and Daughter card.

Note: Simulation tolerances are different than impedance coupon test tolerances. Simulation tolerance should be used to qualify the robustness of a design, test tolerances are used as acceptance criteria for a manufactured board (measured on the impedance coupon).



Switch Interconnect for On-Board Device Trace Spec									(10 Gb/s)		
pwb	er	tand		W	Т	S	D	Z0D	Sim tol	Test tol	Max length
	1GHz	1GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches
IS415	3.7	<0.015	mS	7-9	1.5	>1x	>5x	100	20	10	6
IS415	3.7	<0.015	SL	7-9	1	>1x	>3x	100	15	10	5

	Table 5.15 Switch	Interconnect for	On-Board D	Device Trace Spo	ec
--	-------------------	------------------	-------------------	------------------	----

Enhance GBX Teradyne AG822-00001 Teradvne AG822-00002



Info: This switch has the SERDES down on the board. (Ex/ GbE, FC, IBA)

Table 5.16 Daughter Car	d Interconnect Trace Spec
-------------------------	---------------------------

Daughter Card Interconnect for Trace Spec								(10 Gb/s)			
pwb	er	tand		W	Т	S	D	Z0D	Sim tol	Test tol	Max length
	1GHz	1GHz		mils	oz	x hj	x h	ohm	+/-%	+/-%	inches
IS415	3.7	<0.015	mS	7-9	1.5	>1x	>5x	100	20	10	4
IS415	3.7	<0.015	SL	7-9	1	>1x	>3x	100	15	10	3



Info: The Daughter Card has a SERDES device driving to the mezzanine connector. (Ex/ GbE, FC, IBA)



Figure 5.1 Board Configuration

4.1.8 Clocks and Internal PLL's

Differential source clocks measured at the SERDES clock input must have less than 100ppm jitter. Spreadspectrum clocking is not allowed because those SERDES operates on two independent clock domains. There are two options for supplying the SERDES clocking. There first is supply a crystal input to the SERDES part. The second is to provide a high quality differential clock signals on the Daughter Card, Blade, or Switch to the respective SERDES chips. This may be accomplished two ways: cascading an oscillator and a differential buffer, or oscillator, PLL, and differential buffer distribution.

The oscillator requirement is outlined in the Table below. The Single Side Band Phase Noise is measured in the (1/Hz) band

Offset Frequency(KHz)	Single Side Band Phase Noise (dB)
10	-75
100	-95
1000	-115
10000	-135

4.1.9 Line Width vs. Impedance

The transmission transfer function, Equation 1, allows for tradeoffs between board line width and impedance. For example, at any impedance, every mil less than 8mils wide on the board reduces the board length by 1-3 inches. In addition to that, trace widths below 5mils wide lose 3-6 inches (per mil of width reduced). The other tradeoff available involves the target impedance design, which is ideally 100 ohms differential. Impedance and trace width are often co-dependent upon each other. But, do note that the benefit of wider traces much outweigh the benefit of setting impedance to 100ohms. Improving differential impedance from 85 ohms to 100 ohms can gain 1-3" of trace length, but will incur 1-3" of penalty for each mil width reduced.

4.1.10 Spacing

The spacing defined below is generic for HSS signals. However, there may be different spacing required specifically for an individual bus. A separate spacing requirement may also be required for separating one bus pair to a different bus pair, known as inter-signal spacing. (Example: Gigabit Ethernet-to-Fibre Channel spacing).



4.1.11 Guard Traces

Do not use guard traces on these signals. Provide adequate spacing to reduce crosstalk. Guard traces may have negative impact due to resonance at these frequencies and require extensive 3 dimensional full wave modeling to analyses.

IBM/Intel Confidential

4.1.12 Additional design guidelines for all 4 X 3.125 Gb/s interfaces

- All Daughter Card and Switches require series DC blocking capacitors on transmit and receive lines. DC blocking requires a broad range of low loss, resonant free, frequency coverage. The upper limit of bandwidth over which insertion loss meets specification is determined by the location of parallel resonances. Avoid placement in the middle of the trace route.
- Length matching between TX and RX pairs is not required; however, signals within a pair need to be kept under 5 mils per pair on segment-by-segment basis at the point of discontinuity.
- Route the TX and RX pairs on different layer.
- Ac coupling capacitors placed on top or bottom layer with out any via. If you have to have a via than minimize the via stub to 25 mil length and create anti-pad surrounding via pad.
- Fiber weave effect is introduction of AC Common Mode voltage on the differential pair. To mitigate the effect of Fiber eve following techniques need to use during Switch, Blade and Daughter Card design. (section 6.3)
 - PCB vendor rotates image
 - Designer Rotates Image
 - Rotate Glass
 - Advanced materials: Nelco SI material (NE glass)
 - Specify tighter (or coarser) weaves
 - Floor plan design for 45 degrees
- Remove all unused pads from connector vias and any associated vias on high speed nets.
- Connector via pad need to have anti-pad created to improve signal quality. Here is the optimized drawing indicating anti-pad for Fiber Channel Switch Card connector.



Here is the drawing indicating anti-pad for Switch connector.



4.1.13 Additional Design Guideline for 8.5Gb/s Fiber Interface

- All Daughter Card and Switches require series DC blocking capacitors on transmit and receive lines. DC blocking requires a broad range of low loss, resonant free, frequency coverage. The upper limit of bandwidth over which insertion loss meets specification is determined by the location of parallel resonances. Avoid placement in the middle of the trace route.
- Length matching between TX and RX pairs is not required; however, signals within a pair need to be kept under 5 mils per pair on segment-by-segment basis at the point of discontinuity.
- Route the TX and RX pairs on different layer.
- Ac coupling capacitors placed on top or bottom layer with out any via. If you have to have a via than minimize the via stub to 25 mil length and create anti-pad surrounding via pad.
- Fiber weave effect is introduction of AC Common Mode voltage on the differential pair. To mitigate the effect of Fiber eve following techniques need to use during Switch, Blade and Daughter Card design. (section 6.3)
 - PCB vendor rotates image
 - Designer Rotates Image
 - Rotate Glass
 - Advanced materials: Nelco SI material (NE glass)
 - Specify tighter (or coarser) weaves
 - Floor plan design for 45 degrees
- Remove all unused pads from connector vias and any associated vias on high speed nets.
- Connector via pad need to have anti-pad created to improve signal quality. Here is the optimized drawing indicating anti-pad for Fiber Channel Switch Card connector.





• The reduction in the power system noise is evident with thinner dielectrics with an increase in the capacitance and decrease in the inductance of the power planes, which leads to an overall reduction of the impedance of the planes. Thinner dielectrics can also reduce the EMI radiation at the edges of the power plane structure. Here is the example of layer stack up configuration which can be used for switch design.

L							
T	S1			S1			
Γ		Nelco 4000-13SI			LT <0.015		
Γ	PWR2			PWR2			
Γ		HK04 Material	1mil		HK04 Material	1mil	
Γ	GND3			GND3			
Γ		Nelco 4000-13SI			LT <0.015		
	S4			S4			
		Nelco 4000-13SI			LT <0.015		
L	GND5			GND5			
		Nelco 4000-13SI			LT <0.015		
	S6			S6			
L		Nelco 4000-13SI			LT <0.015	_	
L	GND7			GND7			
L		Nelco 4000-13SI			LT <0.015		
L				S8			
L		Nelco 4000-13SI			LT <0.015		
L	S 9			PWR9			
L		Nelco 4000-13SI			LT <0.015		
L	GND10			GND10			
L		Nelco 4000-13SI			LT <0.015		
L	S11			S11			
L		Nelco 4000-13SI			LT <0.015		
Ļ	GND12			GND12			
L		Nelco 4000-13SI			LT <0.015		
Ļ	S13			S13			
Ļ		Nelco 4000-13SI			LT <0.015		
Ļ	GND14			GND14			
1		HK04 Material	1mil		rik04 Material	1mil	
1	PWR15			PWR15			
L		Nelco 4000-13SI			LT <0.015		
1	S16			S16			
1							
1						1	

58

BladeServer Base Specification

4.1.13.1 EMC Rule Checking

Proper EMC design for printed circuit boards (PCB) can make a significant difference in the final product's performance in the EMC measurement laboratory. Implementing the proper EMC practices during the design phase of the product is critical and can have little or no impact on the product cost when done correctly. Differential signals should obey the same EMC rules as single-ended signals, to minimize the EMC problems due to common mode.

There are rule checking programs like, **EMSAT** are available in the industry. These rule checking programs that is optimized to provide fast and accurate reporting of any EMC violations on a printed circuit board design. The tool can flag following violations easily:

- * Signal reference
- * Wiring and crosstalk
- * Decoupling
- * Component placement

The switch and daughter card fabric requires to analyze using these tools to avoid EMI problems in BladeCenter products.

4.1.14 Split Planes, Plane Referencing, Voids

Crossing split planes is not recommended, but sometimes unavoidable. Only 2 split crossings are allowed. This reduces the total line length by 1 inch.

4.1.14.1 90 degree crossing

This guideline is to reduce EMI radiation and plane noise susceptibility. Use 10nF 0402 or 0603 caps. Only groups of common duplex may cross a plane together. In other words, neither TX nor RX may be routed over a split in the same group.





4.1.14.3 Parallel Ground / Power Splits

Never route high-speed traces in parallel with a voided ground/power plane regardless of how far away the power plane is.



4.1.14.4 Referencing to Power

Follow these rules:

Microstrip - High-Speed signals must not be referenced purely to power in a microstrip case.



Stripline - Make sure that there are no parallel spits above the stripline signals



Asymmetric Stripline - No signals are allow vertically above SERDES signals. Also, keep 4*x spacing from one bus to nearby alternative bus.



4.1.14.5 Plane Voids



4.1.15 Metal Fastener Spacing



BladeServer Base Specification 4.1.16 Corners

SERDES Design

Avoid all 90 or 180-degree corners outside of BGAs or connectors.



Instead combine 2, 45-degree corners, or use a half-circle:



When 90 or 180-degree turns are necessary (within a BGA or connector field):

4.1.17 Pin Grid Array Routing



4.1.17.1 Serpentines (Length Matching)

4.1.17.2 45-degree Angles, Height Restrictions



4.1.17.3 Large-Bump Serpentine



4.1.17.4 Length Matching High Density Connector Escape

When the pins of the pair are very deep, the best way to escape is to length-match the 90-degree turn with serpentines.



4.1.17.5 Widen single-ended traces to yield equivalent Z_{oDiff}

Don't widen traces in connector fields. Only widen traces when parallel routing, due to component connection, is not possible.

BladeServer Base Specification

4.1.18 Thermal Relief

As server technologies move into more dense connectors, there is a reduced likelihood that through-hole mount connectors need to be used. Therefore, the need for thermal relief is becoming antiquated. However, it is still recommended to avoid thermal relief for all High-Speed Serial signaling.



4.1.19 Do not route TX or RX pairs parallel to a near edge of the board



4.1.20 Vias, Anti-Vias

Vias should be used conservatively. Changing layers has an adverse and varying effect on signal quality. There are several reasons to avoid changing layers, including via coupling and return path issues. Only via to other layers during breakout and for the connection of parallel/series components.

The max number of vias on a RX or TX channel is 4 per board if the DC blocking cap is on the board. The max number of vias on either a RX or a TX channels is 2 per board if the DC blocking cap is not on the board.

4.1.20.1 Anti-via Pin Grid Trace Overlap – 3 via rule

Outside of the initial breakout region, do not to route signals through the more than 3 anti-pads. The main issues are susceptibility not transmissibility. I.e. if the signals in the vias are synchronized, significant differential coupling can occur.



4.1.20.2 Stitching Vias for Changing Layers

Make sure that any time a differential pair is routed around a via, the via must not be a signal. Power and ground are OK.



4.1.20.3 Stitching vias for Changing Layers with Multiple Lanes

When changing layers, a ground stitch via is needed every RX or TX lane to avoid via coupling.



4.1.21 DC Blocking Caps

All Blades and Switches require 10nf to100nF series DC blocking capacitors on transmit and receive lines. Avoid placement in middle of the trace route.

Capacitor Dimension	Performance	
402	Best	
603	Adequate	
805	Poor	

4.1.22 Capacitor Placement



4.1.23 Stubs and Test Pads

No Stubs! For external termination where stubs are unavoidable, keep total stub length less than 20mils. Where possible, <u>re-</u>use vias for probe points. Do not exceed maximum via count for adding test pads.

Example:



4.1.24 Power Delivery.

Do not use devices with low PSSR. However, many PECL devices have low PSSR. Refer to vendor guidelines but caution should be observed with PECL devices.

4.1.25 Termination schemes



4.1.26 Connectors

- Electrically short preferred
 - \circ < 100 ps delay
- Insertion Loss < -0.3 dB up to 1.5 GHz
- Near End X-talk < -35 dB up to 1.5 GHz

4.1.27 Net Naming Convention



- Start with a common prefix (ex/ "IB_" for InfiniBand, "FC_" for Fibre Channel, etc.)
- High Speed Serial nets should have a suffix "_P" or "_N" to indicate polarity and to facilitate differential pair grouping.
- Board nets should have a direction association. In other words: TX nets should have a "TX", RX nets should have an "RX". The direction of the naming should be blade-centric, but it's not critical.
- For AC coupled nets, the section between the controller and the AC caps should have a separation denominator such as "CAP_".
- Add a "_CLK" suffix to any critical HSS clock lines
- Example: IBx_TXy_P
 - \circ IBx Infiniband, bus number 'x'
 - TXy Transmit from the blade, pair number 'y' of that bus
 - P Positive (+) signal

4.2 Marginal Layout Rules (Layout Suggestions)

Most of these rules, although desirable for HSS, have less than ¹/₄ inch impact on total trace lengths.

4.2.1 Symmetrical Via "Route-By"



Note: Design H1 and H2 to be greater than double the trace width.



4.2.2 Breakout (Length Matching)

Here are some typical breakout cases on a dense package. Simulations show that these breakouts all exhibit similar insertion loss performance near 1.25 GHz.



L1-



L6-

4.2.3 Symmetrical Route Around a Non-Signal Via

Note: Do not route any differential pair around a signal via.

L2-

L3- L4- L5-
5 Appendix

5.1 Blade Configuration Diagrams

5.1.1 Blade w/ Daughter Card and Repeater



5.1.2 Blade w/ Repeater, but no Daughter Card





5.1.3 Blade w/ no Daughter Card and no Repeater



5.1.4 Double Wide Blade w/ Daughter Card and Repeater





5.2 Backplane (Mid-board) System Concept

5.3 Dielectric Weave-Effect Reduction

When analyzing electro-magnetic interference (EMI) and Common Mode, dielectric weave effects (which are inherent within standard FR-4 under normal routing conditions) are significantly larger than all other length matching issues. Therefore, the importance of minimizing weave effects is important to layout. This effect will contribute to wider ranges of differential impedance than expected from single ended measurements.



To reduce dielectric weave effects, there are several routing rules to follow:

- 1. Route signals with 5-degree stagger. This option averages out the effect because all of the signals in a bus are exposed to the same amount of dielectric variability.
- 2. Use specialized board material that is manufactured without the thick varying glass weave (ex/ silk weave)

5.4 External Termination Schemes

It is desirable for all TX and RX parts (SERDES and repeaters) to have internal termination. However, if external termination is unavoidable, cap size is preferably size 0402, and at most size 0603. This is consistent with the requirements of a DC blocking cap. Below are some layout recommendations to properly terminate either parallel or series networks. The tolerance on all parts should be 1%. External termination increase sensitivity to power/ground noise. Historically detrimental noise effects of this choice may not be observed until a large number of boards are produced and tested in systems. Hence this all external termination is risky

The following is a simple example, but all external termination schemes require extensive analysis.

5.4.1 Terminator Resistor to Ground - One simple example

Example of termination ground at TX pins:



77

5.4.2 Parallel Termination - One simple example

A low ESR decoupling capacitor is required for each termination pair located with in 50 mils of pull-up pad. Example of parallel termination at Rx pins of repeater.

