PMC-Sierra, Inc.

LCS-MUX

LCS Multiplexer Application Note

Preliminary

Issue 1: September 2000

PMC-Sierra, Inc.

PMC-Sierra, Inc.

Revision History

Issue Number	Issue Date	Details Of Change
1	September 2000	Creation of document


PMC-Sierra, Inc.

PMC-Sierra, Inc.


PMC-Sierra, Inc.


PMC-Sierra, Inc.

1 SUMMARY

This document describes the required functionality of an LCS Multiplexer (LCS-MUX) that interfaces to the ETT1 Chip Set. This application note is provided only as an aid to PMC-Sierra's partners and customers.

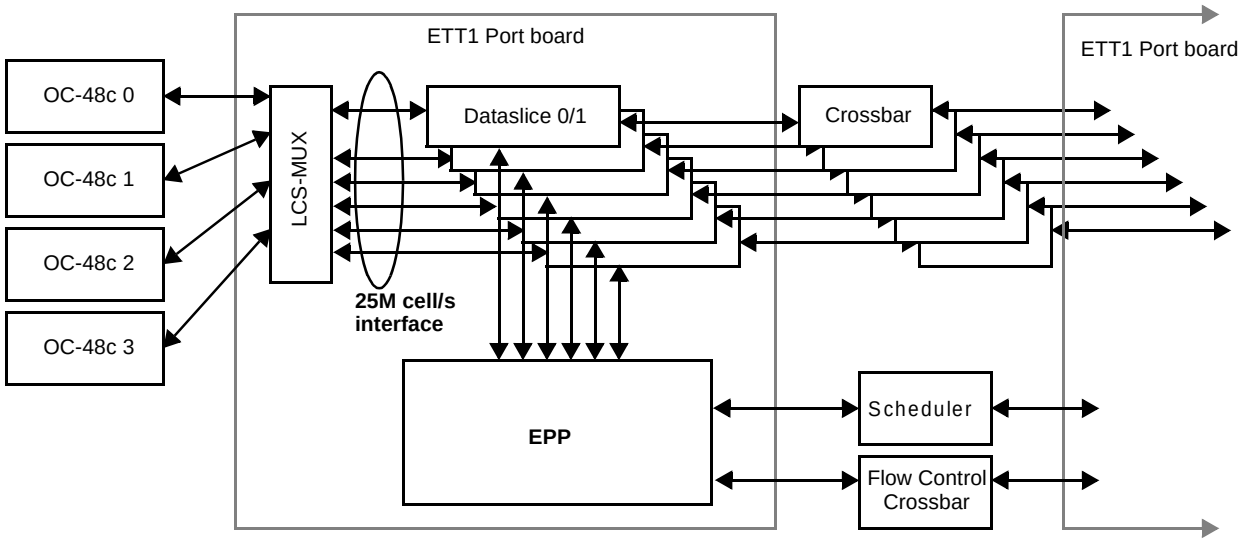
2 OVERVIEW

The LCS-MUX is an interface between four OC-48c linecards and one OC-192c ETT1 port. In the ingress direction, this device multiplexes four OC-48c streams to one OC-192c stream. In the egress direction, it de-multiplexes one OC-192c stream to four OC-48c streams. The cells in the OC-192c stream must contain two bits in the header that indicate the source/destination OC-48c sub-port. The OC-48c streams are time division multiplexed onto the OC-192c stream.

Physically, the LCS-MUX may reside on the ETT1 Port board, or on a separate multiplexer device.

Figure 1 shows how the LCS-MUX communicates with the other ETT1 devices.

Figure 1. The ETT1 Chip Set Including the EPP



A critical design goal for the LCS-MUX is low latency. The round trip between the EPP and the linecards is constrained, and so latency incurred by the LCS-MUX is time that is no longer available for processing cells on the EPP or linecard.

The remainder of this document describes possible implementations of the LCS-MUX. The reader is assumed to be familiar with the ETT1 Chip Set and the LCS Protocol - Protocol Version 2, descriptions of which are available from PMC-Sierra, Inc.

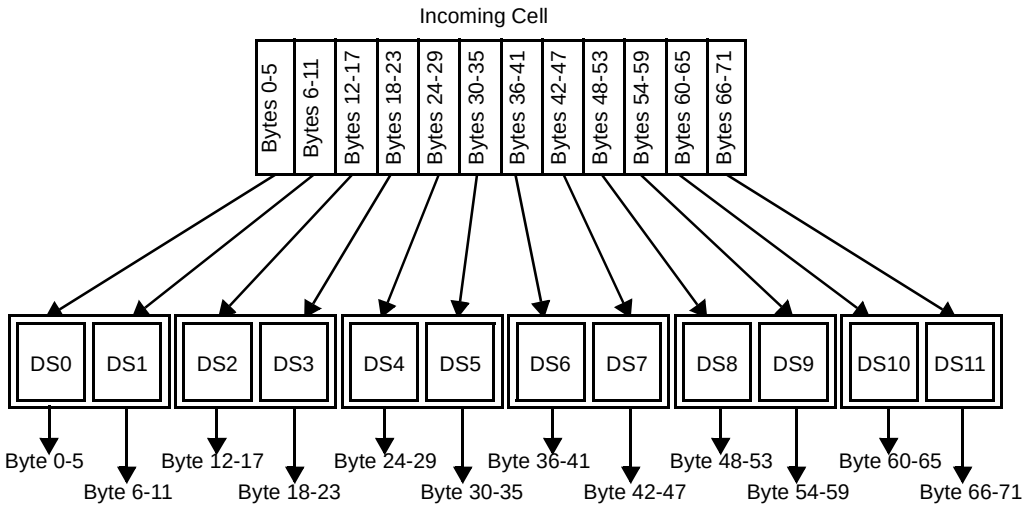
3 INTERFACES

3.1 LCS-MUX to Dataslice Interconnect

On the Dataslice side, the LCS-MUX must interface to 12 Serdes channels that, when combined, provide bandwidth in excess of 10Gb/s. Each channel carries 8b/10b [3] encoded data at 150Mhz.

Each LCS cell consists of 72 data bytes. Each cell is “sliced” across twelve Serdes interfaces, as shown in Figure 2, with each slice carrying six bytes of the cell.

Figure 2. Slicing of a Data Cell into Logical Dataslices

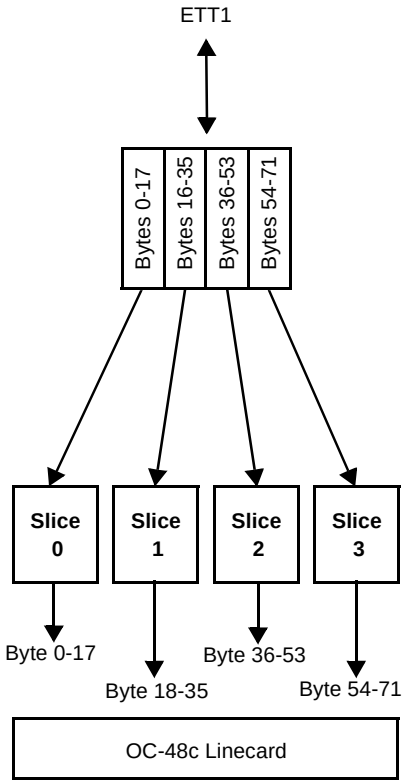


NOTE: See the ETT1 Databook, Chapter 2, Dataslice, available from PMC-Sierra, Inc., for more details on this interface.

3.2 OC-48c Linecards to LCS-MUX Interconnect

The duplex connection between the LCS-MUX and each OC-48c linecard must support at least 4.5 Gbaud of 8b/10b encoded bandwidth. In this document, it is assumed that each OC-48c channel uses the four channels of a quad Serdes device. This will result in the data cell being sliced as shown in Figure 3.

Figure 3. Slicing of Data Cell into the Four Serdes Interfaces



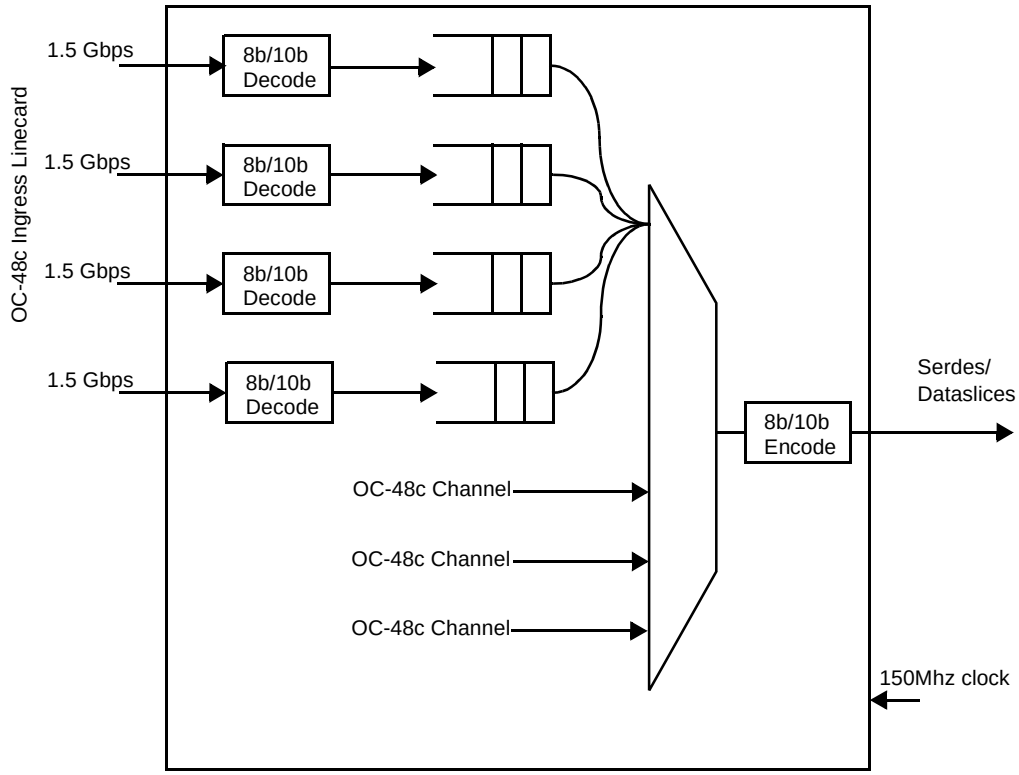
Therefore, the connection between the LCS-MUX and the four OC-48c linecards provides a total of 16 parallel Serdes interfaces, four per OC-48c linecard.

The OC-48c quad Serdes *must* be clocked at a minimum of 112.5MHz in order to sustain the desired cell rate. It is suggested that a faster clock be used in order to support a framing protocol on each channel that would aid in synchronization as well as error recovery. See Appendix A for more information.

4 INGRESS FLOW

First, examine the flow from one OC-48c ingress linecard. In the following, the use of a 150MHz clock for the OC-48c Serdes devices is assumed. Figure 4 describes the OC-48c ingress FIFO. The receive clock from each OC-48c ingress linecard is determined independently for each of the four Serdes interfaces. All four of the individual interfaces are assumed to have established link synchronization before the OC-48c channel is considered ready.

Figure 4. Ingress FIFO for Cells Arriving from a Single OC-48c Linecard

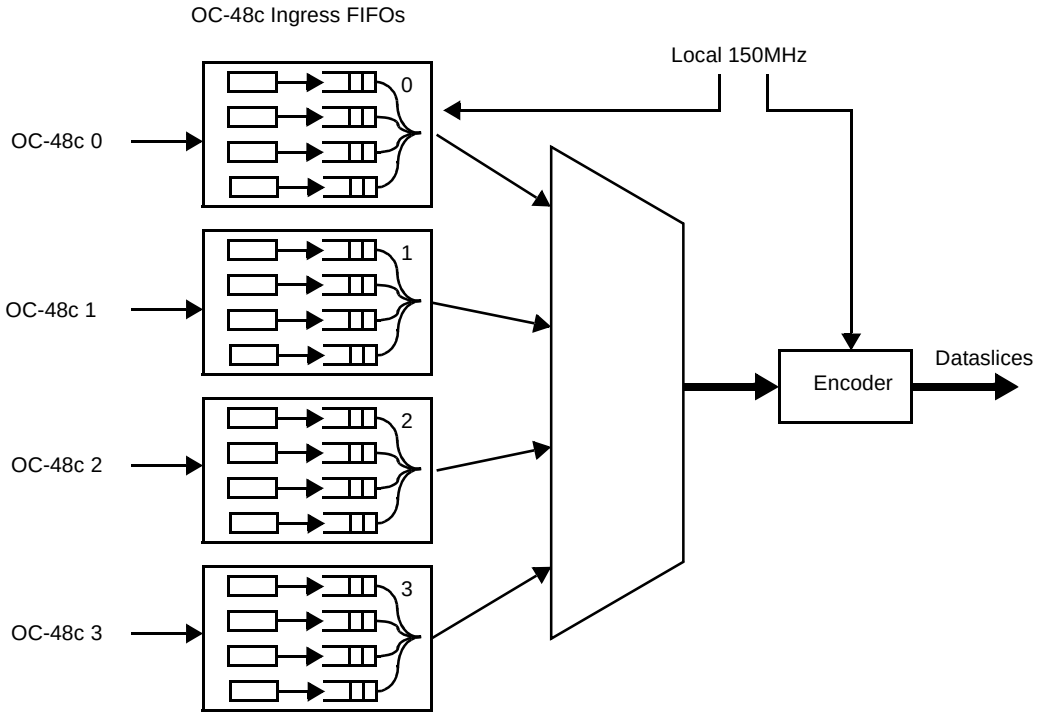


At this point, it is necessary to define some terminology. Each Serdes link consists of a stream of 10 bit words; each word is either a control word or a data word. As shown in Figure 2, each Dataslice only sees six words of each cell, and there are no start-of-frame delimiters. Each Dataslice considers a sequence of six words to be an *idle frame* if all six words are control words (the Dataslice discards idle frames). On the other hand, a frame consisting of only six data words is considered to be *non-idle*. A problem arises if link errors cause a data word to be transformed into a control word, or vice versa. Some decision logic must determine if the incoming frame is idle or non-idle. If an idle frame is translated to a non-idle frame (due to bit errors) or vice-versa, then the entire OC-192 connection (all four OC-48c channels) must be restarted.

After the incoming stream has been decoded, each Serdes interface writes the decoded data into its own FIFO. All four streams in the same OC-48c channel will have identical write frequencies (assuming a common Tx clock at the linecard) but streams will have arbitrary phase offsets relative to each other. An 18-byte frame is written into a receive FIFO only if it is a non-idle frame; the LCS-MUX must make this decision on a frame-by-frame basis. A local 150Mhz clock is used to realign the four incoming streams from the same OC-48c ingress linecard. The four ingress FIFOs associated with an OC-48c channel are considered non-empty only if all four FIFOs are non-empty. The LCS-MUX must also check for the unlikely event of an 18-word idle frame being corrupted to look like an 18-word non-idle frame, and vice versa.

PMC-Sierra, Inc.

Figure 5. Ingress Flow Showing All Four OC-48c Channels and Input FIFOs.



In every OC-48c cell time (~160ns), the multiplexer rotates among all of the four OC-48c channels. If a channel is empty, then an idle frame is sent to the Dataslices in that cell time. If the channel is not empty, then a cell is dequeued and sent to the Dataslices. If an OC-48c linecard is not operating, then the corresponding channel is always empty and an idle frame will be sent to the Dataslices.

The input FIFO queues for each OC-48c Serdes must be deeper than 18 words. If the read mechanism has just “missed” reading an OC-48c cell, then that cell may have to wait for an additional four cells at OC-192c rate (160ns assuming a 40ns OC-192c cell time) before the read mechanism will consider it again.

Before the cell is sent to the Dataslices, it must first be re-encoded.[1]

NOTE: The LCS-MUX must insert two bits in the LCS header to specify which OC-48c linecard was the source of the incoming cell. See the LCS Protocol Specification - Protocol Version 2, available from PMC-Sierra, Inc., for more details.

The total ingress latency is determined by the input FIFO delay (18 clocks) and the wait period for the next TDM slot for this OC-48c channel (24 clocks). Allowing for internal retiming delays, encoding/decoding and

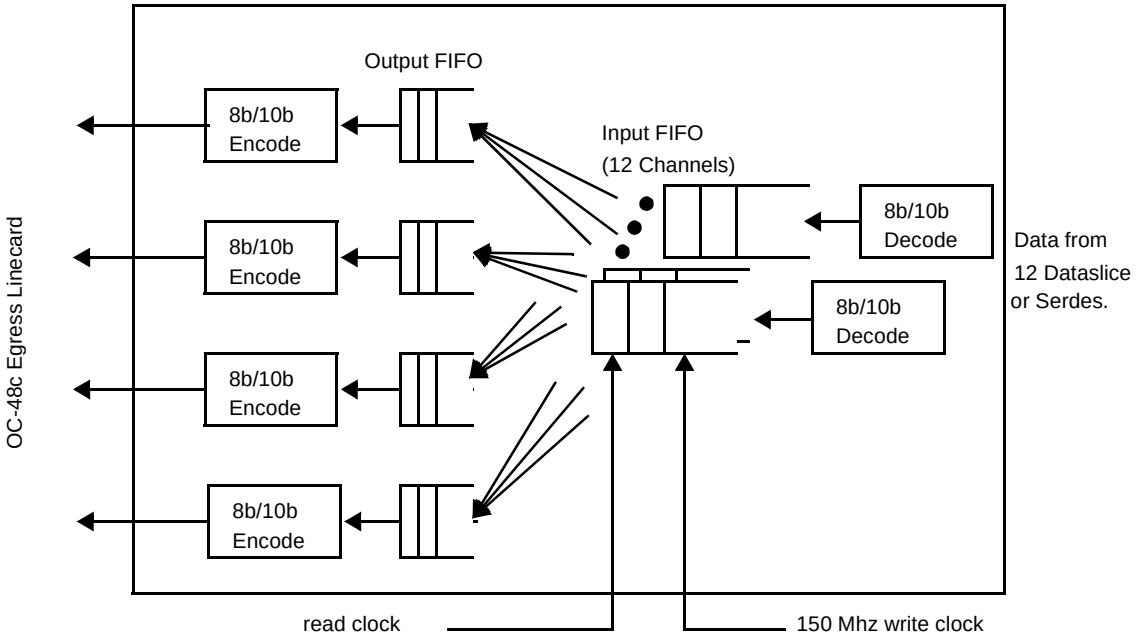
1. Any 8b/10b encoding scheme which supports control words can be used. The Dataslice uses a 1024 entry lookup table to convert incoming 10-bit words into a byte. The Dataslice requires a minimum of 2 control words and 256 data words.

the output FIFO, then the ingress latency is about 48 cycles of 150MHz. If the OC-48c channels do not use a 150MHz clock then this latency is greater.

5 EGRESS FLOW

Now consider the egress flow, as shown in Figure 6.

Figure 6. Egress FIFO for Cells Destined to a Single OC-48c Linecard



There is a single logical input FIFO consisting of 12 physical FIFOs, one for each Dataslice stream. The twelve incoming data streams are at the same frequency but have arbitrary phase with respect to each other. Also, a cell may be skewed across the 12 slices by up to one OC-192c cell time. For example, bytes 12 to 17 (slice 2) could arrive at the LCS-MUX six clocks before bytes 0 to 5 (slice 0) of the *same* cell.

Whenever the input FIFO is non-empty, a 72-byte frame is copied to the four output FIFOs that form one OC-48c output port. The six bytes from each of three input slices are copied to each output FIFO (18 bytes). The four output FIFOs are needed as rate matching FIFOs. Whenever an output FIFO is non-empty it can read the byte at the head of the FIFO and forward it to an 8b/10b encoder. If the FIFO goes empty, then control words are transmitted.

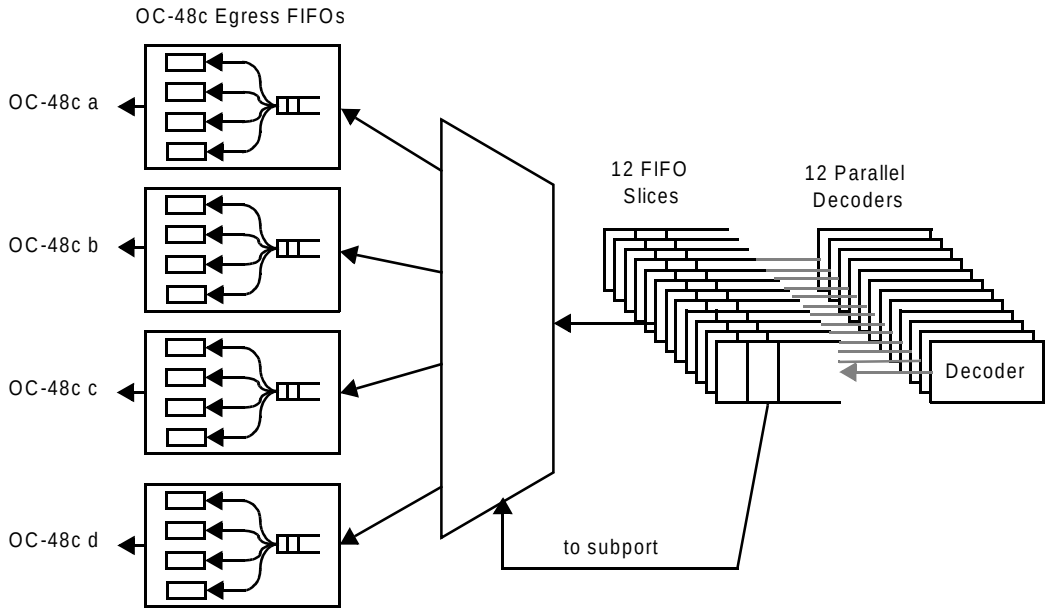
The Dataslice Serdes transmit interfaces are driven by the 150 Mhz clock in the ETT1 system. Since there can be skew between the twelve logical Dataslices, the LCS-MUX considers each of the twelve received streams as being asynchronous. Thus, all the 12 FIFOs must be synchronized within the LCS-MUX before cells can be received from the Dataslices. Each of the Dataslice streams must go through the decoding process. The code can be any 8b/10b code that supports control words as well as data words.

PMC-Sierra, Inc.

Only non-idle frames must be stored in the 12 input FIFOs. Each of the 12 decoders should check for a valid stream of words: either 6 data words or 6 control words. A combination of data words and control words in the same cell would indicate that the cell has been corrupted. The entire cell should be discarded across all 12 channels.

When all 12 FIFOs are non-empty, then the 72-byte frame can be forwarded. The subport field in the LCS header of this frame determines the destination OC-48c egress linecard. The de-multiplexer sends the frame to the appropriate OC-48c egress FIFO. The LCS-MUX must also zero the two bits that indicate the destination subport.

Figure 7. Egress Flow



The ETT1 EPP guarantees that each OC-48c egress FIFO will receive at most one cell every OC-48c cell time. The EPP also implements the 1-in-N protocol so that the OC-48c egress FIFOs never overflow even if the LCS-MUX uses a local clock that is fractionally slower than that used by the Dataslices.

Each OC-48c stream must be encoded before being placed on the output line. Once again, the encoding scheme is dictated by the Serdes components.

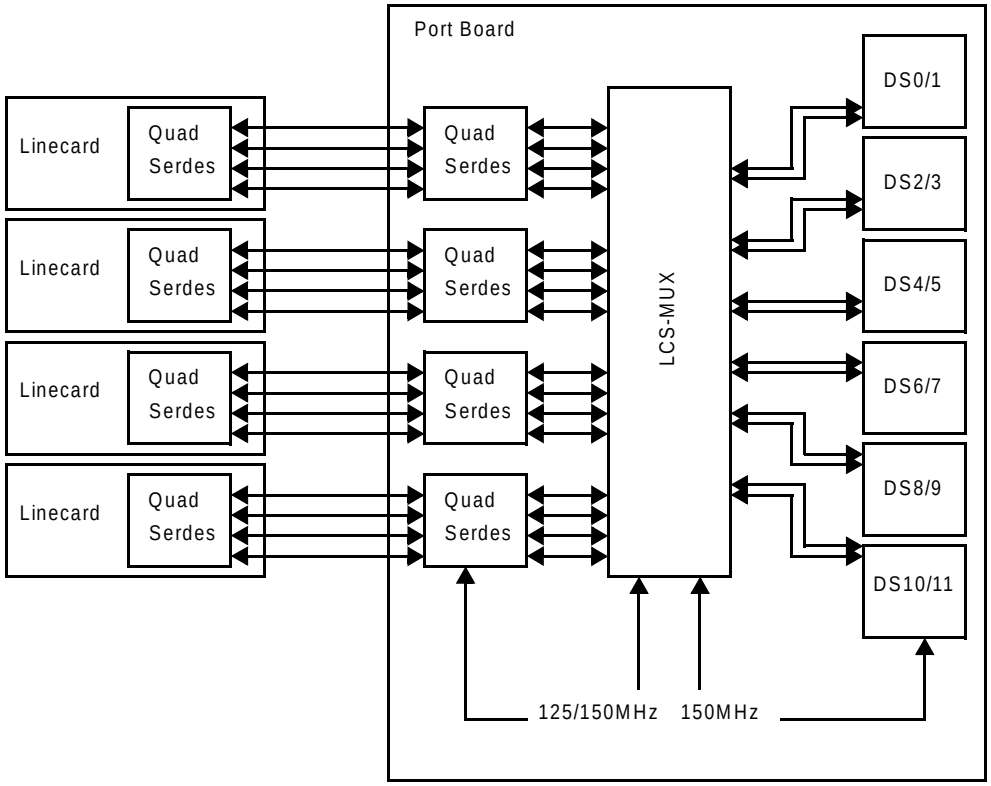
The latency for the egress direction is the time for a cell to arrive in the input FIFO (up to 9 clocks) plus the time to go through the output FIFO and 8b/10b encoder (4 clocks). A couple of extra clocks should be allowed for internal retiming of data, giving a total latency of about 15 clocks at 150MHz.

The total round-trip latency for a cell to traverse the LCS-MUX in both directions is then about 63 clocks at 150MHz or about 420ns.

6 LCS-MUX ON THE PORT BOARD

Figure 8 illustrates the component sequence used if the LCS-MUX is co-located with the Dataslices on the ETT1 Port board.

Figure 8. LCS-MUX Located On the ETT1 Port Board.



When the LCS-MUX is located on the TT1 Port board, it is connected to the Dataslices by parallel busses without the need for Serdes or fiber optical components.

The benefits are two-fold:

- The LCS-MUX/Dataslice transfers are synchronous
- The LCS-MUX/Dataslice transfers are reliable without bit errors

Synchronous transfers allow for the latency in the egress flow from the Dataslices to the LCS-MUX to be reduced as there is no longer a need for clock domain decoupling FIFOs to be used on the OC-192c receiver interface of the LCS-MUX.

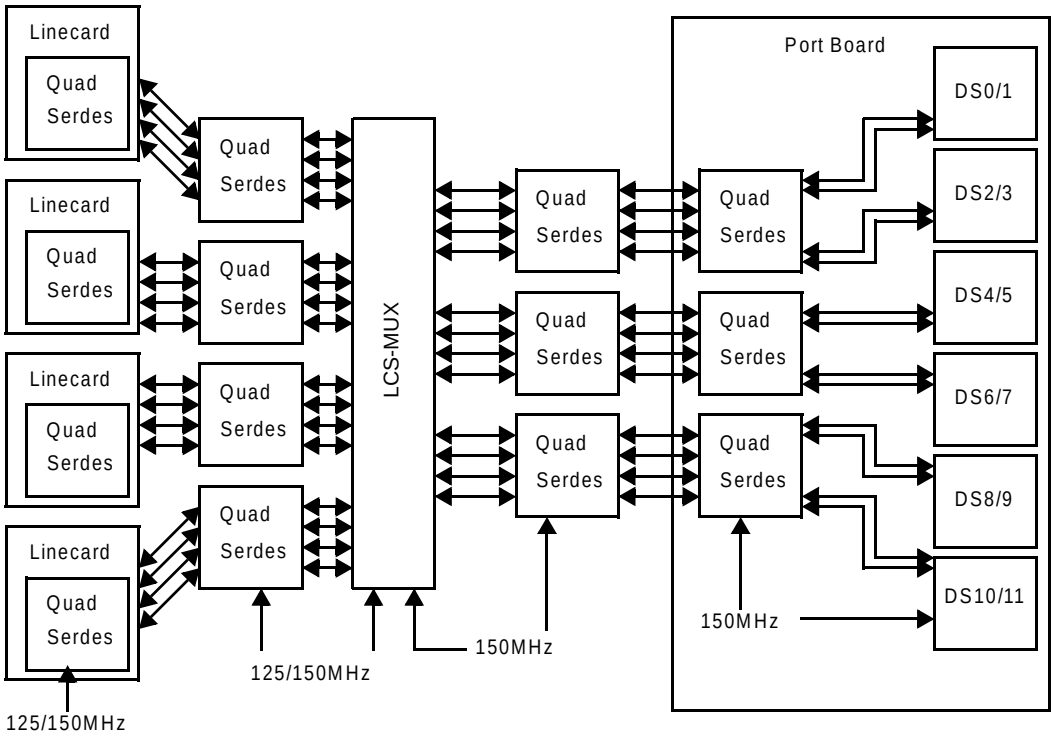
The absence of bit errors in the egress flow from the Dataslices to the LCS-MUX enables the OC-192c receiver interface of the LCS-MUX to decide immediately after receiving the first byte of a cell whether or not the cell is idle without the need for a majority vote mechanism as used on the Dataslice [1]. It also

implies that it is not possible for cells to be corrupted to become idle when they were in fact non-idle and vice-versa. As described in [1], in extreme situations, this problem can cause the 12 fiber optical links to the Dataslices to fall out of alignment and every subsequent data cell to be misaligned. Appendix A of this document describes an OC-48c linecard framing protocol that protects the OC-48c interfaces of the LCS-MUX from the out-of-alignment problem.

7 LCS-MUX ON A SEPARATE DEVICE

Figure 9 illustrates the components used if the LCS-MUX is located on a separate device from the ETT1 Port board. The left hand pairs of Serdes are not needed if the four linecards are co-located with the LCS-MUX.

Figure 9. LCS-MUX Located On a Separate Device



When the LCS-MUX is not co-located with the Dataslices, the LCS-MUX must deal with the issues described in Section 6, “LCS-MUX on the Port Board”, regarding latency and reliability.

8 REFERENCES

[1] ETT1 Databook, Chapter 2, Dataslice, available from PMC-Sierra, Inc.
[2] LCS Protocol Specification - Protocol Version 2, available from PMC-Sierra, Inc.
[3] “A DC-Balanced, Partitioned-Block, 8B/10B Transmission Code”, A.X. Widmer and P.A. Franaszek, IBM J. Res. Develop., Vol. 27 No. 5, September 1983.

APPENDIX A: SUGGESTED OC-48C LINECARDS TO LCS-MUX FRAMING

The following is one possible framing sequence, based on the Gigabit Ethernet 16-bit delimiter set shown in Table 1.

Table 1. Proposed 16-bit Delimiter Set

Description	Code	8b/10b Sequence
Start-of-Frame	-SOF-	-K28.5+ +D5.0-
End-of-Frame (neg)	-EOF1-	-K28.5+ +D5.4-
End-of-Frame (pos)	+EOF2-	+K28.5- -D5.1-
Idle	-IDLE-	-K28.5+ +D2.1-

The + (-) indicates positive (negative) running disparity. Note that the running polarity is negative after all 16-bit delimiters.

The benefit of running the OC-48c linecard quad Serdes at 150MHz lies in the fact that only 18 of the 24 clock cycles are used to communicate Dxx.x 8b/10b data characters (i.e. cell data). The remaining 6 cycles are available to provide explicit start-of-frame alignment (SOF) and end-of-frame alignment (EOF1, EOF2). On average, there are two unused cycles that can be filled with the idle sequence (IDLE). Whether there are zero, two, or more unused cycles between frames is determined by the difference in actual operating frequency of the 150MHz clocks between the OC-192 and OC-48c interfaces to the LCS-MUX.

If the OC-48c link uses a 125MHz clock, then fewer cycles are available for framing and an alternative framing sequence are necessary.

NOTE: These sequences are suggestions only.


PMC-Sierra, Inc.


PMC-Sierra, Inc.


PMC-Sierra, Inc.


PMC-Sierra, Inc.


PMC-Sierra, Inc.


PMC-Sierra, Inc.


PMC-Sierra, Inc.