POS-PHY™ LEVEL 4

A SATURN PACKET AND CELL INTERFACE SPECIFICATION FOR OC192 SONET/SDH AND 10 GIGABIT ETHERNET

ISSUE 6: FEBRUARY 2001

CONTENTS

LIST OF TABLES

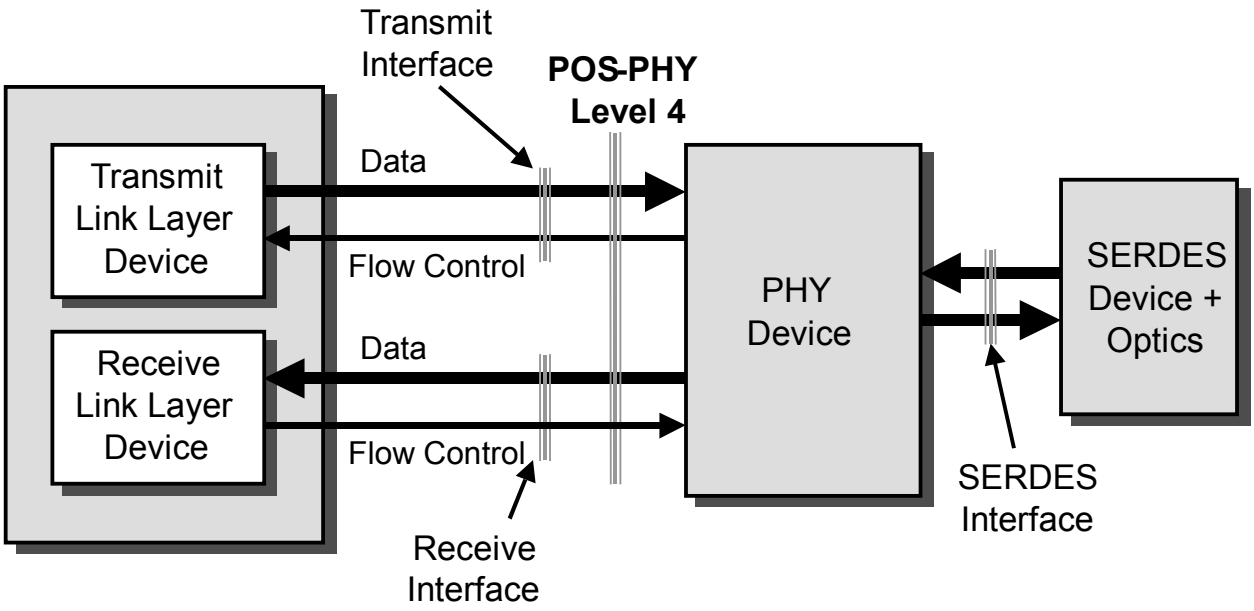
LIST OF FIGURES

1 INTRODUCTION

POS-PHY Level 4 is an interface for packet and cell transfer between a physical layer (PHY) device and a link layer device, for aggregate bandwidths of OC-192 ATM and Packet over SONET/SDH (POS), as well as 10 Gigabit Ethernet applications. This section provides a general overview of the interface. The next section contains more detailed descriptions of the signals and associated operations, data structures, start-up and AC timing parameters. Appendices A through E have been included to provide guidance on implementations; they are for information only and are not normative parts of this specification.

The following is a general synopsis of the POS-PHY Level 4 interface. For reference, a general block diagram is shown in Figure 1.1. POS-PHY Level 4 is the system packet interface for data transfer between the link layer and the PHY device; it is designed to meet requirements of this particular application, although it may be used in other applications as well. “Transmit” and “Receive” refer, respectively, to data flow and associated control/status information for the Link Layer to PHY, and the PHY to Link Layer directions.

Figure 1.1 System Reference Model.



On both the transmit and receive interfaces, FIFO status information is sent separately from the corresponding data path. By taking FIFO status information out-of-band, it is possible to decouple the transmit and receive interfaces so that each operates independently of the other. Such an arrangement makes POS-PHY L4 suitable not only for bidirectional but also for unidirectional link layer devices.

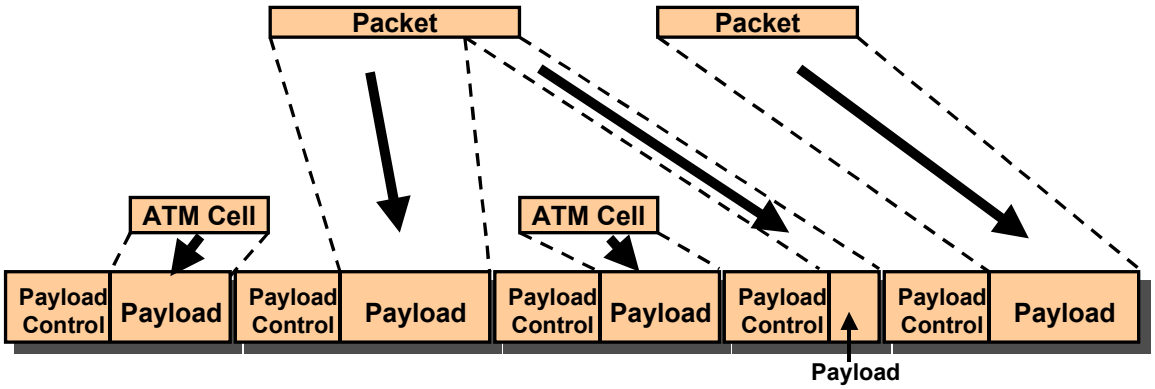
In both the transmit and receive interfaces, the packet's address, delineation information and error control coding is sent in-band with the data.

POS-PHY L4 has the following general characteristics:

- Point-to-point connection (i.e., between single PHY and single Link Layer device).
- Support for 256 ports (suitable for STS-1 granularity in SONET/SDH applications (192 ports) and Fast Ethernet granularity in Ethernet applications (100 ports)).
- Transmit / Receive Data Path:
    - 16 bits wide.
    - In-band port address, start/end-of-packet indication, error-control code.
    - LVDS I/O (IEEE 1596.3 – 1996 [1], ANSI/TIA/EIA-644-1995 [2]).
    - 622 Mb/s minimum data rate per line.
    - Source-synchronous double-edge clocking, 311 MHz minimum.
- Transmit / Receive FIFO Status Interface:
    - Maximum 1/4 data path clock rate for LVTTL I/O.
    - CMOS LVTTL I/O (3.3V) or optional LVDS I/O.
    - 2-bit parallel FIFO status indication.
    - In-band Start-of-FIFO Status signal.
    - Source-synchronous clocking.

Segmentation of data is in multiples of 16 bytes (for example, 48, 128, or 528-byte segments) with the exception of transfers that terminate with an EOP. Information associated with each transfer (port address, start/end-of-packet indication and error-control coding) is sent in 16-bit control words described later in this document. Figure 1.2 shows how ATM cells and variable-length packets map onto the data stream.

Figure 1.2 Mapping of packets and ATM cells onto payload stream.



2 INTERFACE DESCRIPTION

Section 2.1 contains signal definitions for the transmit and receive directions. Section 2.2 describes the signal operation along with the data structures for payload data and in-band control/status information. Section 2.3 describes start-up parameters. Sections 2.4 and 2.5 specify AC timing and DC parameters.

2.1 Signals

A block diagram depicting the interface signals is shown in Figure 2.1. The transmit and receive data paths include, respectively, (TDCLK, TDAT[15:0], TCTL) and (RDCLK, RDAT[15:0], RCTL). The transmit and receive FIFO status channels include (TSCLK, TSTAT[1:0]) and (RSCLK, RSTAT[1:0]) respectively. Table 1 provides an interface signal summary.

Figure 2.1 POS-PHY L4 Interface

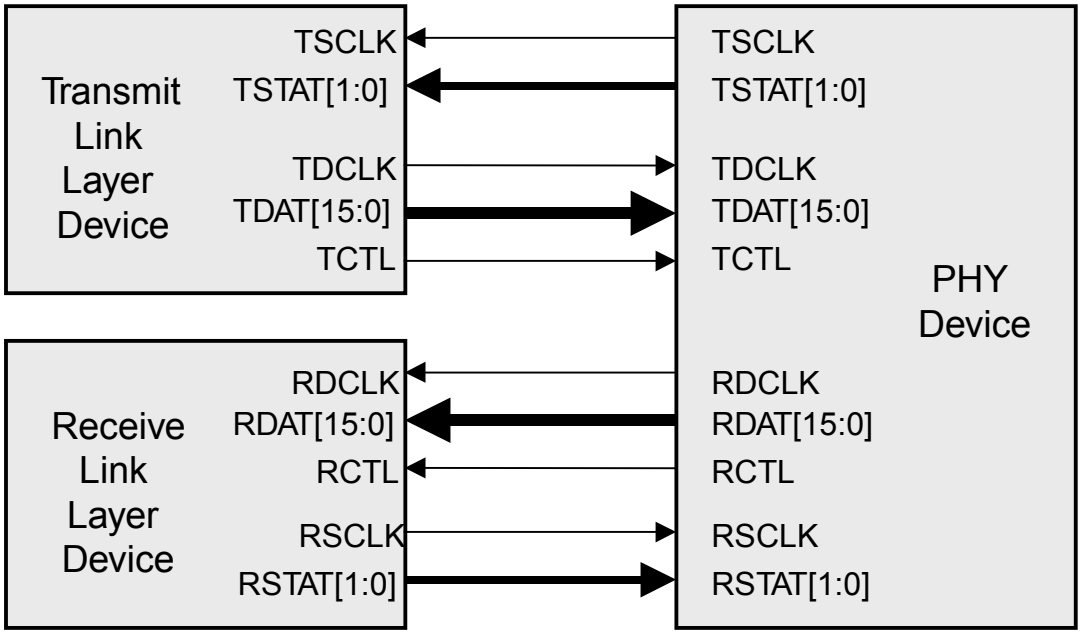


Table 1 POS-PHY Level 4 Interface Signal Summary.

Signal	Direction	Description
TDCLK	Link to PHY	Transmit Data Clock. Clock associated with TDAT and TCTL. Data and control lines are driven off the rising and falling edges of the clock.
TDAT [15:0]	Link to PHY	Transmit Data. Used to carry payload data and in-band control words from the Link Layer to the PHY device. The control word format is described in Section 2.2.
TCTL	Link to PHY	Transmit Control. TCTL is high when a control word is present on TDAT[15:0]. It is low otherwise.
TSCLK	PHY to Link	Transmit Status Clock. Clock associated with TSTAT.
TSTAT [1:0]	PHY to Link	Transmit FIFO Status. Used to carry round-robin FIFO status information, along with associated error detection and framing.
RDCLK	PHY to Link	Receive Data Clock. Clock associated with RDAT and RCTL. Data and control lines are driven off the rising and falling edges of the clock.

Signal	Direction	Description
RDAT [15:0]	PHY to Link	Receive Data. Carries payload data and in-band control from the PHY to the Link Layer device. The control word format is described in Section 2.2.
RCTL	PHY to Link	Receive Control. RCTL is high when a control word is present on RDAT[15:0]. It is low otherwise.
RSCLK	Link to PHY	Receive Status Clock. Clock associated with RSTAT.
RSTAT [1:0]	Link to PHY	Receive FIFO Status. Used to carry round-robin FIFO status information, along with associated error detection and framing.

2.2 Interface Operation and Data Structures

2.2.1 Data Path

Complete packets or shorter bursts may be transferred, as shown in Figure 1.2. The maximum configured payload data transfer size must be a multiple of 16 bytes. Control words are inserted only between burst transfers; once a transfer has begun, data words are sent uninterrupted until end-of-packet or a multiple of 16 bytes is reached, whichever comes first. The interval between the end of a given transfer and the next payload control word (marking the start of another transfer) consists of zero or more idle control words and training patterns (Section 2.2.3).

The minimum and maximum supported packet lengths are determined by the application. For ease of implementation however, successive start-of-packets must occur not less than 8 cycles apart, where a cycle is one control or data word. The gap between shorter packets is filled with idle control words.

Figure 2.2 shows cycle-by-cycle behavior of the data path for valid state transitions. The states correspond to the type of words transferred on the data path. Transitions from the “Data Burst” state (to “Payload Control” (if data is available for transfer) or “Idle Control” (otherwise)) are possible only on integer multiples of 8 cycles (corresponding to multiple of 16 byte segmentation) or upon end-of-packet. A data burst must follow a payload control word immediately on the next cycle. Arcs not annotated correspond to single cycles.

Figure 2.2 Data Path State Diagram.

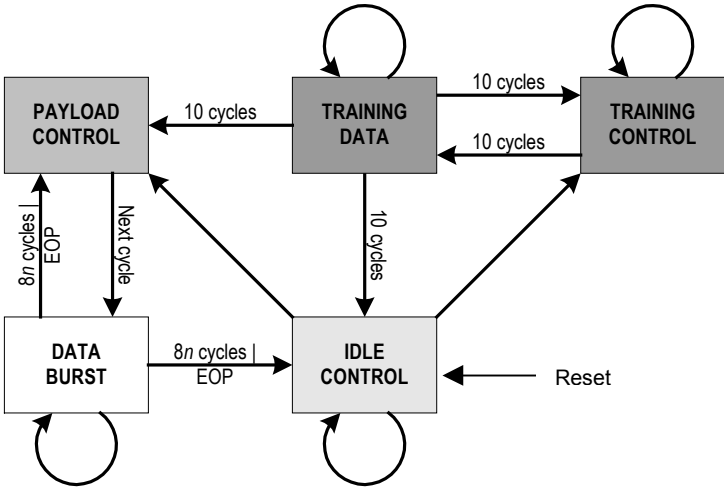
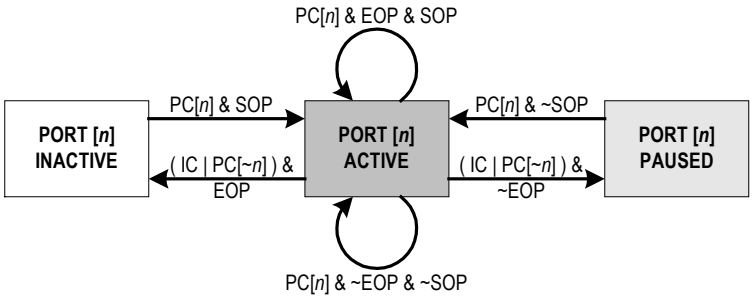


Figure 2.3 shows per-port state transitions at control word boundaries. At any given time, a given port may be active (sending data), paused (not sending data, but pending the completion of an outstanding packet), or inactive (not sending data, no outstanding packet).

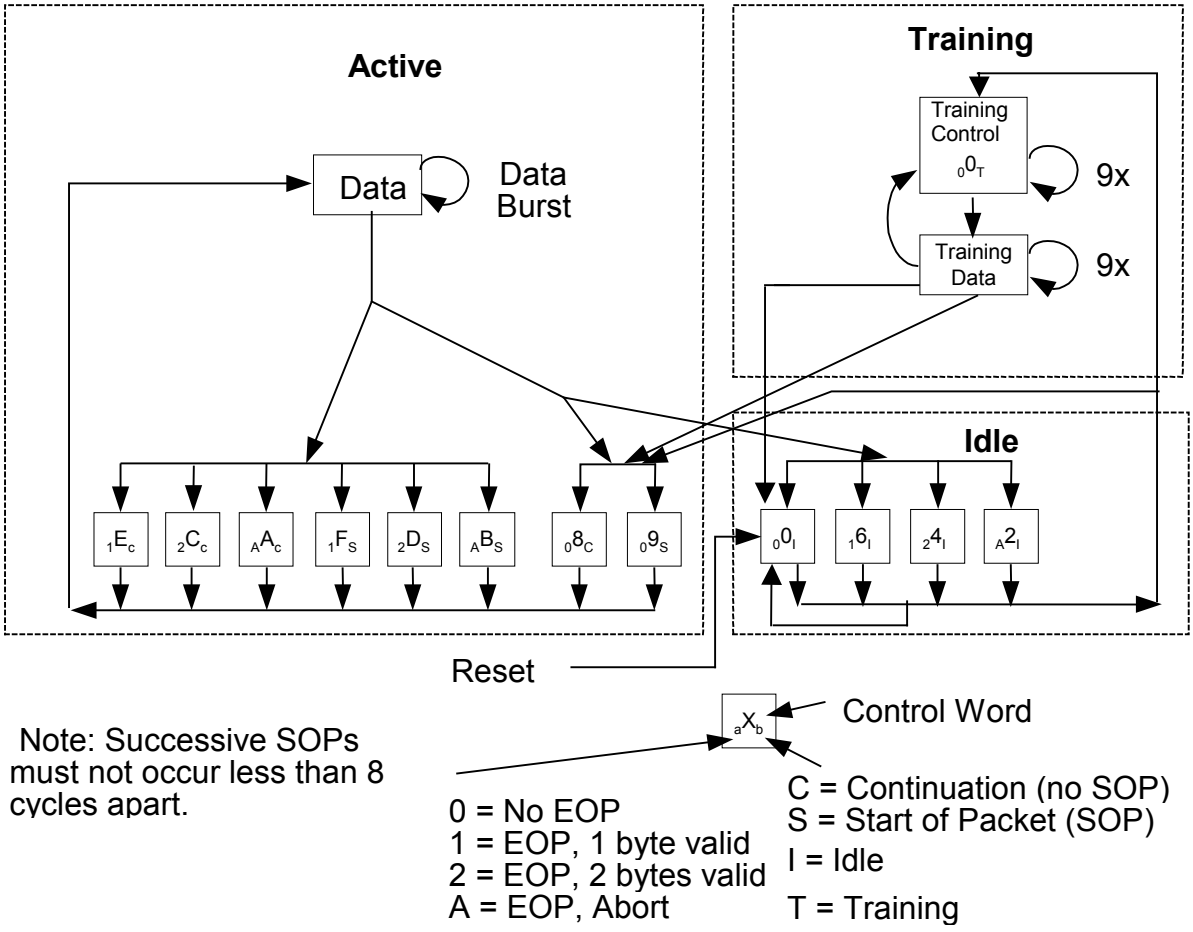
Figure 2.3 Per-Port State Diagram with Transitions at Control Words.



KEY:	
IC:	Idle Control Word
PC[*n*]:	Payload Control Word for port *n*
PC[~*n*]:	Payload Control Word for a port other than port *n*
SOP:	Start-of-Packet in Payload Control Word
~SOP:	No Start-of-Packet in Payload Control Word
EOP:	End-of-Packet in Control Word
~EOP:	No End-of-Packet in Control Word
&	*AND*
|	*OR*

Figure 2.4 shows valid sequences of data/control words for both “data burst” transfers and “idle” or “training” intervals. Valid control word encodings on bit positions [15:12] are shown on the bottom row. The left subscript indicates EOPS status (of the preceding transfer), while the right subscript indicates SOP status (of the next transfer).

Figure 2.4 Sequences of Data / Control Words.



Payload data bytes are transferred over the interface in the same order as they would be transmitted or received on the line side. The order of bytes within a word is shown in Figure 2.5 and Figure 2.6 (for transfer of arbitrary packets that end on odd and even byte boundaries, respectively). The most significant bits (MSBs) of the constituent bytes correspond to bits 15 and 7. The order in which the bits within each byte are transmitted and received on the line is defined by the Physical Layer and is beyond the scope of this specification. On payload transfers that do not end on an even byte boundary, the unused byte (after the last valid byte on bit positions 7 through 0) is set to all zeroes, as shown in Figure 2.5.

Figure 2.5 Example of Payload Transfer Data Structure (43-byte packet, * = set to all zeroes).

	Bit 15 Bit 8	Bit 7 Bit 0
Data Word 1	Byte 1	Byte 2
Data Word 2	Byte 3	Byte 4
Data Word 3	Byte 5	Byte 6
Data Word 4	Byte 7	Byte 8
Data Word 21	Byte 41	Byte 42
Data Word 22	Byte 43	XX*

Figure 2.6 Example of Payload Transfer Data Structure (52-byte packet).

	Bit 15 Bit 8	Bit 7 Bit 0
Data Word 1	Byte 1	Byte 2
Data Word 2	Byte 3	Byte 4
Data Word 3	Byte 5	Byte 6
Data Word 4	Byte 7	Byte 8
Data Word 25	Byte 49	Byte 50
Data Word 26	Byte 51	Byte 52

A common control word format is used in both the transmit and receive interfaces. Table 2 describes the fields in the control word. When inserted in the data path, the control word is aligned such that its MSB is sent on the MSB of the transmit or receive data lines. A payload control word that separates two adjacent burst transfers contains status information pertaining to the previous transfer and the following transfer.

Table 3 shows a list of control word types. Table 4 shows some examples of valid control words. Usage of reserved bits is beyond the scope of this specification. The transmitter shall not send control words with Reserved bit patterns; receivers may ignore these control words and may optionally report an error condition when such control words are observed.

Table 2 Description of Fields in the Control Words.

Bit Position	Label	Description
15	Type	Control Word Type. Set to either of the following values: 1: payload control word (payload transfer will immediately follow the control word). 0: idle or training control word (otherwise).
14:13	EOPS	End-of-Packet (EOP) Status. Set to the following values below according to the status of the immediately preceding payload transfer. 0 0: Not an EOP. 0 1: EOP Abort (application-specific error condition). 1 0: EOP Normal termination, 2 bytes valid. 1 1: EOP Normal termination, 1 byte valid. EOPS is valid in the first control word following a burst transfer. It is ignored and set to “0 0” otherwise.
12	SOP	Start-of-Packet.

Bit Position	Label	Description
		Set to 1 if the payload transfer immediately following the control word corresponds to the start of a packet. Set to 0 otherwise. Set to 0 in all idle and training control words.
11:4	ADR	Port Address. 8-bit port address of the payload data transfer immediately following the control word. None of the addresses are reserved (all are available for payload transfer). Set to all zeroes in all idle control words. Set to all ones in all training control words.
3:0	DIP-4	4-bit Diagonal Interleaved Parity. 4-bit odd parity computed over the current control word and the immediately preceding data words (if any) following the last control word.

Table 3 Control Word Type List.

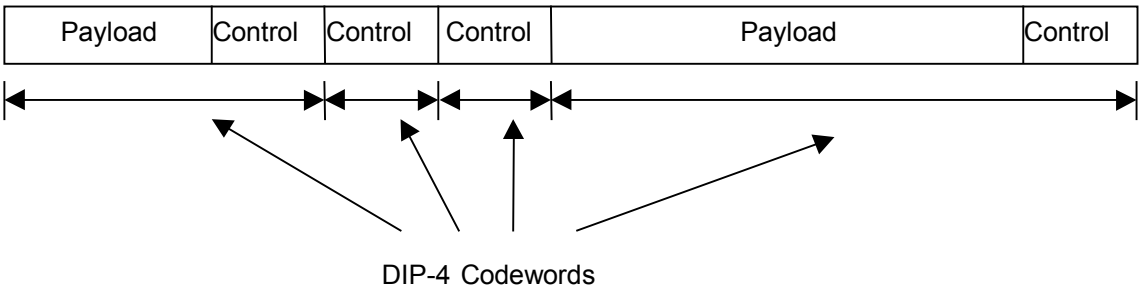
	Bit [15:12]	Prior Word Status (if Data)	Next Word Status	Meaning
0	0000	Not EOP	Control	Idle (not EOP), or training control word
1	0001	Reserved	Reserved	Reserved
2	0010	EOP w/ abort	Control	Idle, Abort last packet
3	0011	Reserved	Reserved	Reserved
4	0100	EOP w/ 2 bytes	Control	Idle, EOP with 2 bytes valid
5	0101	Reserved	Reserved	Reserved
6	0110	EOP w/ 1 byte	Control	Idle, EOP with 1 byte valid
7	0111	Reserved	Reserved	Reserved
8	1000	Not EOP	Data	Payload, no EOP, no SOP
9	1001	Not EOP	Data (SOP)	Payload, no EOP, SOP
A	1010	EOP w/ abort	Data	Payload, abort last packet, no SOP
B	1011	EOP w/ abort	Data (SOP)	Payload, abort last packet, SOP
C	1100	EOP w/ 2 bytes	Data	Payload, EOP with 2 bytes valid, no SOP
D	1101	EOP w/ 2 bytes	Data (SOP)	Payload, EOP with 2 bytes valid, SOP
E	1110	EOP w/ 1 byte	Data	Payload, EOP with 1 byte valid, no SOP
F	1111	EOP w/ 1 byte	Data (SOP)	Payload, EOP with 1 byte valid, SOP

Table 4 Some Examples of Control Words.

Control Word	Description
0110 00000000 abcd	Idle control word following end of transfer. End-of-packet, normal termination, 1 byte valid in last data word. (Note: abcd bits depend on contents of this control word and preceding transfer.)
0000 00000000 1111	Idle control word preceded by another (idle) control word.
1101 00000101 abcd	Payload control word following end of transfer. End-of-packet, normal termination, 2 bytes valid in last data word of preceding transfer (abcd bits depend on contents of this control word and preceding transfer). Start-of-packet in next transfer to port 5.

Figure 2.7 shows the range over which the DIP-4 parity bits are computed. In the presence of random errors, the DIP-4 code offers the same error protection capability as a comparable BIP code, but has an additional advantage of spreading single-column errors (as might occur in a single defective line) across the parity bits. Appendix A discusses the error detection performance of this code.

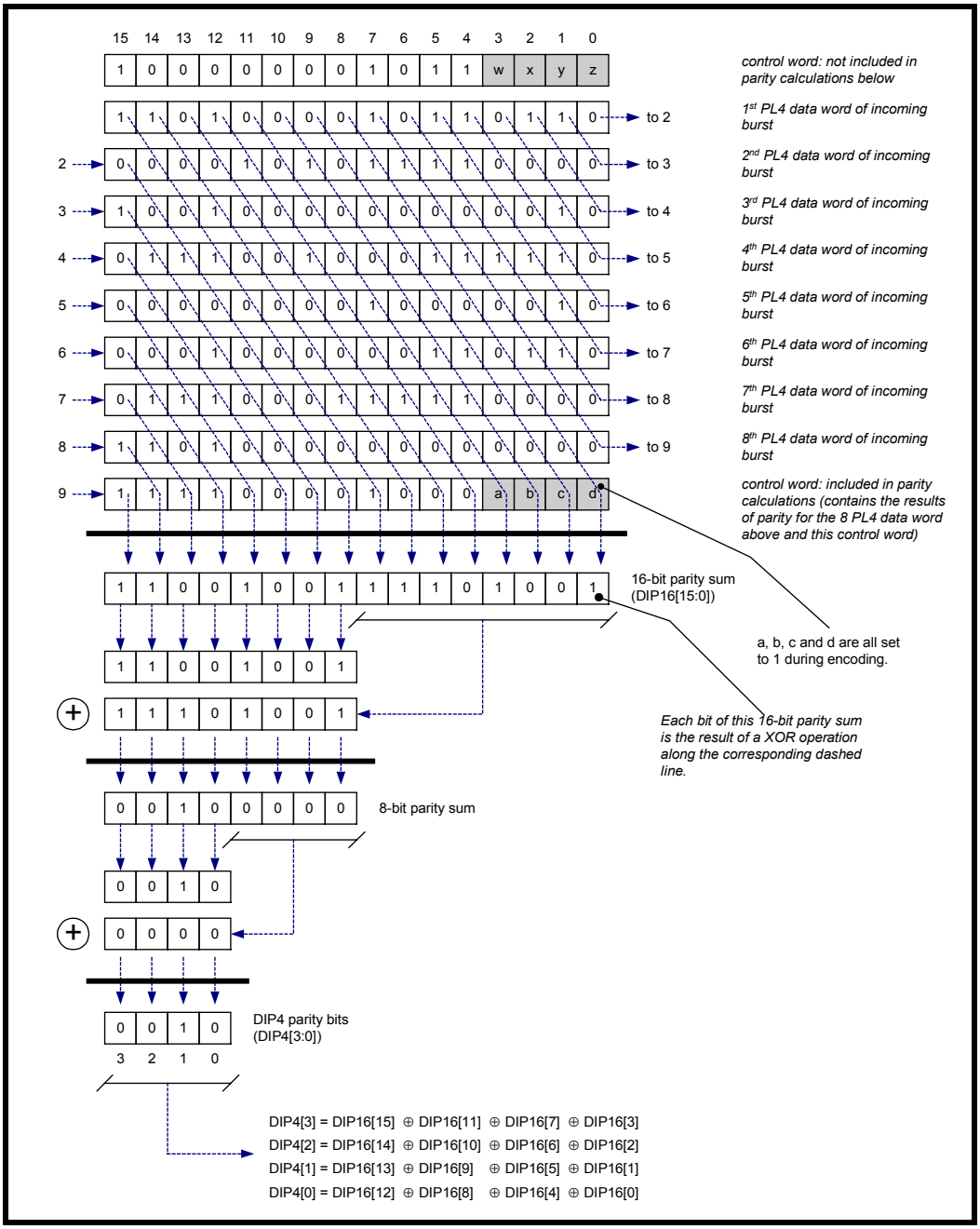
Figure 2.7 Extent of DIP-4 Code-words.



A functional description of calculating the DIP-4 code is given as follows. Assume that the stream of 16-bit data words are arranged as shown Figure 2.8, MSB at the leftmost column, time moving downward. (The first word received is at the top of the figure; the last word is at the bottom of the figure.) The parity bits are generated by summing diagonally (in the control word, the space occupied by the DIP-4 code (bits a,b,c,d) is set to all 1's during encoding). The first 16-bit checksum is split into two bytes, which are added to each other modulo-2 to produce an 8-bit checksum. The 8-bit checksum is then divided into two 4-bit nibbles, which are added to each other modulo-2 to produce the final DIP-4

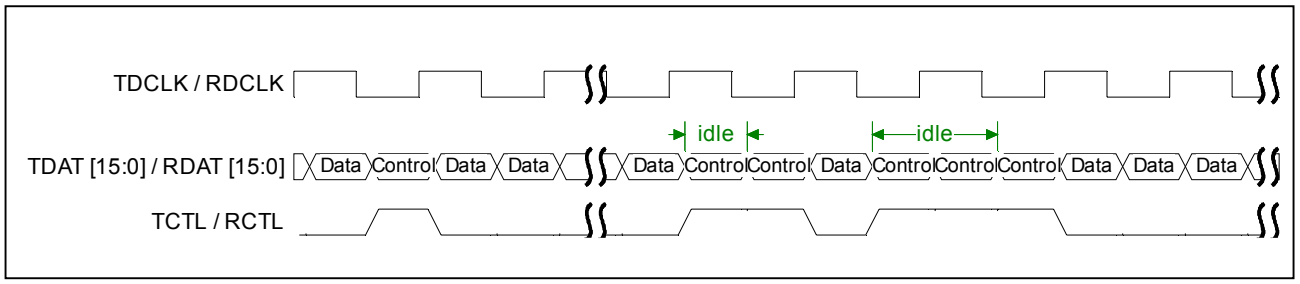
code. The procedure described applies to either parity generation on the egress path or to check parity on the ingress path.

Figure 2.8 Example of DIP-4 Encoding (Odd Parity).



A timing diagram of the data path signals is shown in Figure 2.9. This diagram is applicable to either the transmit or the receive interface. TCTL/RCTL is high when TDAT/RDAT contain control words. Idle periods correspond to back-to-back control words.

Figure 2.9 Data Path Timing Diagram.



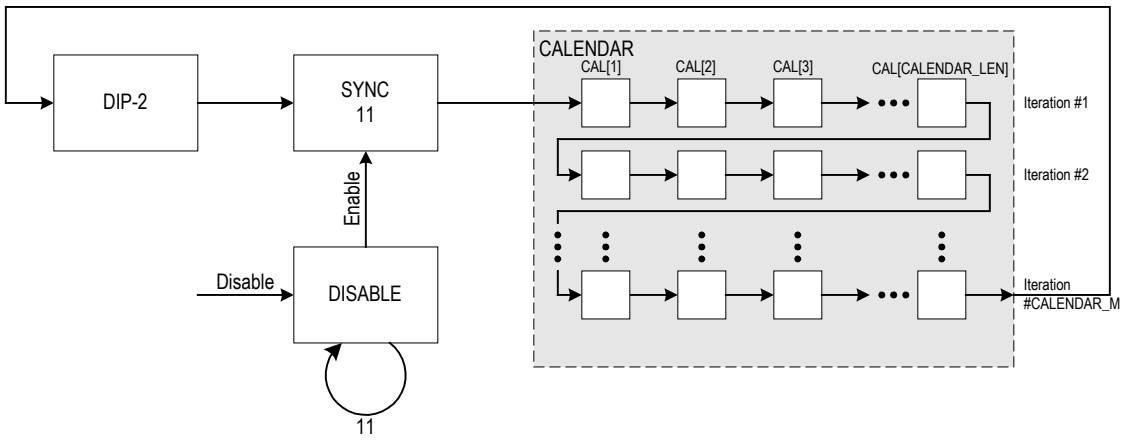
The actual clock rate used in practice is determined by the application at hand. A discussion of the minimum data clock rates required for certain applications is given in Appendix B.

2.2.2 FIFO Status Channel.

FIFO status information is sent periodically over the TSTAT link from the PHY to the Link Layer device, and over the RSTAT link from the Link Layer to the PHY device. Implementation of the FIFO status channel for the transmit interface is mandatory; the corresponding implementation for the receive interface is optional. If both status channels are implemented, they shall operate independently of each other.

Figure 2.10 shows the operation of the FIFO status channel. The sending side of the FIFO status channel is initially in the DISABLE state and sends the “1 1” pattern repeatedly. When FIFO status transmission is enabled, there is a transition to the SYNC state and the “1 1” framing pattern is sent. FIFO status words are then sent according to the calendar sequence (of length CALENDAR_LEN), repeating the sequence CALENDAR_M times followed by the DIP-2 code. FIFO status reporting may be disabled.

Figure 2.10 FIFO Status State Diagram (Sending Side).



The FIFO status of each port is encoded in a 2-bit data structure, whose format is defined in Table 5. The most significant bit of each port status is sent over TSTAT / RSTAT [1], while the least significant bit is sent over TSTAT / RSTAT [0]. Section 2.3 discusses configuration of the FIFO status port sequence. The port sequences on the transmit and receive interfaces may be configured differently from each other. The “1 1” pattern is reserved for in-band framing; it must be sent once prior to the start of the FIFO status sequence.

A DIP-2 odd parity checksum is sent at the end of each complete sequence, immediately before the “1 1” framing pattern. The DIP-2 code is computed over all preceding FIFO status indications sent after the last “1 1” framing pattern, diagonally over TSTAT / RSTAT [1] and TSTAT / RSTAT [0], as shown in Figure 2.11. The first word is at the top of the figure, while the last word is at the bottom of the figure. The parity bits are computed by summing diagonally. Bits a and b in line 9 correspond to the space occupied by the DIP-2 parity bits; these bits are set to 1 during encoding. Note also that the “1 1” framing pattern is not included in the parity calculation. The procedure described applies to either parity generation on the egress path or to check parity on the ingress path.

While the parity bits can mimic the “1 1” pattern, the receiving end can still frame successfully by syncing onto the last cycle in a repeated “1 1” pattern. A timing diagram of the FIFO status channel is shown in Figure 2.12.

To permit more efficient FIFO utilization, the MaxBurst1 and MaxBurst2 credits are granted and consumed in increments of 16-byte blocks. For any given port, these credits correspond to the most recently received FIFO status. They are not cumulative and supersede previously granted credits for the given port. A burst transfer shorter than 16 bytes (e.g., end-of-packet fragment) will consume an entire 16-byte credit.

A continuous stream of repeated “1 1” framing patterns indicates a disabled status link. For example, it may be sent to indicate that data path deskew (Section 2.2.3) has not yet been completed or confirmed. When a repeated “1 1” pattern is detected, all outstanding credits are cancelled and set to zero.

Figure 2.11 Example of DIP-2 Encoding (Odd Parity).

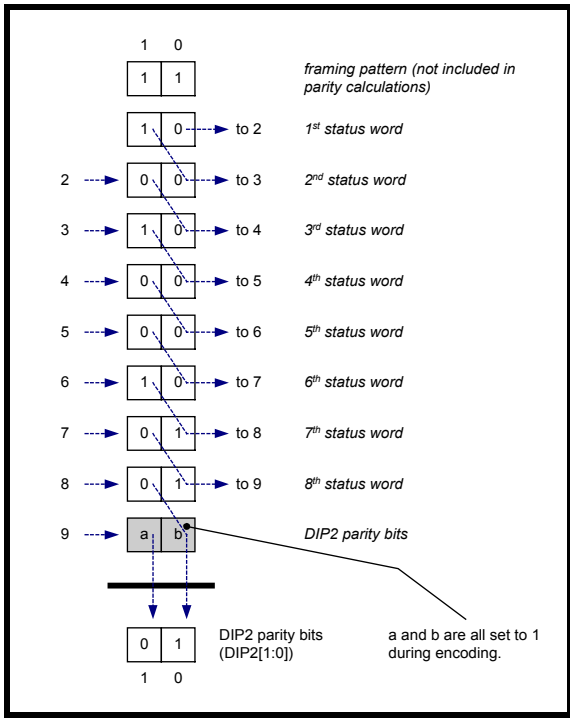


Figure 2.12 FIFO Status Channel Timing Diagram.

(Note: round-robin period = framing + (CALENDAR_LEN * CALENDAR_M) + DIP-2.)

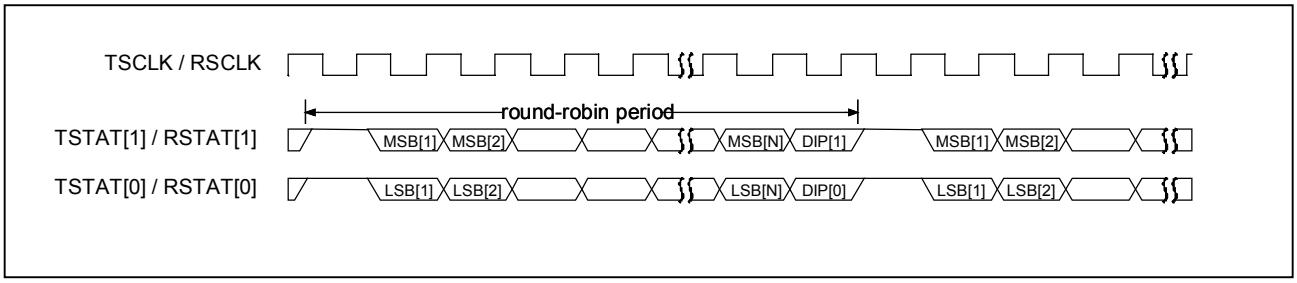


Table 5 FIFO Status Format.

MSB	LSB	Description
1	1	Framing pattern or disabled link indication.
1	0	SATISFIED Indicates that the corresponding port's FIFO is almost full. When SATISFIED is received, only transfers using the remaining previously granted 16-byte blocks (if any) may be sent to the corresponding port until the next status update. No additional transfers to that port are permitted while SATISFIED is indicated.
0	1	HUNGRY When HUNGRY is received, transfers for up to MaxBurst2 16-byte blocks or the remainder of what was previously granted (whichever is greater) may be sent to the corresponding port until the next status update.
0	0	STARVING Indicates that buffer underflow is imminent in the corresponding PHY port. When STARVING is received, transfers for up to MaxBurst1 16-byte blocks may be sent to the corresponding port until the next status update.

The indicated FIFO status is based on the latest available information. A STARVING indication provides additional feedback information, so that transfers can be scheduled accordingly. Applications which do not need to distinguish between HUNGRY and STARVING, may only examine the most significant FIFO status bit.

A further discussion on the required clock rate for the FIFO Status Channel is given in Appendix C.

Higher Bandwidth Operation

Applications that require higher FIFO Status Channel bandwidths than feasible with LVTTL I/O, may optionally use LVDS I/O instead. If LVDS I/O is used, double-edge clocking is used on TSCLK and RSCLK, running at the same rate as the corresponding data path rate. The framing structure and operation of TSTAT[1:0] and RSTAT[1:0] remain unchanged. AC timing parameters for an LVDS FIFO Status Channel are defined in Section 2.4.1.

A training sequence is scheduled to be sent at least once every preconfigured bounded interval (FIFO_MAX_T) on both the transmit and receive FIFO Status interfaces. These training sequences may be used by the receiving end of each interface for deskewing bit arrival times on the FIFO status and control lines.[1] Training sequences may only be inserted between the DIP-2 code of the preceding cycle and the SYNC (1 1) word of the next calendar sequence. The training sequence consists of α repetitions of the training pattern, each of which consists of ten words of “0 0” followed by ten words of “1 1”.

The training sequence is chosen so that it can be distinguished from a valid FIFO Status channel message, and so that the receiving end can correct for relative skew differences of up to +/- 1 bit time. In the absence of bit errors in the training pattern, a receiver should be able to successfully deskew the FIFO Status lines with one training pattern.

Setting FIFO_MAX_T equal to zero will disable the training sequence.

2.2.3 Training Sequence for Data Path Deskew

A training sequence is scheduled to be sent at least once every preconfigured bounded interval (DATA_MAX_T) on both the transmit and receive data paths.

[1] Some specifications based on POS-PHY L4 may require CALENDAR_LEN * CALENDAR_M to be not less than 16 when training sequences are used. POS-PHY L4 does not impose this restriction.

These training sequences may be used by the receiving end of each interface for deskewing bit arrival times on the data and control lines. The sequence defined in this section is designed to allow the receiving end to correct for relative skew differences of up to +/- 1 bit time. The training sequence consists of 1 idle control word followed by one or more repetitions of a 20-word training pattern consisting of 10 (repeated) training control words and 10 (repeated) training data words. The initial idle control word removes dependencies of the DIP-4 in the training control words from preceding data words. Assuming a maximum of +/- 1 bit time in bit alignment jitter on each line, and a maximum of +/- 1 bit time relative skew between lines, there will be at least 8 bit times during which a receiver can detect a training control word prior to deskew. The training data word is chosen to be the complement of the training control word. In the absence of bit errors in the training pattern, a receiver should be able to successfully deskew the data and control lines with one training pattern.

The sending side of the data path on both the transmit and receive interfaces must schedule the training sequence in Table 6 (from cycles 1 through $20\alpha+1$) at least once every DATA_MAX_T cycles, where DATA_MAX_T and α are configurable. Training sequences at the transmit and receive interfaces are scheduled independently. They must not be inserted within a payload burst transfer (i.e., not inserted between a payload control word and any of the subsequent data words until the end of transfer). The training sequence may be optionally disabled by setting DATA_MAX_T equal to zero. Note that the DIP-4 code of the first control word following the training pattern is not affected by the preceding training data words, because an even number of training data words produces no net DIP-4 contribution.

Table 6 Training Sequence.

Cycle	TCTL / RCTL	TDAT[i] / RDAT[i]															
		15	**14**	**13**	**12**	**11**	**10**	**9**	**8**	**7**	**6**	**5**	**4**	**3**	**2**	**1**	**0**
1	1	0	e	f	0	0	0	0	0	0	0	0	0	a	b	c	d
2 to 11	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
12 to 21	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0
20α -18 to 20α-9	1	0	0	0	0	1	1	1	1	1	1	1	1	1	1	1	1
20α-8 to 20α+1	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0

(Note: In cycle 1, "ef" and "abcd" depend on the contents of the interval after the previous control word.)

2.3 Start-Up Parameters

The sequence of ports at a FIFO status channel is defined in a data structure called CALENDAR, where CALENDAR[i], i = 1,…,CALENDAR_LEN, refers to the *i*th port in the repeating sequence. CALENDAR_LEN typically corresponds to the number of ports with the lowest data rate that can be accommodated in the total data rate of the given application. CALENDAR_LEN must be at least as large as the number of active ports in the system. The calendar sequence (of length CALENDAR_LEN) is repeated CALENDAR_M times before the DIP-2 parity and "1 1" framing words are inserted. CALENDAR_LEN and CALENDAR_M are both greater than zero.

Examples:

1. Single OC-192 or 10 Gigabit Ethernet port: CALENDAR_LEN = 1, CALENDAR[1] = 1.

2. Four OC-48 ports: CALENDAR_LEN = 4, CALENDAR[i] = 1, 2, 3, 4.

3. Two OC-48 channels (ports 1 and 2), eight OC-12 channels (ports 3 through 10): CALENDAR_LEN = 16, CALENDAR[i] = 1, 2, 3, 4, 1, 2, 5, 6, 1, 2, 7, 8, 1, 2, 9, 10,…. Other combinations are feasible: CALENDAR[i] = 1, 3, 2, 4, 1, 5, 2, 6, 1, 7, 2, 8, 1, 9, 2, 10, ….

4. Ten 1 Gb/s Ethernet ports: CALENDAR_LEN = 10, CALENDAR[i] = 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

The values of CALENDAR_LEN, CALENDAR_M and CALENDAR[i], i = 1,…,CALENDAR_LEN, must be identical in both the PHY and Link Layer devices for each interface. They need not be identical on both the transmit and receive FIFO status channels. The maximum supported value of CALENDAR_LEN is contained in the parameter MAX_CALENDAR_LEN, whose upper bound is implementation-specific. MAX_CALENDAR_LEN need not be identical on either side of the transmit or receive FIFO status channels. Users however, must ensure that the value of CALENDAR_LEN on the sending side of a FIFO status channel must not exceed MAX_CALENDAR_LEN on the receiving side.

For the FIFO Status channel(s), MaxBurst1 and MaxBurst2 may be configured to apply globally over all ports, or to apply on a per-port basis. In either case, both parameters must be consistently configured at the PHY and Link Layer devices for each interface, but need not be identical between the transmit and receive interfaces. MaxBurst1 must not be less than the corresponding MaxBurst2 (at the same port and interface).

For the data path deskew procedure, DATA_MAX_T is configured only on the sending side of the data paths on the transmit and receive interfaces. DATA_MAX_T need not be identical over both interfaces.

Start-up parameters are listed below in Table 7.

Table 7 Summary of Start-up Parameters.

(P = Provisionable (Yes/No), CH = per-channel [C] or per-interface [I])

Parameter	Definition	P	CH	Units
CALENDAR[i]	Port address at calendar location i.	Y	I	(N/A)
CALENDAR_LEN	Length of the calendar sequence.	Y	I	(N/A)
CALENDAR_M	Number of times calendar sequence is repeated between insertions of framing pattern.	Y	I	(N/A)
MAX_CALENDAR_LEN	Maximum supported value of CALENDAR_LEN.	N	I	(N/A)
MaxBurst1	The maximum number of 16 byte blocks that the FIFO can accept when FIFO Status channel indicates Starving	Y	C/I	16 byte blocks
MaxBurst2	The maximum number of 16 byte blocks that the FIFO can accept when FIFO Status channel indicates Hungry. MaxBurst2 <= MaxBurst1	Y	C/I	16 byte blocks
α	Number of repetitions of the training pattern in the training sequence that must be scheduled every DATA_MAX_T cycles (for data path) and FIFO_MAX_T cycles (for optional LVDS FIFO status channel).	Y	I	(N/A)
DATA_MAX_T	Maximum interval between scheduling of training sequences on Data Path interface.	Y	I	Cycles
FIFO_MAX_T	Maximum interval between scheduling of training sequences on FIFO Status Path interface.	Y	I	Cycles

Start-Up and Recovery Procedures

This section provides a general description of the interface during start-up prior to or upon reset during normal operation. From a data path perspective, normal operation refers to the case where payload data can be transferred over the interface (i.e., transmitter is enabled), and where FIFO status information, properly framed by a DIP-2 and “1 1” framing pattern, is received over the corresponding FIFO Status Channel. For the sake of brevity, the behavior of the FIFO Status Channel is described only for LVTTL implementations. The corresponding behavior for LVDS implementations is similar, except that continuous training patterns are sent instead of continuous “1 1” framing patterns to indicate loss of synchronization.

When the interface is disabled, the data path receiver empties all its FIFOs. Any outstanding credits in the data path transmitter are cleared. The FIFO Status Channel will continuously send “1 1” framing patterns. The data path transmitter shall send continuous training patterns until it has acquired synchronization on the FIFO Status Channel. Interface parameters can be (re)configured during this time. Declaration of acquisition or loss of synchronization on the FIFO Status Channel is implementation-specific and beyond the scope of this specification (e.g., synchronization may be declared after consecutive correct DIP-2 codewords are detected, and loss of synchronization may be declared after multiple incorrect DIP-2 codewords are detected).

At the same time, the data path receiver monitors for the training pattern and synchronizes to signals on the interface. The data path receiver shall ignore all incoming data until it has acquired synchronization. Declaration of acquisition or loss of synchronization on the data path is implementation-specific and beyond the scope of this specification (e.g., synchronization may be declared after consecutive correct codewords or training patterns are detected, and loss of synchronization may be declared after multiple incorrect codewords are detected). The corresponding FIFO Status Channel will begin sending valid status when it is enabled and the associated data path receiver has acquired synchronization.

In the case where the data path receiver is disabled but the data path transmitter remains active, events at the receiver follow the same behavior as above. The receiver will send on its FIFO Status Channel continuous “1 1” framing patterns, causing the data path transmitter to clear all previously granted credits and to start sending training patterns continuously. Hence, the data path transmitter behaves as though it too had been disabled.

In the case where the data path transmitter is disabled but the data path receiver remains active, events at the transmitter follow the same behavior as above. The

data path transmitter will clear all its credits and continuously send training patterns. The data path receiver may also have lost synchronization, and have consequently sent “1 1” framing patterns. The data path transmitter resumes normal operation when it is enabled and when it receives valid status information over the FIFO Status Channel.

2.4 AC Timing

System level reference points for specified parameters in this section are shown in Figure 2.13. Corresponding reference points with respect to the clock edge for specified parameters are shown in Figure 2.14 and Figure 2.15.

Figure 2.13 System-Level Reference Points.

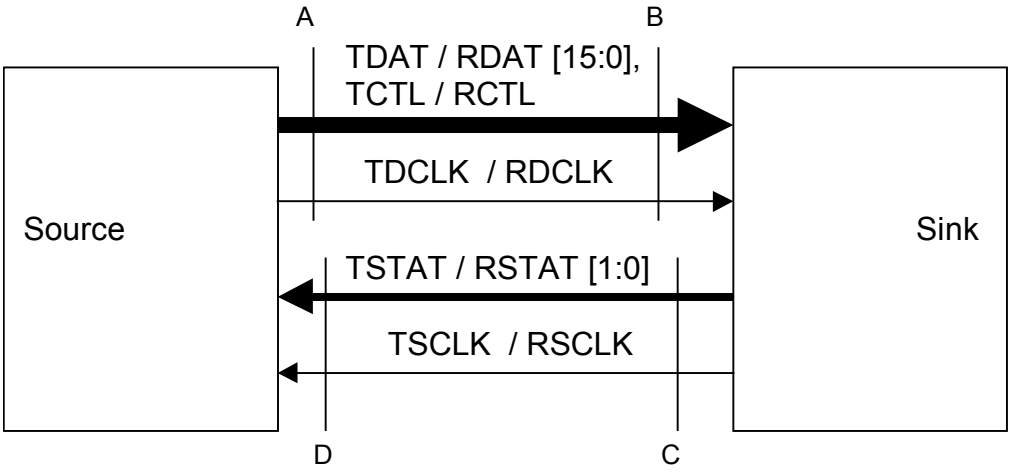


Figure 2.14 Reference Points for Data Path Timing Parameters.

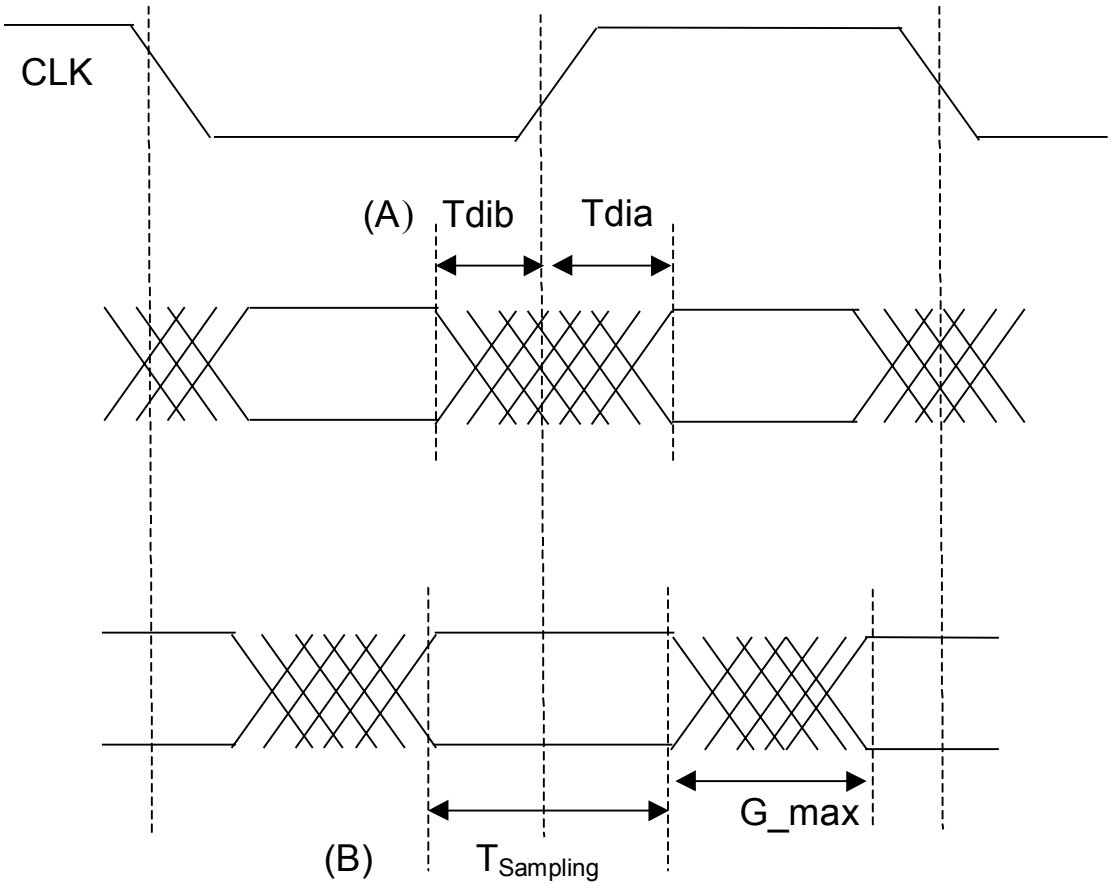
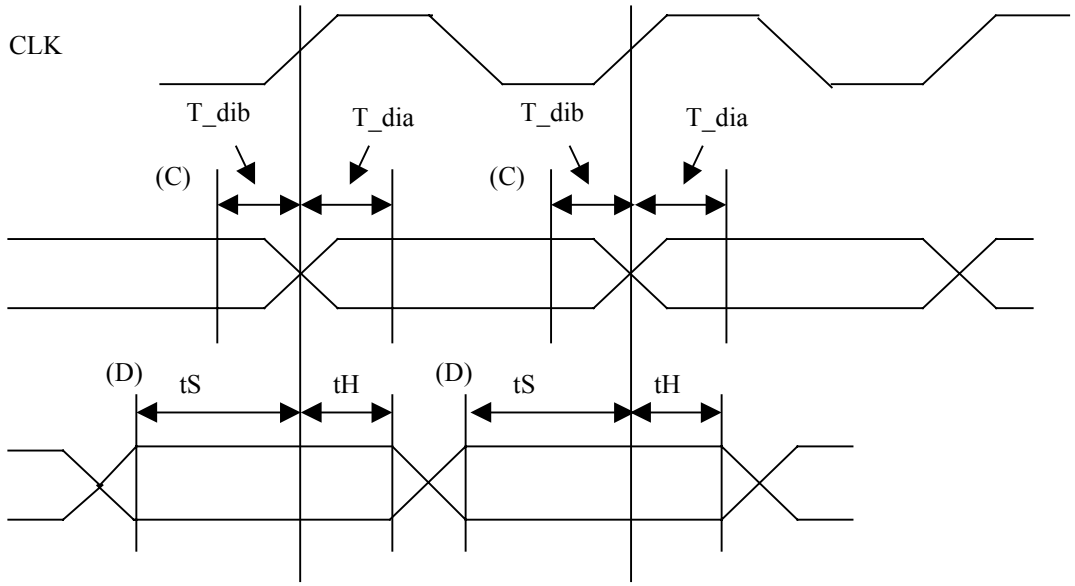


Figure 2.15 Reference Points for FIFO Status Path Timing Parameters.



2.4.1 Data Path

Two sets of data path timing parameters are specified to support different bit alignment schemes at the receiver. Table 8 gives the corresponding parameters for the case of “static alignment”, in which the receiver latches data at a fixed point in time relative to clock (requiring a more precisely specified sampling window). Table 9 gives the corresponding parameters for the case of “dynamic alignment”, in which the receiver has the capability of centering the data and control bits relative to clock. From an AC timing perspective, a compliant interface only needs to meet the parameters at the data path for either static or dynamic alignment, but may also comply to both sets of parameters. A compliant driver must meet both timing specifications to be interoperable with both types of receivers. For the case of static alignment, a sample timing budget suitable for up to fD = 350 MHz in Table 8 is shown in Appendix D. Also shown in that section is a sample budget for the case of dynamic alignment.

Table 8 Data Path Interface Timing (Static Alignment).

Symbol	Description		Min	Max	Units
fD	TDCLK / RDCLK Frequency		311		MHz
	TDCLK / RDCLK Duty Cycle (Reference point A)		45	55	%
T_dia, T_dib	Data invalid window with respect to clock edge. (Reference point A)			280	ps
G_max	Worst-case cumulative skew and jitter. (Reference point B)			790	ps
$T_{Sampling}$	Data valid window with respect to clock edge (Reference point B)			$\frac{1}{2fD} - G\_\max$	ps
	20% - 80% rise and fall times (UI = 1/2fD)	(Reference point A)	100 ps	0.30 UI	
		(Reference point B)	100 ps	0.36 UI	

Table 9 Data Path Interface Timing (Dynamic Alignment).

Symbol	Description		Min	Max	Units
fD	TDCLK / RDCLK Frequency		311		MHz
	TDCLK / RDCLK Jitter (Reference point A)			0.10	UI
	TDAT / RDAT / TCTL / RCTL Jitter (Reference point A)			0.24	UI
	20% - 80% rise and fall times (UI = 1/2fD)	(Reference point A)	100 ps	0.30 UI	
		(Reference point B)	100 ps	0.36 UI	

Notes:

1. Rise and fall times assume nominal 100-ohm termination and exclude reflections.

2. All timing parameters are measured relative to the differential crossing point of the corresponding clock signal.

3. Jitter parameters are peak-to-peak, measured above fD / 1000 and below fD.

4. Receiver sensitivity is assumed to be less than or equal to 100 mV.

5. Assumes a 5 pF output load at reference point A, a 10 pF load at reference point B, and a 50-ohm transmission line in between.

6. Assumes up to 20 ps skew between traces of a differential pair.

2.4.2 FIFO Status Channel

The following section describes AC timing parameters for a FIFO status channel implemented using LVTTL I/O. For LVDS FIFO status channel implementations, the reader is referred to the LVDS data path parameters in Section 2.4.1. As noted in Table 10, the maximum clock frequency of the LVTTL FIFO Status Channel shall not exceed one quarter of the selected data path clock rate.

Table 10 FIFO Status Channel Interface Timing (LVTTL I/O).

Symbol	Description		Min	Max	Units
fS	TSCLK Frequency			fD / 4	
	TSCLK Duty Cycle		40	60	%
T_dia	Data invalid window with respect to clock edge.	(Reference point C)	2.5		ns
T_dib		(Reference point C)	1		
tS_{SCLK}	TSTAT Setup time to TSCLK , RSTAT Setup time to RSCLK.	(Reference point D)	2		ns
tH_{SCLK}	TSTAT Hold time to TSCLK, RSTAT Hold time to RSCLK.	(Reference point D)	0.5		ns

Notes on I/O Timing (Vdd = 3.3V nominal):

1. When a set-up time is specified between an input and a clock, the set-up time is the time from the 1.4 Volt point of the input to the 1.4 Volt point of the clock.

2. When a hold time is specified between an input and a clock, the hold time is the time from the 1.4 Volt point of the clock to the 1.4 Volt point of the input.

3. Assumes a 25 pF, 500 Ohm load.

2.5 DC Parameters

Table 11 lists applicable DC thresholds for the LVTTL FIFO Status Channel.

Table 11 LVTTL DC Thresholds.

Parameter	Symbol	Min (V)	Max (V)
Output High Voltage	V_OH	2.4	3.6
Input High Voltage	V_IH	2.0	
Output Low Voltage	V_OL	0.0	0.4
Input Low Voltage	V_IL		0.8

3 REFERENCES

[1] IEEE Std 1596.3 – 1996, “IEEE Standard for Low-Voltage Differential Signals (LVDS) for Scalable Coherent Interface (SCI)”, approved March 21, 1996.

[2] ANSI/TIA/EIA-644-1995, “Electrical Characteristics of Low Voltage Differential Signaling (LVDS) Interface Circuits”, approved November 15, 1995.

APPENDIX A. ERROR PROTECTION CAPABILITY OF THE DIP-4 CODE.

In this section, the undetected error probability of the DIP-4 code is evaluated for a burst transfer length of 64 bytes, in the presence of random bit errors.

With the DIP-4 code, each constituent parity bit is computed over 66*8/4-1 information bits. These information bits, together with the parity bit, constitute a single-parity codeword. In the assumed case of odd parity, the value of the parity bit is set such that the total number of ones in the codeword is an odd number. It can easily be seen that any odd number of errors can be detected in a single-parity codeword. Hence, an undetected error event for a given codeword occurs whenever an even number of errors falls on that codeword. An undetected error event on a given burst transfer occurs when an undetected error event occurs on at least one of the constituent codewords.

Let P_{UC} and P_{UT} denote, respectively, the undetected error probabilities in a single-parity codeword and in a DIP-4 burst transfer. Let *L* denote the number of bits in a codeword and *p* denote the probability of a bit error. Enumerating all combinations of even-numbered errors in the codeword, we have,

$$P_{UC} = \sum_{\substack{i=2,\\ i\,even}}^{L} \binom{L}{i} p^i (1-p)^{L-i}.$$

Since an undetected error in a burst transfer corresponds to an undetected error in at least one codeword,

$$P_{UT} = 1 - (1 - P_{UC})^4.$$

Table A.1 shows the corresponding probabilities of undetected error in a burst transfer, for the DIP-4 code, over a range of bit error rates. Also shown for comparison are corresponding probabilities without an error-detection code. It can be seen that a DIP-4 code reduces the undetected error probability by several orders of magnitude. While the bit error rates in well-designed implementations may already be very low, the DIP-4 code can reduce the undetected error probability to extremely negligible levels with minimal added complexity to the interface implementation. Note however that longer burst transfers increase the DIP-4 extent, which increases the probability of an undetected error.

Table A.1 Error Detection Capability of DIP-4 Code.

Bit Error Rate	Probability of Undetected Error	
(*p*)	Without Error Detection	With DIP-4 Code
1.0E-6	5.28E-4	3.458E-8
1.0E-7	5.28E-5	3.458E-10
1.0E-8	5.28E-6	3.459E-12
1.0E-9	5.28E-7	3.464E-14
1.0E-10	5.28E-8	3.441E-16
1.0E-11	5.28E-9	(tiny)

APPENDIX B MINIMUM DATA PATH BANDWIDTH REQUIREMENTS FOR TYPICAL APPLICATIONS.

This section discusses the minimum bus frequencies required for a number of applications, assuming maximum 64-byte payload data transfers. The numerical results can be used to provide guidance on the minimum operating frequency of the bus. To provide design margin, the actual operating frequency would be higher to account for training and other overhead. Calculations in this section assume that the training bandwidth is insignificant and not included. For the sake of brevity, the word SONET is used to refer to either SONET or SDH.

Since the PHY device may remove or insert framing overhead to data received from or transmitted to the line side, the actual data rate over the POS-PHY interface may be less than the operating line rate. For a given line rate, the actual packet data rate becomes smaller with the packet length, since the framing overhead becomes a larger proportion of the line bandwidth. The POS-PHY interface, however, inserts a 16-bit control word between payload burst transfers. Hence, the bus tends to operate less efficiently with shorter packet lengths (and therefore tends to require a proportionally higher bus bandwidth for a given data transfer rate). The minimum operating frequency of the bus is the result of an interplay between the control word and the line overhead.

The following cases are considered in this section:

1. ATM cells over SONET.

2. (HDLC-framed) Packet over SONET.

3. 10 Gb/s Ethernet LAN PHY Framing.

In case 1, ATM cells are transported back-to-back in the payload area of a SONET frame. There is an 8-bit HEC field in the ATM cell header, which may or may not be transferred over the POS-PHY interface. In case 2, back-to-back HDLC-framed packets are transported in the SONET payload area. Apart from inter-frame flags, the HDLC framing overhead is assumed to consist of an 8-bit Type field, an 8-bit Address field, and a 32-bit frame check sequence (FCS). Byte-stuffing events will increase the line overhead, but these contributions are ignored since the worst-case in this discussion corresponds to the absence of byte stuffing. For cases 1 and 2, the SONET overhead is assumed to be 3.7% of the line rate. In case 3, the packets are encapsulated in Ethernet frames. In this case, jumbo frames (to the order of 8192 bytes long) are also considered, as these frames may be encountered in some applications even though they are longer than the maximum specified by the IEEE.

In general, the (per line) data path rate of the interface, f_I, can be expressed as,

$$f_I = f_S ABC / W,$$

Where

f_S = line rate (9.953 Gb/s for SONET, 10 Gb/s for Ethernet 10 Gigabit LAN PHY).

A = line (Layer 1) efficiency,

$$= \begin{cases} 0.963, & \text{SONET/SDH} \\ 1, & \text{10 Gb/s Ethernet LAN PHY} \end{cases}$$

B = Layer 2 framing efficiency

$= L / L'$, where

$$L' = \begin{cases} L, & \text{53 - byte ATM cells} \\ L+1, & \text{52 - byte ATM cells, POS incl HDLC header/FCS} \\ L+7, & \text{POS excl HDLC header/FCS} \\ L+12, & \text{10 Gb/s Ethernet LAN PHY} \end{cases}$$

C = packet interface overhead

$$= (L + (\lceil L/W \rceil - L/W) \times W + \lceil L/M \rceil \times N) / L$$

L = packet length, including packet overhead transferred over interface,

W = interface width (number of parallel data lines),

M = maximum burst transfer size,

N = additional interface overhead per transfer.

$\lceil x \rceil$ = ceiling function (the smallest integer not less than x).

The first, second and third terms in the equation for *C* correspond to contributions, respectively, from the packet itself, the unused portion (if any) on the last word upon EOP, and the overhead from segmentation. It is assumed in the formula for *C* that *M* (in units of bits) is an integer multiple of *W*. Setting *W* = 16 bits, *M* = 64 bytes, *N* = 2

bytes, the minimum bus frequency is obtained by finding the maximum of f_I over the applicable range of *L*.

For case 1, the minimum data path line rate, $f_{\min}$, for 52-byte cell transfer on average is 610.37 Mb/s. The corresponding $f_{\min}$ for 53-byte cells on average is 632.97 Mb/s. The peak data path line rates (between SONET framing overhead) correspond to 633.82 Mb/s and 657.29 Mb/s for 52- and 53-byte cells respectively.

For case 2, considering HDLC header and FCS transfer over the interface, the maximum occurs at *L* = 65 bytes, where the corresponding $f_{\min}$ is on average 635.37 Mb/s. In the more common situation where the header and FCS are not transferred over the interface, $f_{\min}$ increases with *L* (though not monotonically), approaching an asymptotic limit of roughly 618 Mb/s on average. The corresponding peak data path line rates are 659.78 Mb/s (header and FCS also transferred) and 642 Mb/s (header and FCS not transferred).

Due to the relatively long inter-frame gap (12 bytes) for case 3, the required f_I is fairly small for short *L* but increases with *L* as the gap becomes a smaller proportion of the frame length. Hence, the worst case corresponds to jumbo frame transfers. These frames are much longer than the maximum specified by IEEE, but have been used in some applications. An upper bound to $f_{\min}$ can be obtained by assuming that the overhead due to the preamble and the inter-frame gap is negligible for very long frames. Hence, $L' \approx L$, and $f_{\min} \leq$ 644.53 Mb/s.

APPENDIX C FIFO STATUS BANDWIDTH REQUIREMENTS AND FIFO THRESHOLD ISSUES.

It is not a trivial task to estimate a universally applicable required bandwidth on the status channel, as plausible worst-case scenarios are likely to be application-specific, and in any case, dependent on other implementation details such as FIFO depth. However, some general statements can be made without delving too deeply into specifics of particular implementations.

Given the nominal operating frequency of the data path, it is reasonable to expect that the status channel can be run at 1/8th the serial rate of the data path. For the case of very short packets sent in 16-byte bursts, there is roughly one status update opportunity for each minimum transfer period, as eight cycles will have elapsed in the data path for each cycle in the status channel. The overhead in the status channel (from parity and framing) can be made sufficiently small with a suitably long calendar length.

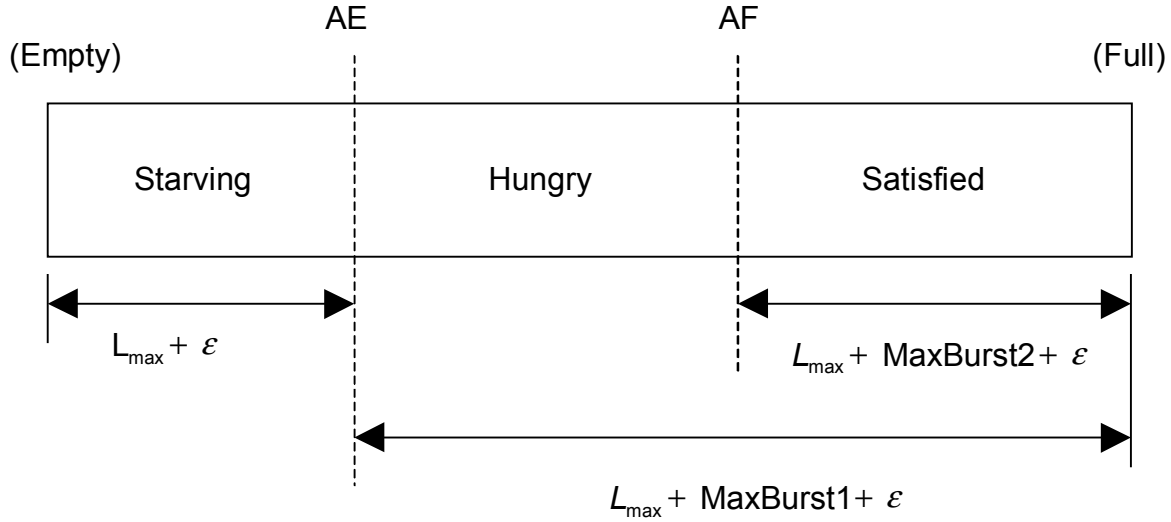
In the more typical cases corresponding to Internet traffic traces, the shortest packets will be roughly 40 bytes long (requiring 21 cycles to transfer, including control overhead). For a 65-byte packet segmented into two transfers, an average of 35 cycles will have elapsed between those transfers (33 cycles for payload and two for control). Hence, such transfers are likely to be in the range of 17.5 ~ 21 cycles, ample time for at least two status update opportunities within the transfer interval. This result does not fundamentally change in multi-port configurations, even with widely different line rates, as each constituent line rate will ultimately bound the amount of data path bandwidth allocated to a given port in the long run. For example, a port running at only a tenth of the total bandwidth supported by the interface can only send and receive data by that same fraction of bandwidth in the long run (excluding transient intervals). By weighting the calendar such that this port has an average of one update out of every ten opportunities, the corresponding status channel bandwidth for this port is scaled accordingly.

The thresholds for STARVING and HUNGRY are set such that the FIFO can accept at least MaxBurst1 and MaxBurst2 (16-byte) blocks respectively, plus an additional amount to account for feedback-response delay. In order to guard against potential buffer underflow, the lowest threshold must be set high enough to allow the other end to respond to transitions to the state of lowest FIFO occupancy in a reasonable length of time (to the order of the status update interval, plus scheduler response time). MaxBurst2 and MaxBurst1 (if applicable) must be provisioned to allow adequate utilization of transfer bandwidth between status updates for the given port.

Figure C.1 shows one possible way for relating FIFO thresholds to MaxBurst1 and MaxBurst2 as well as the response latency. L_max corresponds to the worst-case response time, from the delay in receiving a status update over the FIFO status channel, until observing the reaction to that update on the corresponding data path. The quantity, ε , corresponds to the difference between the granted credit and the actual data transfer length. This difference arises from various protocol overheads and quantization errors in the packet scheduler and the data path.

Figure C.1 Sample FIFO Thresholds.

(AE = Almost Empty waterline, AF = Almost Full waterline)



APPENDIX D SAMPLE LVDS TIMING BUDGETS

The calculations in this appendix are conservative; they assume linear addition of timing error sources. Reference points noted are shown in Figure 2.13.

D.1. Static Alignment at the Receiver

A sample timing budget suitable for operation up to fD=350 MHz (Table 8) is shown in Table D.1.

Table D.1 Timing Budget (Static Alignment).

Description		Value (peak-to-peak, ps)
Before Reference Point A	Clock to Data Skew (fixed) [note 1]	200
	Clock Duty Cycle Distortion (random)	140
	Data Duty Cycle Distortion (random)	140
	Data Jitter (random) [note 2]	80
	Subtotal (= T_dia + T_dib)	**560**
Between Reference Points A and B	Clock to Data Skew (fixed) [note 3]	150
	Relative Jitter (random) [note 4]	80
	Subtotal (= G_max)	**790**
After Reference Point B	Sampling Error (fixed) [note 5]	550
	Relative Jitter (random)	80
	Total	**1420**

Notes:

1. Includes variations in clock-to-data outputs across LVDS drivers, and across process, voltage and temperature.

2. Includes jitter from all sources, including data-dependent contributions.

3. Includes trim / data-eye centering error (assumes clock offset achieved on board). Also includes 20 ps for sum of package and board length mismatch.

4. Relative jitter between reference points A and B includes contributions from the board as well as connector.

5. Includes setup and hold times of D flip-flop.

D.2. Dynamic Alignment at the Receiver

A sample timing budget is shown in Table D.2. The margin remaining is 1 – 0.90 = 0.10 UI.

Table D.2 Timing Budget (Dynamic Alignment).

Description		Value (UI, peak-to-peak)
Clock jitter at reference point A		0.10
Data jitter at reference point A		0.24
Data jitter between reference points A and B		0.20
After reference point B	On-chip jitter due to routing.	0.01
	SYNTH Jitter	0.05
	Sampling granularity	0.25
	Variation in sampling position.	0.05
Total		0.90

Notes:

1. Sampling granularity (quantization error) assumes 4 samples per bit period.

NOTES

CONTACTING PMC-SIERRA, INC.

PMC-Sierra, Inc.
105-8555 Baxter Place Burnaby, BC
Canada V5A 4V7

Tel: (604) 415-6000

Fax: (604) 415-6200

Document Information: document@pmc-sierra.com
Corporate Information: info@pmc-sierra.com
Application Information: apps@pmc-sierra.com
Web Site: http://www.pmc-sierra.com

PMC-1991635 (P6) Issue date: February 2001