Document Number ENG-46158
Revision Revision 1.8
Author Yi-Chin Chu
Project Manager JR Rivers

Serial-GMII Specification

The Serial Gigabit Media Independent Interface (SGMII) is designed to satisfy the following requirements:

- Convey network data and port speed between a 10/100/1000 PHY and a MAC with significantly less signal pins than required for GMII.
- Operate in both half and full duplex and at all port speeds.

Change History

Revision	Date	Description
1.8	April 27, 2005	Add shim to the PHY transmit datapath to suppress TX_ER when TX_EN is not asserted
1.7	July 20, 20001	Clarify data sampling and also the possible loss of the first byte of preamble.
1.6	Jan 4, 20001	Added specifications for Cisco Systems Intellectual Property.
1.5	Aug 4, 2000	Specified the data pattern for the beginning of the frame (preamble, SFD) for the frames sent from the PHY to make the PCS layer work properly.
1.4	June 30, 2000	Took out Jabber info, changed tx_Config_Reg[0] from 0 to 1 to make Auto-Negotiation work
1.3	April 17, 2000	Increased allowable input and output common mode range. The output high and low voltages were also increased appropriately. Added specification for output over/undershoot. Added note about AC coupling and clock recovery.
1.2	Feb 8, 2000	Added timing budget analysis and reduced LVDS input threshold to +/- 50 mV.
1.1	Nov 10, 1999	Incoporated Auto-Negotiation Process for update of link status
1.0	Oct. 14, 1999	Initial Release

Definitions

MII – Media Independent Interface: A digital interface that provides a 4-bit wide datapath between a 10/100 Mbit/s PHY and a MAC sublayer. Since MII is a subset of GMII, in this document, we will use the term “GMII” to cover all of the specification regarding the MII interface.

GMII – Gigabit Media Independent Interface: A digital interface that provides an 8-bit wide datapath between a 1000 Mbit/s PHY and a MAC sublayer. It also supports the 4-bit wide MII interface as defined in the IEEE 802.3z specification. In this document, the term “GMII” covers all 10/100/1000 Mbit/s interface operations.

Overview

SGMII uses two data signals and two clock signals to convey frame data and link rate information between a 10/100/1000 PHY and an Ethernet MAC. The data signals operate at 1.25 Gbaud and the clocks operate at 625 MHz (a DDR interface). Due to the speed of operation, each of these signals is realized as a differential pair thus providing signal integrity while minimizing system noise.

Figure 1 illustrates the simple connections in a system utilizing SGMII.

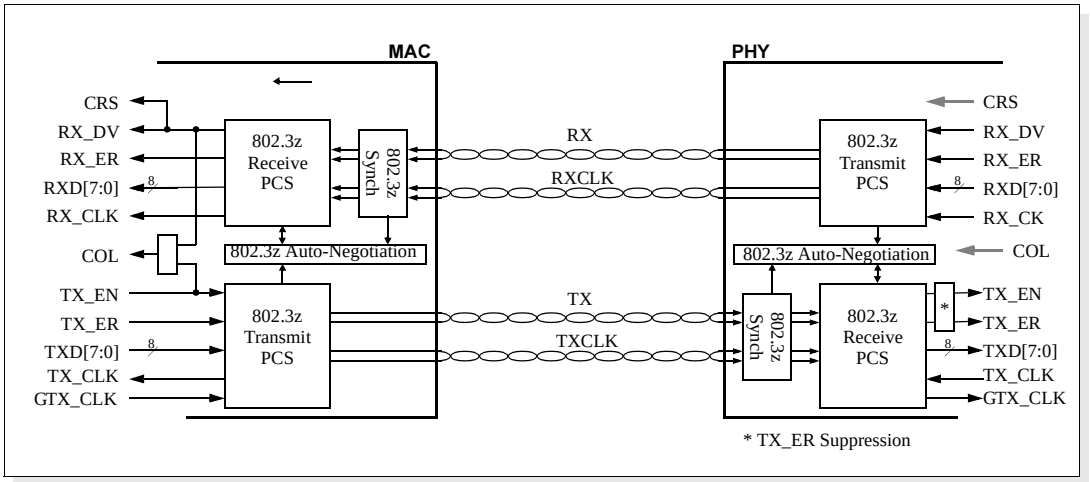

Figure 1 SGMII Connectivity

The transmit and receive data paths leverage the 1000BASE-SX PCS defined in the IEEE 802.3z specification (clause 36). The traditional GMII data signals (TXD/RXD), data valid signals (TX_EN/RX_DV), and error signals (TX_ER/RX_ER) are encoded, serialized and output with the appropriate DDR clocking. Thus it is a 1.25 Gbaud interface with a 625 MHz clock. Carrier Sense (CRS) is derived/inferred from RX_DV, and collision (COL) is logically derived in the MAC when RX_DV and TX_EN are simultaneously asserted. There is a small block in the PHY transmit path to suppress TX_ER in full duplex mode when TX_EN is not asserted.

Control information, as specified in Table 1, is transferred from the PHY to the MAC to signal the change of the control information. This is achieved by using the Auto-Negotiation functionality defined in Clause 37 of the IEEE Specification 802.3z. Instead of the ability advertisement, the PHY sends the control information via its tx_config_Reg[15:0] as specified in Table 1 whenever the control information changes. Upon receiving control information, the MAC acknowledges the update of the control information by asserting bit 14 of its tx_config_reg{15:0] as specified in Table 1.

SGMII details source synchronous clocking; however, specific implementations may desire to recover clock from the data rather than use the supplied clock. This operation is allowed; however, all sources of data must generate the appropriate clock regardless of how they clock receive data.

The link_timer inside the Auto-Negotiation has been changed from 10 msec to 1.6 msec to ensure a prompt update of the link status.

Bit Number	tx_config_Reg[15:0] sent from the PHY to the MAC	tx_config_Reg[15:0] sent from the MAC to the PHY
15	Link: 1 = link up, 0 = link down	0: Reserved for future use
14	Reserved for Auto-Negotiation acknowledge as specified in 802.3z	1
13	0: Reserved for future use	0: Reserved for future use
12	Duplex mode: 1 = full duplex, 0 = half duplex	0: Reserved for future use
11:10	Speed: Bit 11, 10: 1 1 = Reserved 1 0 = 1000 Mbps: 1000BASE-TX, 1000BASE-X 0 1 = 100 Mbps: 100BASE-TX, 100BASE-FX 0 0 = 10 Mbps: 10BASET, 10BASE2, 10BASE5	0: Reserved for future use
9:1	0: Reserved for future use	0: Reserved for future use
0	1	1

table 1 Definition of Control Information passed between links via tx_config_Reg[15:0]

Clearly, SGMII's 1.25 Gbaud transfer rate is excessive for interfaces operating at 10 or 100 Mbps. When these situations occur, the interface "elongates" the frame by replicating each frame byte 10 times for 100 Mbps and 100 types for 10 Mbps. This frame elongation takes place "above" the 802.3z PCS layer, thus the start frame delimiter only appears once per frame. The 802.3z PCS layer may remove the first byte of the "elongated" frame.

Implementation Specification

This section discusses how this SGMII interface shall be implemented by incorporating and modifying the PCS layer of the IEEE Specification 802.3z.

Signal Mapping at the PHY side

Figure 2 shows the PHY functional block diagram. It illustrates how the PCS layer shall be modified and incorporated at the PHY side in the SGMII interface.

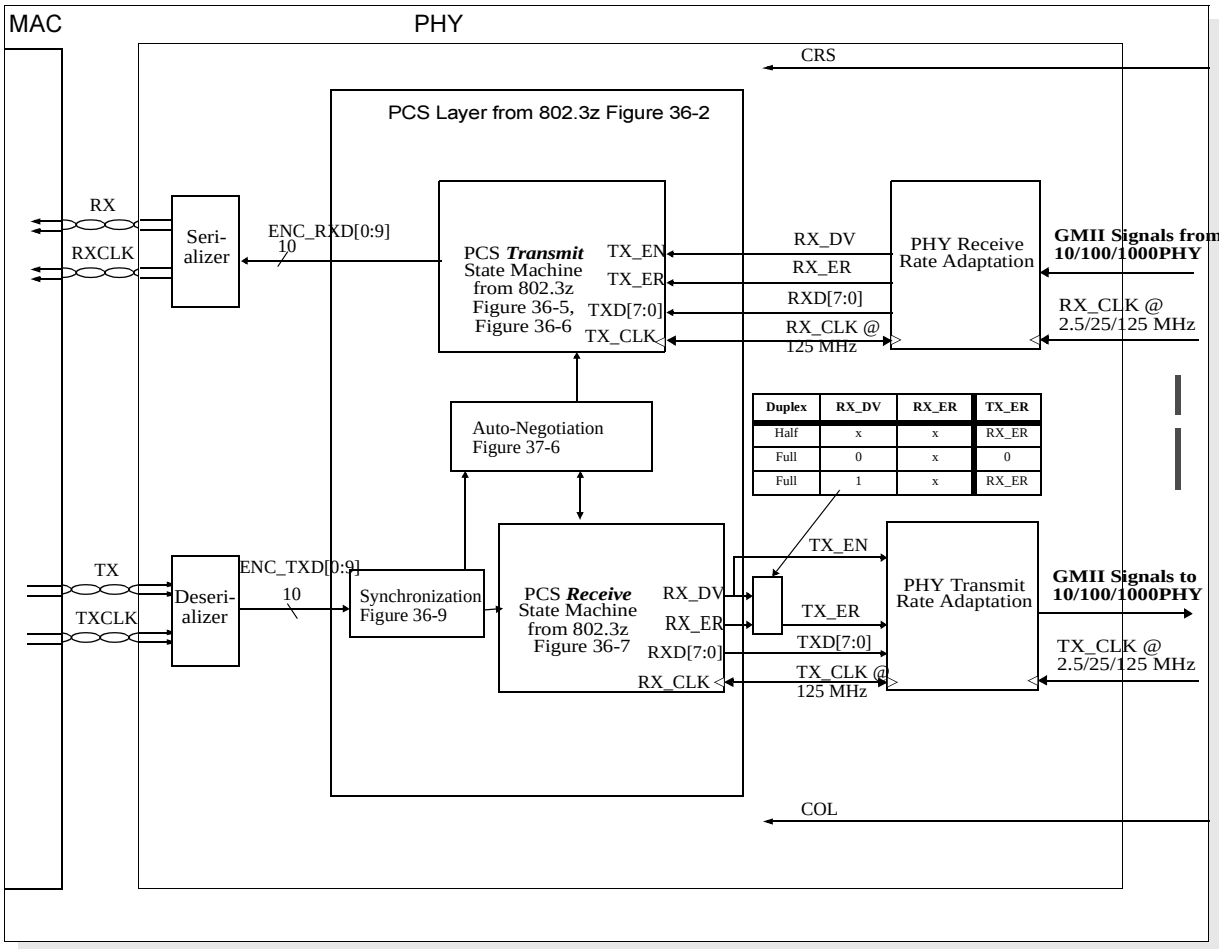

Figure 2 PHY Functional Block

At the receive side, GMII signals come in at 10/100/1000 Mbps clocked at 2.5/25/125 MHz. The PHY passes these signals through the PHY Receive Rate Adaptation to output the 8-bit data RXD[7:0] in 125MHz clock domain. RXD is sent to the PCS Transmit State Machine to generate an encoded 10-bit segment ENC_RXD[0:9]. The PHY serializes ENC_RXD[0:9] to create RX and sends it to the MAC at 1.25 Gbit/s data rate along with the 625 MHz DDR RXCLK.

At the transmit side, the PHY deserializes TX to recover encoded ENC_TXD[0:9]. The PHY passes ENC_TXD[0:9] through the PCS Receive State Machine to recover the GMII signals. In the mean time, Synchronization block checks ENC_TXD[0:9] to determine the synchronization status between links, and to realign if it detects the loss of synchronization.

The decoded GMII signals have to pass the PHY Transmit Rate Adaptation block to output data segments according to the PHY port speed.

To make the PCS layer from 802.3z work properly, the PHY must provide a frame beginning with at least two preamble symbols followed by a SFD symbol. To be more specific, at the beginning of a frame, RXD[7:0] in Figure 2 shall be *{8'h55, 8'h55, (8'h55.....), 8'hD5}* followed by valid frame data.

Some legacy end points have drop frames when RX_ER asserts during the first clock after a frame ends. The Receive PCS state machine generates this signalling at the end of certain frames. To avoid this problem, there is a small block in the PHY transmit path to surppress TX_ER in full duplex mode when RX_DV (from the Receive PCS state machine) is not asserted.

Signal Mapping at the MAC Side

Figure 3 shows the MAC functional block diagram. It illustrates how the PCS layer shall be modified and incorporated at the MAC side in the SGMII interface.

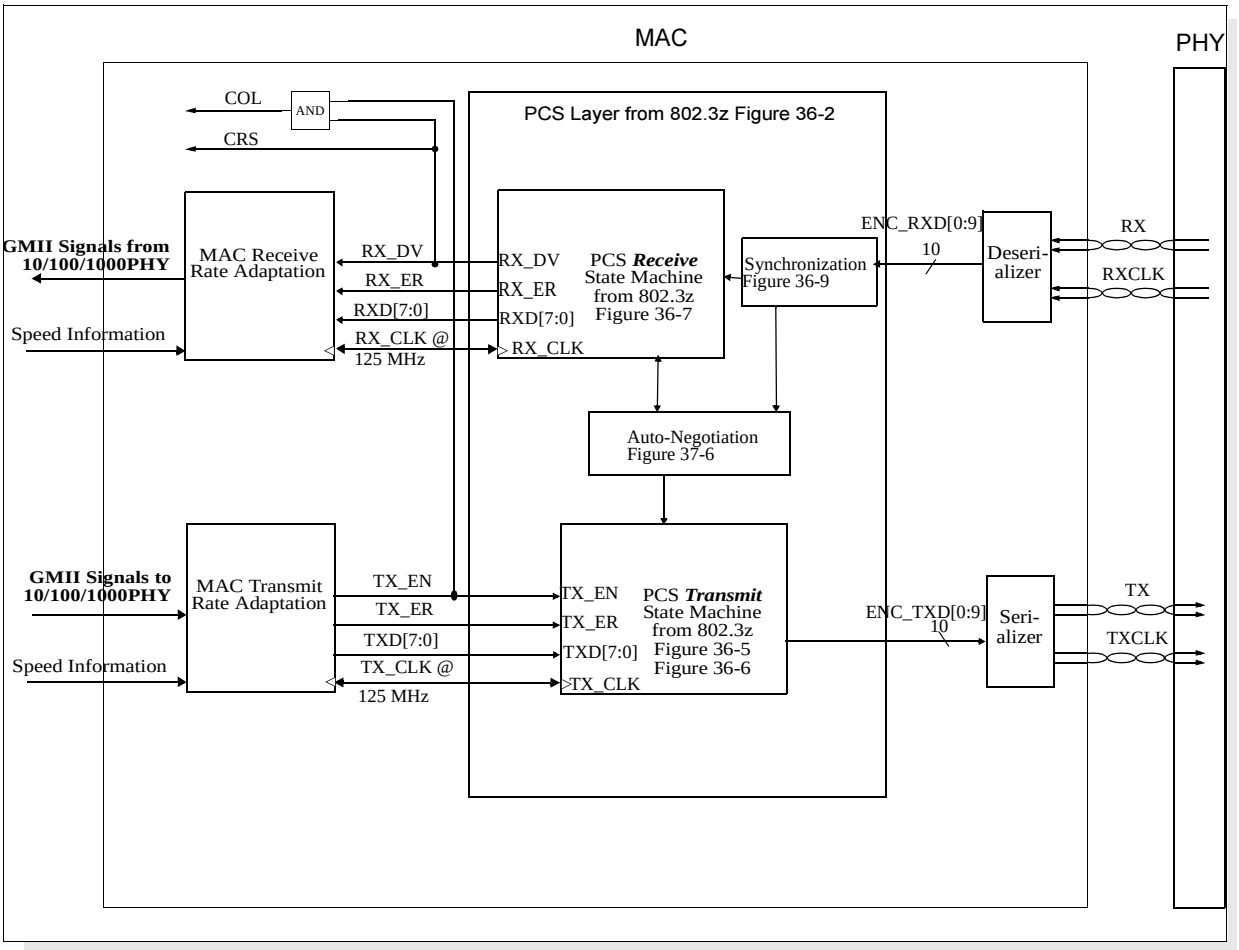

Figure 3 MAC Functional Block

At the receive side, the MAC deserializes RX to recover encoded ENC_RXD[0:9]. The MAC passes ENC_RXD[0:9] through the PCS Receive State Machine to recover the GMII signals. In the mean time, Synchronization block checks ENC_RXD[0:9] to determine the synchronization status between links, and to realign once it detects the loss of synchronization.

The decoded GMII signals have to pass the MAC Receive Rate Adaptation block to output data segments according to the PHY port speed, passed from the PHY to MAC via Auto-Negotiation process.

At the transmit side, GMII signals come in at 10/100/1000 Mbps data clocked at 2.5/25/125 MHz. The MAC passes these signals through the MAC Transmit Rate Adaptation to output the 8-bit data TXD[7:0] in 125MHz clock domain. TXD is sent to the PCS Transmit State Machine to generate an encoded 10-bit segment ENC_TXD[0:9]. The MAC serializes ENC_TXD[0:9] to create TX and sends it to the PHY at 1.25 Gbit/s data rate along with the 625 MHz DDR TXCLK.

Control Information Exchanged Between Links

As described in Overview, it is necessary for the PHY to pass control information to the MAC to notify the change of the link status. SGMII interface uses Auto-Negotiation block to pass the control information via tx_config_Reg[15:0].

If the PHY detects the control information change, it starts its Auto-Negotiation process, switching its Transmit block from “data” to “configuration” state and sending out the updated control information via tx_config_Reg[15:0]. The Receive block in the MAC receives and decodes control information, and starts the MAC’s Auto-Negotiation process. The Transmit block in the MAC acknowledges the update of link status via tx_config_Reg[15:0] with bit 14 asserted, as specified in Table 1. Upon receiving the acknowledgement from the MAC, the PHY completes the auto-negotiation process and returns to the normal data process.

As specified in Overview, inside the SGMII interface, the Auto-Negotiation link_timer has been changed from 10 msec to 1.6 msec, ensuring a prompt update of the link status. The expected latency for the update of link is 3.4 msec (two link_timer time + an acknowledgement process).

Data Information Transferred Between Links

Below we briefly describe at receive side how GMII signals get transferred across from the PHY and recovered at the MAC by using the 8B/10B transmission code. The same method applies to the transmit side.

According to the assertion and deassertion of RX_DV, the PHY encodes the Start_of_Packet delimiter (SPD /S/) and the End_Of_Packet delimiter (EPD) to signal the beginning and end of each packet. The MAC recovers RX_DV signal by detecting these two delimiters.

The PHY encodes the Error_Propagation(/V/) ordered_set to indicate a data transmission error. The MAC asserts RX_ER signal whenever it detects this ordered_set.

CRS is not directly encoded and passed to the MAC. To regenerate CRS, the MAC shall uses signal RX_DV before it is being passed to the MAC Receive Rate Adaptation block as shown in Figure 3.

The MAC decodes ENC_RXD[0:9] to recover RXD[7:0].

Figure 4 illustrates how the MAC samples data in 100 Mbit/s mode. As signals shown in Figure 2, the GMII data in 100 Mbit/s mode get replicated ten times after passing through the PHY Receive Rate Adaptation to generate RXD[7:0]. The modified PCS Transmit State Machine encodes RXD[7:0] to create ENC_RXD[0:9]. As noted in the Overview, the SPD(/S/) only appears once per frame. SAMPLE_EN is a MAC *internal* signal to enable the MAC sampling of data starting at the first data segment (/S/) once every ten data segments in 100 Mbit/s mode.

A note to Figure 4: there is *no* fixed boundary for the data sampling. Also the first byte of preamble might be only repeated 9/99 instead of 10/100 times due to the algorithm of the 802.3z PCS Transmit State Machine.

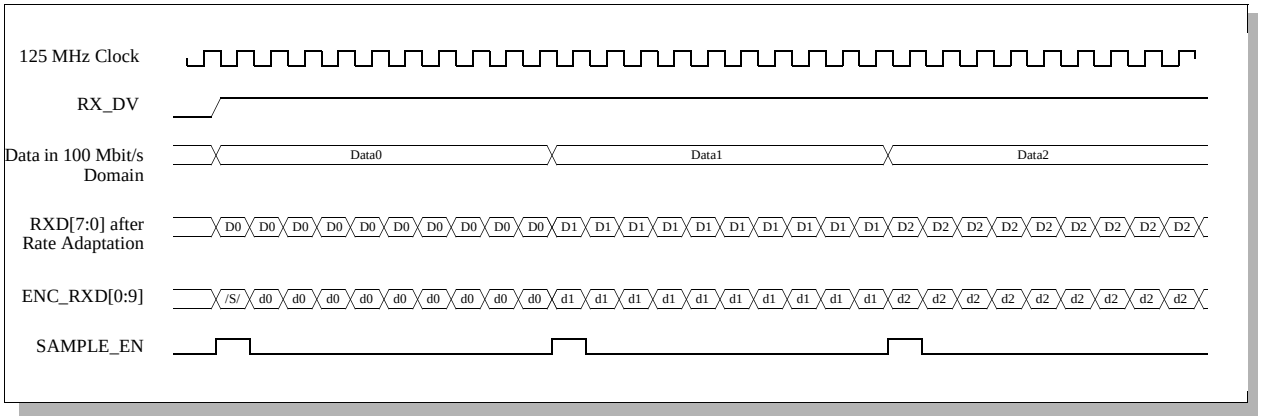

Figure 4 Data Sampling in 100 Mbit/s mode

LVDS AC/DC Specification

The basis of the LVDS and termination scheme can be found in IEEE1596.3-1996. Some parameters have been modified to accommodate the 1.25Gb/s requirements. SGMII consists of the most lenient DC parameters between the general purpose and reduced range LVDS.

Both the data and clock signals are DC balanced; therefore, implementations that meet the AC parameters but fail to meet the DC parameters may be AC coupled.

Figure 5 shows the DDR circuit at the source of the LVDS. The circuit passes data and clock with a 90 degree phase difference. The receiver samples data on both edges of the clock.

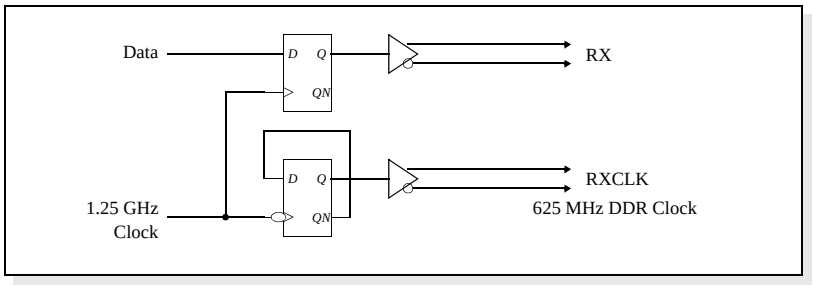

Figure 5 Reference data and clock circuit

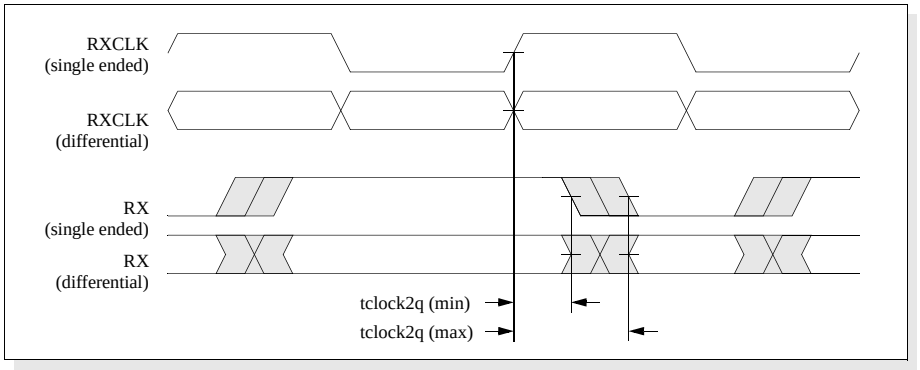

Figure 6 Driver Clock and Data Alignment

Symbol[a]	Parameter[b]	Min	Max	Units
Voh	Output voltage high,		1525	mV
Vol	Output voltage low	875		mV
Vring	Output ringing		10	%
|Vod|	Output Differential Voltage	150	400	mV
Vos	Output Offset Voltage	1075	1325	mV
Ro	Output impedance (single ended)	40	140	ohms
ΔRo	Mismatch in a pair		10	%
Δ|Vod|	Change in Vod between “0” and “1”		25	mV
ΔVos	Change in Vos between “0” and “1”		25	mV
Isa, Isb	Output current on Short to GND		40	mA
Isab	Output current when a, b are shorted		12	mA
Ixa, Ixb	Power off leakage current		10	mA

table 2 Driver DC specification

a. For a detailed description of the symbols please refer to the IEEE1596.3-1996 standard
b. All parameters measured at R_{load} = 100ohms +-1% load

Symbol	Parameter	Min	Max	Units
Vi	Input Voltage range a or b	675	1725	mV
Vidth	Input differential threshold	-50	+50	mV
Vhyst	Input differential hysteresis	25		mv
Rin	Receiver differential input impedance	80	120	ohms

table 3 Receiver DC specification

Symbol[a]	Parameter	Min	Max	Units
clock	Clock signal duty cycle @ 625MHz	48	52	%
t_{fall}	Vod fall time (20%-80%)	100	200	pSec
t_{rise}	Vod rise time (20%-80%)	100	200	pSec
t_{skew1}[b]	Skew between two members of a differential pair - $\|tp_{HLA}-tp_{LHB}\|$ or $\|tp_{LHA} - tp_{HLB}\|$		20	pSec
$t_{clock2q}$[c]	Clock to Data relationship: from either edges of the clock to valid data	250	550	pSec

table 4 Driver AC specification

a. For a detailed description of the symbols please refer to the IEEE1596.3-1996 standard
b. Skew measured at 50% of the transition
c. Skew measured at 0v differential

Symbol	Parameter	Min	Max	Units
t_{setup}[a]	setup time	100		pSec
t_{hold}	hold time	100		pSec

table 5 Receiver AC specification

a. Measured at 50% of the transition

Representative Timing Budget

A transmit and receive path timing budget provided in the table below. The exact allocation of the budget is implementation specific; however, verification of overall requirements are tested against the specifications in the tables external to the packaged integrated circuit.

Element	Value	Units
Effective clock period	800	ps
Cycle to cycle clock jitter	100	ps peak-peak
Imperfect duty cycle	30	ps peak-peak
Data dependent jitter	70	ps peak-peak
Static package skew	100	ps peak-peak
Remaining window	**500 (250 ps clock2q)**	**ps peak-peak**

table 6 Timing budget for driver requirements

Element	Value	Units
Driver window	500	ps peak-peak
Static package skew	100	ps peak-peak
Receiver setup time	100	ps peak-peak
Remaining window	**300**	**ps peak-peak**

table 7 Timing budget for receiver requirements

This budget shows the driver generating a data signal with a 500 ps eye centered around the sampling clock edge (see Figure 6 “Driver Clock and Data Alignment” on page 9). The receiver will add additional skew, leaving 300 ps of margin.

Cisco Systems Intellectual Property

Cisco Systems has released its proprietary rights to information contained in this document for the express purpose of implementation of this specification to encourage others to adopt this interface as an industry standard. Any company wishing to use this specification may do so if they will in turn relinquish their proprietary rights to information contained or referenced herein. Any questions concerning this release should be directed to the Robert Barr, World Wide Patent Counsel, Cisco Systems, 300 East Tasman Drive, San Jose, CA.