Working Group: **Physical and Link Layer**

TITLE:

SERDES Framer Interface Level 4 (SFI-4) Phase 2: Implementation Agreement for 10Gb/s Interface for Physical Layer Devices.

SOURCE:	Kevin Voegele Technical Editor	Mike Lerer PLL Chair	Jeffrey Lynch PLL Vice Chair and Chair of the PLL Electrical Track
	Applied Micro Circuits 7701 France Avenue Suite 200 Edina, MN 55435 Phone: 1-952-837-6206 Email: kvoegele@amcc.com	Avici Systems Inc 101 Billerica Ave North Billerica, MA 01862 USA Phone: 1-978-964-2058 Email: mlerer@avici.com	IBM P.O. Box 12195 Research Triangle Park, NC 27709 USA Phone: +1 919-254-4454 Email: jjlynch@us.ibm.com

DATE: **September 12, 2002**

ABSTRACT:

SFI-4 phase 2 Implementation Agreement for an interface between the SERDES and Framer devices within the Physical Layer (SERDES Framer Interface or SFI). SFI-4 is an interface for aggregate data bandwidths of OC-192 ATM and Packet over SONET/SDH (POS), as well as other applications at the 10 Gb/s data rate.

1 Document Revision History

OIF2002.166.00	Apr. 10th 2002	Initial revision of SFI-4 phase 2
OIF2002.166.01	Apr. 23rd 2002.	Remove RXDCK. Change TXREFCK and RXREFCK to REFCK and remove redundancies.
OIF2002.166.02	Jul. 12th 2002.	Incorporate straw ballot comments.
OIF2002.166.03	Jul. 31st 2002.	Incorporate comment resolutions as defined in oif2002.376.02
OIF2002.166.04	Aug. 1st 2002.	Editorial Corrections of math errors and missing section header
OIF2002.166.05	Sept. 9th 2002.	Corrections from 2nd round of straw ballot contained in oif2002.446.00

2 Contents

3 Introduction

The typical line interface of a communications system with 10Gb/s optical links may consist of three separate devices: an optical module containing a SERDES component, a forward error correction (FEC) processor and a framer. The interconnection between these devices requires a parallel electrical bus operating significantly slower than the optical data rate. The objectives and requirements for this parallel bus, known as SFI-4 phase 2 are defined below:

3.1 SFI-4 Phase 2 Objectives and Requirements

1. Support up to 12.12Gb/s bi-directional aggregate data throughput such as SONET OC-192, SDH STM-64 [2], Digital Wrapper ITU G.709 [3], 10GbE LAN, 10GbE WAN and other systems operating at the payload data rate in the 10Gb/s range, and including up to 21.7% FEC overhead.

2. The interface can be used for point to point connection between framer and FEC processor, framer and SERDES, FEC processor and SERDES. Requirements for electrical parameters and control are the same between the different components.

3. The Interface is payload agnostic – function of the interface will not depend on any protocol or framing characteristics of the transmitted or received data.

4. Capable of driving at least 8” of FR4 interconnect with one connector

5. Interface should be capable of operating indefinitely, without losing sync.

6. Support both in-band and out-of-band Forward Error Correction (FEC). A flexible clocking arrangement accounts for the additional FEC overhead.

7. The interface is independent of the type of optics – serial, WWDM or parallel, SMF or MMF.

8. Support of skew compensation over the signals comprising the interface bus. The de-skew algorithm used should minimize the complexity of the SERDES component.

9. Support of simple and robust clock and data recovery of the signals comprising the interface bus.

10. The interface may support independent status monitoring and received loss of signal detection through an unspecified management interface.

11. The receive timing reference operates continuously, independent of link status.

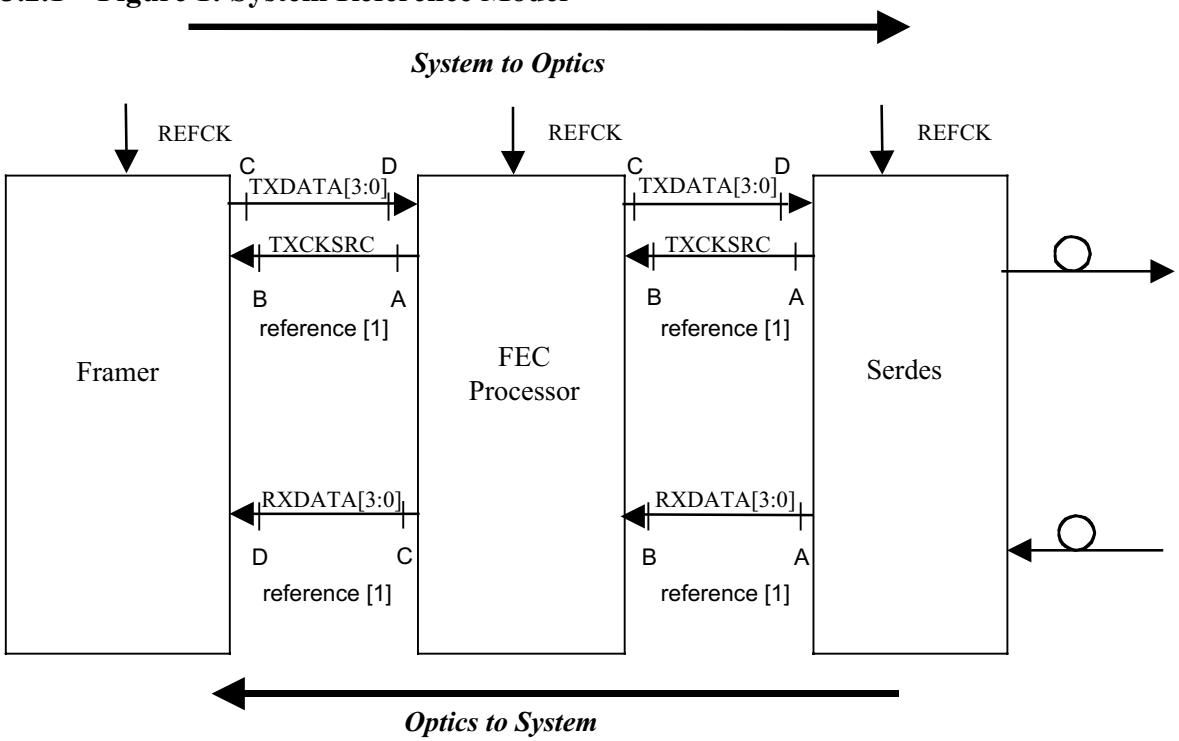
12. Any SFI-4 phase 2 data receiver will tolerate being driven while powered down without damage.

13. The interface may support data path link verification for component qualification and system integration test purposes through an unspecified management interface.

3.2 SFI-4 Phase 2 system reference model

The following is a general synopsis of the SFI-4 phase 2 interface. For reference, a general block diagram is shown in Figure 1. SFI-4 phase 2 is the interface between the SERDES component, the forward-error-correction (FEC) processor, and the framer. It is designed to meet requirements of this particular application, although it may also be used in other applications. "Receive" and "transmit" refer to data flow from the optics-to-system direction, and from the system-to-optics direction, respectively. Note that the two instantiations of the SFI-4 phase 2 bus in Figure 1 are independent and may operate at different frequencies. For a description of test points A, B, C and D see reference [1].

3.2.1 Figure 1: System Reference Model



3.3 SFI-4 phase 2 general description

On both the receive and transmit interfaces, control and status information is sent separately from the corresponding data path. To accommodate the expected minute frequency offsets between the data stream in the two directions, the receive and transmit interfaces operate independently.

The SFI-4 phase 2 bus has the following general characteristics:

1. Point-to-point connections applicable for connection of the SERDES component to the FEC processor, the FEC processor to the SONET/SDH framer or the SERDES component directly to the SONET/SDH framer.

2. 4-bit wide data bus with each channel operating at up to 3.125 Gb/s.

3. Electrical parameters are defined in the common electrical specification oif.2001.149 [1]. Parameters specific to SFI-4 phase 2 are included in this document.

4. The maximum bus bandwidth of 12.12 Gb/s is sufficient to support SONET OC-192, SDH STM-64 [2], Digital Wrapper ITU G.709 [3] and other systems operating at the payload data rate in the 10Gb/s range, with up to 21.7% FEC overhead.

5. De-skew algorithm on the data lines will be operating continuously to monitor skew tracking using transitions in the 64b/66b coded data.

6. Data transferred will be explicitly scrambled. [reference 4]

7. Minimize power consumption and the number of I/O signals to simplify PC board manufacture.

8. Maximize commonality of the receive and transmit directions and of instantiations in different applications of the bus.

4 Signal definition

4.1 Receive signals

The receive signals are related to the transport of data from the optics towards the system. They are applicable to conducting data from the SERDES component to the FEC processor, from the FEC processor to the framer, or from the SERDES component directly to the framer. All receive signals, unless otherwise specified below, are differential CML as defined in the common electrical implementation agreement oif2001.149 [1].

4.1.1 Table 1: SFI-4 phase 2 Receive Signal Summary

Signal Name	Direction	Function
RXDATA[3:0]	Optics to System	The **Receive Data** (RXDATA[3:0]) signals carry the data in the optics to system direction. Serial optical data, in 64-bit data blocks, is scrambled and 64b/66b encoded and the new 66-bit blocks are striped onto RXDATA[3:0] in round-robin fashion. RXDATA[3] contains the first 66-bit block, RXDATA[2] contains the next 66-bit block and so on. When RXDATA[3:0] is generated, adjacent 66-bit blocks have a relative offset of 16-bits relative to each other. There will be 18-bits of relative offset between the previous lane 0 and the next lane 3 due to the additional 01 sync header prepended on lane 0 prior to the next lane 3 transmission.Each RXDATA[X] signal is a 2.566 Gb/s to 3.125 Gb/s stream to reflect the additional 64b/66b de-skew overhead.

4.2 Transmit Signals

The transmit signals are related to the transport of data from the system towards the optics. They are applicable to conducting data from the framer to the FEC processor, from the FEC processor to the SERDES component, or from the framer directly to the SERDES component. All transmit signals, unless otherwise specified below, are differential CML as defined in the Common Electrical Implementation Agreement OIF2001.149 [1]

4.2.1 Table 2: SFI-4 phase 2 Transmit Signal Summary

Signal Name	Direction	Function
TXDATA[3:0]	System to Optics	The **Transmit Data** (TXDATA[3:0]) signals carry data in the system to optics direction. Data on TXDATA[3:0] are placed on the transmit optical stream in a round-robin fashion. TXDATA will be octet aligned. TXDATA[3] contains the first 64-bit data block to be transmitted while TXDATA[0] the last 64-bit data block to be transmitted. The 64-bit data block is part of the coded 66-bit block which includes the prepended 01 sync header. The 66-bit blocks have a 16-bit relative offset to each other. There will be 18-bits of relative offset between the previous lane 0 and the next lane 3 due to the additional 01 sync header prepended on lane 0 prior to the next lane 3 transmission. Each TXDATA[X] signal is a 2.566 Gb/s to 3.125 Gb/s stream to reflect the additional 64b/66b de-skew overhead.
TXCKSRC	Optics to System	The **Transmit Clock Source** (TXCKSRC) signal provides timing reference for the transmit data path signals, TXDATA. TXCKSRC is nominally a 50% duty cycle clock with a frequency that is one-quarter of the parameter: TXDATA rate minus the 64b/66b overhead coding rate, which is 622.08 MHz with a serial input data rate of 9.95328 Gbps. TXCKSRC is frequency locked to REFCK in the sink device. It is mandatory for the TXDATA source device in the transmit interface to be able to receive TXCKSRC into the REFCK and to use this input as a frequency reference for TXDATA. REFCK is frequency locked to TXDATA in that an edge of REFCK/32 is synchronous to every 66th bit of TXDATA. Note: Although defined as separate functions by the reference model, this signal description allows an implementation to use a single device pin on the source device as a common TXCKSRC or REFCK input.

4.3 Reference Clock

The frequency reference clock provides the frequency reference for transmit data path timing. All frequency references, unless otherwise specified below, are differential CML as defined in the Common Electrical Implementation Agreement OIF2001.149 [1].

4.3.1 Table 3: Reference Clock Description

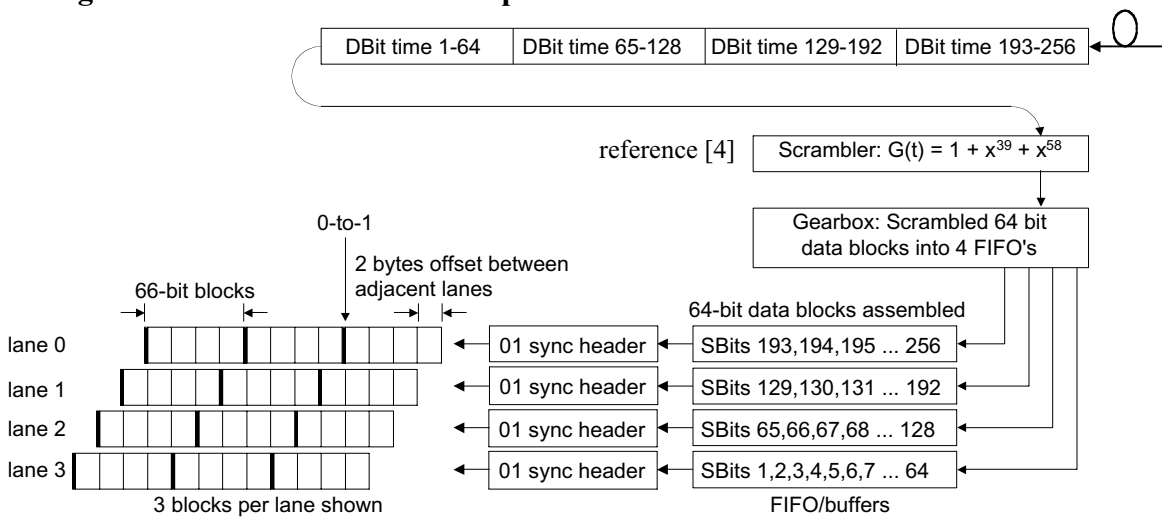
Signal Name	Direction	Function
REFCK	System to Optics or Optics to System	The **Frequency Reference** (REFCK) signal provides a frequency reference for transmit data path timing (framer, FEC, and SERDES). REFCK is nominally a 50% duty cycle clock with a frequency that one-quarter of the parameter: TXDATA or RXDATA rate minus the 64b/66b overhead coding rate, which is 622.08 MHz with a serial input data rate of 9.95328 Gbps.

5 Logical Reference Models

5.1.1 SERDES component receive interface

Figure 2 below shows a logical model of the receive interface in a SERDES component. The intent is to show the conversion of the serial data stream to the scrambled and striped data lanes. No limit is placed on device implementation, and this model is not intended to represent an actual design.

5.1.2 Figure 2: Model of SERDES component - Receive I/F Source



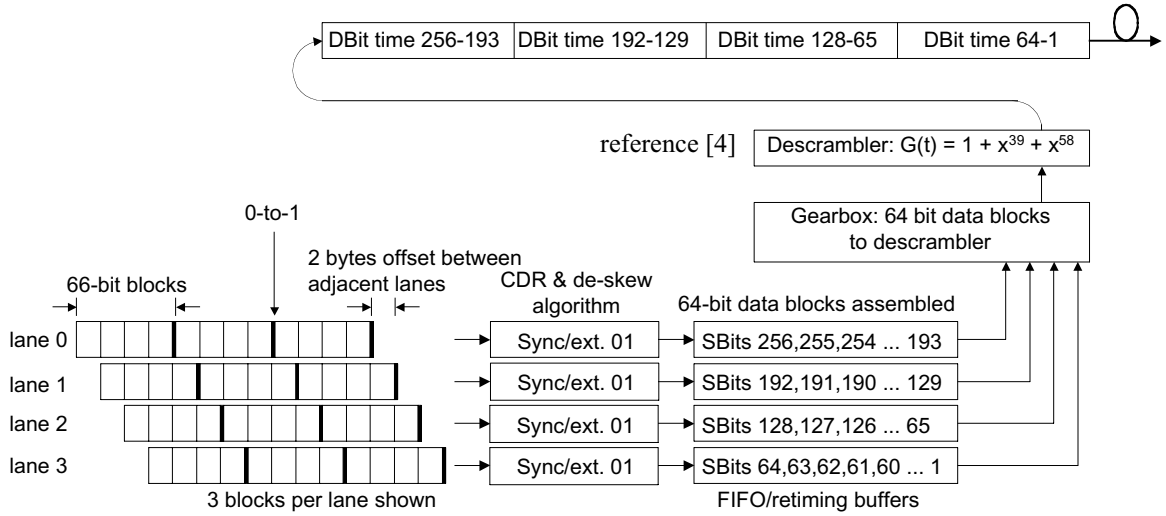
Data in the optical stream is scrambled [reference 4] and 64-bit data block striped across the 4 bit lanes of the receive data bus (RXDATA[3:0]) in a round-robin fashion. The first 64-bits received is written into the buffer associated with RXDATA[3] and the last into that associated with RXDATA[0]. The buffers act as a set of FIFOs to bridge between the input timing domain and the receive interface timing domain. A 01 sync header is prepended on each 64-bit data block to construct the 66-bit block prior to transmission. Each lane is transmitted 2 bytes (16 bits) after the previous lane so that the 4 lanes can start transmission in a 66-bit time period.

The SERDES component includes functionality to ensure that RXDATA is clocked at a continuous rate to the FEC processor or framer device. The lock range of the optical receiver shall exceed 100ppm over temperature and supply range. The clock recovered from the optical signal is compared with a clock derived from the derived reference frequency, REFCK. If the recovered clock deviates from the reference frequency by more than 1000ppm, then RXDATA shall be clocked from a source derived from REFCK (plus the 64B66B overhead). Somewhere between 100ppm and 1000ppm, an out of lock condition will be declared via an unspecified management interface, and RXDATA shall switch from being clocked derived from the recovered clock to being clocked derived from the reference.

5.1.3 SERDES Component transmit interface

Figure 3 below shows a logical model of the transmit interface in a SERDES component. The intent is to show the conversion of the parallel and scrambled data lanes to the unscrambled, synchronized and serial data stream. No limit is placed on device implementation, and this model is not intended to represent an actual design.

5.1.4 Figure 3: Model of SERDES component - Transmit I/F Sink



Data on the transmit data bus (TXDATA[3:0]) is recovered. Data from each bit lane is written into an associated re-timing buffer. Data from the re-timing buffer associated with TXDATA[3] is transmitted first while data in that associated with TXDATA[0] is transmitted last. The re-timing buffers act as a set of FIFOs to bridge between the transmit interface timing domain and the next downstream timing domain. Wander between the bit lanes is absorbed by the re-timing buffers using the 01 sync headers to eliminate skew.

The function of the de-skew algorithm is to recognize the 01 sync header in the DATA overhead. These identify the start of the reference data in each of RXDATA[X]. Each of the RXDATA[X] channels is then compared with respect to the 01 sync headers. The de-skew algorithm performs a pattern match regarding the 01 sync headers. Where there is a match, the relative delay of RXDATA[X], in relation to the 01 sync headers is found. It is possible to compensate for the skew by adjusting the delay elements specific to each channel.

A delay chain is associated with each interface signal that allows the de-skew algorithm to use the set of 4 relative delays to compensate for the skew in the SFI-4 phase 2 interface and to reconstruct the original alignment of the transmit data TXDATA[3:0]. By default there is 16 bits of relative offset between adjacent lanes.

There will be 18-bits of relative offset between lane 0 and lane 3 due to the additional 01 sync header prepended on lane 0 prior to the next lane 3 transmission.

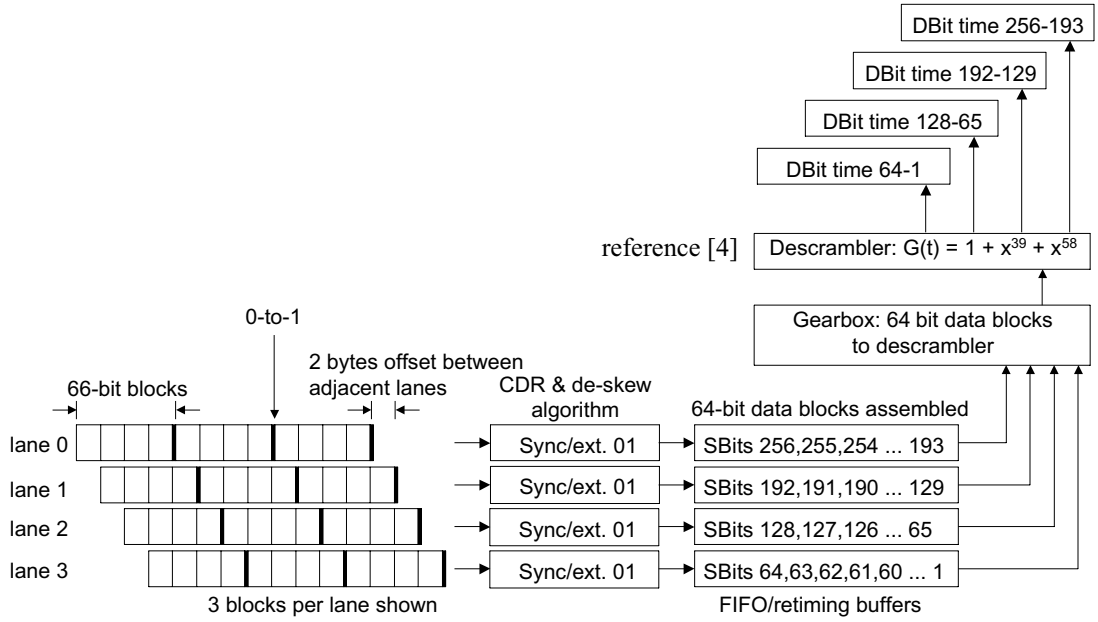
TXOOA is set if a match has not been found on any of the 4 data channels. When a data match has been found on all 4 data channels, and stable skew data derived, then the de-skew algorithm is considered as locked, and the TXOOA is cleared. Skew compensation is monitored continuously, and the TXOOA alarm remains cleared as long as consistent skew data is generated. Reference [4], figure 49-12 and 49-13, Lock State Machine. TXOOA is active low and may be supported by an unspecified management interface.

The function of the de-skew algorithm is to monitor the 01 sync headers on the respective data channels TXDATA[3:0]. Any mis-match errors can be detected, which would represent errors generated over the SFI-4 phase 2 interface. These errors should be reported as part of the minimum test requirements for the interface. These test requirements are described in section 7.2.

5.1.5 Receive interface of FEC processor / framer

Figure 4 below shows a model of the receive interface in the FEC processor or in the framer. The intent is to show the conversion of the 4 scrambled serial data streams to the unscrambled and de-skewed data. No limit is placed on device implementation, and this model is not intended to represent an actual design.

5.1.6 Figure 4: Model of FEC Processor / SONET Framer – Receive I/F Sink



The CDR recover the data on the receive data bus (RXDATA[3:0]). The re-timing buffers bridge between the receive interface timing domain and the system timing domain. Wander between the bit lanes is absorbed by the re-timing buffers.

The function of the de-skew algorithm block is to recognize the 01 sync headers on the DATA channels. Each of the RXDATA[X] channels is then analyzed for the 01 sync headers. The de-skew algorithm performs a pattern match between the data and the 01 sync headers. Where there is a match, the relative delay of RXDATA[X], in relation to other data channels, is found. It is possible to compensate for the skew by adjusting the delay elements specific to each channel.

The de-skew algorithm uses the set of 4 relative delays to construct the compensated receive data (RXDATA[3:0]) that recovers the original alignment, prior to timing distortions. By default there is 16 bits of relative offset between adjacent lanes. There will be 18-bits of relative offset between lane 0 and lane 3 due to the additional 01 sync header prepended on lane 0 prior to the next lane 3 transmission.

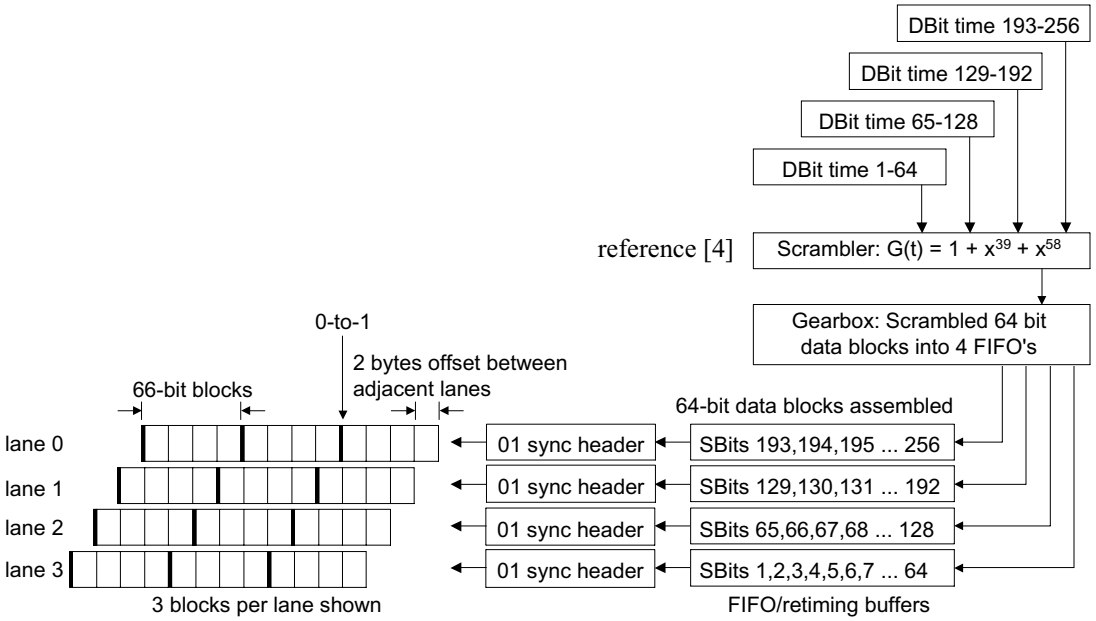
RXOOA is set if a match has not been found on any of the 4 data channels. When a data match has been found on all 4 data channels, and stable skew data derived, then

the de-skew algorithm is considered as locked, and the RXOOA is cleared. Skew compensation is monitored continuously, and the RXOOA alarm remains cleared as long as consistent skew data is generated. Reference [4], figure 49-12 and 49-13, Lock State Machine. RXOOA is active low and may be supported by an unspecified management interface.

5.1.7 Transmit interface in FEC Processor / Framer

Figure 5 below shows a model of the transmit interface in the FEC processor or in the framer. The intent is to show the conversion of the data streams to the scrambled and striped data lanes. No limit is placed on device implementation, and this model is not intended to define an actual design.

5.1.8 Figure 5: Model of FEC Processor / SONET Framer – Transmit I/F Source



Data from the core logic of the FEC processor or framer is 64-bit data block striped across the 4 bit lanes of the transmit data bus (TXDATA[3:0]) in a round-robin fashion. The first bit received is written into the re-timing buffer associated with TXDATA[3] and the last into that associated with TXDATA[0]. The re-timing buffers act as a set of FIFOs to bridge between the core logic timing domain and the transmit interface timing domain.

6 Component Functionality

6.1.1 SERDES component receive functionality

The SERDES component performs clock and data recovery (CDR) on the received optical data stream. A simple round-robin serial to parallel conversion is performed on the recovered data to construct 64-bit data blocks. No heed is paid to the underlying octet alignments within the optical data stream. The first bit received is placed in the most significant bit of the word while the last bit received is placed in the least significant bit.

6.1.2 SERDES component transmit functionality

The SERDES component performs data recovery (CDR) on the transmit data (TXDATA[3:0]) signals. The signals are frequency locked but not phase locked to each other and can have un-correlated jitter characteristics. Thus, the CDR process must be performed independently on each signal. Data recovered at each signal is written into one of 4 re-timing buffers and is read out in parallel every 4 bit times of the transmit optical data stream. The purpose of the re-timing buffers is to compensate for relative jitter and arrival times on TXDATA signals as well as transient clock differences between TXDATA and the transmit stream. A simple round-robin parallel to serial conversion process stripes transmitted data from the re-timing buffer servicing TXDATA[3] first and data from that servicing TXDATA[0] last.

6.1.3 FEC processor receive functionality

The FEC processor performs data recovery (CDR) on the receive data (RXDATA[3:0]). The signals are frequency locked but not phase locked and can have un-correlated jitter characteristics. Thus, the CDR process must be performed independently. The purpose of the re-timing buffer is to compensate for relative jitter and arrival times on RXDATA.

The FEC processor compares the 01 sync headers on the RXDATA signals to determine the relative skew between the signals. It then adjusts the delay through the re-timing buffers to compensate. The receive out-of-alignment alarm RXOOA is set until all 4 channels are locked. Reference [4], figure 49-12 and 49-13, Lock State Machine.

As a result of the de-skewing process, the order of arrival of the bits at the output of the re-timing buffers is known. I.e., output of FIFO #3 is the first bit received while output of FIFO #0 is the last bit received. The FEC processor will search the data stream constructed from the output of the re-timing buffers for the framing pattern associated with the FEC frame.

6.1.4 FEC processor transmit functionality

The transmit side of the FEC processor generates data for transmission by the SERDES component. It also continuously generates 66-bit blocks on the data channels (TXDATA[3:0]) to enable the de-skew algorithm in the downstream SERDES component, to compensate for timing skew in the SFI-4 phase 2 transmit interface. The reference 66-bit blocks consists of a 01 sync header and a 64-bit data block taken from the transmit data bus (TXDATA[3:0]).

6.1.5 Framer Receive Functionality

The functionality of the receive side of the 66-bit block striper/synchronizer/descrambler is identical to that of the FEC processor.

6.1.6 Framer Transmit Functionality

The functionality of the transmit side of the 66-bit block striper/synchronizer/scrambler is identical to that of the FEC processor.

7 64b/66b CODEC de-skew

The data signals may encounter different delays in transit from the SFI-4 phase 2 source device to the sink device. The maximum relative skew introduced by the interconnect is specified at 20 UI.

The bit-lane de-skew function is shared between the SFI-4 phase 2 source and sink devices at either end of the receive and transmit interfaces. In the source device, data is scrambled and 01 sync headers prepended onto each of the 4 data channels. The de-skew “overhead” is then sent with the 4 data channels to the sink device over the SFI-4 phase 2 interface. Data input to the sink device will have a relative offset by at least 16-bit delays in each of the adjacent data channels. This is an intentional offset between adjacent data lanes. Relative to the earliest data lane, each of the remaining signals is at least 16 bits, or is up to 48 unit intervals, late, i.e. lane 3 to lane 0. There will be 18-bits of relative offset between lane 0 and lane 3 due to the additional 01 sync header prepended on lane 0 prior to the next lane 3 transmission.

It is the function of the 64b/66b de-skew algorithm operating in the sink device to compensate the amount of skew on each data channel relative to the 01 sync headers. The earliest arriving data lane may lead the latest arriving by 68 bits, which is 48 bits relative offset from the transmit process and up to 20 bits skew from the interconnect. The de-skew process shall compensate for intra-lane skew during the synchronization phase. External conditions may cause the skew variations which require additional compensation. Compliant devices are required to compensate skew up to 20 UI.

7.1 De-skew Functionality

Table 4.1 below shows a reference for one of the four data channels, either RXDATA[X] or TXDATA[X].

The 01 sync header is prepended on each 64-bit data block to construct a 66-bit block. A 66-bit block is generated in the source device, consisting of 8 bytes of data per lane plus the prepended 01 sync header. Each reference block is delimited by the synchronization header, which is the 01 sync header. The data samples are transferred from the data bus RXDATA[3:0] or TXDATA[3:0] in 8-byte sets plus the 01 sync header per block, starting with RXDATA[3] or TXDATA[3] and ending with RXDATA[0] or TXDATA[0]. Each lane has 64-bit data blocks which are offset by 2 bytes on adjacent lanes so that subsequent frames on a specific lane can be transmitted 66 bit times later. After all the data lanes have been sampled, new 66-bit blocks are initiated and generated and the cycle repeats with the next string of data.

7.1.1 Table 4: Sequence for sending serial data to per lane data

Bit time	Value	Value	Value	Value	Comments
1,2	01 (2 bits)	-	-	-	64b/66b 01 sync header
3-34	Sbit1-Sbit8 (1st byte lane 3)	Sbit9-Sbit16 (2nd byte lane 3)	Sbit17-Sbit24 (3rd byte lane 3)	Sbit25-Sbit32 (4th byte lane 3)	Scrambled versions of the original serial data stream Dbit1-Dbit32
35-66	Sbit33-Sbit40 (5th byte lane 3)	Sbit41-Sbit48 (6th byte lane 3)	Sbit49-Sbit56 (7th byte lane 3)	Sbit57-Sbit64 (8th byte lane 3)	Scrambled versions of the original serial data stream Dbit33-Dbit64
67,68	01 (2 bits)	-	-	-	64b/66b 01 sync header
69-100	Sbit65-Sbit72 (1st byte lane 2)	Sbit73-Sbit80 (2nd byte lane 2)	Sbit81-Sbit88 (3rd byte lane 2)	Sbit89-Sbit96 (4th byte lane 2)	Scrambled versions of the original serial data stream Dbit65-Dbit96
101-132	Sbit97-Sbit104 (5th byte lane 2)	Sbit105-Sbit112 (6th byte lane 2)	Sbit113-Sbit120 (7th byte lane 2)	Sbit121-Sbit128 (8th byte lane 2)	Scrambled versions of the original serial data stream Dbit97-Dbit128
133,134	01 (2 bits)	-	-	-	64b/66b 01 sync header
135-166	Sbit129-Sbit136 (1st byte lane 1)	Sbit137-Sbit144 (2nd byte lane 1)	Sbit145-Sbit152 (3rd byte lane 1)	Sbit153-Sbit160 (4th byte lane 1)	Scrambled versions of the original serial data stream Dbit129-Dbit160
167-198	Sbit161-Sbit168 (5th byte lane 1)	Sbit169-Sbit176 (6th byte lane 1)	Sbit177-Sbit184 (7th byte lane 1)	Sbit185-Sbit192 (8th byte lane 1)	Scrambled versions of the original serial data stream Dbit161-Dbit192
199,200	01 (2 bits)	-	-	-	64b/66b 01 sync header
201-232	Sbit193-Sbit200 (1st byte lane 0)	Sbit201-Sbit208 (2nd byte lane 0)	Sbit209-Sbit216 (3rd byte lane 0)	Sbit217-Sbit224 (4th byte lane 0)	Scrambled versions of the original serial data stream Dbit193-Dbit224
233-264	Sbit225-Sbit232 (5th byte lane 0)	Sbit233-Sbit240 (6th byte lane 0)	Sbit241-Sbit248 (7th byte lane 0)	Sbit249-Sbit256 (8th byte lane 0)	Scrambled versions of the original serial data stream Dbit225-Dbit256

Note: Transmission is from left to right then top to bottom.

Until the de-skew algorithm finds lock on all 4 data channels, the interface is in an out-of-alignment state, and the alarm RXOOA or TXOOA is set. The out-of-alignment alarms RXOOA and TXOOA are signals communicated over an unspecified management interface, and are not SFI-4 phase 2 signal pins. Reference [4], figure 49-12 and 49-13, Lock State Machine. RXOOA and TXOOA are active low and may be supported by an unspecified management interface.

The function of the de-skew algorithm is to recognize the 01 sync header in the data channels RXDATA[X]/TXDATA[X], to identify the start of the 64-bit blocks. Each of the RXDATA[X]/TXDATA[X] channels is then compared with respect to the 01 sync header. The de-skew algorithm performs a pattern match on the 01 sync header. Where there is a match, the relative delay between RXDATA[X]/TXDATA[X] data lanes, in relation to the 01 sync header, is found. The de-skew algorithm shall adjust the delay of each RXDATA[3:0]/TXDATA[3:0] signal, such that the delay from source device to the output of the delay chain at the sink device is identical for all 4 signals.

When the skew of all 4 data channels is compensated the RXOOA/TXOOA alarm is cleared. The de-skew algorithm will continue to operate after the RXOOA/TXOOA is removed. Under normal circumstances, the interface will operate continuously with skew being monitored, and the RXOOA/TXOOA alarm remains off. If failure of the de-skew algorithm occurs, and any of the 4 channels fall out of alignment then the RXOOA/TXOOA alarm will be set. Reference [4], figure 49-12 and 49-13, Lock State Machine.

7.2 Interface test requirements

A minimum set of test requirements is defined to indicate that the interface is set up correctly and that it monitors itself for continuous reliable operation. The following functions are necessary to provide a minimum set of test functions:

1. Confirm that each of the 4 data channels are connected correctly in both receive and transmit directions. The 64b/66b CODEC and striper/synchronization [reference 4] and re-timing circuits in each direction can implement this function. If persistent mis-match occurs with any of the 4 channels, then the RXOOA or TXOOA alarm is raised relating to the respective receive or transmit directions. The minimum requirement is to monitor each alarm. It would be possible to include functionality to indicate which channel is faulty – this would be within the scope of individual implementations and not mandatory in this implementation agreement.

2. Verify that the de-skew function correctly compensates for skew over the SFI-4 phase 2 interface. It is the function of the de-skew algorithm to compensate the level of skew on each data channel in both the receive and transmit directions. When a stable skew compensation is made on each channel, the interface is considered in lock, and the RXOOA and TXOOA alarms are cleared. These

alarms should indicate that the de-skew compensation is working correctly after initial power up and during continuous operation.

3. It is possible to use the 64b/66b CODEC and striper/synchronization [reference 4] circuit in both receive and transmit directions to detect gross errors over the interface. Mismatch errors can then be detected in the data sample. It is a requirement of the interface to continuously monitor for these potential errors. The procedure for reporting errors is not part of the SFI-4 phase 2 implementation agreement, but should be included as part of the management function.

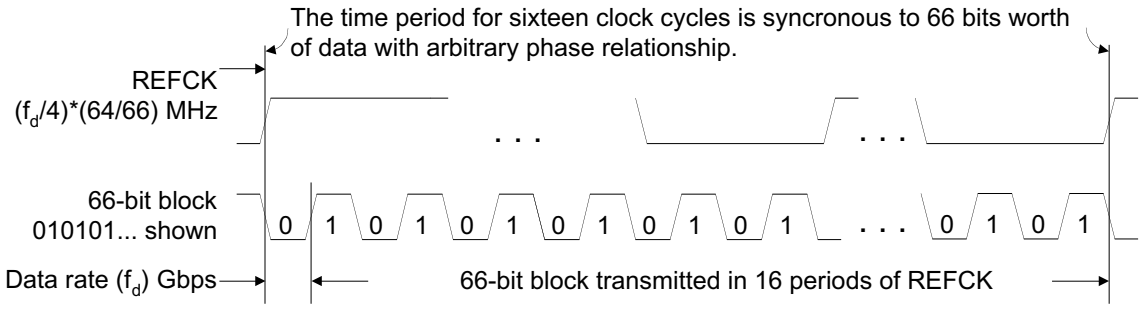
8 AC Characteristics

The majority of the AC characteristic of SFI-4 phase 2 is detailed in the common electrical parameters in contribution oif2001.149 [1]. Only those that differ are listed below. Skew tolerance is also different since it is increased to 20 UI.

8.1.1 Table 5: Data Path Interface Timing

Symbol	Description	Min	Max	Units
f_D	RXDATA[3:0], TXDATA[3:0] data rate	2.566	3.125	Gb/s
f_R	REFCK, TXCKSRC frequency	f_D(MIN)/4 * (64/66)	f_D(MAX)/4 * (64/66)	MHz

8.1.2 Figure 6: AC Timing Diagram



9 References

[1] Optical Internetworking Forum OIF2001.149, “SxI-5: Electrical Characteristics for 2.488 – 3.125 Gbps parallel interfaces.”

[2] ITU, Recommendation G.707 - "Network Node Interface For The Synchronous Digital Hierarchy", 1996.

[3] ITU, G.709v2-0110 – Interfaces for the optical transport network (OTN), Oct., 2001.

[4] IEEE P802.3ae, 10 GbE. “Supplement to Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method & Physical Layer Specifications (Section 49 useful in this implementation agreement)

10 Appendix A: Sample System Configurations

10.1 Abstract

This appendix describes some examples of different system configurations where the SFI-4 phase 2 interface will be implemented.

10.2 Introduction

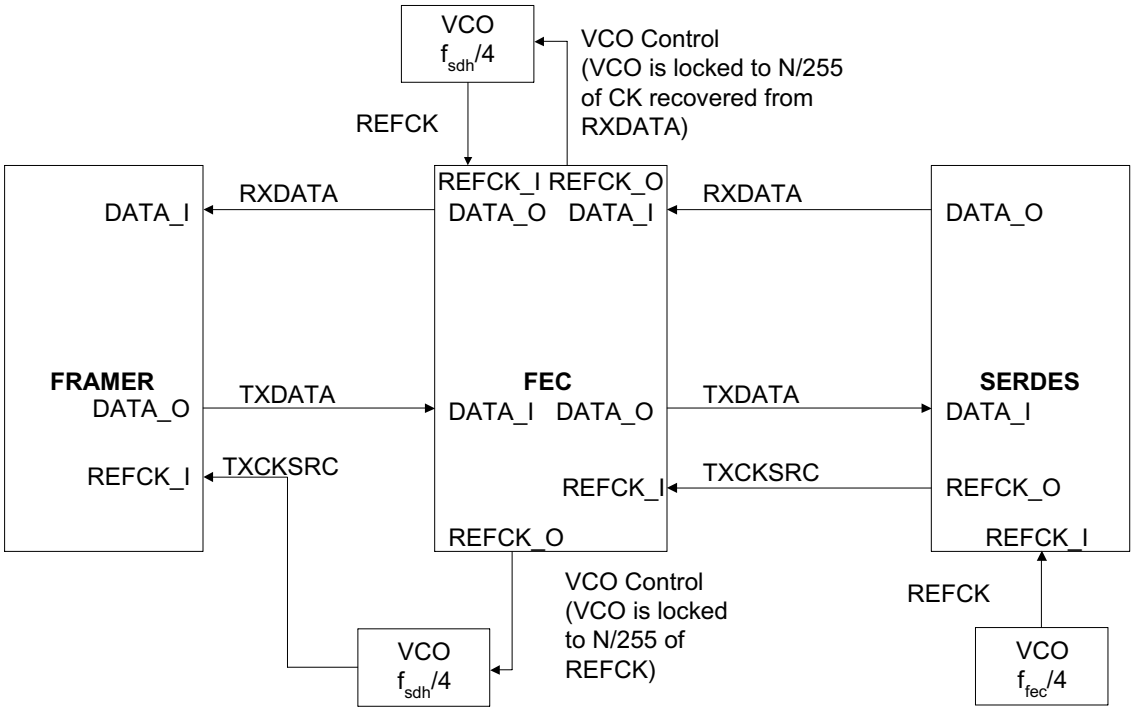
There are multiple ways to connect two sides of the SFI-4 phase 2 interface together, specifically the clocking architecture. The SFI-4 phase 2 interface reference model is meant to provide a guideline for this without restricting different configurations. This appendix provides example system configurations in which the reference model has been applied to show the implementation flexibility. This appendix does not restrict or dictate any specific configuration and is to be used solely as a reference.

10.3 Sample System Configurations

10.3.1 Reverse clock example

In this model the system sources all clocks from the SERDES. The system reference is fed into the SERDES and in the case of the receive (RX) path the reference clock is passed in the same direction as the data, where in the transmit path the reference clock is passed in the opposite direction as the data.

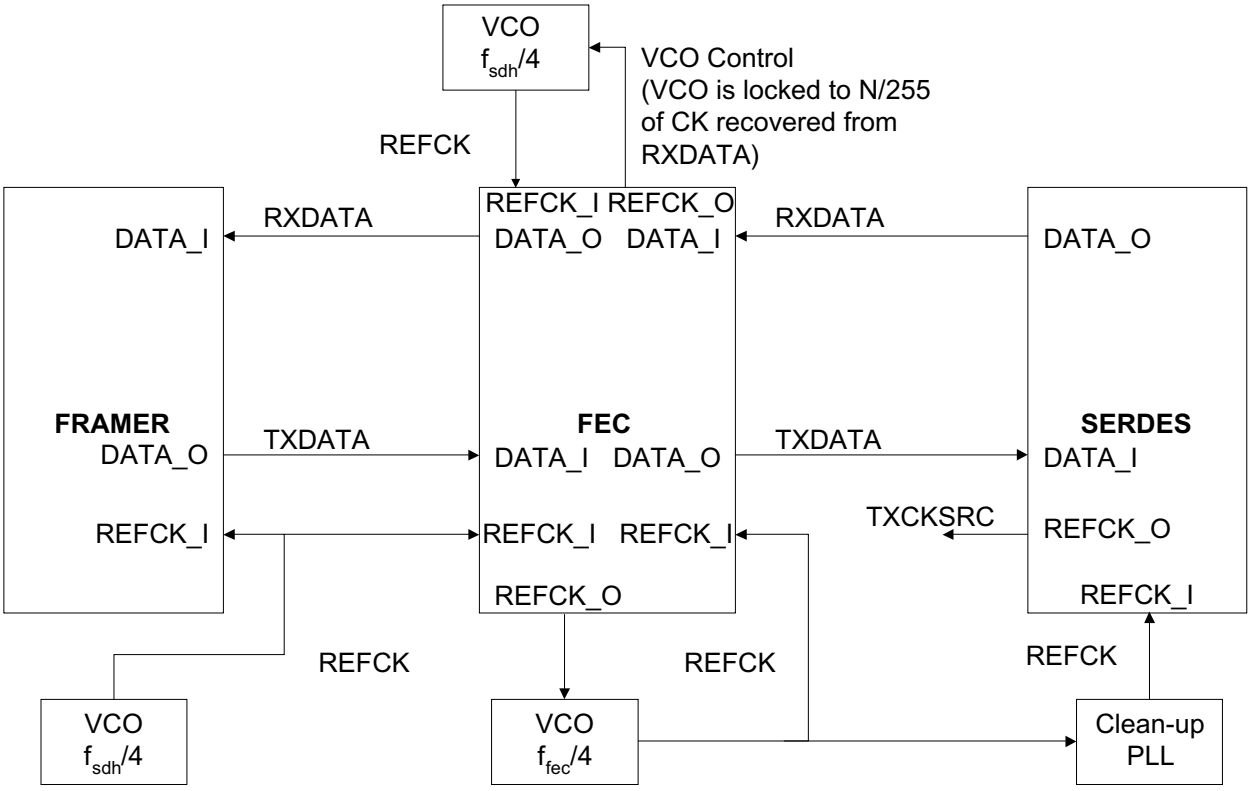
10.3.2 Figure A.1: Reverse Clock Reference Model



10.3.3 Forward clock example

In this model the system sources all clocks from the source of the data. This indicates that the reference clock is passed in the same direction as the data. In the case of the receive (RX) path the system reference is fed into the SERDES and is passed downstream to the FEC/framer, where in the transmit (TX) path the reference clock is fed into the FEC/framer and passed downstream to the SERDES.

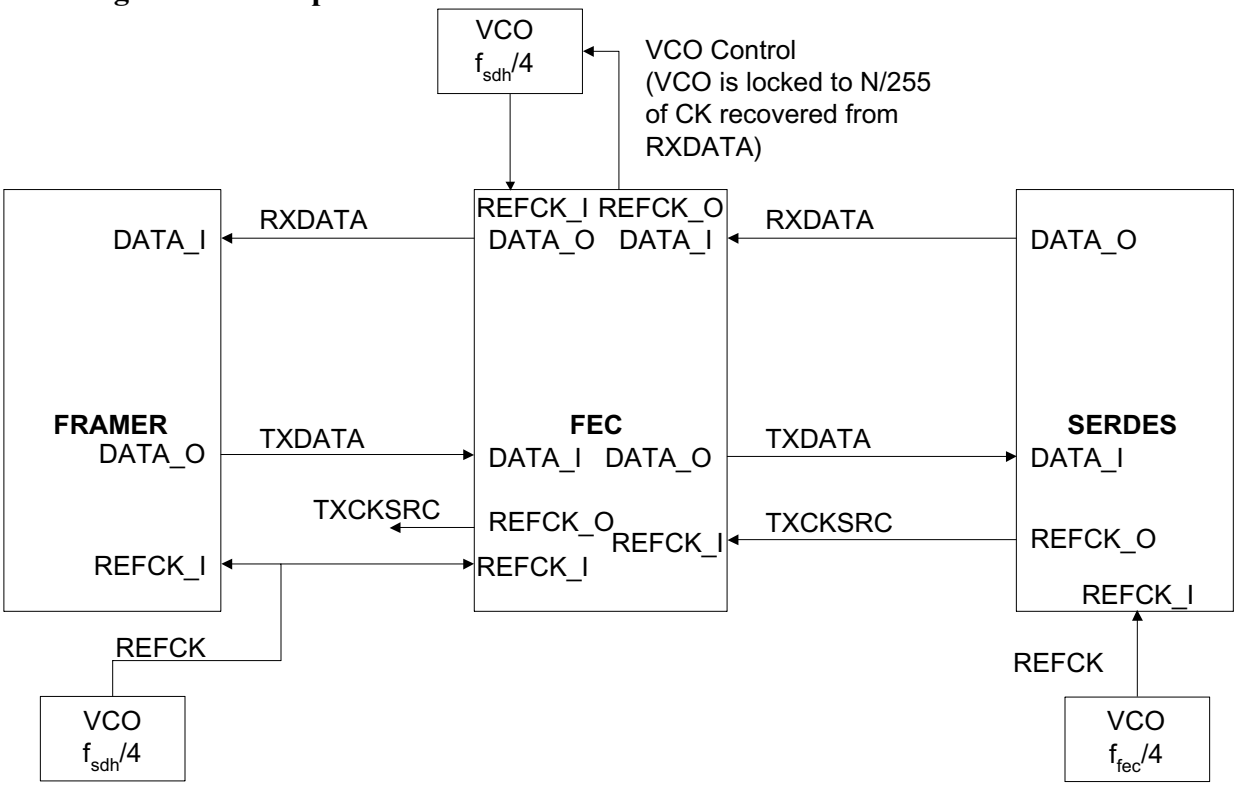
10.3.4 Figure A.2: Forward Clock Reference Model



10.3.5 Independent reference example

The transmit path have independent reference clocks at different frequencies to both the framer and the SERDES. The FEC is responsible for byte stuffing to accommodate for any differences in the two reference clocks. In this case the SFI-4 phase 2 interface is treated independently on either side of the FEC device. The receive path has the reference clock passed in the same direction as the data as described in the previous two examples.

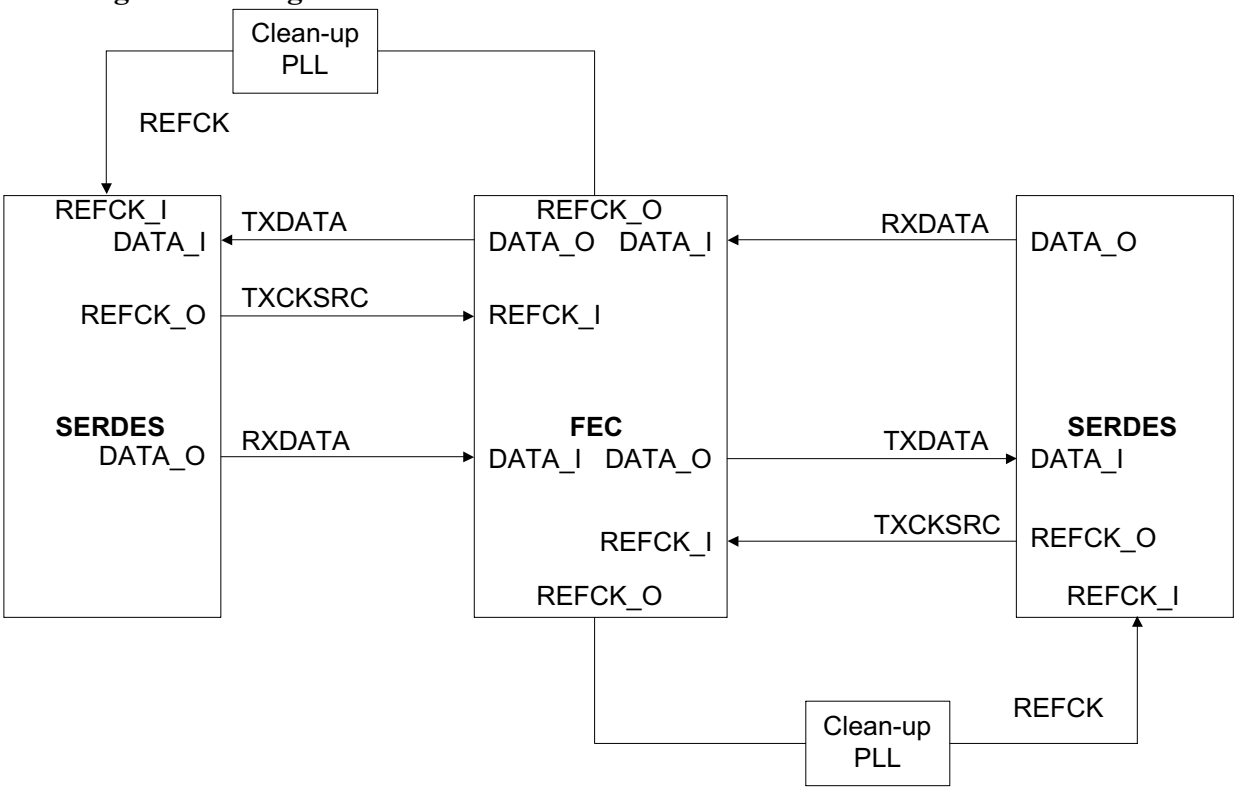
10.3.6 Figure A.3: Independent Reference Model



10.3.7 Regenerator example

The reference clock in this configuration must be derived from the receive data. A "clean-up PLL" is used to provide a clean reference clock to the output SERDES to achieve high quality transmission. The clean-up PLL can be of unity gain or include a frequency converter depending on whether data is being encoded or decoded by the FEC device.

10.3.8 Figure A.4: Regenerator Reference Model



11 Appendix: List of companies belonging to OIF when document is approved

Accelerant Networks
Accelight Networks
Actel
Acterna Eningen GmbH
ADC Telecommunications
Aeluros
Agere Systems
Agilent Technologies
Agility Communications
Alcatel
All Optical Networks, Inc.
Altamar Networks

Altera
Alvesta Corporation
AMCC
America Online
Ample Communications
Analog Devices
ANDO Corporation
Anritsu
Aralight
ASTRI
AT&T
Atrica Inc.
Avici Systems
Axiowave Networks
Bandwidth9
Bay Microsystems
Big Bear Networks
Bit Blitz Communications
Blaze Network Products
Blue Sky Research
Bookham Technology
Booz-Allen & Hamilton
Broadcom
Cable & Wireless
Cadence Design Systems
Calient Networks
Calix Networks
Caspian Networks
Celion Networks
Centellax
Centillium Communications
Ceyba
Chiaro Networks
Chunghwa Telecom Labs
Ciena Communications
Cisco Systems
Coherent Telecom
Conexant
CoreOptics
Coriolis Networks
Corrigent Systems
Cortina Systems
Corvis Corporation
Cypress Semiconductor
Data Connection
Department of Defense

DeriveIt
E2O Communications
ELEMATICS
Elisa Communications
Emcore
Equant Telecommunications SA
Equipe Communications
Ericsson
ETRI
Extreme Networks
EZChip Technologies
Fiberhome Telecommunications
Fiberspace
Finisar Corporation
Flextronics
Force 10 Networks
France Telecom
Free Electron Technology
Fujikura
Fujitsu
Furukawa Electric Technologies
Galazar Networks
General Dynamics
Glimmerglass Networks
Harris Corporation
Harting Electro-Optics GmbH
Helix AG
Hi/fn
Hitachi
Huawei Technologies
IBM Corporation
Ignis Optics
Industrial Technology Research Institute
Infineon Technologies
Infinera
Innovance Networks
Inphi
Integrated Device Technology
Intel
Internet Machines
Interoute
Intune Technologies, Ltd.
Iolon
Japan Telecom
JDS Uniphase
Jennic

Juniper Networks
KDDI R&D Laboratories
Kirana Networks
KT Corporation
Larscom
Lattice Semiconductor
LSI Logic
Lucent
Lumentis
LuxN
LYNX - Photonic Networks
Mahi Networks
Marconi Communications
MathStar
Maxim Integrated Products
MergeOptics GmbH
Meriton
Metro-OptiX
Mintera
Mitsubishi Electric Corporation
Multilink Technology Corporation
Multiplex
MultiWave Networks
Myrica Networks
Mysticom
National Semiconductor
Nayna Networks
NEC
NetTest
Network Elements
NIST
Nortel Networks
NTT Corporation
NurLogic Design
OpNext
Optical Datacom
Optillion
Optium
Optix Networks
Optobahn
OptronX
PacketLight Networks
Parama Networks
Paxonet Communications
Peta Switch Solutions
PhotonEx

Photuris, Inc.
Phyworks
Picarro
Pine Photonics Communications
PMC Sierra
Polaris Networks, Inc.
Princeton Optronics
Procket Networks
Quake Technologies
Qwest Communications
RedClover Networks
RF Micro Devices
RHK
Sandia National Laboratories
Santec Corporation
Santel Networks
Santur
SBC
Siemens
Sierra Monolithics
Silicon Access Networks
Silicon Labs
Silicon Logic Engineering
Sky Optix
Solidum
Southampton Photonics
Spirent Communications
StrataLight Communications
Stratos Lightwave
Sumitomo Electric Industries
Sun Microsystems
Sycamore Networks
TDK Semiconductor
Tektronix
Telcordia Technologies
Telecom Italia Lab
Tellabs
Tellium
Tenor Networks
TeraBurst Networks
TeraConnect
Teradiant Networks, Inc.
Texas Instruments
T-Networks, Inc.
Toshiba Corporation
Transpectrum

Transpera Networks
TriQuint Semiconductor
Tropic Networks Inc.
Tsunami Photonics
T-Systems Nova
Turin Networks
Tyco Electronics
US Conec
Velio Communications
Velocium (TRW)
Verizon
Vitesse Semiconductor
VSK Photonics
W.L. Gore & Associates
Wavecrest Corporation
Wavium AB
West Bay Semiconductor
Xanoptix
Xelerated
Xignal Technologies
Xilinx
Xindium
Xlight Photonics
Zagros Networks
Zarlink Semiconductor

End of document